# DISTRIBUTION OF FREQUENCIES OF DIGITS VIA MULTIFRACTAL ANALYSIS

L. BARREIRA, B. SAUSSOL, AND J. SCHMELING

Abstract. We study the Hausdorff dimension of a large class of sets in the real line defined in terms of the distribution of frequencies of digits for the representation in some integer base. In particular, our results unify and extend classical work of Borel, Besicovitch, Eggleston, and Billingsley in several directions. Our methods are based on recent results concerning the multifractal analysis of dynamical systems and often allow us to obtain explicit expressions for the Hausdorff dimension. This work is still another illustration of the role that the theory of dynamical systems can play in number theory.

## 1. Introduction

Instead of formulating general statements at this point, we want to discuss explicit examples (although already nontrivial), which illustrate well the nature of our work.

Given an integer $m > 1$, for each number $x \in [0,1]$ we shall denote by $0.x_1x_2\cdots$ a base-$m$ representation of $x$. It is easy to see that this representation is unique except for countably many points. We remark that since countable sets have zero Hausdorff dimension, the nonuniqueness of the representation does not interfere with our study.

For each $k \in \{0, \dots, m-1\}$, $x \in [0,1]$, and $n \in \mathbb{N}$ set

$$\tau_k(x,n) = \text{card}\{i \in \{1, \dots, n\} : x_i = k\}.$$

Whenever there exists the limit

$$\tau_k(x) = \lim_{n\to\infty} \frac{\tau_k(x,n)}{n} \tag{1}$$

it is called the *frequency* of the number $k$ in the base-$m$ representation of $x$. When we write the symbol $\tau_k(x)$ we are already assuming the existence of the limit in (1).

A classical result of Borel [6] says that for Lebesgue-almost every $x \in [0,1]$ we have $\tau_k(x) = 1/m$ for every $k$. Furthermore, for $m = 2$, Hardy and Littlewood [10] showed that for Lebesgue-almost every $x \in [0,1]$, $k = 0$, 1, and all sufficiently large $n$,

$$\left|\tau_k(x,n) - \frac{1}{2}\right| < \sqrt{\frac{\log n}{n}}.$$

---

2000 *Mathematics Subject Classification.* Primary: 11K55, 37C45.

*Key words and phrases.* frequencies of digits, Hausdorff dimension, multifractal analysis.

L. Barreira was partially supported by the Center for Mathematical Analysis, Geometry, and Dynamical Systems, through FCT's Funding Program.

In particular, Lebesgue-almost all numbers are normal in every integer base. This remarkable result (even though it is today straightforward to prove in a variety of ways) does not mean that all numbers are normal. In fact we shall see that quite the opposite happens.

Consider now the set

$$F_m(\alpha_0, \dots, \alpha_{m-1}) = \{x \in [0,1] : \tau_k(x) = \alpha_k \text{ for } k = 0, \dots, m-1\}, \quad (2)$$

whenever $\alpha_0 + \cdots + \alpha_{m-1} = 1$ with $\alpha_i \in [0,1]$ for each $i$. It is composed of the numbers in $[0,1]$ having a ratio $\alpha_k$ of digits equal to $k$ in its base-$m$ representation for each $k$. A precursor result concerning the size of these sets from the point of view of dimension theory is due to Besicovitch [4]. For $m = 2$, he showed that if $\alpha \in (0, \frac{1}{2})$ then

$$\dim_H \left\{ x \in [0,1] : \limsup_{n\to\infty} \frac{\tau_1(x,n)}{n} \le \alpha \right\} = -\frac{\alpha \log \alpha + (1-\alpha)\log(1-\alpha)}{\log 2},$$

where $\dim_H Z$ denotes the Hausdorff dimension of the set $Z$. More detailed information was later obtained by Eggleston [8], who showed that

$$\dim_H F_m(\alpha_0, \dots, \alpha_{m-1}) = -\frac{\sum_{k=0}^{m-1} \alpha_k \log \alpha_k}{\log m}. \quad (3)$$

An immediate consequence is that if $\alpha_i \in (0,1)$ for some $i$, then the set $F_m(\alpha_0, \dots, \alpha_{m-1})$ is nonempty (and thus dense in $[0,1]$), with uncountable many points and even positive Hausdorff dimension. The work of Eggleston was further generalized by Billingsley (see his book [5] for details and references; see also Section 5.4).

We now consider sets of points for which the limit in (1) does not exist. For each $k \in \{0, \dots, m-1\}$ we define the set

$$M_k = \left\{ x \in [0,1] : \liminf_{n\to\infty} \frac{\tau_k(x,n)}{n} < \limsup_{n\to\infty} \frac{\tau_k(x,n)}{n} \right\}. \quad (4)$$

Notice that $M_k$ has zero Lebesgue measure, due to the above-mentioned result of Borel. Clearly,

$$[0,1] = \bigcup_{(\alpha_0, \dots, \alpha_{m-1}) \in L_m} F_m(\alpha_0, \dots, \alpha_{m-1}) \cup \bigcup_{k=0}^{m-1} M_k, \quad (5)$$

where

$$L_m = \{(\alpha_0, \dots, \alpha_{m-1}) \in [0,1]^m : \alpha_0 + \cdots + \alpha_{m-1} = 1\}. \quad (6)$$

In this paper we provide further nontrivial information about the decomposition in (5). In particular we prove the following statement.

**Theorem 1.** *For each $k \in \{0, \dots, m-1\}$ the set $M_k$ contains a dense $G_\delta$ set in $[0,1]$, and*

$$\dim_H \bigcap_{k=0}^{m-1} M_k = 1. \quad (7)$$

Theorem 1 implies that $\bigcup_{k=0}^{m-1} M_k$ has Hausdorff dimension equal to 1, and thus, from the point of view of dimension theory, it is as large as the interval $[0,1]$. On the other hand, the union $\bigcup_{k=0}^{m-1} M_k$ has not only zero

Lebesgue measure but also zero measure with respect to any measure which is invariant under the map $x \mapsto mx \pmod 1$ (see Section 2 for details), and thus the set $\bigcup_{k=0}^{m-1} M_k$ is rather small from the point of view of measure theory.

We can also consider sets more complicated than those in (2), and in particular sets defined by linear or even nonlinear relations among the numbers $\tau_k(x)$. These generalizations are described in the remaining sections. Here we shall give a simple but nontrivial example. Let $m = 4$, and define the set

$$F = \{x \in [0, 1] : \tau_1(x) = 5\tau_0(x)\}.$$

This is the set of numbers in $[0, 1]$ such that its base-4 representation has a ratio of ones which is five times the ratio of zeros. The ratios of twos and threes is arbitrary. Again, the nonuniqueness of the representation is not an issue in the study of Hausdorff dimension. It is easy to see that

$$F \supset \bigcup_{\alpha\in[0,1/6]} \bigcup_{\beta\in[0,1-6\alpha]} F_4(\alpha, 5\alpha, \beta, 1 - 6\alpha - \beta). \tag{8}$$

We emphasize that the inclusion is proper since $F$ contains points for which $\tau_2(x)$ and $\tau_3(x)$ are not well-defined (this fact substantially complicates the problem of computing $\dim_H F$ since a priori this dimension may not be entirely carried by the union in (8)). We shall prove that

$$\dim_H F = \frac{\log(2 + 6/5^{5/6})}{\log 3} \approx 0.91779 \cdots \quad \text{(in base 10)}. \tag{9}$$

We first remark that it is easy to show that this last number is a lower bound for $\dim_H F$. Namely, it follows from (3) and (8) that

$$\begin{aligned} \dim_H F &\geq \max_{\alpha\in[0,1/6]} \max_{\beta\in[0,1-6\alpha]} \dim_H F_4(\alpha, 5\alpha, \beta, 1 - 6\alpha - \beta) \\ &= \max_{\alpha\in[0,1/6]} -\frac{\alpha \log \alpha + 5\alpha \log(5\alpha) + (1 - 6\alpha) \log \frac{1-6\alpha}{2}}{\log 4} \end{aligned} \tag{10}$$

The maximum is attained at $\alpha = 1/(2 \cdot 5^{5/6} + 6)$ and it is a straightforward computation to show that it is equal to the constant in (9). This establishes the lower bound.

The corresponding upper bound is more delicate, since the union in (8) is composed of an uncountable number of nonempty pairwise disjoint sets. Moreover the inclusion in (8) is proper. This is where the theory of multifractal analysis comes into the play. Namely, using what is called a *conditional variational principle* we can show that although the inclusion is proper, the Hausdorff dimension of $F$ is carried by exactly *one* set in the union in (8). We now formulate a particular case of the conditional variational principle that is sufficient for the purpose of the present example.

**Theorem 2.** *For each $k \neq \ell$ and $\beta \geq 0$ we have*

$$\begin{aligned} &\dim_H\{x \in [0, 1] : \tau_k(x) = \beta\tau_\ell(x)\} \\ &\quad = \max\left\{ -\frac{\sum_{j=0}^{m-1} \alpha_j \log \alpha_j}{\log m} : (\alpha_0, \ldots, \alpha_{m-1}) \in L_m \text{ and } \alpha_k = \beta\alpha_\ell \right\} \\ &\quad = \frac{\log(m - 2 + (\beta + 1)/\beta^{\beta/(\beta+1)})}{\log m}. \end{aligned}$$

An easy consequence of Theorem 2 and (3) is that for each $k \neq \ell$ and $\beta \geq 0$, there exists $(\alpha_0, \dots, \alpha_{m-1}) \in L_m$ such that

$$F_m(\alpha_0, \dots, \alpha_{m-1}) \subset \{x \in [0,1] : \tau_k(x) = \beta\tau_\ell(x)\}$$

and

$$\dim_H F_m(\alpha_0, \dots, \alpha_{m-1}) = \dim_H\{x \in [0,1] : \tau_k(x) = \beta\tau_\ell(x)\}.$$

In particular, setting $m = 4$, $k = 1$, $\ell = 0$, and $\beta = 5$, this observation implies that the inequality in (10) is in fact an identity, thus establishing the claim in (9).

The main advantage of our approach is that the classes of problems that we consider (which a priori could seem of very different nature) can be treated in a unified manner, as an application of the theory of multifractal analysis. Another advantage of this approach is that the value of the Hausdorff dimension does not need to be guessed a priori. In some works this a priori guess is crucial in order to construct auxiliary measures sitting on the set. These measures are then used to establish, rigorously, the value of the Hausdorff dimension.

The statements formulated above are consequences of more general statements established in this paper. As mentioned before, our results are based on recent work concerning the multifractal analysis of dynamical systems. On the other hand, we emphasize that the paper is self-contained. In particular, we assume no former knowledge of multifractal analysis.

The description of the state-of-the-art of multifractal analysis in 1997 can be found in the book by Pesin [12]. For some later developments (and in particular those concerning the results that we use here) the reader should look at the papers in the bibliography. In particular our approach requires a multidimensional version of the classical multifractal analysis. Such a version was first introduced by the authors in [2].

The structure of the paper is as follows. The necessary notions from ergodic theory are briefly recalled in Section 2. In Section 3 we establish Theorem 1 and several related results. In particular, we show that the set $M_k$ can be further decomposed in a natural way into an uncountable union of pairwise disjoint sets with positive Hausdorff dimension. The above-mentioned conditional variational principle (of which Theorem 2 is a particular case) is described in Section 4. Further applications to number-theoretical problems are given in Section 5.

*Acknowledgment.* We would like to thank the referee for a number of good suggestions that helped us improve the exposition. We also would like to thank Christian Wolf for his comments.

## 2. Basic notions

Fix a positive integer $m$ and consider the map $g_m \colon [0,1] \to [0,1]$ defined by $g_m x = mx \pmod 1$. Observe that if $0.x_1x_2\cdots$ is a base-$m$ representation of $x \in [0,1]$, then $g_m x = 0.x_2x_3\cdots$.

Let $\mu$ be a $g_m$-invariant probability measure on $[0,1]$, i.e., a probability measure such that $\mu({g_m}^{-1}A) = \mu(A)$ for every measurable set $A \subset [0,1]$.

The *entropy* of $g_m$ with respect to $\mu$ is defined by

$$h_\mu(g_m) = \inf_{n\geq 1} -\frac{1}{n} \sum_{i_1\cdots i_n} \mu(I_{i_1\cdots i_n}) \log \mu(I_{i_1\cdots i_n}), \tag{11}$$

where

$$I_{i_1\cdots i_n} = [0.i_1 \cdots i_n,\, 0.i_1 \cdots i_n + m^{-n}). \tag{12}$$

We need the following basic result of ergodic theory (see for example [11, Chapters 1 and 2] for details).

**Proposition 3.** *We have*

$$\begin{aligned} h_\mu(g_m) &= \inf_{n\geq 1} - \sum_{i_1\cdots i_n} \mu(I_{i_1\cdots i_n}) \log \frac{\mu(I_{i_1\cdots i_n})}{\mu(I_{i_2\cdots i_n})} \\ &= \int_0^1 \lim_{n\to\infty} \frac{\log \mu(I_{x_1\cdots x_n})}{-n}\, d\mu(x), \end{aligned}$$

*where the limit exists for $\mu$-almost every $x = 0.x_1x_2\cdots \in [0,1]$.*

It follows from (11) that

$$h_\mu(g_m) \leq - \sum_{k=0}^{m-1} \mu(I_k) \log \mu(I_k). \tag{13}$$

Given $(\alpha_0, \dots, \alpha_{m-1}) \in L_m$ (see (6)), consider the $g_m$-invariant probability measure $\mu = \mu_{\alpha_0,\dots,\alpha_{m-1}}$ such that $\mu(I_{i_1\cdots i_n}) = \alpha_{i_1} \cdots \alpha_{i_n}$ for each $I_{i_1\cdots i_n}$. It is called a Bernoulli measure. It follows easily from the first formula for $h_\mu(g_m)$ in Proposition 3 that

$$h_{\mu_{\alpha_0,\dots,\alpha_{m-1}}}(g_m) = - \sum_{k=0}^{m-1} \alpha_k \log \alpha_k. \tag{14}$$

Let now $\mathcal{M}$ be the family of $g_m$-invariant probability measures on $[0,1]$. By (13) and (14) we obtain

$$\max\{h_\mu(g_m) : \mu \in \mathcal{M} \text{ and } \mu(I_k) = \alpha_k \text{ for each } k\} = h_{\mu_{\alpha_0,\dots,\alpha_{m-1}}}(g_m). \tag{15}$$

Recall that a probability measure $\mu$ is *ergodic* if any $g_m$-invariant set has either zero or full $\mu$-measure (a set $A$ is $g_m$-invariant if ${g_m}^{-1}A = A$). For example, Bernoulli measures are ergodic. We define the *Hausdorff dimension* $\dim_H \mu$ of the measure $\mu$ by

$$\dim_H \mu = \inf\{\dim_H Z \colon \mu(Z) = 1\}.$$

We now consider the relation between entropy and Hausdorff dimension.

**Proposition 4.** *For any measure $\mu \in \mathcal{M}$ we have*

$$\dim_H \mu \geq h_\mu(g_m)/\log m, \tag{16}$$

*with equality when $\mu$ is ergodic.*

*Proof.* Assume first that $\mu$ is ergodic. In this case (see for example [12]) it is known that

$$\dim_H \mu = \lim_{n\to\infty} \frac{\log \mu(I_{x_1\cdots x_n})}{-n \log m}$$

for $\mu$-almost every $x = 0.x_1x_2\cdots \in [0,1]$. It follows from the second formula for $h_\mu(g_m)$ in Proposition 3 that $\dim_H \mu = h_\mu(g_m)/\log m$.

Consider now an arbitrary measure $\mu \in \mathcal{M}$ and an ergodic decomposition $(\mu_\alpha)_{\alpha\in S}$ of $\mu$ (it always exists; see for example [14, Section 6.2]). This means that for each $\alpha \in S$ the measure $\mu_\alpha$ is ergodic and there exists a measure $\nu$ on $S$ such that $\mu(Z) = \int_S \mu_\alpha(Z)\,d\nu(\alpha)$ for every measurable set $Z$. If $\mu(Z) = 1$ then $\mu_\alpha(Z) = 1$ for $\nu$-almost every $\alpha \in S$. Hence,

$$\dim_H Z \geq \dim_H \mu_\alpha = h_{\mu_\alpha}(g_m)/\log m.$$

Integrating over $\alpha$ we obtain (see for example [14, Theorem 8.4])

$$\dim_H Z \geq \int_S h_{\mu_\alpha}(g_m)\,d\nu(\alpha)/\log m = h_\mu(g_m)/\log m.$$

Taking the infimum over all the sets $Z$ with $\mu(Z) = 1$ we obtain (16). $\square$

We note that the inequality in (16) is in general strict in the case of nonergodic measures. Explicit examples can be readily obtained for instance from the following properties: if $\mu_1$ and $\mu_2$ are ergodic measures and $\mu = c_1\mu_1 + c_2\mu_2$ with $c_1 + c_2 = 1$ and $c_1$, $c_2 > 0$ then

$$\dim_H \mu = \max\{\dim_H \mu_1, \dim_H \mu_2\}$$

and $h_\mu(g_m) = c_1 h_{\mu_1}(g_m) + c_2 h_{\mu_2}(g_m)$.

## 3. Irregular sets

In this section we establish Theorem 1 on the "size" of the sets $M_k$ (see (4)) from the points of view of topology and dimension theory. It is in fact a particular case of stronger statements proved in this section.

For each $\underline{\alpha} < \overline{\alpha}$ we consider the set

$$M_k^{\underline{\alpha},\overline{\alpha}} = \left\{ x \in [0,1] : \liminf_{n\to\infty} \frac{\tau_k(x,n)}{n} = \underline{\alpha} \text{ and } \limsup_{n\to\infty} \frac{\tau_k(x,n)}{n} = \overline{\alpha} \right\}.$$

Notice that $M_k^{\underline{\alpha},\overline{\alpha}}$ has zero measure with respect to any $g_m$-invariant probability measure. We have

$$M_k = \bigcup_{0\leq\underline{\alpha}<\overline{\alpha}\leq 1} M_k^{\underline{\alpha},\overline{\alpha}}, \tag{17}$$

and hence, the interval $[0,1]$ can be decomposed into the disjoint union

$$[0,1] = \bigcup_{(\alpha_0,\dots,\alpha_{m-1})\in L_m} F_m(\alpha_0,\dots,\alpha_{m-1}) \cup \bigcup_{k=0}^{m-1} \bigcup_{0\leq\underline{\alpha}<\overline{\alpha}\leq 1} M_k^{\underline{\alpha},\overline{\alpha}}.$$

This type of decomposition is often called a multifractal decomposition. The theory of multifractal analysis has mainly been concerned with sets such as $F_m(\alpha_0,\dots,\alpha_{m-1})$ that consist of points for which certain limit or limits exist (see (1) and (2)). On the other hand, it was recently observed that the "irregular parts" of certain multifractal decompositions (of the type of those in (17)) may be (and often are, in a very precise sense) rather large from the points of view of topology and dimension theory. This allows us to have a much more detailed information about multifractal decompositions.

We are also interested in the irregular sets $M_k^{\underline{\alpha},\overline{\alpha}}$. We shall prove that even though they are rather small from the point of view of measure theory they have positive Hausdorff dimension. Furthermore, one of them is residual (i.e., it contains a dense $G_\delta$ set).

**Proposition 5.** *For each $k \in \{0, \dots, m-1\}$ the set $M_k^{0,1}$ is residual.*

*Proof.* Choose $\ell \in \{0, \dots, m-1\}$ different from $k$. Let $x = 0.x_1x_2 \cdots \in [0,1]$ and fix an integer $n \in \mathbb{N}$. We consider the set $U_n(x)$ of points $y = 0.y_1y_2 \cdots \in [0,1]$ such that

$$y_i = \begin{cases} x_i & \text{if } 1 \le i \le n \\ k & \text{if } n < i \le n^2 \\ \ell & \text{if } n^2 < i \le n^3 \end{cases}.$$

Observe that if $y \in U_n(x)$ then

$$\frac{\tau_k(y, n^2)}{n^2} \ge 1 - \frac{1}{n} \quad \text{and} \quad \frac{\tau_k(y, n^3)}{n^3} \le \frac{1}{n}. \tag{18}$$

Note that only the first $n^3$ digits of $y$ are specified. The open set

$$V_m = \bigcup_{n>m} \bigcup_{x \in [0,1]} \operatorname{int} U_n(x)$$

is dense in $[0,1]$. It follows from (18) that the dense $G_\delta$ set $\bigcap_{m=1}^\infty V_m$ is contained in $M_k^{0,1}$. This completes the proof. ∎

One can easily verify that the $G_\delta$ set $\bigcap_{m=1}^\infty V_m$ constructed in the proof of Proposition 5 (and that is dense in $M_k^{0,1}$) has zero Hausdorff dimension. On the other hand we shall see below (see Theorem 7) that $\dim_H M_k^{\underline{\alpha},\overline{\alpha}} > 0$ whenever $0 < \underline{\alpha} < \overline{\alpha} < 1$ (and for each $0 \le \underline{\alpha} < \overline{\alpha} \le 1$ whenever $m > 2$).

In view of (17), Proposition 5 implies that the set $M_k$ is residual. Thus, in order to establish Theorem 1 it remains to prove the identity (7). This will be obtained as a consequence of a more general statement.

Given functions $\varphi$, $\psi \colon [0,1] \to \mathbb{R}$ such that $\psi > 0$ we consider the set

$$K_\alpha(\varphi, \psi) = \left\{ x \in [0,1] : \lim_{n\to\infty} \frac{\sum_{i=0}^n \varphi({g_m}^i x)}{\sum_{i=0}^n \psi({g_m}^i x)} = \alpha \right\}, \tag{19}$$

where $g_m$ is defined by $g_m x = mx \pmod 1$. For example, if

$$\varphi_k = \chi_{[k/m,(k+1)/m)} \text{ and } \psi_k = 1 \text{ for } k = 0, \dots, m-1 \tag{20}$$

then

$$\bigcap_{k=0}^{m-1} K_{\alpha_k}(\varphi_k, \psi_k) = F_m(\alpha_0, \dots, \alpha_{m-1}).$$

We shall also consider some discontinuous functions. We call a function $\varphi\colon [0,1] \to \mathbb{R}$ *$m$-Hölder continuous* if it is piecewise Hölder continuous with finitely many discontinuities and at most at negative powers of $m$.

**Theorem 6.** *Let $\varphi_1, \dots, \varphi_d$, $\psi_1, \dots, \psi_d$ be $m$-Hölder continuous functions with $\psi_i > 0$ for each $i$. We have*

$$\dim_H \left( [0,1] \setminus \bigcup_{i=1}^d \bigcup_{\alpha \in \mathbb{R}} K_\alpha(\varphi_i, \psi_i) \right) = 1$$

*provided that $K_\alpha(\varphi_i, \psi_i) = [0,1]$ for no $i$ and no $\alpha$.*

*Proof.* We use the approach developed in [3] based on the notion of distinguishing family of measures. We recall that a family of invariant probability measures $\mu_1, \ldots, \mu_\ell$ is a distinguishing family for the sequences of functions $(f_{1n})_n, \ldots, (f_{dn})_n$ provided that for $i = 1, \ldots, d$ there exist measures $\nu_{i1}$ and $\nu_{i2}$ in $\{\mu_1, \ldots, \mu_\ell\}$ and constants $a_{i1} \neq a_{i2}$ such that for each $j = 1, 2$ and $\nu_{ij}$-almost every $x$,

$$\lim_{n\to\infty} f_{in}(x) = a_{ij}.$$

Set

$$\Lambda = \left\{ x \in [0,1] : \liminf_{n\to\infty} f_{in}(x) < \limsup_{n\to\infty} f_{in}(x) \text{ for } i = 1, \ldots, d \right\}. \tag{21}$$

The following statement is a particular case of Theorem 7.6 in [3].

**Lemma 1.** *If the ergodic $g_m$-invariant probability measures $\mu_1, \ldots, \mu_\ell$ form a distinguishing family for the sequences of $m$-Hölder continuous functions $(f_{1n})_n, \ldots, (f_{dn})_n$ then*

$$\dim_H \Lambda \geq \min\{\dim_H \mu_1, \ldots, \dim_H \mu_\ell\}.$$

Consider the sequences

$$f_{in}(x) = \frac{\sum_{j=0}^n \varphi_i({g_m}^j x)}{\sum_{j=0}^n \psi_i({g_m}^j x)}$$

for $i = 1, \ldots, d$. It is easy to verify that

$$\Lambda = [0,1] \setminus \bigcup_{i=1}^d \bigcup_{\alpha\in\mathbb{R}} K_\alpha(\varphi_i, \psi_i).$$

By Theorem 7.3 in [3], for each $i = 1, \ldots, d$ and $\varepsilon > 0$ there exist ergodic $g_m$-invariant probability measures $\nu_{i1}$ and $\nu_{i2}$ satisfying

$$\min\{\dim_H \nu_{i1}, \dim_H \nu_{i2}\} > 1 - \varepsilon$$

for each $i$, such that the family $\{\nu_{i1}, \nu_{i2} : i = 1, \ldots, d\}$ is a distinguishing family of measures for the sequences of functions $(f_{1n})_n, \ldots, (f_{dn})_n$. More precisely, for $j = 1, 2$ and $\nu_{ij}$-almost every $x \in [0,1]$ we have

$$\lim_{n\to\infty} f_{in}(x) = \frac{\int_0^1 \varphi_i \, d\nu_{ij}}{\int_0^1 \psi_i \, d\nu_{ij}}.$$

Furthermore

$$\frac{\int_0^1 \varphi_i \, d\nu_{i1}}{\int_0^1 \psi_i \, d\nu_{i1}} \neq \frac{\int_0^1 \varphi_i \, d\nu_{i2}}{\int_0^1 \psi_i \, d\nu_{i2}}$$

for each $i$, and Birkhoff's ergodic theorem (see for example [11, Chapter 1]) shows that $\{\nu_{i1}, \nu_{i2} : i = 1, \ldots, d\}$ is a distinguishing family of measures for the sequences of functions $(f_{1n})_n, \ldots, (f_{dn})_n$.

This allows us to apply Lemma 1 and conclude that

$$\dim_H \Lambda \geq \min\{\dim_H \nu_{i1}, \dim_H \nu_{i2} : i = 1, \ldots, d\} > 1 - \varepsilon.$$

The arbitrariness of $\varepsilon$ yields the desired result. □

One can show that in the case of $m$-Hölder continuous functions $\varphi$ and $\psi$ the following conditions are equivalent (see [1] and [3] for details):

1. $K_\alpha(\varphi, \psi) = [0,1]$ for no $\alpha$;
2. $K_\alpha(\varphi, \psi)$ is a nonempty proper subset of $[0,1]$ for some $\alpha$;
3. $K_\alpha(\varphi, \psi)$ is a nonempty proper dense subset of $[0,1]$ for every $\alpha$ in some interval;
4. there exists no constant $c \in \mathbb{R}$ such that $\varphi - c\psi = a - a \circ g_m$ for some bounded function $a \colon [0,1] \to \mathbb{R}$.

We can now establish Theorem 1.

*Proof of Theorem 1.* The first statement follows from Proposition 5. Setting $\varphi_k$ and $\psi_k$ as in (20), one can easily verify that $K_\alpha(\varphi_k, \psi_k) = [0,1]$ for no $k$ and no $\alpha$ (note that one can explicitly determine a point in $K_\alpha(\varphi_k, \psi_k)$ for each $\alpha \in [0,1]$). Furthermore, since

$$\bigcap_{k=0}^{m-1} M_k = [0,1] \setminus \bigcup_{k=0}^{m-1} \bigcup_{\alpha \in \mathbb{R}} K_\alpha(\varphi_i, \psi_i),$$

the second statement follows from Theorem 6. $\square$

We now study the Hausdorff dimension of the sets $M_k^{\underline{\alpha}, \overline{\alpha}}$ and more generally of intersections of these sets. Given a rectangle

$$R = [\underline{\alpha}_0, \overline{\alpha}_0] \times \cdots \times [\underline{\alpha}_{m-1}, \overline{\alpha}_{m-1}] \subset [0,1]^m,$$

we consider the decomposition of $\partial R$ into the $m$ pairs $\underline{R}_k$, $\overline{R}_k$ of closed faces of $R$, where $\underline{R}_k$ corresponds to $\underline{\alpha}_k$ and $\overline{R}_k$ corresponds to $\overline{\alpha}_k$. For each $k$ we define the numbers

$$\underline{d}_k = \max\{\dim_H F_m(z) : z \in \underline{R}_k \cap L_m\}$$

and

$$\overline{d}_k = \max\{\dim_H F_m(z) : z \in \overline{R}_k \cap L_m\}.$$

We also define the set

$$M_R = \bigcap_{k=0}^{m-1} M_k^{\underline{\alpha}_k, \overline{\alpha}_k}.$$

The following statement establishes an explicit formula for the Hausdorff dimension of the irregular set $M_R$ in terms of the numbers $\underline{d}_k$ and $\overline{d}_k$. In particular, although the set $M_R$ is totally unrelated to the "regular" part of the multifractal decomposition, this formula indicates that the Hausdorff dimension of $M_R$ is entirely carried by a single "regular" set $F_m(z)$, where $z$ is some vector in one of the faces of the rectangle $R$.

**Theorem 7.** *If* $\operatorname{int} R \neq \varnothing$ *and* $R \cap L_m \neq \varnothing$ *then*

$$\dim_H M_R = \min\{\underline{d}_k, \overline{d}_k : k = 0, \ldots, m-1\}.$$

*Proof.* Set $\varphi_k$ and $\psi_k$ as in (20). Let $\underline{z}_k \in \underline{R}_k \cap L_m$ and $\overline{z}_k \in \overline{R}_k \cap L_m$ be vectors such that

$$\dim_H F_m(\underline{z}_k) = \underline{d}_k \quad \text{and} \quad \dim_H F_m(\overline{z}_k) = \overline{d}_k.$$

Since $\underline{z}_k \in \underline{R}_k \subset R$ and $\overline{z}_k \in \overline{R}_k \subset R$, the (ergodic) Bernoulli measures $\mu_{\underline{z}_k}$ and $\mu_{\overline{z}_k}$ satisfy

$$\int_0^1 \varphi_k \, d\mu_{\underline{z}_k} = \underline{\alpha}_k, \quad \int_0^1 \varphi_k \, d\mu_{\overline{z}_k} = \overline{\alpha}_k, \tag{22}$$

and

$$\int_0^1 \varphi_\ell \, d\mu_{\underline{z}_k} \in [\underline{\alpha}_\ell, \overline{\alpha}_\ell], \quad \int_0^1 \varphi_\ell \, d\mu_{\overline{z}_k} \in [\underline{\alpha}_\ell, \overline{\alpha}_\ell] \tag{23}$$

for every $k$ and $\ell$. Furthermore, by Eggleston's result, Proposition 4, and (14) we have

$$\dim_H \mu_{\underline{z}_k} = \underline{d}_k \quad \text{and} \quad \dim_H \mu_{\overline{z}_k} = \overline{d}_k. \tag{24}$$

It follows from Birkhoff's ergodic theorem and (22) that the measures $\mu_{\underline{z}_k}$, $\mu_{\overline{z}_k}$ for $k = 0, \ldots, m-1$ form a distinguishing family (see the proof of Theorem 6 for the definition) for the sequences

$$f_{kn} = \sum_{j=0}^{n} \varphi_k \circ {g_m}^j \tag{25}$$

for $k = 0, \ldots, m-1$.

We now require a version of Lemma 1 for the set $\Gamma_R$ defined by

$$\left\{ x \in [0,1] : \liminf_{n\to\infty} f_{in}(x) = \underline{\alpha}_i \text{ and } \limsup_{n\to\infty} f_{in}(x) = \overline{\alpha}_i \text{ for } i = 1, \ldots, d \right\}. \tag{26}$$

Notice that if $\operatorname{int} R \neq \varnothing$ then $\Gamma_R \subset \Lambda$ (see (21)). The following statement is obtained from a straightforward modification of the proof of Theorem 7.6 in [3] (see also the observations at the end of this section).

**Lemma 2.** *Let $\mu_1, \ldots, \mu_\ell$ be a distinguishing family of ergodic $g_m$-invariant probability measures for the sequences $(f_{1n})_n, \ldots, (f_{dn})_n$. Assume that for each $i$, $j = 1, \ldots, d$ there exist measures $\nu_{i1}$ and $\nu_{i2}$ in $\{\mu_1, \ldots, \mu_\ell\}$ and constants $a_{ij}$ and $b_{ij}$ such that for each $i$ and $j$:*

1. *$a_{ii} = \underline{\alpha}_i$, $b_{ii} = \overline{\alpha}_i$, and $a_{ij}$, $b_{ij} \in [\underline{\alpha}_i, \overline{\alpha}_i]$;*
2. $$\lim_{n\to\infty} f_{in}(x) = a_{ij} \text{ for } \nu_{j1}\text{-almost every } x,$$ $$\lim_{n\to\infty} f_{in}(x) = b_{ij} \text{ for } \nu_{j2}\text{-almost every } x.$$

*If $\operatorname{int} R \neq \varnothing$ and $R \cap L_m \neq \varnothing$ then*

$$\dim_H \Gamma_R \geq \min\{\dim_H \mu_1, \ldots, \dim_H \mu_\ell\}.$$

Using (22) and (23), we can apply Lemma 2 to the sequences in (25) to conclude that

$$\dim_H M_R \geq \kappa \stackrel{\text{def}}{=} \min\{\underline{d}_k, \overline{d}_k : k = 0, \ldots, m-1\}.$$

We now obtain an upper bound for $\dim_H M_R$. Given $x \in [0,1]$ we denote by $V(x)$ the set of accumulation points (in the weak-$*$ topology) of the sequence of measures

$$\frac{1}{n}\sum_{j=0}^{n-1}\delta_{g_m{}^j x}, \tag{27}$$

where $\delta_y$ denotes the probability measure with $\delta_y(\{y\}) = 1$. We shall use the following statement.

**Lemma 3.** *If*

$$X_t = \{x \in [0,1] : \dim_H \mu \le t \text{ for some } \mu \in V(x)\}$$

*then* $\dim_H X_t \le t$.

*Proof of the lemma.* The lemma is a version of a statement of Bowen in [7] saying that (in our setup) if

$$Y_t = \{x \in [0,1] : h_\mu(g_m) \le t \text{ for some } \mu \in V(x)\}$$

then $\dim_H Y_t \le t/\log m$. It should be noted that the number $\log m \cdot \dim_H Z$ coincides with the topological entropy of $g_m$ on the set $Y_t$ (and since this set may in general be noncompact we are referring to the notion of topological entropy for noncompact sets introduced by Bowen in [7]). We shall deduce the lemma from Bowen's result. By Proposition 4, we have $h_\mu(g_m)/\log m \le \dim_H \mu$. Therefore, $X_t \subset Y_{t\log m}$ and hence, using Bowen's result, $\dim_H X_t \le t$. □

When $x \in M_R$, for each $k = 0, \ldots, m-1$ the sequence of measures in (27) has accumulation points $\mu_{\underline{w}_k}$ and $\mu_{\overline{w}_k}$ for some $\underline{w}_k \in \underline{R}_k$ and $\overline{w}_k \in \overline{R}_k$. It follows from (24) that

$$\min\{\dim_H \mu_{\underline{w}_k}, \dim_H \mu_{\overline{w}_k} : k = 0, \ldots, m-1\} \le \kappa.$$

Therefore $M_R \subset X_\kappa$ and Lemma 3 implies that $\dim_H M_R \le \kappa$. This completes the proof. □

The following is now an immediate consequence of Theorem 7.

**Corollary 8.** *For each* $k \in \{0, \ldots, m-1\}$ *and* $\underline{\alpha}$, $\overline{\alpha} \in [0,1]$ *such that* $\underline{\alpha} < \overline{\alpha}$, *the set* $M_k^{\underline{\alpha},\overline{\alpha}}$ *has positive Hausdorff dimension.*

As we observed, Lemma 2 is obtained from a straightforward modification of the proof of Theorem 7.6 in [3]. Nevertheless, the proof itself (as is already the case in [3]) involves considerable technical difficulties, related to the study of the pointwise dimension of noninvariant measures (sitting on sets that have zero measure with respect to any invariant measure). On the other hand, it is easy to describe why the assumptions in Lemma 2 are crucial when compared to those in Lemma 1 and we shall do it now. Notice first that although the sets $\Lambda$ and $\Gamma_R$ (see (21) and (26)) satisfy $\Gamma_R \subset \Lambda$ this inclusion is in general proper. The proof of Lemma 2 (that simply follows the proof of Theorem 7.6 in [3]) starts with the juxtaposition of cylinder sets at the level of symbolic dynamics which are successively typical with respect to $\nu_{i1}$ and $\nu_{i2}$ for $i = 1, \ldots, d$, in this order. We recall that, for a point $x \in [0,1]$, the cylinder sets corresponding to the intervals

$C_n(x) = (x - m^{-n}, x + m^{-n}) \cap [0,1]$ are said to be *typical* with respect to the measure $\mu$ if

$$\lim_{n\to\infty} -\frac{\log \mu(C_n(x))}{n \log m} = \dim_H \mu.$$

Then we repeat the same procedure ad infinitum, choosing cylinder sets which are successively typical with respect to $\nu_{i1}$ and $\nu_{i2}$ for $i = 1, \ldots, d$. Since $\operatorname{int} R \neq \varnothing$ we obtain with this construction a subset $\Lambda'$ of $\Lambda$ (see (21)). When $d = 1$ one has $\Lambda' \subset \Gamma_R$ (see (26)) under the assumptions of Lemma 2, and in fact also under those of Lemma 1. However, when $d > 1$ the assumptions in Lemma 1 are in general not enough to construct a subset of $\Gamma_R$ using this procedure. This is due to the fact that the cylinder sets juxtaposed *between* two pairs with the same fixed $i$ (i.e., the pairs of cylinder sets corresponding to $\nu_{i1}$ and $\nu_{i2}$, which occur infinitely often) may change the values of $\liminf_{n\to\infty} f_{in}(x)$ and $\limsup_{n\to\infty} f_{in}(x)$. More precisely, we can construct a set $\Lambda' \subset \Lambda$ such that for every $x \in \Lambda'$ and $i = 1, \ldots, d$,

$$\liminf_{n\to\infty} f_{in}(x) \leq \underline{\alpha}_i \quad \text{and} \quad \limsup_{n\to\infty} f_{in}(x) \geq \overline{\alpha}_i,$$

but a priori these inequalities may be strict (unless $d = 1$). On the other hand, under the additional assumptions of Lemma 2, we can construct $\Lambda'$ in such a way that for every $x \in \Lambda'$ and $i = 1, \ldots, d$,

$$\liminf_{n\to\infty} f_{in}(x) = \underline{\alpha}_i \quad \text{and} \quad \limsup_{n\to\infty} f_{in}(x) = \overline{\alpha}_i,$$

and thus $\Lambda' \subset \Gamma_R$.

## 4. Conditional variational principle

In this section we describe the conditional variational principle mentioned in the introduction.

We first recall the concept of topological pressure. The *topological pressure* of a continuous function $\varphi \colon [0,1] \to \mathbb{R}$ with respect to $g_m$ is defined by

$$P(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{i_1\cdots i_n} \exp \sup_{I_{i_1\cdots i_n}} \sum_{k=0}^{n-1} \varphi \circ {g_m}^k. \tag{28}$$

For an $m$-Hölder continuous $\varphi$ (i.e., a piecewise Hölder continuous function with finitely many discontinuities and at most at negative powers of $m$) the limit in (28) also exists and we still call it topological pressure of $\varphi$. We are particularly interested in locally constant functions. More precisely, consider a function $\varphi$ such that

$$\varphi(0.x_1x_2\cdots) = a_{x_1\cdots x_\kappa}$$

for some constants $a_{i_1\cdots i_\kappa} \in \mathbb{R}$ for $i_1, \ldots, i_\kappa \in \{0, \ldots, m-1\}$ and some fixed positive integer $\kappa$. These are called *$\kappa$-locally constant functions.* Note that they are $m$-Hölder continuous. In particular, it follows easily from (28) that if $\varphi$ is 1-locally constant then

$$P(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{i_1\cdots i_n} \prod_{j=1}^{n} \exp a_{i_j} = \log \sum_{k=0}^{m-1} \exp a_k. \tag{29}$$

Given functions $\varphi_k$, $\psi_k\colon [0,1] \to \mathbb{R}$ with $\psi_k > 0$ for $k = 1, \ldots, d$, and a vector $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, we write

$$K_\alpha = \bigcap_{k=1}^{d} K_{\alpha_k}(\varphi_k, \psi_k), \tag{30}$$

where each set $K_{\alpha_k}(\varphi_k, \psi_k)$ is defined by (19). For each $\mu \in \mathcal{M}$ we set

$$\mathcal{P}(\mu) = \left( \frac{\int_0^1 \varphi_1 \, d\mu}{\int_0^1 \psi_1 \, d\mu}, \ldots, \frac{\int_0^1 \varphi_d \, d\mu}{\int_0^1 \psi_d \, d\mu} \right).$$

Barreira, Saussol and Schmeling [2] established the following conditional variational principle.

**Theorem 9.** *Let $\varphi_1, \ldots, \varphi_d$, $\psi_1, \ldots, \psi_d$ be $m$-Hölder continuous functions with $\psi_i > 0$ for each $i$. If $\alpha \in \operatorname{int} \mathcal{P}(\mathcal{M})$, then*

$$\begin{aligned} \dim_H K_\alpha &= \frac{1}{\log m} \max \{ h_\mu(g_m) : \mu \in \mathcal{M} \text{ and } \mathcal{P}(\mu) = \alpha \} \\ &= \frac{1}{\log m} \inf \left\{ P\left( \sum_{k=1}^{d} q_k(\varphi_k - \alpha_k \psi_k) \right) : (q_1, \ldots, q_d) \in \mathbb{R}^d \right\}. \end{aligned} \tag{31}$$

We emphasize that the identities in (31) express the dimension spectrum $\mathcal{D}(\alpha) = \dim_H K_\alpha$ in two different ways. The first formula for $\mathcal{D}$ in (31) is a maximum over the closed set of measures $\mu$ for which $\mathcal{P}(\alpha) = \mu$ (and thus the name "conditional variational principle"). This first formula was obtained independently by Fan, Feng and Wu [9]. Unfortunately it is in general not easily amenable to explicit computations due to the fact that we have to know $h_\mu(g_m)$ for all such measures. This is why we are especially interested in the second formula.

The second formula for $\mathcal{D}$ in (31) is an infimum of a real valued function involving the topological pressure. In view of applications, and in particular those in this paper, it is crucial to have this formula. For example, to know an *explicit* expression for the dimension spectrum it is often enough to study the derivative of the function

$$q \mapsto P\left( \sum_{k=1}^{d} q_k(\varphi_k - \alpha_k \psi_k) \right).$$

We emphasize that even when the infimum in (31) does not allow one to obtain an explicit formula it is still crucial in several respects. In particular, under the assumptions of Theorem 9, it can be used to show that $\mathcal{D}$ is analytic (see [1, 2] for details). Furthermore, the first example of a nonconvex spectrum was given in [1] also as an application of this formula.

It was also shown in [2, Theorem 14] that for each fixed Hölder exponent $\theta$ and a residual vector $(\varphi_1, \ldots, \varphi_d, \psi_1, \ldots, \psi_d)$ in the space of $\theta$-Hölder continuous functions we have:

1. $\overline{\operatorname{int} \mathcal{P}(\mathcal{M})} = \mathcal{P}(\mathcal{M})$;
2. $\dim_H K_\alpha = 0$ for every $\alpha \in \partial \mathcal{P}(\mathcal{M})$.

We now describe a generalization of the first identity in (31), or more precisely of the formula (replacing the maximum by a supremum in (31))

$$\dim_H K_\alpha = \frac{1}{\log m} \sup \{h_\mu(g_m) : \mu \in \mathcal{M} \text{ and } \mathcal{P}(\mu) = \alpha\}.$$

In order to describe this generalization, let us consider continuous functions $\zeta\colon [0,1] \to U \subset \mathbb{R}^r$ and $\eta\colon U \to \mathbb{R}^p$. For each $\alpha \in \mathbb{R}^p$ we define the set

$$K_\alpha^{(\eta,\zeta)} = \left\{ x \in [0,1]\colon \lim_{n\to\infty} \eta \left( \frac{1}{n} \sum_{j=0}^{n-1} \zeta({g_m}^j x) \right) = \alpha \right\}.$$

When

$$\zeta = (\varphi_1, \dots, \varphi_d, \psi_1, \dots, \psi_d) \quad \text{and} \quad \eta(r_1, \dots, r_{2d}) = \left( \frac{r_1}{r_{d+1}}, \dots, \frac{r_d}{r_{2d}} \right)$$

we obtain $K_\alpha^{(\eta,\zeta)} = K_\alpha$ (see (30)).

The following is an immediate consequence of work of Takens and Verbitskiy [13], where we make the convention that $\sup \varnothing = 0$.

**Theorem 10.** *Let $\zeta\colon [0,1] \to U \subset \mathbb{R}^r$ and $\eta\colon U \to \mathbb{R}^p$ be piecewise continuous functions with finitely many discontinuities and at most at negative powers of $m$. For each $\alpha \in \mathbb{R}^p$ we have*

$$\dim_H K_\alpha^{(\eta,\zeta)} = \frac{1}{\log m} \sup \left\{ h_\mu(g_m) : \mu \in \mathcal{M} \text{ and } \eta \left( \int_0^1 \zeta \, d\mu \right) = \alpha \right\}.$$

We observe that the classes of dynamical systems for which the statements in Theorems 9 and 10 were respectively established in [2] and [13] are more general than the situation considered here. However, none of these two papers includes the class of dynamical systems considered in the other one.

The identity in Theorem 10 says that for each $\alpha$ with $K_\alpha^{(\eta,\zeta)} \neq \varnothing$ and each $\varepsilon > 0$ there exists a measure $\mu_{\alpha,\varepsilon} \in \mathcal{M}$ for which

$$\eta \left( \int_0^1 \zeta \, d\mu_{\alpha,\varepsilon} \right) = \alpha \quad \text{and} \quad h_{\mu_{\alpha,\varepsilon}}(g_m)/\log m > \dim_H K_\alpha^{(\eta,\zeta)} - \varepsilon.$$

Unfortunately, in general the measure $\mu_{\alpha,\varepsilon}$ may not sit on $K_\alpha^{(\eta,\zeta)}$. Such a property would be crucial for example as a departure point to obtain similar results to those in Section 3 in this situation. On the other hand, under the assumptions of Theorem 9 it is shown in [2] that there exists an ergodic measure $\mu_\alpha \in \mathcal{M}$ such that $\mathcal{P}(\mu_\alpha) = \alpha$ with the additional properties

$$\mu_\alpha(K_\alpha) = 1 \quad \text{and} \quad h_{\mu_\alpha}(g_m)/\log m = \dim_H K_\alpha. \tag{32}$$

## 5. Applications to frequencies of digits

In this section we show how the conditional variational principle can be used in a number of problems in the study of the Hausdorff dimension of sets defined in terms of frequencies of digits.

5.1. **Linear relations.** When dealing with frequencies of digits it suffices to consider 1-locally constant functions (sometimes called digit-functions). In this case we can combine Theorems 9 and 10 to obtain a more explicit statement. In this section we always take $\psi_i = 1$ for $i = 1, \ldots, d$ in Theorem 9, and thus for each $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ the set $K_\alpha$ in (30) is given by

$$K_\alpha = \left\{ x \in [0,1] : \lim_{n\to\infty} \sum_{j=0}^{n} \varphi_i({g_m}^j x) = \alpha_i \text{ for } i = 1, \ldots, d \right\}.$$

For each $\mu \in \mathcal{M}$ set

$$\mathcal{Q}(\mu) = \left( \int_0^1 \varphi_1 \, d\mu, \ldots, \int_0^1 \varphi_d \, d\mu \right).$$

For simplicity we shall also write $\varphi_{ik} = \varphi_i([k/m, (k+1)/m))$.

**Theorem 11.** *Let $\varphi_i \colon [0,1] \to \mathbb{R}$ be 1-locally constant functions for $i = 1, \ldots, d$. The following properties hold:*

1. *if $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathcal{Q}(\mathcal{M})$ then*

$$\dim_H K_\alpha = \frac{1}{\log m} \max \left\{ -\sum_{k=0}^{m-1} \beta_k \log \beta_k : (\beta_0, \ldots, \beta_{m-1}) \in \Delta_m \right\}, \tag{33}$$

*where*

$$\Delta_m = \left\{ (\beta_0, \ldots, \beta_{m-1}) \in L_m : \sum_{k=0}^{m-1} \beta_k \varphi_{ik} = \alpha_i \textit{ for } i = 1, \ldots, d \right\};$$

2. *if, in addition, $\alpha = (\alpha_1, \ldots, \alpha_d) \in \operatorname{int} \mathcal{Q}(\mathcal{M})$ then*

$$\dim_H K_\alpha = \inf \left\{ \log_m \sum_{k=0}^{m-1} \exp \sum_{i=1}^{d} q_i(\varphi_{ik} - \alpha_i) : (q_1, \ldots, q_d) \in \mathbb{R}^d \right\}; \tag{34}$$

3. *the function $\alpha \mapsto \dim_H K_\alpha$ is continuous on $\mathcal{Q}(\mathcal{M})$.*

*Proof.* Setting $\zeta = (\varphi_1, \ldots, \varphi_d)$ and $\eta = \mathrm{id}$ it follows from Theorem 10 that

$$\dim_H K_\alpha = \frac{1}{\log m} \sup\{h_\mu(g_m) : \mu \in \mathcal{M} \text{ and } \mathcal{Q}(\mu) = \alpha\} \tag{35}$$

for each $\alpha \in \mathcal{Q}(\mathcal{M})$. Let $\mu$ be a $g_m$-invariant measure on $[0,1]$ and write $\beta_k = \mu(I_k)$ for each $k$. The condition $\mathcal{Q}(\mu) = \alpha$ is equivalent to

$$\left( \sum_{k=0}^{m-1} \beta_k \varphi_{1k}, \ldots, \sum_{k=0}^{m-1} \beta_k \varphi_{dk} \right) = \alpha. \tag{36}$$

Therefore, using (14) and (15) we can replace the supremum in (35) by the maximum in (33), and we obtain the first statement in the theorem.

By (29) we have

$$P\left( \sum_{k=1}^{d} q_k(\varphi_k - \alpha_k \psi_k) \right) = \log \sum_{k=0}^{m-1} \exp \sum_{i=1}^{d} q_i(\varphi_{ik} - \alpha_i).$$

Applying Theorem 9 we obtain the second statement in the theorem.

To complete the proof it is enough to show that the maximum in (33) varies continuously with $\alpha$. Observe first that the condition in (36) defines a plane in $\mathbb{R}^m$ varying continuously with $\alpha$. The strict convexity of the function

$$(\beta_0, \dots, \beta_{m-1}) \mapsto -\sum_{k=0}^{m-1} \beta_k \log \beta_k$$

guarantees that the maximum in (33) also varies continuously with $\alpha$. This completes the proof. □

Theorem 11 offers two methods to compute the Hausdorff dimension of the set $K_\alpha$, although of different nature. The first method involves the computation of the maximum in (33). Unfortunately it consists of a problem of conditional extrema. In applications we may try to use, for example, the method of Lagrange multipliers. This often leads to less explicit formulas. On the other hand, the second method to compute the Hausdorff dimension, based on the computation of the infimum in (34), should in general be more amenable to computation. Essentially it amounts to determine the extrema of a function involving the topological pressure, without any extra condition. Although the identity in (34) is only known to hold in $\operatorname{int} \mathcal{Q}(\mathcal{M})$ (conjecturally it holds in $\mathcal{Q}(\mathcal{M})$), the continuity of the function $\alpha \mapsto \dim_H K_\alpha$ on $\mathcal{Q}(\mathcal{M})$ allows us to obtain $\dim_H K_\alpha$ for $\alpha \in \overline{\operatorname{int} \mathcal{Q}(\mathcal{M})}$ from the knowledge of (34) on $\operatorname{int} \mathcal{Q}(\mathcal{M})$. It is shown in [2] (see Section 4), that for each fixed Hölder exponent $\theta$ and a residual vector $(\varphi_1, \dots, \varphi_d)$ in the space of $\theta$-Hölder continuous functions (not necessarily 1-locally constant) we have $\overline{\operatorname{int} \mathcal{Q}(\mathcal{M})} = \mathcal{Q}(\mathcal{M})$.

Theorem 11 (and its straightforward generalization to $\kappa$-locally constant functions) allows us to provide a unified and simple approach to substantially complicated problems. Moreover, it follows from work in [2] that for each choice of functions $\varphi_i$ one can explicitly exhibit a measure sitting on the level set $K_\alpha$, in the sense that (32) holds. In the case of 1-locally constant functions these measures are always Bernoulli measures.

A first consequence of Theorem 11 is the classical result of Eggleston [8] described in the introduction.

**Corollary 12.** *For every* $(\alpha_0, \dots, \alpha_{m-1}) \in L_m$*,*

$$\dim_H F_m(\alpha_0, \dots, \alpha_{m-1}) = -\frac{\sum_{k=0}^{m-1} \alpha_k \log \alpha_k}{\log m}.$$

*Proof.* Setting $\varphi_k$ and $\psi_k$ as in (20) for $k = 0, \dots, m-1$, we obtain $K_\alpha = F_m(\alpha_0, \dots, \alpha_{m-1})$. The statement follows immediately from Theorem 11. □

Theorem 11 allows us to complete the proof of the results in the introduction.

*Proof of Theorem 2.* Set $d = 1$ and

$$\varphi_1 = \chi_{[k/m,(k+1)/m)} - \beta \chi_{[\ell/m,(\ell+1)/m)}.$$

The first identity in the theorem follows from Theorem 11, (14), and (15). Furthermore, since $0 \in \operatorname{int} \mathcal{Q}(\mathcal{M}) = (-\beta, 1)$, Theorem 11 shows that

$$\dim_H\{x \in [0,1] : \tau_k(x) = \beta\tau_\ell(x)\} = \inf_{q\in\mathbb{R}} \frac{\log(\mathrm{e}^q + \mathrm{e}^{-\beta q} + m - 2)}{\log m}.$$

It is straightforward to verify that the infimum is attained at $q = \frac{\log\beta}{1+\beta}$ and an easy computation yields the desired result. $\square$

For example, setting $\beta = 1$ in Theorem 2 we obtain

$$\dim_H\{x \in [0,1] : \tau_k(x) = \tau_\ell(x)\} = 1, \tag{37}$$

independently of $k$, $\ell$, and $m$. It is also interesting to observe that for a fixed $\beta \geq 0$ the Hausdorff dimension in Theorem 2 tends to 1 when $m \to \infty$. This corresponds to the fact that the disjoint union

$$\bigcup_{\alpha_k=\beta\alpha_\ell} F_m(\alpha_0, \dots, \alpha_{m-1}) \subset \{x \in [0,1] : \tau_k(x) = \beta\tau_\ell(x)\}$$

contains sets with larger and larger Hausdorff dimension as $m \to \infty$.

We emphasize that the formula with the maximum in (33) can also be used to obtain the explicit value of the Hausdorff dimension in Theorem 2. Although the computation is not as immediate as with the infimum in (34), in this special case it can be effected in a simple manner. This alternative approach reflects the fact that the entropy of an invariant measure is maximized by a mass distribution as uniform as possible (see (38) below).

We assume for simplicity that $k = 0$ and $\ell = 1$ in Theorem 2. Set $\rho = 1 - \sum_{j=2}^{m-1} \alpha_j$. We have $\alpha_0 = \rho/(1+\beta)$ and $\alpha_1 = \rho\beta/(1+\beta)$. One can easily verify that the function

$$\begin{aligned}(\alpha_2, \dots, \alpha_{m-1}) \mapsto &-\sum_{j=0}^{m-1} \alpha_j \log\alpha_j \\ &= -\frac{\rho}{1+\beta}\log\frac{\rho}{1+\beta} - \frac{\rho\beta}{1+\beta}\log\frac{\rho\beta}{1+\beta} - \sum_{j=2}^{m-1}\alpha_j\log\alpha_j\end{aligned}$$

attains its maximum when

$$\alpha_2 = \cdots = \alpha_{m-1} = \frac{1-\rho}{m-2}. \tag{38}$$

Therefore, the Hausdorff dimension in Theorem 2 is given by

$$\max_{0\leq\alpha\leq\frac{1}{1+\beta}} -\frac{\alpha\log\alpha + \alpha\beta\log(\alpha\beta) + (1-\alpha-\alpha\beta)\log\frac{1-\alpha-\alpha\beta}{m-2}}{\log m}.$$

One can verify that this maximum is attained at $\alpha = 1/(2\beta^{\beta/(\beta+1)} + \beta + 1)$, and hence obtain its explicit value.

A similar approach yields the following statement.

**Corollary 13.** *Given integers $i_1 < \cdots < i_k$ in $\{0, \dots, m-1\}$ and numbers $\beta_1, \dots, \beta_k \in [0,1]$ such that $\beta \overset{\text{def}}{=} \sum_{j=1}^k \beta_k \leq 1$, we have*

$$\begin{aligned}\dim_H\{x \in [0,1] : (\tau_{i_1}(x), \dots, \tau_{i_k}(x)) = (\beta_1, \dots, \beta_k)\} \\ = -\frac{\sum_{j=1}^k \beta_j\log\beta_j + (1-\beta)\log\frac{1-\beta}{m-1}}{\log m}.\end{aligned}$$

For example, setting $k = 1$ we obtain

$$\dim_H\{x \in [0,1]\colon \tau_i(x) = \alpha\} = -\frac{\alpha \log \alpha + (1-\alpha)\log \frac{1-\alpha}{m-1}}{\log m}.$$

Note that this tends to $1-\alpha$ when $m \to \infty$. In particular, the Hausdorff dimension of the set of numbers having 99.99% of zeros in their base-$m$ representation is uniformly bounded away from zero as $m \to \infty$. This should be contrasted with the behavior in Theorem 2 when $m \to \infty$.

In view of the existence of nontrivial irregular sets with large Hausdorff dimension (see Section 3) the following is another nontrivial application of Theorem 11. Namely, we want to show that the set of points for which the frequency of some digits is zero and the set of points where the same digits do not occur have equal dimension.

**Corollary 14.** *The Hausdorff dimension of the set of points for which the frequency of a number $k$ of fixed digits in the base-$m$ representation is zero equals the Hausdorff dimension of the set of points where the same digits do not appear in the base-$m$ representation. The common value is* $\log_m(m-k)$.

*Proof.* Let $E$ be the set of points where the frequency of some digits in the base-$m$ representation is zero. Without loss of generality we assume that the first $k$ digits have frequency zero. Setting $d = 1$ and $\varphi_1 = \chi_{[0,(k-1)/m)}$ it follows from Theorem 11 that $\dim_H E = \log_m(m-k)$. Let now $F \subset E$ be the set of points where the same digits do not appear in their base-$m$ representation. Let $\mu$ be the (ergodic) Bernoulli measure on $m-k$ symbols with equal probabilities. We have $\mu(F) = 1$ (with the natural identification of a base-$(m-k)$ representation with a base-$m$ representation) and by Proposition 4, $\dim_H \mu = \log_m(m-k)$. Therefore $\dim_H F \geq \dim_H \mu$ and $\dim_H E = \dim_H F$. ∎

We can also consider more involved problems. Let $A = (a_{ij})$ be an $d \times m$ matrix and $b = (b_1, \ldots, b_d) \in \mathbb{R}^d$. Consider the set $K(A, b)$ of numbers $x \in [0,1]$ such that $A\tau_m(x) = b$, where $\tau_m(x) = (\tau_0(x), \ldots, \tau_{m-1}(x))$.

**Corollary 15.** *If $b \in \operatorname{int} A(L_m)$ then*

$$\begin{aligned}\dim_H K(A,b) &= \max\left\{-\sum_{k=0}^{m-1}\alpha_k \log_m \alpha_k : (\alpha_0, \ldots, \alpha_{m-1}) \in L_m \cap A^{-1}b\right\}\\ &= \inf\left\{\log_m \sum_{k=0}^{m-1}\exp\sum_{i=1}^{d} q_i(a_{ik} - b_i) : (q_1, \ldots, q_d) \in \mathbb{R}^d\right\}.\end{aligned}$$

*Proof.* For $i = 1, \ldots, d$, set

$$\varphi_i = \sum_{j=0}^{m-1} a_{ij}\chi_{[j/m,(j+1)/m)},$$

and note that $\mathcal{Q}(\mathcal{M}) = A(L_m)$. The desired statement follows immediately from Theorem 11. ∎

We remark that $\overline{\operatorname{int} A(L_m)} = A(L_m)$ if and only if $\det(AA^t) \neq 0$, where $A^t$ denotes the transpose of $A$.

We now give several nontrivial examples as applications of Corollary 15 (the computations are straightforward and are not reproduced here):

1. given integers $a_0, \ldots, a_{m-1}$, not all zero, the set of points $x \in [0,1]$ such that
$$a_0\tau_0(x) + \cdots + a_{m-1}\tau_{m-1}(x) = \beta$$
has Hausdorff dimension
$$\log_m \min\left\{\sum_{i=0}^{m-1} r^{a_i-\beta} : r > 0 \text{ and } \sum_{i=0}^{m-1}(a_i-\beta)r^{a_i} = 0\right\}$$
(note that this amounts to determine the positive real roots of the polynomial $\sum_{i=0}^{m-1}(a_i-\beta)r^{a_i-\min_i a_i}$);
2. the set of points $x \in [0,1]$ such that $\tau_0(x) + 2\tau_1(x) = \beta$ has Hausdorff dimension
$$\log_m\left[\frac{r^{1-\beta}+2r^{-\beta}}{2-\beta}\right],$$
where
$$r = \frac{\beta - 1 + \sqrt{(\beta-1)^2 + 4\beta(\beta-2)(m-2)}}{4-2\beta};$$
3. the set of points $x \in [0,1]$ such that $\tau_0(x)+\cdots+\tau_{j-1}(x) = j\beta$ (i.e., the average of the frequencies of the first $j$ digits equals $\beta$) has Hausdorff dimension
$$\log_m\left[j\left(\frac{(m-j)\beta}{1-j\beta}\right)^{1-j\beta} + (m-j)\left(\frac{1-j\beta}{(m-j)\beta}\right)^{j\beta}\right];$$
4. given $\gamma \geq 0$, the set of points $x \in [0,1]$ such that $\tau_0(x)+\tau_1(x) = \gamma\tau_2(x)$ has Hausdorff dimension
$$\log_m\left[2\left(\frac{\gamma}{2}\right)^{1/(\gamma+1)} + \left(\frac{2}{\gamma}\right)^{\gamma/(\gamma+1)} + m - 3\right]$$
(for example, this is $\log_m(2\sqrt{2}+m-3)$ when $\gamma = 1$).

Using results in [2] we can also compute the Hausdorff dimension of sets that are obtained from the intersection of irregular sets (see Section 3) with those in this section. In particular, we can consider sets of points for which some fixed frequencies are not well-defined and for which the remaining ones satisfy some linear relations.

5.2. **Nonlinear relations.** Theorem 10 allows us to study continuous nonlinear relations between frequencies of digits. In this case there is in general no appropriate generalization of the formula given by Theorem 9 in terms of the topological pressure, and thus we have to use the variational principle in Theorem 10. This often makes the computations much harder when we study nonlinear relations between frequencies of digits. We shall concentrate here on nontrivial examples.

Fix $m > 2$. For each $b > 0$, we set
$$F_b = \{x \in [0,1] : \tau_1(x) = \mathrm{e}^{1-b\tau_0(x)}/b\}.$$

**Corollary 16.** *If $b > 2$ then*

$$\dim_H F_b = \log_m b + \frac{b-2}{b} \log_m \left( \frac{m-2}{b-2} \right).$$

*Proof.* Define $f(x) = \mathrm{e}^{1-bx}/b$. It follows from Theorem 10 that $\dim_H F_b$ is equal to the supremum of

$$-\alpha_0 \log_m \alpha_0 - f(\alpha_0) \log_m f(\alpha_0) - \sum_{i=2}^{m-1} \alpha_i \log_m \alpha_i$$

under the assumption that $\alpha_0 + f(\alpha_0) + \alpha_2 + \cdots + \alpha_{m-1} = 1$ and $\alpha_i \in [0, 1]$. It is thus enough to determine the infimum of the function

$$G(\alpha) = \alpha \log \alpha + f(\alpha) \log f(\alpha) + (1 - \alpha - f(\alpha)) \log \frac{1 - \alpha - f(\alpha)}{m-2} \tag{39}$$

over all $\alpha \in [0, 1]$ such that $\alpha + f(\alpha) \leq 1$. The derivative is

$$G'(\alpha) = \log \alpha + f'(\alpha) \log f(\alpha) - (1 + f'(\alpha)) \log \frac{1 - \alpha - f(\alpha)}{m-2}. \tag{40}$$

For $\alpha = 1/b$ we have $f(\alpha) = \alpha$ and $f'(\alpha) = -1$, and hence, $G'(\alpha) = 0$. Some elementary calculus shows that $1/b$ is indeed a global minimum of $G$. We conclude that $\dim_H F_b = -G(1/b)/\log m$ and the desired statement amounts now to an elementary computation. □

The approach used in the proof of the corollary applies in a similar way to other types of sets. In particular, given a continuous function $f \colon \mathbb{R} \to \mathbb{R}$ the set

$$F_f = \{x \in [0, 1] : \tau_1(x) = f(\tau_0(x))\}$$

has Hausdorff dimension

$$\dim_H F_f = -\frac{1}{\log m} \inf\{G(\alpha) : \alpha \in [0, 1] \text{ and } \alpha + f(\alpha) \leq 1\}, \tag{41}$$

where $G$ is the function defined by (39). In particular, when $f$ is differentiable, if $\alpha \in (0, \frac{1}{2})$ is such that $f(\alpha) = \alpha$ and $f'(\alpha) = -1$ it follows from (40) that $G'(\alpha) = 0$ and we should verify if it is a global minimum.

The formula (41) allows us to determine how the dimension is affected by small perturbations of $f$. For example, let $f_\varepsilon(x) = x + \varepsilon x^2$ (and $m > 2$). With similar arguments to those in [1, 2] one can show that the map $\varepsilon \mapsto \dim_H F_{f_\varepsilon}$ is analytic. Furthermore, one can look for an analytic function $\alpha = \alpha(\varepsilon)$ such that $\dim_H F_{f_\varepsilon} = -G(\alpha(\varepsilon))/\log m$. Setting $\alpha(\varepsilon) = \alpha_0 + \alpha_1 \varepsilon + o(\varepsilon)$ we obtain

$$\begin{aligned} G'(\alpha(\varepsilon)) =& 2 \log \frac{\alpha_0(m-2)}{1 - 2\alpha_0} + \left( 2\alpha_0 \log \frac{\alpha_0(m-2)}{1 - 2\alpha_0} \right) \varepsilon \\ &+ \left( \alpha_0 + \frac{2{\alpha_0}^2(m-2)}{1 - 2\alpha_0} + \alpha_1 \left( \frac{1}{\alpha_0} + \frac{4(m-2)}{1 - 2\alpha_0} \right) \right) \varepsilon + o(\varepsilon) \end{aligned}$$

and thus, solving $G'(\alpha(\varepsilon)) = 0$,

$$\alpha(\varepsilon) = \frac{1}{m} - \frac{3\varepsilon}{5m^2} + o(\varepsilon).$$

Therefore

$$\dim_H F_{f_\varepsilon} = -\frac{G(\alpha(\varepsilon))}{\log m} = 1 + \frac{(m-1)\varepsilon}{5m^3 \log m} + o(\varepsilon). \tag{42}$$

Note that the term 1 in (42) could have also been obtained from (37), which corresponds to the case $\varepsilon = 0$.

5.3. **Forbidden blocks.** We now consider frequencies of digits on sets for which some fixed blocks of digits are forbidden. These can be modeled by topological Markov chains.

Let $A = (a_{ij})$ with $i, j = 0, \ldots, m-1$ be an $m \times m$ matrix with each entry either 0 or 1. Using base-$m$ representations, we define

$$X_A = \{0.x_1x_2\cdots \in [0,1] : a_{x_n x_{n+1}} = 1 \text{ for all } n \in \mathbb{N}\}.$$

Recall that the nonuniqueness of the representation does not affect the study of the Hausdorff dimension.

Denote by $\rho(B)$ the spectral radius of the matrix $B$, and by $\Lambda_{a_0,\dots,a_{m-1}}$ the diagonal matrix with entries $a_0, \ldots, a_{m-1}$. Consider a 1-locally constant function $\varphi\colon X_A \to \mathbb{R}$ with $\varphi(0.x_1x_2\cdots) = a_{x_1}$. It is well known that

$$P(\varphi) = \log \rho(\Lambda_{a_0,\dots,a_{m-1}} A).$$

Consider an $m \times m$ stochastic matrix $\Pi = (\pi_{ij})$, with $i, j = 0, \ldots, m-1$, i.e., a matrix with nonnegative entries such that $\sum_{j=0}^{m-1} \pi_{ij} = 1$ for every $i$. Given a probability vector $p = (p_0, \dots, p_{m-1}) \in L_m$ such that $p\Pi = p$ we define a $g_m$-invariant probability measure $\mu_{\Pi,p}$ by

$$\mu_{\Pi,p}(I_{i_1\cdots i_n}) = p_{i_1} \prod_{k=1}^{n-1} \pi_{i_k i_{k+1}}.$$

If $A$ satisfies $a_{ij} = 0$ if and only if $\pi_{ij} = 0$, then the support of $\mu_{\Pi,p}$ is $X_A$. We have

$$h_{\mu_{\Pi,p}}(g_m) = -\sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p_i \pi_{ij} \log \pi_{ij}. \tag{43}$$

Using similar arguments to those in the proof of Theorem 11 one can establish the following statement.

**Theorem 17.** *Assume that some power of $A$ has only positive entries and let $\varphi_i\colon X_A \to \mathbb{R}$ be 1-locally constant functions for $i = 1, \ldots, d$. The following properties hold:*

1. *if $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathcal{Q}(\mathcal{M})$ then*

$$\dim_H K_\alpha = \frac{1}{\log m} \max h_{\mu_{\Pi,p}}(g_m),$$

*where the maximum is taken over all measures $\mu_{\Pi,p}$ such that*

$$\left( \sum_{k=0}^{m-1} p_k \varphi_{1k}, \dots, \sum_{k=0}^{m-1} p_k \varphi_{dk} \right) = \alpha;$$

2. *if* $\alpha = (\alpha_1, \dots, \alpha_d) \in \operatorname{int} \mathcal{Q}(\mathcal{M})$ *then*

$$\dim_H K_\alpha = \inf\{\log_m \rho(\Lambda_{a_0(q),\dots,a_{m-1}(q)}A) : q \in \mathbb{R}^d\},$$

*where*

$$a_k(q_1, \dots, q_d) = \exp \sum_{i=1}^{d} q_i(\varphi_{ik} - \alpha_i) \text{ for } k = 0, \dots, m-1.$$

One can also consider $\kappa$-locally constant functions with an arbitrary $\kappa$. In this case the most convenient approach consists in reducing the problem to the study of 1-locally constant functions. This can be done by considering a topological Markov chain in an appropriate larger space composed of blocks with a certain fixed length. This idea is used in the following section.

5.4. **Frequencies of blocks.** We now discuss how to obtain versions of the above results for frequencies of blocks. This is related to the study of $\kappa$-locally constant functions for an arbitrary $\kappa$. In this case (29) and (15) have to be replaced by less explicit formulas. This makes it often difficult to compute the Hausdorff dimension explicitly.

Consider the interval $I_{i_1\cdots i_n}$ in (12) for a fixed integer $m$. Whenever there exists the limit

$$\tau_{[i_1\cdots i_n]}(x) = \lim_{m\to\infty} \frac{\operatorname{card}\{j \in \{1, \dots, m\} : (x_{j+1}\cdots x_{j+n}) = (i_1\cdots i_n)\}}{m} \tag{44}$$

it is called the *frequency* of the block $[i_1\cdots i_n]$ in the base-$m$ representation of $x$. When we write the symbol $\tau_{[i_1\cdots i_n]}(x)$ we are already assuming the existence of the limit in (44).

We first consider a particular case in which one is able to obtain an explicit value for the Hausdorff dimension. Given a nonnegative $m \times m$ matrix $P = (p_{ij})$ we define the set

$$F_m(P) = \{x \in [0,1] : \tau_{[ij]}(x) = p_{ij} \text{ for every } i, j = 0, \dots, m-1\}.$$

Write $p_i = \sum_{j=0}^{m-1} p_{ij}$. We shall assume that

$$\sum_{i=0}^{m-1} p_i = 1 \quad \text{and} \quad p_j = \sum_{i=0}^{m-1} p_{ij} \tag{45}$$

(otherwise the set $F_m(P)$ would be empty). Then the matrix $\Pi = (\pi_{ij})$ with entries $\pi_{ij} = p_{ij}/p_i$ is a stochastic matrix (see Section 5.3).

**Theorem 18.** *If* $P = (p_{ij})$ *is a nonnegative matrix satisfying* (45) *for which some power has only positive entries, then*

$$\dim_H F_m(P) = -\frac{1}{\log m} \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p_i \pi_{ij} \log \pi_{ij}.$$

*Proof.* Consider the functions $\varphi_{ij} = \chi_{I_{ij}}$ for $i, j = 0, \dots, m-1$. It follows from Theorems 9 and 10 that

$$\dim_H F_m(P) = \frac{1}{\log m} \max_\mu h_\mu(g_m),$$

where the maximum is taken over all $g_m$-invariant probability measures $\mu$ on $[0,1]$ such that $\int_0^1 \varphi_{ij}\,d\mu = \mu(I_{ij}) = p_{ij}$ for every $i$ and $j$. By the first formula for $h_\mu(g_m)$ in Proposition 3, we have

$$h_\mu(g_m) \le -\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} \mu(I_{ij}) \log \frac{\mu(I_{ij})}{\mu(I_j)} = -\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} p_i\pi_{ij} \log \frac{p_i\pi_{ij}}{p_j}. \tag{46}$$

Note that

$$\begin{aligned}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} p_i\pi_{ij}\log\frac{p_i}{p_j} &= \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} p_i\pi_{ij}\log p_i - \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} p_i\pi_{ij}\log p_j \\ &= \sum_{i=0}^{m-1} p_i\log p_i - \sum_{j=0}^{m-1} p_j \log p_j = 0.\end{aligned}$$

It follows from (46) and (43) that $h_\mu(g_m) \le h_{\mu_{\Pi,p}}(g_m)$. This implies that the Hausdorff dimension of $F_m(P)$ is given by $h_{\mu_{\Pi,p}}(g_m)/\log m$. $\square$

The identity in Theorem 18 is obtained by Billingsley in [5] in the special case when all entries of the matrix $P$ are positive. On the other hand, using the techniques that we develop in this paper one can consider much more involved problems. In particular we can study sets of points for which not all frequencies of blocks are known. Furthermore, we can consider sets defined in terms of frequencies of blocks of different length. We emphasize that the technique developed by Billingsley cannot be applied to these situations without further changes. This is due to the fact that it is not known, at least a priori, whether the dimension of each of these sets is carried by a single subset for which all frequencies are known. It turns out, as a consequence of the theory that we develop in [2], that this special subset always exists, but to show that this happens amounts to compute the Hausdorff dimension of the initial set and thus it is a problem of similar difficulty.

For example, consider the set of points

$$F_\alpha = \{x \in [0,1] \colon \tau_{[00]}(x) = \alpha\}.$$

In a similar way to that in Section 5.1 in the case of 1-locally constant functions one can show that the entropy of $g_m$-invariant probability measures satisfying $p_{00} = \mu(I_{00}) = \alpha$ is maximized by the measure $\mu_{\Pi,p}$ obtained from the matrix $P$ with entries $p_{ij} = (1-\alpha)/(m^2-1)$ for each $(ij) \ne (00)$. This occurs precisely when the ratio $1-p_{00}$ is equally distributed over the remaining entries. In this case we have $p_0 = (m\alpha+1)/(m+1)$ and $p_i = m(1-\alpha)/(m^2-1)$ for each $i \ne 0$. Applying Theorems 9 and 10, a straightforward computation allows one to conclude that

$$\begin{aligned}\dim_H F_\alpha &= -\frac{1}{\log m}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} p_i\pi_{ij}\log\pi_{ij} \\ &= -\frac{1}{\log m}\left(p_{00}\log\pi_{00} + \sum_{j=1}^{m-1} p_{0j}\log\pi_{0j} + \sum_{i=1}^{m-1}\sum_{j=0}^{m-1} p_{ij}\log\pi_{ij}\right) \\ &= \alpha\log_m\frac{m+1}{m\alpha+1} + \frac{1-\alpha}{m+1}\log_m\frac{(m-1)(m\alpha+1)}{1-\alpha} + \frac{m(1-\alpha)}{m+1}.\end{aligned}$$

Note that $\dim_H F_\alpha \to 1-\alpha$ when $m \to \infty$. Therefore, the Hausdorff dimension of the set of numbers having 99.99% of pairs of zeros in their base-$m$ representation is uniformly bounded away from zero as $m \to \infty$.

We also want to consider sets defined in terms of frequencies of blocks of different length. Instead of presenting the general theory, which would hide the principles behind our approach, we choose to consider a specific example. Consider the set

$$F = \{x \in [0,1] \colon \tau_{[000]}(x) = \tau_{[11]}(x)\},$$

with respect to the base-2 representation of $x$. To compute the Hausdorff dimension of $F$ it is convenient to transform the related 3-locally constant functions into 1-locally constant functions on a new space and then apply the results of the former sections. We consider an extended representation of each number (sometimes called higher block representation). When $x = 0.x_1x_2\cdots$ is a base-2 representation, we substitute the digit $x_k$ by the block $b_k = [x_kx_{k+1}x_{k+2}]$. In this way each number is now represented by an infinite sequence $x = 0.b_1b_2\cdots$ on 8 symbols. The original representation can be recovered from this one by simply looking at the first symbol of each block $b_k$.

There is however a strong contrast between the original representation and the new one: not all sequences of blocks of 3 symbols are allowed in the new representation. For example, the symbol [011] cannot follow [000]. This relation is encoded in the transition matrix:

$$A = (a_{ij}) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix},$$

where $a_{ij} = 1$ if and only if the $j$th block in the lexicographic order can follow the $i$th one, i.e.,

$$a_{[x_1x_2x_3][y_1y_2y_3]} = 1 \text{ if and only if } x_2 = y_1 \text{ and } x_3 = y_2.$$

Any function that is 3-locally constant with respect to the original representation becomes 1-locally constant in the new representation. By Theorem 17, $\log 2 \cdot \dim_H F$ is equal to the infimum over $q \in \mathbb{R}$ of

$$P(q(\chi_{[000]} - \chi_{[11]})) = \log \rho \left( \begin{pmatrix} \mathrm{e}^q & \mathrm{e}^q & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{e}^{-q} & \mathrm{e}^{-q} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathrm{e}^{-q} & \mathrm{e}^{-q} \end{pmatrix} \right).$$

The characteristic polynomial of the matrix is $x^5(x^3-(\mathrm{e}^q+\mathrm{e}^{-q})x^2+(\mathrm{e}^q-1))$ and one can deduce that

$$\dim_H F = \frac{1}{\log 2}\log\inf_{y>0}\left(\frac{1+y^2}{3y}+\frac{(1+y^2)^2}{3y\sqrt[3]{p(y)/2}}+\frac{\sqrt[3]{p(y)/2}}{3y}\right),$$

where

$$\begin{aligned}p(y) =&2y^6-21y^4+27y^3+6y^2+2\\&+\sqrt{-27y^3(4y^7-4y^6-15y^5+42y^4-15y^3-12y^2+4y-4)}.\end{aligned}$$

After the acceptance of this paper we noticed that the numerical value announced for $\dim_H F$ in [2] was unfortunately obtained implementing incorrectly the formulas in the computer (and thus it should be revised).

A similar approach allows us to consider any other set defined in terms of linear relations between frequencies of blocks with an arbitrary length.

## References

1. L. Barreira and B. Saussol, *Variational principles and mixed multifractal spectra*, Trans. Amer. Math. Soc. **353** (2001), 3919–3944.
2. L. Barreira, B. Saussol and J. Schmeling, *Higher-dimensional multifractal analysis*, J. Math. Pures Appl. **81** (2002), 67–91.
3. L. Barreira and J. Schmeling, *Sets of "non-typical" points have full topological entropy and full Hausdorff dimension*, Israel J. Math. **116** (2000), 29–70.
4. A. Besicovitch, *On the sum of digits of real numbers represented in the dyadic system*, Math. Ann. **110** (1934), 321–330.
5. P. Billingsley, *Ergodic theory and information*, Wiley and Sons, 1965.
6. E. Borel, *Sur les probabilités dénombrables et leurs applications arithmétiques*, Rend. Circ. Mat. Palermo **26** (1909), 247–271.
7. R. Bowen, *Topological entropy for noncompact sets*, Trans. Amer. Math. Soc. **184** (1973), 125–136.
8. H. Eggleston, *The fractional dimension of a set defined by decimal properties*, Quart. J. Math. Oxford Ser. **20** (1949), 31–36.
9. A. Fan, D. Feng and J. Wu, *Recurrence, dimension and entropy*, J. London Math. Soc. **64** (2001), 229–244.
10. G. Hardy and J. Littlewood, *Some problems on Diophantine approximations*, Acta Math. **37** (1914), 155–190.
11. W. Parry, *Topics in ergodic theory*, Cambridge Tracts in Mathematics 75, Cambridge University Press, 1981.
12. Ya. Pesin, *Dimension theory in dynamical systems: contemporary views and applications*, Chicago Lectures in Mathematics, Chicago University Press, 1997.
13. F. Takens and E. Verbitskiy, *On the variational principle for the topological entropy of certain non-compact sets*, preprint.
14. P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics 79, Springer, 1981.

Departamento de Matemática, Instituto Superior Técnico, 1049-001 Lisboa, Portugal
*E-mail address*: luis.barreira@math.ist.utl.pt
*URL*: http://www.math.ist.utl.pt/~barreira/

LAMFA/CNRS FRE 2270, Université de Picardie Jules Verne, 33 rue Saint Leu, 80039 Amiens, France
*E-mail address*: benoit.saussol@mathinfo.u-picardie.fr
*URL*: http://www.mathinfo.u-picardie.fr/saussol/

Centre for Mathematical Sciences, LTH, Lund University, Box 118, SE-221 00 Lund, Sweden
*E-mail address*: joerg@maths.lth.se