J. London Math. Soc. (2) 99 (2019) 707–732
doi:10.1112/jlms.12197

Regularity of Minkowski's question mark measure, its inverse and a class of IFS invariant measures

Giorgio Mantica and Vilmos Totik

Abstract

We prove the recent conjecture that Minkowski's question mark measure is regular in the sense of logarithmic potential theory. The proof employs: an Iterated Function System composed of Möbius maps, which yields the classical Stern–Brocot sequences, an estimate of the cardinality of large spacings between numbers in these sequences and a criterion due to Stahl and Totik. We also generalize this result to a class of balanced measures of Iterated Function Systems in one dimension.

1. *Introduction and statement of the main results*

1.1. *Minkowski's question mark function and measure*

A remarkable function was introduced by Hermann Minkowski in 1904, to map algebraic numbers of second degree to the rationals, and these latter to binary fractions, in a continuous, order preserving way [36]. This function is called the *question mark function* and is indicated by $?(x)$, perhaps because of its enigmatic — yet captivating, multi-faceted personality. In fact, it is linked to continued fractions, to the Stern–Brocot tree and to the theory of numbers [12, 43]. It also appears in the theory of dynamical systems, in relation with the Farey shift map [9, 11, 26] and in the coding of motions on manifolds of negative curvature [7, 18, 19, 24, 44].

Let us define Minkowski's question mark function following [43]. Consider the interval $I = [0,1]$ and let $x \in I$. Write this latter in its continued fraction representation, $x = [n_1, n_2, \ldots]$, set $N_j(x) = \sum_{l=1}^{j} n_l$, and define $?(x)$ as the sum of the series

$$?(x) = \sum_{j=1}^{\infty} (-1)^{j+1} 2^{-N_j(x)+1}. \tag{1.1}$$

To deal with rational values $x \in I$, we also stipulate that terminating continued fractions correspond to finite sums in the above series.

The analytical properties of the question mark function are so interesting that its graph has been named the *slippery devil's staircase* [19]: it is continuous and Hölder continuous of order $\log 2/(1+\sqrt{5})$ [43]. It can be differentiated almost everywhere; its derivative is almost everywhere null [12, 43] and yet it is strictly increasing: $?(y) - ?(x) > 0$ for any $x, y \in I$, $x < y$. The fractal properties of the level sets of the derivative of $?(x)$ have been studied via the multi-fractal formalism [19, 24].

Since $?(x)$ is monotone non-decreasing, it is the distribution function of a Stieltjes measure μ:

$$?(x) = \mu([0, x)), \tag{1.2}$$

Received 23 March 2018; revised 9 November 2018; published online 27 November 2018.

2010 *Mathematics Subject Classification* 42C05, 31A15 (primary), 11A55, 11B57, 37D40 (secondary).

V. T. acknowledges support by NSF grant DMS 1564541.

which, because of the above, turns out to be singular continuous with respect to Lebesgue measure. We call μ the Minkowski's question mark measure and we always indicate it by this letter. A result by Kinney [25] asserts that its Hausdorff dimension can be expressed in terms of the integral of the function $\log_2(1+x)$ with respect to the measure μ itself. Very precise numerical estimates of this dimension have been obtained with high precision arithmetics [1]; rigorous numerical lower and upper bounds derived from the Jacobi matrix of μ place this value between 0.874716305108207 and 0.874716305108213 [32]. Further analytical properties of μ have been recently studied, among others, by the authors of [2, 3, 53].

Since Minkowski's $?(x)$ is invertible, it is natural to also consider its inverse, $?^{-1}(x)$, sometimes called Conway Box function, and the associated measure, which we will denote by μ^{-1}:

$$?^{-1}(x) = \mu^{-1}([0,x)), \tag{1.3}$$

or $\mu^{-1}([0,?(x))) = x$. This measure is also singular continuous [37].

1.2. *Potential theoretic regularity*

In this paper, we are concerned with additional fine properties of Minkowski's question mark measure μ and its inverse μ^{-1}, stemming from logarithmic potential theory in the complex plane [40, 42]. In this context, Dresse and Van Assche [13] asked whether μ is regular, in the sense defined below. Their numerical investigation suggested a preliminary negative answer, but their method was successively refined by a more powerful technique by the first author in [32], to provide compelling numerical evidence in favor of regularity of this measure. We now provide a rigorous proof of this result, which further unveils the intriguing nature of Minkowski's question mark function. The stronger conjecture that μ belongs to the so-called Nevai class, also supported by numerical investigation [32], still lies open.

The notion of *regularity* of a measure that we consider originated from [14, 51] and it concerns the asymptotic properties of its orthonormal polynomials $p_j(\mu;x)$ — recall the defining property: $\int p_j(\mu;x)p_m(\mu;x)d\mu(x) = \delta_{jm}$, where δ_{jm} is the Kronecker delta. We need the definition of regularity only when the support of the measure μ is the interval [0,1], in which case the regularity of μ (we write $\mu \in \mathbf{Reg}$ for short) means that for large orders its orthogonal polynomials $p_j(\mu;x)$ somehow mimic Chebyshev polynomials (that are orthogonal with respect to the equilibrium measure on [0,1] and extremal with respect to the infinity norm) both in *root asymptotics* away from [0,1] and in the *asymptotic distribution of their zeros* in [0,1].

Formally, letting γ_j be the (positive) coefficient of the highest order term, $p_j(\mu;x) = \gamma_j x^j + O(x^{j-1})$, *regularity* is defined in [45, 46] as the fact that $\gamma_j^{1/j}$, when the order j tends to infinity, tends to the logarithmic capacity of [0,1], that is, to $\frac{1}{4}$. In this case, we write $\mu \in \mathbf{Reg}$, and in what follows regularity of measures is always understood in this sense. An equivalent property is that the jth root limit of the *sup* norms of the orthogonal polynomials $p_j(\mu;x)$ on the support of μ is one, see [45, Theorem 3.2.3]. Further equivalent definitions of regularity can be found in [46], collected in Definition 3.1.2. A wealth of potential-theoretic results follow from regularity, as discussed in [41] and in [46, Chapter 3], so that assessing whether this property holds is a fundamental step in the analysis of a measure.

Notwithstanding this relevance and the time-honored history of Minkowski's question mark measure, proof of its regularity has not been achieved before. The asymptotic behavior of its orthogonal polynomials have been investigated theoretically and numerically in [32], with detailed pictures illustrating the abstract properties. This investigation continues in this paper from a slightly different perspective: we do not prove regularity of μ directly from the definition, that is, orthogonal polynomials play no rôle herein, but we use a purely measure-theoretic criterion, which translates the idea that a regular measure is not too thin on its support. This is Criterion λ^*:

Criterion 1.1 (λ^*, [**46**, Theorem 4.2.7]). If the support of μ is [0,1] and if for every $\eta > 0$ the Lebesgue measure of

$$\Lambda(\eta; s) = \{x \in [0,1] \text{ s.t. } \mu([x - 1/s, x + 1/s]) \geqslant e^{-\eta s}\} \tag{1.4}$$

tends to one, when s tends to infinity, then $\mu \in \mathbf{Reg}$.

Our fundamental result is therefore

Theorem 1.2. *Minkowski's question mark measure satisfies Criterion λ^* and hence is regular.*

The same can be asserted about the inverse question mark measure:

Theorem 1.3. *Minkowski's inverse question mark measure satisfies Criterion λ^* and hence is regular.*

Let us now describe the tools employed for the proof of these results, and let us place them into a wider perspective.

1.3. *Balanced measures of Iterated Function Systems and their regularity*

The main set-up of this investigation is that of Iterated Function Systems (in short IFS) and their balanced measures, of which Minkowski's question mark is an example. In its simplest form, an IFS is a finite collection of continuous maps φ_i, $i = 0, \ldots, M$ of $\mathbf{R}^n$ into itself. A set $\mathcal{A}$ that satisfies the equation $\mathcal{A} = \bigcup_{i=0}^{M} \varphi_i(\mathcal{A})$ is an *attractor* of the IFS. A family of measures on $\mathcal{A}$ can be constructed in terms of a set of parameters $\{\pi_i\}_{i=0}^{M}$, $\pi_i > 0$, $\sum_i \pi_i = 1$. Define the operator T on the space of Borel probability measures on $\mathcal{A}$ via

$$(T\nu)(A) = \sum_{i=0}^{M} \pi_i \nu(\varphi_i^{-1}(A)),$$

where A is any Borel set and ν is any such measure. A fixed point of this operator, $\nu = T\nu$ is called a *balanced* (or *invariant*) measure of the IFS. We shall see in Section 2 that Minkowsky's question mark measure is the invariant measure of an IFS with two maps φ_i that are contractions on $\mathcal{A} = [0,1]$. It follows from standard theory that such fixed point (as well as the attractor) is unique when the maps are strict contractions, that is, there is a $\delta < 1$ such that $|\varphi_i(x) - \varphi_i(y)| \leqslant \delta|x - y|$ for all $x, y \in \mathcal{A}$, and also when they are 'contractive on average' (see [**6**, **35**]). Minkowski's question mark measure falls in this second class. The contractions in the corresponding IFS are not strict contractions, but they rather satisfy $|\varphi_i(x) - \varphi_i(y)| < |x - y|$ if $x \neq y$: we call such maps *weak contractions*. Moreover, two different sets $\varphi_i(\mathcal{A})$ intersect each other at a single point, which is of zero measure, this measure being continuous. We call such an IFS *just touching* (or *disconnected* when the intersection is empty). In this case, the above relation defining an invariant measure ν can be shown to be equivalent to

$$\nu(\varphi_i(A)) = \pi_i \nu(A), \ i = 0, \ldots, M, \tag{1.5}$$

for any Borel set $A \subseteq \mathcal{A}$. This simple characterization will be used throughout the paper. We will first prove a general theorem for strictly contractive IFS:

Theorem 1.4. *If the maps φ_i, $i = 0, \ldots, M$, are strict contractions in $\mathbf{C}$ and μ (with support $\mathcal{A}$) is invariant with respect to the disconnected or just-touching IFS $\{\varphi_i\}_{i=0}^{M}$, $\{\pi_i\}_{i=0}^{M}$, then $\mu \in \mathbf{Reg}$.*

We will then show that Minkowski's question mark measure is the invariant measure of an IFS composed of *weak* contractions, so that regularity does not follow from the above Theorem 1.4. Nonetheless, it belongs to a larger family in which strict contractivity can be replaced by a combination of monotonicity and convexity. We will prove regularity also in this wider situation: see Theorem 8.1 in Section 8.

1.4. *Outline of the paper and additional results*

First we need a more transparent definition of Minkowski's question mark function than equation (1.1): this is provided by the symmetries of $?(x)$, which permit to regard it as *the invariant of an Iterated Function System* (*IFS*) *composed of Möbius maps*, following [**7**, **30**]. We review this approach in Section 2. In Lemma 2.1, we show that such Möbius IFS can be used to define *a countable family of partitions of* [0,1] *in a finite number of intervals*, I_σ, with elements labeled by words σ in a binary alphabet. These intervals are called *cylinders* in a dynamical approach, a term that we will also use frequently in this paper. The notable characteristic of any of these partitions is that all its elements have the same μ-measure, while obviously they have different lengths. The statistical distribution of these lengths is of paramount importance in assessing regularity.

In Section 3, we exploit the relation of Minkowski's question mark function with the Farey tree and Stern–Brocot sequences. In fact, in Lemma 3.1, we show that *these sequences coincide with the ordered set of endpoints in the Möbius IFS partitions of* [0,1]. None of these results is new, but we present them in a coherent and concise set-up, that of IFS, which is both elegant and renders sequent analysis easier. We build our theory on this approach, so that the paper is fully self-contained and the reader has no need of external material.

In Section 4, we apply the previous techniques to prove that the inverse question mark measure is regular: Theorem 2. The proof is rather concise: it follows from the λ^* Criterion and Hölder continuity of Minkowski's question mark function, which permits to bound *from below* the measure of intervals. This property does *not* hold for Conway's box function, so that such an easy proof is *not* available for the *inverse* of Conway's, that is, Minkowski's measure.

To use criterion λ^* in this wider context, we replace Hölder continuity of the inverse function by a combination of geometric and measure properties, composing Proposition 5.1, described in Section 5. One of the three hypotheses of this general proposition — perhaps the most important — is tailored on a remarkable characteristics of the cylinders of Minkowski's question mark measure. This characteristics is given by Proposition 7.2: *for any real positive* α, *the cardinality of intervals in the* n*th IFS partition, whose length is larger that* $\alpha/(n+1)$, *is bounded by a constant independent of* n. In Section 6, we present the first proof or regularity, which is based on these propositions. While this approach is sufficient to prove regularity and it hints at the generalization in Section 8, much more detail can be obtained on the distribution of the above intervals.

In fact, in Section 7, we focus our attention on the set of 'α-large' IFS / Stern–Brocot cylinders just defined. There are at least three reasons behind this interest. The first is that Proposition 7.2 is loosely related to the pressure function appearing in the so-called thermodynamical formalism, which gauges the exponential growth rate of sums of the partition interval lengths, raised to a real power. These sums, in the present case of Stern–Brocot intervals, have been studied in [**4**, **22**, **23**]. In this context, it is important to obtain precise estimates on the Lebesgue measure of 'α–large' cylinders. Secondly, as we will discuss momentarily, further conjectures on Minkowski's question mark measure have been formulated and numerically tested in [**32**]. The rigorous proof of these conjectures should presumably require such fine control. Finally, in this endeavor, we obtain a result that we believe to be relevant by its own merit. Proposition 7.1 fully characterizes the set of α-large Stern–Brocot cylinders, putting them in relation to the Farey series $\mathcal{F}^m$ (where $m = \lfloor 1/\sqrt{\alpha} \rfloor$). Figure 1 graphically exemplifies the situation.

The paper then continues in Section 8 with a broader discussion of regularity of IFS measures. We first prove Theorem 1.4 described above, which deals with the case of IFS composed of strict contractions. Regularity is here obtained via a further criterion from the comprehensive list in [**46**]. We then characterize a new family of weakly contracting IFS, whose invariant measure is supported on [0,1], for which Proposition 7.2 holds, which permits to prove regularity. This result is Theorem 8.1, whose proof occupies the last part of the paper. Minkowski's IFS belongs to this larger class, which can therefore regarded as its generalization.

1.5. *Further perspectives*

The fact that Minkowski's question mark function is regular is remarkable in many ways. First, it was not at all obvious how to reveal it numerically: standard techniques failed and specific ones were required [**13**, **30**, **32**]. From the theoretical side, regularity of Minkowski's question mark measure appears in the hypotheses of [**32**, Propositions 1,2], whose implications are therefore now rigorously established: these propositions describe and quantify the *local* asymptotic behavior of zeros of the orthogonal polynomials $p_j(\mu; x)$ and of the Christoffel functions associated with μ, linking these behaviors to the Farey/Stern–Brocot organization of the set of rational numbers.

Further conjectures were presented in [**32**], on the speed of convergence in the above asymptotic behaviors and, more significantly, on the fact that Minkowski's question mark might belong to Nevai's class: numerical indication is that its off diagonal Jacobi matrix elements converge to the limit value one-fourth, although slowly. If confirmed, this conjecture will provide us with a further example of a measure in Nevai's class which does not fulfill Rakhmanov's sufficient condition [**34**, **39**]: almost everywhere positivity of the Radon Nikodyn derivative of μ with respect to Lebesgue. It is well known that Nevai's class does contain pure point [**52**] and singular measures [**27**] but these examples do not seem to indicate a general criterion on a par with Rakhmanov's. To the contrary, Minkowski's question mark function might perhaps indicate a widening of such condition, involving the characteristics described here in Section 7.

In conclusion, the picture of Minkowski's question mark measure that emerges from recent investigations is that of a singular continuous measure that nonetheless has many *regular* characteristics: it is regular according to logarithmic potential theory; we conjectured that it belongs to Nevai's class [**32**]; its Fourier transform tends to zero polynomially [**21**, **38**, **54**, **55**] even if it does not fulfill the Riemann–Lebesque sufficient condition. It is, therefore, an interesting direction of further research to study the so-called Fourier–Bessel functions [**31**] generated by Minkowski's question mark measure, to detect whether they display any of the features usually associated with singular continuous measures [**16**, **17**, **33**, **48–50**] with almost-periodic Jacobi matrices [**8**, **28**, **29**, **31**].

2. *Minkowski's question mark measure and Möbius IFS*

In our view, the most effective representation of Minkowski's question mark function is via an Iterated Function System [**5**, **20**] composed of Möbius maps. This is a translation in modern language of the relation between Minkowski's question mark function and modular transformations, already discussed in [**12**]. Let us therefore adopt and develop the formalism introduced in [**7**]. Define maps M_i and P_i, $i = 0, 1$ from [0,1] to itself as follows:

$$\begin{aligned} M_0(x) &= \tfrac{x}{1+x}, \quad P_0(x) = \tfrac{x}{2}, \\ M_1(x) &= \tfrac{1}{2-x}, \quad P_1(x) = \tfrac{x+1}{2}. \end{aligned} \tag{2.1}$$

Then, using the properties of the continued fraction representation of a real number and equation (1.1) (see for example, [**7**]) it is not difficult to show that the following properties hold :

$$?(0) = 0, \ \ ?(1) = 1, \tag{2.2}$$

$$?(M_i(x)) = P_i(?(x)), \ i = 0, 1. \tag{2.3}$$

Note also that M_0 and M_1 play a symmetric role, for the mapping $x \to 1 - x$ maps these functions into each other: $1 - M_0(1 - x) = M_1(x)$.

It is well established that these relations uniquely define the function $?(x)$. Moreover, it was observed in [**7**, **30**] that an Iterated Function System, consisting of the two Möbius maps M_i, $i = 0, 1$, and of the probabilities $\pi_i = \frac{1}{2}$ has Minkowski's question mark measure μ as its invariant measure. This fact has been exploited also in [**32**]. We now start from the following standard construction of the *cylinders* of this measure.

Definition 2.1. Let Σ be the set of finite words in the letters 0 and 1. Denote by $|\sigma|$ the length of $\sigma \in \Sigma$: if $|\sigma| = n$, then σ is the n-letters sequence $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ where σ_i is either 0 or 1. When all σ_i are equal to the same $j = 0$ or 1, then we also write j^n for σ. Let $\emptyset$ be the empty word and assign to it length zero. Denote by Σ^n the set of n-letter words, for any $n \in \mathbf{N}$. Given two words $\sigma \in \Sigma^n$ and $\eta \in \Sigma^m$, the composite word $\sigma\eta \in \Sigma^{n+m}$ is the sequence $(\sigma_1, \ldots, \sigma_n, \eta_1, \ldots, \eta_m)$. Associate to any $\sigma \in \Sigma^n$ the map composition

$$M_\sigma = M_{\sigma_1} \circ M_{\sigma_2} \circ \cdots \circ M_{\sigma_n}, \tag{2.4}$$

when $n > 0$, and let $M_\emptyset$ be the identity transformation. Let I_σ be the basic intervals, or cylinders, of the IFS: $I_\sigma = M_\sigma([0, 1])$. Denote by $|I_\sigma|$ the Lebesgue measure of I_σ.

Because of the aforementioned symmetries, for a given n the set of intervals $\{I_\sigma, \ \sigma \in \Sigma^n\}$ is symmetric with respect to the point $1/2$.

Lemma 2.1. *Let Σ^n, M_σ and I_σ be as in Definition 2.1. Then, for any integer value $n \in \mathbf{N}$, the intervals I_σ, with $\sigma \in \Sigma^n$, are pairwise disjoint except possibly at one endpoint and fully cover* [0,1]*:*

$$[0, 1] = \bigcup_{\sigma \in \Sigma^n} I_\sigma. \tag{2.5}$$

Proof. When $\sigma = \emptyset$, the lemma is obvious. Observe that the functions M_i, $i = 0, 1$ are continuous, strictly increasing and map [0,1] to the two intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, respectively, which are disjoint except for a common endpoint. Then, the same happens for the two intervals $(M_\sigma \circ M_i)([0, 1]) = I_{\sigma i}$, $i = 0, 1$, where σ is any finite word and σi is the composite word. Explicit computation yields

$$I_{\sigma 0} = [M_\sigma(M_0(0)), M_\sigma(M_0(1))] = \left[M_\sigma(0), M_\sigma\left(\frac{1}{2}\right) \right]$$

and

$$I_{\sigma 1} = [M_\sigma(M_1(0)), M_\sigma(M_1(1))] = \left[M_\sigma\left(\frac{1}{2}\right), M_\sigma(1) \right],$$

where we have used a property that will be useful also in the sequel: for any $\sigma \in \Sigma$

$$M_{\sigma 0}(1) = M_{\sigma 1}(0) = M_\sigma\left(\frac{1}{2}\right), \tag{2.6}$$

which is valid since $M_1(0) = M_0(1) = 1/2$. It follows from this that $I_{\sigma 0}$ and $I_{\sigma 1}$ not only are adjacent, but also they exactly cover I_σ:

$$I_{\sigma 0} \cup I_{\sigma 1} = I_\sigma. \tag{2.7}$$

Using induction, one then proves equation (2.5). □

As a consequence of this Lemma, each set Σ^n is associated with a partition of [0,1] produced by the Möbius IFS. Since any word in Σ^n is uniquely associated to an interval of this partition, in the text we will use the terms word and interval as synonyms.

LEMMA 2.2. *Let Σ^n be as in Definition 2.1. For any $n \in \mathbf{N}$ the function*

$$\Theta(\sigma) = \sum_{j=1}^{n} \sigma_j 2^{n-j}, \tag{2.8}$$

induces the lexicographical order $\prec$ in Σ^n, in which the letter 1 follows the letter 0 and we read words from left to right: $\sigma \prec \eta$ precisely when $\Theta(\sigma) < \Theta(\eta)$.

In addition, letting

$$x_\sigma = M_\sigma(0) = M_{\sigma_1} \circ \cdots \circ M_{\sigma_n}(0) \tag{2.9}$$

the set $\{x_\sigma, \sigma \in \Sigma^n\}$ is increasingly ordered: $x_\sigma < x_\eta$ if and only if $\sigma < \eta$. Finally, one has that

$$I_\sigma = [x_\sigma, x_{\hat{\sigma}}] \tag{2.10}$$

where $\hat{\sigma}$ is the successive word of σ when $\sigma \neq 1^n$ and $x_{\hat{\sigma}} = 1$ in the opposite case.

Proof. Observe that when $n = 0$ we have $\sigma = \emptyset$ and $\Theta(\sigma) = 0$ because the sum in (2.8) contains no terms. It is immediate that Θ is bijective from Σ^n to $\{0, \ldots, 2^n - 1\}$ and therefore it induces an order on Σ^n. This coincides with the lexicographical order that we denote by '$\prec$'. To prove this statement, if $\sigma \neq \eta$ we can define $k = \min\{j \text{ s.t. } \sigma_j \neq \eta_j\}$. Then, $\sigma \prec \eta$ happens if and only if $\sigma_k = 0$ and $\eta_k = 1$. But in this case one has

$$\Theta(\sigma) = \sum_{j=1}^{k-1} \sigma_j 2^{n-j} + 0 + \sum_{j=k+1}^{n} \sigma_j 2^{n-j}$$

and

$$\Theta(\eta) = \sum_{j=1}^{k-1} \eta_j 2^{n-j} + 2^{n-k} + \sum_{j=k+1}^{n} \eta_j 2^{n-j}.$$

The first sums at the right-hand sides are equal, since $\sigma_j = \eta_j$ for $j < k$. In addition, the last sum in $\Theta(\sigma)$ is strictly less than 2^{n-k} for any choice of the sequence $\sigma_{k+1}, \ldots, \sigma_n$ and therefore $\Theta(\sigma) < \Theta(\eta)$. The same argument also proves that $\Theta(\sigma) < \Theta(\eta)$ implies that $\sigma \prec \eta$ in the lexicographical order.

Consider now $\sigma \prec \eta$ and x_σ, x_η defined as in equation (2.9). Define k as before and suppose $k < n$. Write $y = M_{\sigma_{k+1}} \circ \cdots \circ M_{\sigma_n}(0)$, $z = M_{\sigma_k}(y)$, so that $x_\sigma = M_{\sigma_1} \circ \cdots \circ M_{\sigma_{k-1}}(z)$. Observe that y is less than, or equal to $M_1^{n-k}(0) = 1 - \frac{1}{n-k+1}$, so that $z \leqslant M_0(1 - \frac{1}{n-k+1}) = \frac{n-k}{2n-2k+1} < \frac{1}{2}$. Equivalently, write $u = M_{\eta_{k+1}} \circ \cdots \circ M_{\eta_n}(0)$, $v = M_{\eta_k}(u)$, so that $x_\eta = M_{\eta_1} \circ \cdots \circ M_{\eta_{k-1}}(v)$. Now, $u \geqslant 0$, so that $v = M_1(u) \geqslant \frac{1}{2}$, and therefore $v > z$. The map composition $M_{\eta_1} \circ \cdots \circ M_{\eta_{k-1}}$ is the same as $M_{\sigma_1} \circ \cdots \circ M_{\sigma_{k-1}}$, since $\sigma_j = \eta_j$ for $j < k$; being composed of strictly increasing maps is itself strictly increasing, so that $z < v$ implies $x_\sigma < x_\eta$. It remains to consider the case $k = n$. In this case, $\sigma = \upsilon 0$, $\eta = \upsilon 1$, with $\upsilon \in \Sigma^{n-1}$. Therefore, $x_\sigma = M_\upsilon(0)$, which is smaller than $x_\eta = M_\upsilon(\frac{1}{2})$.

Let us now prove the third statement of the lemma by induction on n. When $n = 0$, we have that $I_\emptyset = [0,1]$ and $x_\emptyset = M_\emptyset(0) = 0$ (because $M_\emptyset$ is the identity); also $x_{\hat\sigma} = 1$, because $\emptyset$ is 1^0, so that $x_{\hat\sigma} = 1$ by definition, so that equation (2.10) holds. When $n > 0$, $I_\sigma = [M_\sigma(0), M_\sigma(1)] = [x_\sigma, M_\sigma(1)]$: we have to prove that $M_\sigma(1) = x_{\hat\sigma}$. Clearly, when $\sigma = 1^n$ $M_\sigma(1) = 1$ and, by the definition above, $x_{\hat\sigma} = 1$. Suppose that $M_\sigma(1) = x_{\hat\sigma}$ holds for any $\sigma \in \Sigma^n$. This is clearly true for $n = 1$, since either $\sigma = 0$, $\hat\sigma = 1$ and $M_0(1) = M_1(0) = 1/2 = x_1$, or $\sigma = 1$, $M_1(1) = 1$ and by definition $x_{\hat\sigma} = 1$. Consider now a $\sigma \in \Sigma^{n+1}$. Write $\sigma = \eta i$ with $\eta \in \Sigma^n$, $i = 0, 1$. In the first case,

$$M_\sigma(1) = M_\eta M_0(1) = M_\eta M_1(0) = M_{\eta 1}(0) = x_{\eta 1}$$

and clearly $\eta 1 = \hat\sigma$. In the second case, suppose $\eta \neq 1^n$, since the opposite instance means $\sigma = 1^{n+1}$, which was treated above. Then, using the induction hypothesis and the fact that $M_0(0) = 0$ we obtain

$$M_\sigma(1) = M_\eta M_1(1) = M_\eta(1) = M_{\hat\eta}(0) = M_{\hat\eta} M_0(0) = M_{\hat\eta 0}(0) = x_{\hat\eta 0}.$$

Since $\hat\sigma = \widehat{\eta 1} = \hat\eta 0$, the thesis follows. $\square$

Lemma 2.3. *Let Σ^n be as in Definition 2.1 and let x_σ, I_σ, for $\sigma \in \Sigma^n$, be defined as in Lemma 2.2, equations (2.9) and (2.10). Then, for any $n \in \mathbf{N}$, $\sigma \in \Sigma^n$*

$$?(x_\sigma) = \sum_{j=1}^{n} \sigma_j 2^{-j} = 2^{-n}\Theta(\sigma) \tag{2.11}$$

and

$$\mu(I_\sigma) = 2^{-n}. \tag{2.12}$$

Proof. Let us first prove equation (2.11). From equation (2.3), it follows that $?(x_\sigma) = P_\sigma(0)$ for any $\sigma \in \Sigma$. Let us use induction again. For $n = 0$, we have that $\sigma = \emptyset$ and equation (2.8) implies that $\Theta(\emptyset) = 0 = ?(0)$. For $n = 1$, we have that $x_0 = 0$ and $?(0) = 0$; $x_1 = \frac{1}{2}$ and $?(x_1) = \frac{1}{2}$, which again confirms equation (2.11). Next, suppose that equation (2.11) holds in Σ^n and let us compute $?(x_\sigma)$, with $\sigma \in \Sigma^{n+1}$. Clearly, $\sigma = i\eta$, with $i = 0$ or $i = 1$, $\eta \in \Sigma^n$. Therefore,

$$?(x_\sigma) = ?(x_{i\eta}) = P_i(?(x_\eta)) = P_i\left(\sum_{j=1}^{n} \eta_j 2^{-j}\right),$$

Since $P_i(x) = i/2 + x/2$, we find

$$?(x_{i\eta}) = i\, 2^{-1} + \sum_{j=1}^{n} \eta_j 2^{-j-1},$$

which proves formula (2.11).

Let us now compute $\mu(I_\sigma) = \mu([x_\sigma, x_{\hat\sigma}]) = ?(x_{\hat\sigma}) - ?(x_\sigma)$. When $n = 0$, $\sigma = \emptyset$, we have that $I_\sigma = [0,1]$ so that $\mu(I_\sigma) = 1$. When $\sigma \neq 1^n$ we can use equation (2.11), to obtain $?(x_{\hat\sigma}) - ?(x_\sigma) = 2^{-n}[\Theta(\hat\sigma) - \Theta(\sigma)] = 2^{-n}$, where we used that $\Theta(\hat\sigma) = \Theta(\sigma) + 1$, since, by Lemma 2.2, $\Theta(\hat\sigma)$ is the successor of $\Theta(\sigma)$ in $\{0, 1, 2, \ldots, 2^n - 1\}$. If $\sigma = 1^n$, then $x_{\hat\sigma} = 1$ and $?(x_{\hat\sigma}) - ?(x_\sigma) = ?(1) - ?(x_{1^n}) = 1 - 2^{-n}(2^n - 1) = 2^{-n}$ where the value of $?(x_{1^n})$ follows from (2.11) and the fact that 1^n is the last word in the lexicographical ordering $\prec$. Thus, (2.12) holds in this case, as well. $\square$

3. *Stern–Brocot sequences and Möbius IFS*

In this section, we demonstrate that the boundary points of the Möbius IFS partitions described in Section 2 coincide with the classical Stern–Brocot sequences [**10**, **15**, **47**]. We also define Farey sequences and further notations for later usage.

DEFINITION 3.1. The Stern–Brocot sequence $\mathcal{B}^n \subset \mathbf{Q}$ is defined for any $n \in \mathbf{N}$ by induction: $\mathcal{B}^0 = \{0, 1\}$ and $\mathcal{B}^{n+1}$ is the increasingly ordered union of $\mathcal{B}^n$ and the set of mediants of consecutive terms of $\mathcal{B}^n$. The mediant, or Farey sum, of two rational numbers written as irreducible fractions is

$$\frac{p}{q} \oplus \frac{r}{s} = \frac{p+r}{q+s}. \tag{3.1}$$

Denote points in the ordered $\mathcal{B}^n$ sequence as $\mathcal{B}^n = \{x_0^n, x_1^n, \dots\}$

Observe that the mediant of two numbers is intermediate between the two. Moreover, the definition implies that the cardinality of $\mathcal{B}^n$ obeys the rules $\#(\mathcal{B}^0) = 2$, $\#(\mathcal{B}^{n+1}) = 2\#(\mathcal{B}^n) - 1$, so that $\#(\mathcal{B}^n) = 2^n + 1$. Therefore, the induction rule can be written as

$$\mathcal{B}^n = \{x_0^n, x_1^n, x_2^n, \dots, x_{2^n}^n\} \Rightarrow \mathcal{B}^{n+1} = \{x_0^n, x_0^n \oplus x_1^n, x_1^n, x_1^n \oplus x_2^n, x_2^n, \dots, x_{2^n}^n\}. \tag{3.2}$$

LEMMA 3.1. *Let Σ^n be as in Definition 2.1 and let x_σ, I_σ, for $\sigma \in \Sigma^n$, be defined as in Lemma 2.2, equations (2.9) and (2.10). For any $n \in \mathbf{N}$, the increasingly ordered set $\{\{x_\sigma, \sigma \in \Sigma^n\}, 1\}$ coincides with the nth Stern–Brocot sequence $\mathcal{B}^n$.*

Proof. Observe that $\{\{x_\sigma, \sigma \in \Sigma^n\}, 1\}$ is the set of extrema of the intervals I_σ, with $\sigma \in \Sigma^n$, which can be increasingly ordered according to Lemma 2.2. For $n = 0$, one has $\{x_\emptyset, 1\} = \{0, 1\}$, which can also be written as $\mathcal{B}^0 = \{\frac{0}{1}, \frac{1}{1}\}$. It is then enough to show that the induction property (3.2) holds for the sequence of sets $\{\{x_\sigma, \sigma \in \Sigma^n\}, 1\}$. Let $\sigma \in \Sigma^n$. Each $I_\sigma = [x_\sigma, x_{\hat\sigma}]$ splits into $I_{\sigma 0}$ and $I_{\sigma 1}$, as seen above in Lemma 2.1. Because of equation (2.7) the points x_σ and $x_{\hat\sigma}$ of the nth set also belong to the $n+1$-th set: in fact, they coincide with $x_{\sigma 0}$ and $x_{\hat\sigma 0}$. It remains to show that the intermediate point $x_{\sigma 1}$ is a rational number that fulfills the Farey sum rule. We now prove by induction on the length n of σ that

$$M_\sigma(0) = x_\sigma = \frac{p}{q},\ M_\sigma(1) = x_{\hat\sigma} = \frac{\hat p}{\hat q}, \tag{3.3}$$

where p and q, $\hat p$ and $\hat q$ are relatively prime integers with

$$\Delta\left(\frac{p}{q}, \frac{\hat p}{\hat q}\right) = \hat p q - \hat q p = 1, \tag{3.4}$$

and

$$M_\sigma(x) = \frac{(\hat p - p)x + p}{(\hat q - q)x + q}. \tag{3.5}$$

Indeed, this is certainly true for $n = 0$ with $p = 0$, $q = \hat p = \hat q = 1$, and suppose that the claim holds for all σ of length n. Consider a word of length $n+1$, say of the form $\sigma 1$ with $\sigma \in \Sigma^n$. Then,

$$x_{\sigma 1} = M_{\sigma 1}(0) = M_\sigma\left(\frac{1}{2}\right) = \frac{p + \hat p}{q + \hat q},$$

and easy inspection based on explicit computation of (3.4) shows that the Farey sum property (3.4) holds for both pairs $\frac{p}{q}, \frac{p+\hat{p}}{q+\hat{q}}$ and $\frac{p+\hat{p}}{q+\hat{q}}, \frac{\hat{p}}{\hat{q}}$. In particular, $\frac{p+\hat{p}}{q+\hat{q}}$ is in its lowest form, that is, in it $p+\hat{p}$ and $q+\hat{q}$ are relative primes. Finally,

$$M_{\sigma 1}(x) = M_\sigma(M_1(x)) = \frac{(\hat{p}-p)\frac{1}{2-x}+p}{(\hat{q}-q)\frac{1}{2-x}+q} = \frac{(\hat{p}+p)-px}{(\hat{q}+q)-qx} = \frac{(\hat{p}-(p+\hat{p}))x+(p+\hat{p})}{(\hat{q}-(q+\hat{q}))x+(q+\hat{q})}$$

so (3.5) is also preserved. The proof for $(n+1)$-long words of the form $\sigma 0$ is analogous. □

Because of the previous Lemma, the cylinder I_σ, when $\sigma \in \Sigma^n$ and $\Theta(\sigma) = j$, can be equivalently indicated as $[x_\sigma, x_{\hat{\sigma}}]$, following equation (2.10) and $[x^n_j, x^n_{j+1}] = [\frac{p^n_j}{q^n_j}, \frac{p^n_{j+1}}{q^n_{j+1}}]$, with p^n_j and q^n_j, p^n_{j+1} and q^n_{j+1} relatively prime integers. We will find convenient to use the shorthand notation $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$ introduced in equation (3.3) dropping the indices n and j when no confusion can arise.

Closely related objects are the so-called Farey sequences $\mathcal{F}^m$. Let us give their definition, which will come to use in the next sections.

Definition 3.2. The Farey sequence $\mathcal{F}^m \subset \mathbf{Q}$ is the ordered set of irreducible rationals p/q in [0,1] whose denominator is less than, or equal to, $m \in \mathbf{N}$.

4. *Regularity of the Inverse ? measure*

Thanks to the results of the previous sections we can easily prove that the Minkowski's inverse question mark measure is regular, Theorem 1.3. In essence, the proof is an exploitation of the fact that Minkowski's question mark function is Hölder continuous.

Proof. Theorem 1.3. For any $r > 0$, let n be such that $2^{-n} < r \leqslant 2^{-n+1}$. Then, the ball of radius r at any $y \in [0,1]$ contains a dyadic interval D_y of diameter 2^{-n}. Let $\sigma \in \Sigma^n$ be the symbolic word that verifies $?(I_\sigma) = D_y$, the existence of which follows from Lemma 2.2. Clearly, $\mu^{-1}(B_r(y)) \geqslant \mu^{-1}(D_y)$. Since μ^{-1} is the inverse measure of μ, $\mu^{-1}(D_y) = |I_\sigma|$.

According to equations (3.3) and (3.4), $|I_\sigma| = 1/(q\hat{q})$. Furthermore, the recursive rule (3.1) implies that $q, \hat{q} \leqslant 2^n$, so that $|I_\sigma| \geqslant 1/q\hat{q} \geqslant 2^{-2n}$. Hence,

$$\mu^{-1}(B_r(y)) \geqslant 2^{-2n} \geqslant \frac{r^2}{4}. \tag{4.1}$$

Let now $r = 1/s$. Then, for any $\eta > 0$, there exists $\bar{s}$ such that $e^{-\eta s}$ is smaller than $s^{-2}/4$ for $s > \bar{s}$, and so

$$\mu^{-1}(B_{1/s}(y)) \geqslant e^{-\eta s}$$

for all $y \in [0,1]$, thereby proving that Criterion λ^* holds. □

5. *Regularity of a measure via cylinder estimates*

The case of the inverse Minkowski's question mark measure is particularly simple, since we have been able to prove the strong estimate (4.1). When such result is not available, we can have recourse to cylinder estimates, as follows. Assume that we are still in the case when $\mathcal{A} = [0,1]$. Suppose that there is a countable family of partitions of [0,1] by adjacent intervals labeled by words σ with letters in a finite alphabet $\{0,\ldots,M\}$, so that for any n

$$[0,1] = \bigcup_{\sigma \in \Sigma^n} I_\sigma.$$

Define the set $L^n(\alpha) \subset \Sigma^n$, for $n \in \mathbf{N}$ and $\alpha > 0$, as

$$L^n(\alpha) = \left\{ \sigma \in \Sigma^n \text{ s.t. } |I_\sigma| \geqslant \frac{\alpha}{n+1} \right\}. \tag{5.1}$$

Similarly, let $S^n(\alpha)$ be the complement of $L^n(\alpha)$ in Σ^n. Then, we can use the following Proposition.

Proposition 5.1. *Suppose that a measure μ supported in* [0,1] *is such that: (i) There exists $\pi > 0$ such that $\mu(I_\sigma) \geqslant \pi^n$ for all $\sigma \in \Sigma^n$ and all $n \in \mathbf{N}$; (ii) For any $\alpha > 0$, there exists C_α such that $\#(L^n(\alpha)) \leqslant C_\alpha$ for any n; (iii) The maximum length of cylinders in Σ^n is infinitesimal when n tends to infinity: $l_n = \max\{|I_\sigma|, |\sigma| = n\} \to 0$. Then, the measure μ satisfies Criterion λ^* and hence is regular.*

Proof. Let $\alpha > 0$ be small, s large and $n \in \mathbf{N}$ such that $n < \alpha s \leqslant n+1$. Consider points $x \in [0,1]$ which belong to a 'short' interval: there exists $\bar\sigma \in S^n(\alpha)$ such that $x \in I_{\bar\sigma}$. Since $|I_{\bar\sigma}| < \alpha/(n+1)$, this latter is enclosed in the ball of radius $1/s$ at x. Therefore,

$$\mu([x-1/s, x+1/s]) \geqslant \mu(I_{\bar\sigma}) \geqslant \pi^n \geqslant \pi^{\alpha s} = e^{-\alpha \log(\pi^{-1}) s}.$$

Letting $\eta = \alpha \log(\pi^{-1})$, the above means that such x belongs to the set $\Lambda(\eta; s)$ (see definition (1.4)), so that

$$\bigcup_{\sigma \in S^n(\alpha)} I_\sigma \subset \Lambda(\eta; s).$$

Taking the Lebesgue measure of both sets and using (ii), one has

$$|\Lambda(\eta; s)| \geqslant \left| \bigcup_{\sigma \in S^n(\alpha)} I_\sigma \right| = 1 - \left| \bigcup_{\sigma \in L^n(\alpha)} I_\sigma \right| \geqslant 1 - \#(L^n(\alpha)) l_n \geqslant 1 - C_\alpha l_n.$$

Because of (iii) the final expression at right-hand side tends to one as s, hence n, tends to infinity, which proves that Criterion λ^* holds, and so $\mu \in \mathbf{Reg}$. □

Remark 5.1. Note that for partitions $\{I_\sigma, \sigma \in \Sigma^n\}$ generated by an IFS with finitely many maps, condition (i) is always verified setting $\pi = \min_i\{\pi_i\}$. It can also be shown that if the φ_i are contractions in an IFS with $\mathcal{A} = [0,1]$, then (iii) also holds.

6. *First proof of regularity of Minkowski's ? measure*

We are now in position to use Proposition 5.1 to obtain our first proof of regularity of Minkowski's question mark measure, Theorem 1.2. We will also use the results of Lemmas 2.1– 2.3.

Proof of Theorem 1.2. First observe that, by Remark 5.1 we can put $\pi = \frac{1}{2}$ in Proposition 5.1(i). Next, we exploit the fact that, away from the fixed points at zero and one, the IFS maps M_i are strictly contractive. Let s be a positive integer, $s \geqslant 3$ and consider an interval $J = [a,b] \subseteq [\frac{1}{s}, \frac{1}{2}]$. Applying the Möbius transformation M_0 to this interval, we obtain $M_0(J) = [\frac{a}{1+a}, \frac{b}{1+b}] \subseteq [\frac{1}{s+1}, \frac{1}{2}]$ and

$$|M_0(J)| = \frac{|b-a|}{(1+a)(1+b)} \leqslant \frac{|b-a|}{(1+1/s)^2} = |J| \left(\frac{s}{1+s} \right)^2.$$

On the other hand, if $J \subseteq [\frac{1}{2}, 1]$, then $M_0(J) \subseteq [\frac{1}{3}, \frac{1}{2}]$ and similarly as before $|M_0(J)| \leqslant |J|(\frac{2}{3})^2$. By symmetry, if $J \subseteq [\frac{1}{2}, 1-\frac{1}{s}]$, then $M_1(J) \subseteq [\frac{1}{2}, 1-\frac{1}{1+s}]$ and $|M_1(J)| \leqslant |J|(s/(1+s))^2$, while for $J \subseteq [0, \frac{1}{2}]$, we have $M_1(J) \subseteq [\frac{1}{2}, 1-\frac{1}{3}]$, and $|M_1(J)| \leqslant |J|(\frac{2}{3})^2$.

Thus, if $J \subseteq [\frac{1}{s}, \frac{1}{2}]$ or $J \subseteq [\frac{1}{2}, 1-\frac{1}{s}]$, then for $i = 0, 1$ we have that $M_i(J) \subseteq [\frac{1}{1+s}, \frac{1}{2}]$ or $M_i(J) \subseteq [\frac{1}{2}, 1-\frac{1}{1+s}]$, and $|M_i(J)| \leqslant |J|(s/(1+s))^2$. This can be iterated so that, for J in the above conditions and $\sigma = (\sigma_1, \dots, \sigma_k) \in \Sigma^k$

$$|M_\sigma(J)| \leqslant |J| \left(\frac{s}{1+s}\right)^2 \left(\frac{s+1}{1+s+1}\right)^2 \cdots \left(\frac{s+k-1}{s+k}\right)^2 = |J| \left(\frac{s}{s+k}\right)^2. \tag{6.1}$$

We also get in the same way by induction on k that $M_0^k([0,1]) = [0, 1/(k+1)]$, $M_1^k[0,1] = [1-1/(k+1), 1]$, while for all other words in Σ^k

$$I_\sigma = (M_{\sigma_1} \circ \cdots \circ M_{\sigma_k})([0,1]) \subseteq \left[\frac{1}{k+1}, \frac{1}{2}\right] \text{ or } I_\sigma \subseteq \left[\frac{1}{2}, 1 - \frac{1}{k+1}\right].$$

Based on these facts, simple induction yields $|I_\sigma| \leqslant \frac{1}{|\sigma|+1}$ for all σ.

Choose and fix a large integer m. Let $n > 2m^2$ and suppose that for some $\sigma = (\sigma_1, \dots, \sigma_n) \in \Sigma^n$ there is an integer r such that $1 \leqslant r < n - m$ and $\sigma_{n-r} \neq \sigma_{n-r+1}$. Then, according to the above inequalities

$$|(M_{\sigma_{n-r+1}} \circ \cdots \circ M_{\sigma_n})([0,1])| \leqslant \frac{1}{r+1}, \tag{6.2}$$

and the above interval is contained in $I_{\sigma_{n-r+1}}$. Since $\sigma_{n-r} \neq \sigma_{n-r+1}$ it follows that

$$|(M_{\sigma_{n-r}} \circ M_{\sigma_{n-r+1}} \circ \cdots \circ M_{\sigma_n})([0,1])| \leqslant \frac{1}{r+1} \left(\frac{2}{3}\right)^2.$$

Observe that the interval in the last equation is either enclosed in $[\frac{1}{s}, \frac{1}{2}]$ or in $[\frac{1}{2}, 1-\frac{1}{s}]$, according to the value of σ_{n-r}, with $s = 3$. We can, therefore, apply the estimate (6.1) with $k = n - r - 1$, to get

$$|(M_{\sigma_1} \circ \cdots \circ M_{\sigma_{n-r}} \circ \cdots \circ M_{\sigma_n})([0,1])| \leqslant \frac{1}{r+1} \left(\frac{2}{3}\right)^2 \left(\frac{3}{n-r+2}\right)^2 \leqslant \frac{8}{m^2} \frac{1}{n+1}.$$

To obtain the last inequality we used that $2(r+1)(n-r+2)^2 \geqslant m^2(n+1)$ because $n - r + 2 > m$ and $n > 2m^2$ (it just suffices to consider the cases $r \geqslant n/2$ and $r < n/2$ separately).

Hence, if $\sigma \in \Sigma^n$ and $|I_\sigma| > \frac{8}{m^2}\frac{1}{n+1}$, then σ must be either of the form $\sigma = \eta 0^{n-m} = (\eta_1, \dots, \eta_m, 0, \dots, 0)$ or $\sigma = \eta 1^{n-m} = (\eta_1, \dots, \eta_m, 1, \dots, 1)$, with arbitrary $\eta \in \Sigma^{n-m}$. If now we choose m such that $\frac{8}{m^2} < \alpha$ we have that the cardinality of $L^n(\alpha)$ is less than $2 \cdot 2^m$ for all n larger than $2m^2$, and clearly also bounded by a constant for $n < 2m^2$, so that the hypothesis (ii) of Proposition 5.1 holds.

Finally, we employ (6.2) (which we shall re-derive in equation (7.11)) that $|I_\sigma| \leqslant \frac{1}{n+1}$ for all $\sigma \in \Sigma^n$, which implies the remaining condition (iii) in the hypothesis of Proposition 5.1, and the theorem is proven.

Remark 6.1. Observe that letting $m = \lceil \sqrt{8/\alpha}\, \rceil$ for any $\alpha > 0$, when $n > 16/\alpha$, the intervals I_σ, with $\sigma \in \Sigma^n$ for which $|I_\sigma| \geqslant \alpha/(n+1)$ are necessarily labeled by $\eta 1^{n-m}$ or $\eta 0^{n-m}$ with an $\eta \in \Sigma^m$ and their cardinality is therefore bounded by $2^{2+\sqrt{8/\alpha}}$. We show in the following that this estimate, although sufficient for the proof of regularity, fails to describe accurately the words in $L^n(\alpha)$, which on the contrary have a remarkable arithmetical structure.

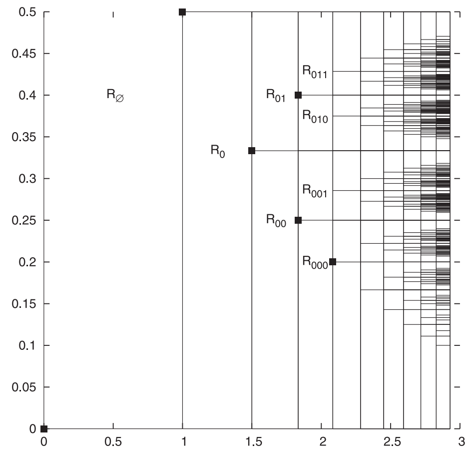

FIGURE 1. *Rectangles $R_\sigma = J_n \times I_\sigma$, where $n = |\sigma|$, bounded by the continuous lines. The Farey sequence $\mathcal{F}^5$ is represented by small boxes: see text for details. Because of symmetry, only half of the figure is displayed.*

7. *Arithmetical properties of Möbius partitions*

Proposition 5.1 shows that regularity of Minkowski's question mark measure can be seen as a consequence of the distribution of 'geometrical' lengths of cylinders. To appreciate fully its subtleties, in this section we examine more deeply the structure of the Möbius partitions of the unit interval, whose extremes compose the Stern–Brocot sequences. The fundamental results of this section are Proposition 7.1 and Corollary 7.2, which describe the set of 'large' intervals $L^n(\alpha)$ (see (5.1)) of these partitions.

We find that $L^n(\alpha)$ is directly determined by an arithmetical set: for any value of $\alpha > 0$, define $\mathbf{Q}_\alpha$ by considering all irreducible fractions with denominator smaller than or equal to $1/\sqrt{\alpha}$:

$$\mathbf{Q}_\alpha = \left\{\zeta \in \mathbf{Q} \cap [0,1] \text{ s.t. } \zeta = \frac{p}{q},\ p,q \in \mathbf{N},\ p,q \text{ relative primes, and } 1 \leqslant q^2 \leqslant \frac{1}{\alpha}\right\}. \tag{7.1}$$

PROPOSITION 7.1. *The set $L^n(\alpha) \subset \Sigma^n$ can be characterized as follows: for any $0 < \alpha < 1$, there exists $\bar{n} \in \mathbf{N}$ such that for any $n \geqslant \bar{n}$*

$$L^n(\alpha) = \{\sigma \in \Sigma^n \text{ s.t. } x_\sigma \in \mathbf{Q}_\alpha \text{ or } x_{\hat{\sigma}} \in \mathbf{Q}_\alpha\}. \tag{7.2}$$

COROLLARY 7.2. *Let $L^n(\alpha) \subset \Sigma^n$, for $n \in \mathbf{N}$ and $\alpha > 0$, be as in definition (5.1). Then, for any $\alpha > 0$, the cardinality of $L^n(\alpha)$ is bounded: there exists $C_\alpha \in \mathbf{N}$ so that for all $n \in \mathbf{N}$*

$$\#(L^n(\alpha)) \leqslant C_\alpha. \tag{7.3}$$

REMARK 7.1. From Definition 3.2, it appears that letting $m = \lfloor 1/\sqrt{\alpha} \rfloor$ one has $\mathbf{Q}_\alpha = \mathcal{F}^m$, the mth Farey sequence. In particular, this implies that the cardinality of $L^n(\alpha)$ is asymptotically $3/(\alpha\pi^2)$ when α tends to zero, for large n [**15**]. This is the optimal estimate, which improves the results of Remark 6.1. In addition, Proposition 7.1 exactly characterizes the words in $L^n(\alpha)$ revealing their arithmetical nature.

The content of Proposition 7.1 is well exemplified in Figure 1: each cylinder I_σ, $\sigma \in \Sigma$, is uniquely associated with a rectangle $R_\sigma = J_n \times I_\sigma$, where $n = |\sigma|$. The horizontal sides

$J_n = [\zeta_n, \zeta_{n+1}]$ are constructed in this way: $J_0 = [0, 1]$, J_{n+1} is adjacent to the right of J_n for any n and $|J_n| = 1/(n+1)$, so that $\zeta_n = -1 + \sum_{l=0}^{n} 1/(l+1)$.

The choice of the horizontal segments J_n implies that $\sigma \in L^n(\alpha)$ if and only if the *vertical* side of the rectangle R_σ is *larger* than α times the horizontal. Therefore, by suitably fixing the ratio of the graphical units of the horizontal and vertical axis in the figure to the constant α, words in $L^n(\alpha)$ *appear to the eye* as tall (taller than wider) rectangles, while rectangles associated with words in $S^n(\alpha)$ are wide. For instance, Figure 1 is such that a unit in the vertical direction has a graphical length of six horizontal units: a rectangle whose vertical size is one-sixth of the horizontal appears as a square. Therefore, here $\alpha = 1/6$ and words in $L^n(\alpha)$ appear as tall rectangles.

According to Proposition 7.1, for large n, such functions I_σ must have a point of $\mathbf{Q}_{1/6} = \mathcal{F}^2$ as extremum. This is clearly observed, since $\mathcal{F}^2 = \{\frac{0}{1}, \frac{1}{2}, \frac{1}{1}\}$. In addition, consider in the figure corners of the rectangles that lie on the vertical segment at abscissa ζ_n. Note that their ordinates compose the Stern–Brocot sequence $\mathcal{B}^n$. In the figure, we have also plotted the sequence $\mathcal{F}^5 \subset \mathcal{B}^5$ as small squares placed at the corresponding ordinate, at the abscissa of the *first* Stern-Brocot sequence in which they appear — which can be earlier than $\mathcal{B}^5$, as explained below by the notion of *depth*.

The remainder of this section contains the proof of these propositions. We have already introduced the complementary set of $L^n(\alpha)$ in Σ^n, which we denoted by $S^n(\alpha)$ (S for *small*):

$$S^n(\alpha) = \{\sigma \in \Sigma^n \text{ s.t. } |I_\sigma| < \alpha/(n+1)\}. \tag{7.4}$$

In the discussion below, we always set $I_\sigma = [x_\sigma, x_{\tilde{\sigma}}]$, $x_\sigma = \frac{p}{q}$, $x_{\hat{\sigma}} = \frac{\hat{p}}{\hat{q}}$, with relative primes p, q and $\hat{p}, \hat{q}$, as described in Section 3. The property of Farey fractions, equation (3.4), imply that

$$|I_\sigma| = x_\sigma - x_{\hat{\sigma}} = \frac{1}{q\hat{q}}. \tag{7.5}$$

This permits to assess the useful condition

$$\sigma \in S^n(\alpha) \iff q\hat{q} > (n+1)/\alpha. \tag{7.6}$$

Finally, we define a subset of Σ by requiring that neither extremum of I_σ belongs to $\mathbf{Q}_\alpha$:

$$\mathcal{E} = \left\{\sigma \in \Sigma \text{ s.t. } q > \sqrt{1/\alpha},\ \hat{q} > \sqrt{1/\alpha}\right\}. \tag{7.7}$$

Note that in this definition we do not let $\mathcal{E}$ be a subset of Σ^n, but rather of the full set Σ: this is necessary to study different IFS partitions.

To prove the above results we will now proceed through several steps, some of which can be considered as sublemmas in their own right, whose proof terminates at a triangle $\triangle$. For simplicity of notation, put $a = 1/\alpha$ in what follows. It might be helpful to follow the proofs with the aid of Figure 1.

SUBLEMMA 7.1. *The class $\mathcal{E}$ is stable under successive partitions:*

$$\sigma \in \mathcal{E} \Rightarrow \sigma\eta \in \mathcal{E}, \qquad \eta \in \Sigma. \tag{7.8}$$

Proof. In fact, the endpoints of $I_{\sigma\eta}$ belong to the Stern–Brocot sequence $\mathcal{B}^{|\sigma|+|\eta|}$. Let $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$. When $|\eta| = 1$, because of the construction rule (3.2), the denominators of the endpoints of $I_{\sigma\eta}$ are $\{q, q+\hat{q}, \hat{q}\}$, which are all larger than $\sqrt{a}$. Induction extends the result to general $\eta \in \Sigma$. $\triangle$ $\square$

A second claim considers words that are in $\mathcal{E}$ and at the same time are associated with 'small' intervals, that is, wide rectangles in Figure 1. This class is also stable under successive partitions:

SUBLEMMA 7.2.

$$\sigma \in [\mathcal{E} \cap S^{|\sigma|}(\alpha)] \Rightarrow \sigma\eta \in [\mathcal{E} \cap S^{|\sigma|+|\eta|}(\alpha)], \qquad \eta \in \Sigma. \tag{7.9}$$

Proof. The implication regarding $\mathcal{E}$ has just been proven in (7.8). In addition, letting $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$, the left-hand side of (7.9) means that $q^2 > a$, $\hat{q}^2 > a$ and $q\hat{q} > a(|\sigma| + 1)$, the last inequality following from (7.6). Then, $I_{\sigma 0} = [\frac{p}{q}, \frac{\hat{p}+p}{\hat{q}+q}]$ and

$$q(\hat{q} + q) = q\hat{q} + q^2 > a(|\sigma| + 1) + a = a(|\sigma 0| + 1).$$

A similar estimate clearly holds for $I_{\sigma 1}$, and by induction (7.9) follows. □

Suppose now that $\sigma \in \mathcal{E}$, but do not require that $|I_\sigma| < \alpha/(|\sigma| + 1)$, that is, σ may *not* belong to $S^{|\sigma|}(\alpha)$, a case that often occurs (like for example, R_{010} in Figure 1). Then, we can prove that there is a subdivision of I_σ whose intervals are all smaller than the threshold in (7.4), that is, the words $\sigma\eta$ in this subdivision all belong to $S^{|\sigma|+|\eta|}(\alpha)$. Actually, we shall prove more:

SUBLEMMA 7.3. *For every $\sigma \in \mathcal{E}$, there exists a $k_1(\sigma) \in \mathbf{N}$ such that*

$$\sigma\eta \in [\mathcal{E} \cap S^{|\sigma|+|\eta|}(\alpha)], \qquad \eta \in \Sigma,\ |\eta| \geqslant k_1(\sigma). \tag{7.10}$$

Proof. The proof of this implication is rather long. The part regarding $\mathcal{E}$ has been already proven, equation (7.8). Let us use again the notation $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$. Suppose that $q < \hat{q}$ without loss of generality: the opposite case can be dealt with similarly, by replacing the symbols q with $\hat{q}$ and 0 with 1 in what follows. In this $q < \tilde{q}$ case, among all intervals $I_{\sigma\eta}$, with $|\eta| = k$ fixed, the largest is $I_{\sigma 0^k}$, as we are going to prove in the next two paragraphs.

Since $I_{\sigma\eta} = M_\sigma(I_\eta)$, let us first study the intervals I_η and prove that $|I_\eta| \leqslant |I_{0^k}|$ for any $\eta \in \Sigma^k$ and for any $k \in \mathbf{N}$. This is clearly true for $k = 0$ and $k = 1$ by direct inspection. Suppose that it holds true for a certain k and let us consider symbolic words of length $k + 1$. As was mentioned in Section 2, the associated intervals are symmetric around the point $\frac{1}{2}$, it is sufficient to study the set $\{I_{0\eta},\ \eta \in \Sigma^k\}$. We now want to use the mean value theorem: the equality $I_{0\eta} = M_0(I_\eta)$ holds by definition, so that $|I_{0\eta}| = M_0'(z_\eta)|I_\eta|$, where z_η is a point in I_η. Now, $M_0'(x) = 1/(x + 1)^2$, so that $M_0'(x)$ is strictly decreasing on [0,1]. Since the ordering of intervals in Lemma 2.2 implies that $\inf I_\eta \geqslant \sup I_{0^k}$ for any $\eta \in \Sigma^k$ different from 0^k, while $|I_\eta| \leqslant |I_{0^k}|$ by the induction hypothesis, it follows that $|I_{0\eta}| = M_0'(z_\eta)|I_\eta| \leqslant M_0'(z_{0^k})|I_{0^k}| = |I_{0^{k+1}}|$, which completes the proof by induction of this part. We can also explicitly compute

$$|I_\eta| \leqslant |I_{0^k}| = |I_{1^k}| = \frac{1}{k+1}, \tag{7.11}$$

which has been used in Section 6 and will be useful again below.

Recall now that $I_{\sigma\eta} = M_\sigma(I_\eta)$. When $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$, the Möbius transformation M_σ is given by equation (3.5), so that $M_\sigma'(x) = [(\hat{q} - q)x + q]^{-2}$, using also equation (3.4). Since $q < \hat{q}$, $M_\sigma'(x)$ is decreasing on [0,1]. Using again the mean value theorem and the fact that $I_{\sigma 1^k}$ is the leftmost interval among all $I_{\sigma\eta}$ with $\sigma \in \Sigma^k$, we can write $|I_{\sigma\eta}| = M_\sigma'(z_\eta^*)|I_\eta|$, with some $z_\eta^* \in I_\eta$, from which we conclude that $|I_{\sigma\eta}| \leqslant |I_{\sigma 0^k}|$ for any $\eta \in \Sigma^k$, $\sigma \in \Sigma$.

The length of $I_{\sigma 0^k}$ can be easily computed from the explicit representation

$$I_{\sigma 0^k} = \left[\frac{p}{q}, \frac{\hat{p} + kp}{\hat{q} + kq}\right],$$

which yields

$$|I_{\sigma 0^k}|^{-1} = q(\hat{q} + kq) = q\hat{q} + kq^2. \tag{7.12}$$

Call $k_1(\sigma)$ the least k such that $k > [a(n+1) - q\hat{q}]/(q^2 - a)$, where $n = |\sigma|$. Since $q^2 - a > 0$ this implies that, for $k \geqslant k_1(\sigma)$,

$$q\hat{q} + kq^2 > a(n+k+1). \tag{7.13}$$

Now inequality (7.13) and the above reasoning imply that

$$|I_{\sigma\eta}| \leqslant |I_{\sigma 0^k}| < \frac{\alpha}{n+k+1} = \frac{\alpha}{|\sigma\eta|+1}$$

for all $\eta \in \Sigma^k$, $k \geqslant k_1(\sigma)$, which proves (7.10). $\square$

We have just proven that starting from a word/interval in $\mathcal{E}$ and taking successive partitions of it we end up in $\mathcal{E} \cap S^m(\alpha)$ for all $m \in \mathbf{N}$ larger than a certain value. We now need to examine the fate of intervals which do *not* belong to $\mathcal{E}$. Let us therefore take a general $\sigma \in \Sigma$, not necessarily in $\mathcal{E}$ and consider the associated interval $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$, where either $q^2 \leqslant a$ or $\hat{q}^2 \leqslant a$ may happen.

Sublemma 7.4. *For all $\sigma \in \Sigma$, there exists $k_2(\sigma) \in \mathbf{N}$ such that*

$$\sigma 0^k 1 \in [\mathcal{E} \cap S^{|\sigma|+k+1}(\alpha)], \qquad k \geqslant k_2(\sigma). \tag{7.14}$$

Proof. By direct computation one gets

$$I_{\sigma 0^k 1} = \left[\frac{\hat{p} + (k+1)p}{\hat{q} + (k+1)q}, \frac{\hat{p} + kp}{\hat{q} + kq}\right]. \tag{7.15}$$

Observe that these intervals approach $x_\sigma = \frac{p}{q}$ when k grows. Actually, $I_{\sigma 0^k 1}$ is the second interval to the right of the point $\frac{p}{q}$ in the family $\{I_\eta,\ |\eta| = |\sigma| + k + 1\}$. It is clear that for sufficiently large k the squares of both denominators in equation (7.15) are larger than a, so that $\sigma 0^k 1 \in \mathcal{E}$. Moreover, since

$$|I_{\sigma 0^k 1}|^{-1} = [\hat{q} + (k+1)q](\hat{q} + kq) = k^2 q^2 + kq^2 + (2k+1)\hat{q}q + \hat{q}^2,$$

it is also clear that, for sufficiently large k, the right-hand side of the above equation is larger than $a(|\sigma| + k + 2)$, so that $\sigma 0^k 1 \in S^{|\sigma|+k+1}$. This proves (7.14). $\square$

We will use the above property for σ such that $I_\sigma = [\frac{p}{q}, \frac{\hat{p}}{\hat{q}}]$, with $q^2 \leqslant a$. We also need a symmetrical property, to be used when $\hat{q}^2 \leqslant a$. The previous technique yields that for all $\sigma \in \Sigma$ there is a $k_3(\sigma) \in \mathbf{N}$ such that

$$\sigma 1^k 0 \in [\mathcal{E} \cap S^{|\sigma|+k+1}(\alpha)], \qquad k \geqslant k_3(\sigma). \tag{7.16}$$

Here, $I_{\sigma 1^k 0}$ is the second interval to the left of the point $\frac{\hat{p}}{\hat{q}}$ in the family $\{I_\eta,\ |\eta| = |\sigma| + k + 1\}$. $\square$

This ends the sequence of sublemmas.

Proof of Proposition 7.1. We now go through a series of three levels $n = n_1, n_2, n_3$ at which we study the partitions Σ^n.

First level, n_1, when all elements of $\mathbf{Q}_\alpha$ appear at even positions in $\mathcal{B}^{n_1}$.

Consider the set $\mathbf{Q}_\alpha = \mathcal{F}^m$, with $m = \lfloor 1/\sqrt{\alpha} \rfloor$. For $\zeta \in \mathbf{Q}_\alpha$, let $n(\zeta)$ be the least n such that $\zeta \in \mathcal{B}^n$ (for every rational number there is such an n since rational numbers are mapped by Minkowski's question mark function into dyadic rationals, so every rational number is one of the x_σ, and then the existence of such an n is a consequence of Lemma 3.1). This number is called the *depth* of ζ in the Stern–Brocot tree [**15**]. It is standard to show that the maximum depth of ζ in $\mathbf{Q}_\alpha$ is m. In Figure 1, the set $\mathcal{F}^5$ is plotted, showing also the depth of different points. Moreover, since $\zeta \in \mathcal{B}^n$ implies that $\zeta \in \mathcal{B}^l$ for any $l \geqslant n$, it follows that $\zeta \in \mathcal{B}^n$ for all $\zeta \in \mathbf{Q}_\alpha$ and for all $n \geqslant m$. Let now $n_1 = m+1$, so that if $x_j^{n_1} \in \mathbf{Q}_\alpha$, then j is even: in fact, for any $\zeta \in \mathbf{Q}_\alpha$, there exists $j \in \{0, \dots, 2^m\}$ such that $\zeta = x_j^m$. At the next level, we have $\zeta = x_{2j}^{m+1}$, so that all elements in $\mathbf{Q}_\alpha$ appear with even indices in $\mathcal{B}^{n_1}$.

Consider now the set F of words in Σ^{n_1}, such that one endpoint of I_σ belongs to $\mathbf{Q}_\alpha$. Because of what we have just proven, no more than one endpoint of any such interval can belong to $\mathbf{Q}_\alpha$. Part these words in two disjoint groups, $F = F_l \uplus F_r$, according to whether the left or the right endpoint of $I_\sigma = [x_\sigma, x_{\hat{\sigma}}]$ lies in $\mathbf{Q}_\alpha$:

$$F_l = \{\sigma \in \Sigma^{n_1} \text{ s.t. } x_\sigma \in \mathbf{Q}_\alpha\},$$

$$F_r = \{\sigma \in \Sigma^{n_1} \text{ s.t. } x_{\hat{\sigma}} \in \mathbf{Q}_\alpha\}.$$

The symbol $\uplus$ indicates the disjoint union of two sets.

Apply now sublemmas (7.14) and (7.16) to define $K_2(F) = \max\{k_2(\sigma), \sigma \in F_l\}$ and $K_3(F) = \max\{k_3(\sigma), \sigma \in F_r\}$. Let $\kappa = \max\{K_2(F), K_3(F)\} + 1$. This defines the second level, $n_2 = n_1 + \kappa$.

Second level, $n_2 = n_1 + \kappa$, when properties (7.14) and (7.16) are realized for all words σ in $F \subset \Sigma^{n_1}$.

Among all words of length $n_2 = n_1 + \kappa$ we start by considering those that originate from a word $\sigma \in F_l$. They are written as $\sigma\eta$, where η is any word in Σ^κ. All of these belong to $\mathcal{E}$, except for $\sigma 0^\kappa$. Equally, when $\sigma \in F_r$, the words $\sigma\eta$, where η is any word in Σ^κ, belong to $\mathcal{E}$, except for $\sigma 1^\kappa$ since no refinement of I_σ can contain any point of Q_α other than p/q as an endpoint of one of its subintervals. Because of the argument above, these two exceptions yield all words of Σ^{n_2} that are *not* in $\mathcal{E}$. We can, therefore, write Σ^{n_2} as the union of three sets that are pairwise disjoint:

$$\Sigma^{n_2} = (\mathcal{E} \cap \Sigma^{n_2}) \uplus \{\sigma 0^\kappa,\ \sigma \in F_l\} \uplus \{\sigma 1^\kappa,\ \sigma \in F_r\}. \tag{7.17}$$

Consider now the first set in the disjoint union above, call it $E = \mathcal{E} \cap \Sigma^{n_2}$. It corresponds to intervals I_ω, with $\omega \in \Sigma^{n_2}$, such that neither endpoint of I_ω belongs to $\mathbf{Q}_\alpha$. Within this family, two cases are possible: small and large intervals, $E = E_s \uplus E_l$,

$$E_s = E \cap S^{n_2}(\alpha),\ E_l = E \cap L^{n_2}(\alpha).$$

In the first case, that is when $\omega \in E_s$, $\omega\eta$ belongs to $\mathcal{E} \cap S^{n_2+k}(\alpha)$ for any $k \geqslant 0$, $\eta \in \Sigma^k$, in force of (7.9). In the second case, $\omega \in E_l$, we use (7.10): for any such $\omega \in E_l$ there exists $k_1(\omega)$ such that $\omega\eta \in \mathcal{E} \cap S^{n_2+k}(\alpha)$ for any $k \geqslant k_1(\omega)$, $\eta \in \Sigma^k$. Since the cardinality of E_l is finite, there exists the maximum of $\{k_1(\omega),\ \omega \in E_l\}$: call it $K_1(E)$. This defines a new level, $n_3 = n_2 + K_1(E)$.

Third level, $n_3 = n_2 + K_1(E)$, where we make the final separation between small and large intervals.

We have just proven that for any $n \geqslant n_3$ the words $\omega\eta$, with $\omega \in \mathcal{E} \cap \Sigma^{n_2}$ and $\eta \in \Sigma^{n-n_3}$ belong to $\mathcal{E} \cap S^n(\alpha)$. It remains to consider words in Σ^n, with $n \geqslant n_3$, which originate from $\{\sigma 0^\kappa,\ \sigma \in F_l\}$ and $\{\sigma 1^\kappa,\ \sigma \in F_r\}$. Recall that these two sets are included in Σ^{n_2}; we need to consider their successive partitions. Let us show how to proceed by induction. Consider the first case and the word $\sigma 0^\kappa$, with $\sigma \in F_l$. Its partition yields the two words $\sigma 0^{\kappa+1}$ and $\sigma 0^\kappa 1$. Because of (7.14) and because $\kappa > k_2(\sigma)$, the latter belongs to both $\mathcal{E}$ and $S^{n_2+1}(\alpha)$. The property (7.9) implies that all of its successive partitions $\sigma 0^\kappa 1\eta$, with $\eta \in \Sigma^m$, for every $m \in \mathbf{N}$ belong to $\mathcal{E} \cap S^{n_2+1+m}(\alpha)$. We iterate the procedure on $\sigma 0^{\kappa+i}$, $i = 1, \dots$, so that induction proves that for any $m \in \mathbf{N}$ all words of the set

$$\{\sigma 0^\kappa \eta,\ \sigma \in F_l, \eta \in \Sigma^m\}$$

also belong to $S^{n_2+m}(\alpha)$, except possibly for the words in the subset $\{\sigma 0^{\kappa+m},\ \sigma \in F_l\}$. Similarly, we prove that all words of the set

$$\{\sigma 1^{\kappa}\eta,\ \sigma \in F_r, \eta \in \Sigma^m\}$$

also belong to $S^{n_2+m}(\alpha)$, except possibly for the words in the subset $\{\sigma 1^{\kappa+m},\ \sigma \in F_r\}$.

The above classification of intervals shows that for all $n \geqslant n_3$

$$\Sigma^n = S^n(\alpha) \cup (\{\sigma 0^{n-n_1},\ \sigma \in F_l\} \uplus \{\sigma 1^{n-n_1},\ \sigma \in F_r\}), \tag{7.18}$$

so that

$$L^n(\alpha) \subset \{\sigma 0^{n-n_1},\ \sigma \in F_l\} \uplus \{\sigma 1^{n-n_1},\ \sigma \in F_r\}. \tag{7.19}$$

We now prove the reverse inclusion. This happens to be much simpler. Consider F_l, the case of F_r being exactly symmetrical. The word $\sigma \in F_l \subset \Sigma^{n_1}$ can terminate with a certain number of consecutive zeros. This means that the point $x_\sigma = p/q$ has appeared in Stern–Brocot sequences of smaller index than n_1. Let $\tilde{n}$ be the first n for which $\frac{p}{q} = x_\sigma$ belongs to $\mathcal{B}^n$ (we called $\tilde{n}$ its depth in the tree). Then, there exists η in $\Sigma^{\tilde{n}}$ such that $\sigma = \eta 0^j$, with $j = n_1 - \tilde{n}$ and clearly $\frac{p}{q} = x_\sigma = M_\sigma(0) = M_\eta(0) = x_\eta$. We prove that $\eta 0^k$ belongs to $L^{\tilde{n}+k}(\alpha)$ for any $k \geqslant 0$. In fact, because of equation (7.12) this is equivalent to the validity of the inequality $q\hat{q} + kq^2 \leqslant a(|\eta| + k + 1)$. Since $x_\sigma \in \mathbf{Q}_\alpha$, we have $q^2 \leqslant a$. Moreover, if $\tilde{n} \geqslant 1$, then x_σ does *not* belong to $\mathcal{B}^{\tilde{n}-1}$, so the Farey rule (3.1) reads $q = \hat{q} + q'$, where q' is a denominator of an irreducible fraction in $\mathcal{B}^{\tilde{n}-1}$ and therefore $q > \hat{q}$. When $\tilde{n} = 0$, one has $q = \hat{q} = 1$. Combining this information we obtain

$$a(|\eta| + k + 1) \geqslant q^2(|\eta| + k + 1) \geqslant q\hat{q}(|\eta| + 1) + kq^2 \geqslant q\hat{q} + kq^2,$$

for any $k \geqslant 0$, which is what we wanted to show. This proves

$$L^n(\alpha) = \{\sigma 0^{n-n_1},\ \sigma \in F_l\} \uplus \{\sigma 1^{n-n_1},\ \sigma \in F_r\}, \tag{7.20}$$

and, because of the previous discussion, Proposition 7.1. The cardinality of $L^n(\alpha)$ in equation (7.20) is equal to the sum of the cardinalities of the two sets at the right-hand side, which are obviously finite and independent of n, for all $n \geqslant n_3$. Clearly the maximum cardinality of $L^n(\alpha)$ for $n < n_3$ is also finite, so that we obtain also Corollary 7.2. □

Regularity of Minkowski's question mark measure follows again from a combination of the previous Proposition, Remark 5.1 and equation (7.11). As in Section 6, they serve to verify that the hypotheses of Proposition 5.1 hold true.

8. *Regularity of invariant measures of Iterated Function Systems*

In this section, we put the regularity of Minkowski's measure into the more general perspective of the regularity of invariant measures of Iterated Function Systems. To appreciate the difficulty, we first prove a result in the case of strict contractions. We shall use the notations established in Section 1.3.

Proof of Theorem 1.4. Without loss of generality, we may assume that the diameter of $\mathcal{A}$ is 1. Set $\pi = \min_i \pi_i$. For a generic $\sigma \in \Sigma^n$, induction on n gives that $\mu(\varphi_\sigma(\mathcal{A})) \geqslant \pi^n \mu(\mathcal{A})$, and $\operatorname{diam}(\varphi_\sigma(\mathcal{A})) \leqslant \delta^n \operatorname{diam}(\mathcal{A}) \leqslant \delta^n$, where δ is the IFS contraction rate introduced in Section 1.3. For a small $r > 0$, choose n so that $\delta^n < r \leqslant \delta^{n-1}$. For $x \in \mathcal{A}$, let $\bar{\sigma} \in \Sigma^n$ be a word such that $x \in \varphi_{\bar{\sigma}}(\mathcal{A})$. Then, $\varphi_{\bar{\sigma}}(\mathcal{A}) \subset B_r(x)$, where $B_r(x)$ is the ball of radius r about x. Hence,

$$\mu(B_r(x)) \geqslant \mu(\varphi_{\bar{\sigma}}(\mathcal{A})) \geqslant \pi^n = \delta^{n \log \pi / \log \delta} \geqslant r^{(n/(n-1)) \log \pi / \log \delta} \geqslant r^{2 \log \pi / \log \delta},$$

and hence the regularity of μ is a consequence of Criterion Λ^* in [**46**, Theorem 4.2.3]. □

The case when the φ_i are weak contractions is more subtle. We shall deal with this case only under the following assumptions:

A-0. $\mathcal{A} = [0, 1]$ and the functions φ_i are increasing weak contractions. Let the image of [0,1] under φ_i be $[A_i, B_i]$. Assume that the intervals $[A_i, B_i]$ cover [0,1] and are pairwise disjoint —- except possibly for their endpoints, and then, by re-enumeration, we may assume $A_0 = 0$ and so $\varphi_0(0) = 0$, and $B_M = 1$ and so $\varphi_M(1) = 1$.
Each φ_i has a unique fixed point $X_i \in [A_i, B_i]$ (according to the agreement before $X_0 = 0$ and $X_M = 1$). Furthermore:

A-1. When $i = 0$, we assume that the function φ_0 is concave and there is $\rho_0 > 0$ such that

$$\frac{\varphi_0(x)}{x} \leqslant 1 - \rho_0 x. \tag{8.1}$$

A-2. When $i = M$, in symmetry with the previous case, we assume that φ_M is convex and with some $\rho_M > 0$

$$\frac{1 - \varphi_M(x)}{1 - x} \leqslant 1 - \rho_M(1 - x). \tag{8.2}$$

A-3. When $1 \leqslant i \leqslant M - 1$, the fixed point X_i lies in (A_i, B_i). On $[X_i, 1]$, we assume the behavior described in I: φ_i is concave and with some $\rho_i > 0$ it satisfies the inequality

$$\frac{\varphi_i(x) - X_i}{x - X_i} \leqslant 1 - \rho_i(x - X_i). \tag{8.3}$$

Symmetrically, on the interval $[0, X_i]$ we assume the behavior described in II: φ_i is convex and with some $\rho_i^* > 0$

$$\frac{X_i - \varphi_i(x)}{X_i - x} \leqslant 1 - \rho_i^*(X_i - x)). \tag{8.4}$$

A-4. To insure consistency with the requirement that the maps φ_i be increasing, we also impose that $\rho_i < 1$, $\rho_i^* < 1$ for all i. By selecting $\rho = \min_i\{\rho_i, \rho_i^*\}$, we may assume that all ρ_i, ρ_i^* are the same ρ.

Theorem 8.1. *Under assumptions A-0 to A-4, the measure μ on [0,1] which is invariant with respect to the IFS $\{\varphi_i\}_{i=0}^M$, $\{\pi_i\}_{i=0}^M$, is regular: $\mu \in \mathbf{Reg}$.*

Remark 8.1. Since Minkowski's question mark measure is invariant for the system $\{\varphi_0, \varphi_1\}$ where $\varphi_0(x) = x/(1 + x)$ and $\varphi_1(x) = 1/(2 - x)$, and since these maps satisfy the just given conditions, regularity of Minkowski's measure is a consequence of this theorem.

Before moving to the proof, let us briefly discuss the set-up of this theorem.

Remark 8.2. If ρ is the smallest of the numbers ρ_i, ρ_i^*, then conditions (8.3) and (8.4) can be unified as

$$\left|\frac{\varphi_i(x) - X_i}{x - X_i}\right| \leqslant 1 - \rho|x - X_i|. \tag{8.5}$$

Remark 8.3. The pairwise disjointness of the interiors of the image sets can be weakened to the assumption that $X_i \notin [A_j, B_j]$ if $i \neq j$, but we do not go into details.

Remark 8.4. Theorem 8.1 is still true if the φ_i are assumed to be (strictly) monotonic, though not necessarily increasing. When, for instance, a particular φ_i is decreasing, then necessarily $X_i \in (A_i, B_i)$ and this falls under A-3: we need to require convexity from the right of X_i, concavity from the left, and instead of (8.3) and (8.4) we need to use the common form (8.5). The proof requires many formal modifications in this case, but the main ideas remain the same.

Remark 8.5. Some explanations regarding the conditions (8.1)–(8.4) are in order. Consider, for example, A-1. The point 0 is a fixed point for φ_0 and for reasons that will become immediately clear we want φ_0 to be more contractive away from 0 than around 0. The simplest way to achieve this is to require that φ_0 be a concave function — this property could be relaxed somewhat, but we omit details here. Then φ_0, as a concave function on [0,1], has a right derivative φ_0' at every point, which is a decreasing function. Hence, φ_0'' exists almost everywhere. The contractive property of φ_0 then implies $\varphi_0'(0) \leqslant 1$. Two cases are now possible. If $\varphi_0'(0) < 1$, then φ_0 is a strict contraction, described by Theorem 1.4. On the other hand, if $\varphi_0'(0) = 1$, then 0 is a marginally stable fixed point for φ_0. This fact might lead, in the absence of further specification, to an invariant measure that is too thin in its neighborhood, impairing regularity: as an extreme case, let $\varphi_0(x) = x$ on $[0, a]$, so that for this interval the property $\mu(\varphi_0(E)) = \pi_0\mu(E)$ implies that μ is the null measure. We therefore require that $\varphi_0(x)$ is not too close to x as x approaches 0, which is guaranteed by condition (8.1). If $\varphi_i'' \leqslant -c < 0$, property (8.1) is true, so that we can roughly think of the latter as the requirement that $\varphi_0'' \leqslant -c < 0$. Assumption A-2 is the analogue of A-1 for the right endpoint 1 (the mapping $x \to 1 - x$ takes these two cases into each other), and finally if the fixed point X_i is different from 0 and 1, in A-3, we replicate the above assumptions by requiring that to the right of X_i the behavior of φ_i is similar to that of φ_0 around 0 in A-1, while to the left of X_i the behavior is like that around 1 in A-2.

Remark 8.6. We do not know if Theorem 8.1 is true for any Iterated Function System consisting of weak contractions on [0,1] (in other words, if conditions (8.1)–(8.4), as well as the convexity/concavity conditions can be dropped altogether).

Let us now move to the proof of Theorem 8.1. We will obtain it via

Proposition 8.2. *Let the intervals I_σ for $\sigma \in \Sigma$ be generated by an IFS which fulfills the conditions stated above in this section. Then, Corollary 7.2 holds for these intervals.*

To prove this Proposition, we need some properties of the IFS satisfying the above requirements. Define $\beta_{s,i} = \varphi_i^s(0)$, $\gamma_{s,i} = \varphi_i^s(1)$. Then, $\beta_{s,0} = 0$ and $\gamma_{s,M} = 1$ for all $s \in \mathbf{N}$. In all other cases, $\{\beta_{s,i}\}_{s=1}^\infty$ is a strictly increasing sequence and $\{\gamma_{s,i}\}_{s=1}^\infty$ is a strictly decreasing sequence, both converging to X_i. Clearly, $\varphi_i^s([0,1]) = [\beta_{s,i}, \gamma_{s,i}]$ for all s. Note also that if $\sigma = j\eta$, for any $\eta \in \Sigma$, then $\varphi_\sigma([0,1]) \subseteq [A_j, B_j]$. Therefore, when $i \neq j$, we have that $\varphi_\sigma([0,1]) \subseteq [0,1] \setminus (A_i, B_i)$, and hence $\varphi_i \circ \varphi_\sigma([0,1]) \subset [\gamma_{2,i}, \gamma_{1,i}]$ (when $\varphi_\sigma([0,1]) \subseteq [B_i, 1]$) or $\varphi_i \circ \varphi_\sigma([0,1]) \subset [\beta_{1,i}, \beta_{2,i}]$ (when $\varphi_\sigma([0,1]) \subseteq [0, A_i]$), where we used that, for example, $\gamma_{1,i} = B_i$ and $\gamma_{2,i} = \varphi_i(\gamma_{1,i})$.

Now if the interval $J = [a, b]$ is such that $J \subseteq [\gamma_{s,i}, 1]$ for a pair s, i, $i \in \{0, \dots, M\}$, $s \in \mathbf{N}$, then $\varphi_i(J) \subseteq [\gamma_{s+1,i}, 1]$. Using that φ_i is increasing, concave on $[X_i, 1]$ and $\varphi_i(X_i) = X_i$ we arrive at

$$\frac{|\varphi_i(J)|}{|J|} = \frac{\varphi_i(b) - \varphi_i(a)}{b - a} \leqslant \frac{\varphi_i(a) - \varphi_i(X_i)}{a - X_i} \leqslant \frac{\varphi_i(\gamma_{s,i}) - X_i}{\gamma_{s,i} - X_i} = \frac{\gamma_{s+1,i} - X_i}{\gamma_{s,i} - X_i},$$

in which the final ratios are increasing monotonically with s. Therefore, we can iterate this inequality (with s replaced by $s + 1$, then $s + 1$ by $s + 2$, *etc.*) to conclude that for $k \geqslant 1$

$$|\varphi_i^k(J)| \leqslant |J| \frac{\gamma_{s+1,i} - X_i}{\gamma_{s,i} - X_i} \cdots \frac{\gamma_{s+k,i} - X_i}{\gamma_{s+k-1,i} - X_i} = |J| \frac{\gamma_{s+k,i} - X_i}{\gamma_{s,i} - X_i}. \tag{8.6}$$

In a similar manner, if $J \subseteq [0, \beta_{s,i}]$, then $\varphi_i(J) \subseteq [0, \beta_{s+1,i}]$, and for $k \geqslant 1$

$$|\varphi_i^k(J)| \leqslant |J| \frac{X_i - \beta_{s+k,i}}{X_i - \beta_{s,i}}. \tag{8.7}$$

We now prove two lemmas:

Lemma 8.3. *For all* $0 \leqslant i \leqslant M$ *and* $s \in \mathbf{N}$

$$\gamma_{s,i} - X_i \leqslant \frac{C_0}{s+1}, \qquad X_i - \beta_{s,i} \leqslant \frac{C_0}{s+1}, \tag{8.8}$$

with $C_0 = 1/\rho$, *where* ρ *is the number from* (8.5).

Proof. Equation (8.8) is certainly true for $s = 0$, since $C_0 > 1$: recall that $\rho_i \leqslant \rho < 1$. Letting $Z_{s,i} = \gamma_{s,i} - X_i$ and using (8.3) we have

$$Z_{s+1,i} = \varphi_i(\gamma_{s,i}) - X_i \leqslant (\gamma_{s,i} - X_i)(1 - \rho(\gamma_{s,i} - X_i)) = Z_{s,i}(1 - \rho Z_{s,i}) \leqslant Z_{s,i}. \tag{8.9}$$

Suppose for induction that (8.8) is true for a certain s. We need to prove that it holds for $s+1$. We have the chain of inequalities

$$\frac{s+1}{C_0} \leqslant \frac{1}{Z_{s,i}} \leqslant \frac{1 - \rho Z_{s,i}}{Z_{s+1,i}} \leqslant \frac{1 - \rho Z_{s+1,i}}{Z_{s+1,i}} = \frac{1}{Z_{s+1,i}} - \rho.$$

The first inequality is the induction hypothesis; to prove the second, we employ the intermediate inequality in (8.9); the third follows from the full inequality (8.9). Therefore,

$$\frac{1}{Z_{s+1,i}} \geqslant \frac{s+1+C_0\rho}{C_0} = \frac{s+2}{C_0},$$

which proves induction. The second relation in (8.8) follows from the same reasoning if we use (8.4) instead of (8.3). □

Lemma 8.4. *There is a constant* C_1 *such that, for all* $n \in \mathbf{N}$,

$$|I_\sigma| = |\varphi_\sigma([0,1])| \leqslant \frac{C_1}{n+1}, \qquad \sigma \in \Sigma^n. \tag{8.10}$$

Proof. To prove this lemma, we need to define two quantities:

$$\tau := \max\left\{ \max_{0 \leqslant i < M} \frac{\gamma_{2,i} - X_i}{\gamma_{1,i} - X_i}, \max_{0 < i \leqslant M} \frac{X_i - \beta_{2,i}}{X_i - \beta_{1,i}} \right\} \tag{8.11}$$

and

$$\kappa = \min\{ \min_{0 \leqslant i < M} (\gamma_{1,i} - X_i), \min_{0 < i \leqslant M} (X_i - \beta_{1,i}) \}. \tag{8.12}$$

Then, $\tau < 1$ because of (8.1)–(8.4), and $\kappa > 0$ because the IFS maps are increasing functions.

Since $|I_\sigma| \leqslant 1$, when $n = 0$ or $n = 1$ it is enough to take $C_1 \geqslant 2$ for (8.10) to hold. Next, suppose that (8.10) holds for all $n' < n$ and consider a $\sigma \in \Sigma^n$. We separate three cases.

The first case is $\sigma = i^n$, with $i \in \{0, \ldots, M\}$. Then, by (8.8)

$$|I_\sigma| = \gamma_{n,i} - \beta_{n,i} \leqslant 2\frac{C_0}{n+1},$$

which proves the desired inequality (8.10), when $C_1 \geqslant 2C_0$ as before.

If the first case does not hold, then the word $\sigma \in \Sigma^n$ can be written as $\sigma = i^k j\eta$, with $i \neq j$, $1 \leqslant k \leqslant n-1$, with some word $\eta \in \Sigma^{n-k-1}$. Let $\theta \in (0,1)$ to be specified later, and consider separately two alternatives: $k \geqslant \theta n$ and $k < \theta n$. According to the relative value of i and j, we apply either (8.6) or (8.7) with $s = 1$, with identical results. Let us show computations using (8.6). We start from the case $k < \theta n$:

$$|I_\sigma| \leqslant \frac{\gamma_{k+1,i} - X_i}{\gamma_{1,i} - X_i} |I_{j\eta}| \leqslant \frac{\gamma_{2,i} - X_i}{\gamma_{1,i} - X_i} |I_{j\eta}| \leqslant \tau \frac{C_1}{n-k+1} \leqslant \tau \frac{C_1}{n - \theta n + 1},$$

where we have used $\gamma_{2,i} \geqslant \gamma_{k+1,i}$, definition (8.11) and the induction hypothesis. Since $\tau < 1$ if we choose $\theta \leqslant 1 - \tau$ simple algebra shows that $|I_\sigma| \leqslant \frac{C_1}{n+1}$. Note that this does not put bounds on C_1 but only restricts the range of values of θ that can be used in the proof.

In the other alternative, $k \geqslant \theta n$, we use the first inequality above, but after that we continue differently: using (8.8) and definition (8.12) we obtain

$$|I_\sigma| \leqslant \frac{\gamma_{k+1,i} - X_i}{\gamma_{1,i} - X_i} |I_{j\eta}| \leqslant \frac{\gamma_{k+1,i} - X_i}{\gamma_{1,i} - X_i} \leqslant \frac{C_0}{k+2} \frac{1}{\gamma_{1,i} - X_i} \leqslant \frac{C_0}{\theta n + 2} \frac{1}{\kappa} \leqslant \frac{C_0}{\kappa \theta} \frac{1}{n+1}.$$

The optimal choice of θ, which is bound to the interval $(0, 1-\tau]$, to minimize the constant $\frac{C_0}{\kappa\theta}$, is $\theta = 1 - \tau$. In conclusion, with

$$C_1 = C_0 \max\left\{2, \frac{1}{\kappa(1-\tau)}\right\} \tag{8.13}$$

the relation (8.10) is proven. □

Proof of Proposition 8.2. Let us focus our attention on words $\sigma \in \Sigma^n$, with $n > 2m^2$, when m is an integer to be specified later, which are of the form

$$\sigma = i_q^{k_q} i_{q-1}^{k_{q-1}} \cdots i_1^{k_1} j\eta \tag{8.14}$$

for some $m^2 < r < n - m^2$ and $\eta \in \Sigma^{r-1}$, where $i_1 \neq j$, $i_{l+1} \neq i_l$ for $l = 1, \dots, q-1$, the powers k_l are *all* positive, and sum up to $k_1 + k_2 + \cdots + k_l = n - r$. For such σ, we want to estimate the length of the interval I_σ. Using again (8.6) and (8.7) with $s = 1$ we obtain

$$|I_\sigma| \leqslant |I_{j\eta}|\, \omega_{k_1,i_1} \omega_{k_2,i_2} \cdots \omega_{k_q,i_q}, \tag{8.15}$$

where ω_{k_l,i_l} stands for either

$$\omega_{k_l,i_l} = \frac{\gamma_{k_l+1,i_l} - X_{i_l}}{\gamma_{1,i_l} - X_{i_l}} \quad \text{or} \quad \omega_{k_l,j_l} = \frac{X_{i_l} - \beta_{k_l+1,i_q}}{X_{i_l} - \beta_{1,i_l}}, \tag{8.16}$$

possibly independently of each other for different $l = 1, \dots, q$. The first factor at the right-hand side in (8.15) can be controlled as $|I_{j\eta}| \leqslant \frac{C_1}{r+1}$ by (8.10), or simply by $|I_{j\eta}| \leqslant 1$. For the remaining factors, as before, the two alternatives in (8.16) are equivalent, because equations (8.8), (8.11) and (8.12) yield for both the bounds

$$\omega_{k_l,i_l} \leqslant \begin{cases} \omega_{1,i_l} \leqslant \tau < 1 \\[2ex] \dfrac{C_0}{k_l+2}\frac{1}{\kappa} \leqslant \dfrac{C_0}{\kappa k_l}. \end{cases} \tag{8.17}$$

Let $\alpha > 0$ be an arbitrary constant. Recall that $r > m^2$ is the length of the word $j\eta$ and distinguish two cases.

Case 1. $r > n/2$. If $q \geqslant m$, then we obtain from Lemma 8.4 (applied to $I_{j\eta}$) and from equation (8.15)

$$|I_\sigma| \leqslant \frac{C_1}{r+1}\tau^q \leqslant \frac{2C_1}{n+1}\tau^m < \frac{\alpha}{n+1},$$

if m is chosen large enough, $m \geqslant Q_1(\alpha,\tau)$, where we have implicitly defined the quantity $Q_1(\alpha,\tau)$. On the other hand, if $q < m$, then, since $k_1 + k_2 + \cdots + k_q = n - r \geqslant m^2$, there must be an l such that $k_l \geqslant m$. Hence, in this case, bounding all remaining $\omega_{k_{l'},j_{l'}}$ by one we obtain

$$|I_\sigma| \leqslant \frac{C_1}{r+1}\omega_{k_l,j_l} \leqslant \frac{C_1}{n/2+1}\frac{C_0}{\kappa k_l} \leqslant \frac{2C_1}{n+1}\frac{C_0}{\kappa m} < \frac{\alpha}{n+1},$$

if m is sufficiently large. Precisely, this requires that m is larger than $Q_2(\alpha) = 2C_1C_0/\kappa\alpha$.

Case 2. $r \leqslant n/2$. Let p, respectively $\bar{p}$, be the number of those k_l for which $k_l < m$, respectively $k_l \geqslant m$, so that, say, $k_{l_1},\ldots,k_{l_{\bar{p}}} \geqslant m$. We so have

$$pm + k_{l_1} + \cdots + k_{l_{\bar{p}}} \geqslant k_1 + \cdots + k_q = n - r \geqslant n/2,$$

and

$$|I_\sigma| \leqslant \frac{C_1}{r+1}\tau^p \frac{C_0}{\kappa k_{l_1}} \cdots \frac{C_0}{\kappa k_{l_{\bar{p}}}}. \tag{8.18}$$

Each factor on the right can be replaced by 1 at our discretion — recall equation (8.17). There are now two subcases. In the first, $pm \geqslant n/4$ case we can write

$$|I_\sigma| \leqslant \tau^p \leqslant \tau^{n/4m} < \frac{\alpha}{n+1},$$

if n is sufficiently large, that is, when $n \geqslant Q_0(m,\alpha,\tau)$. It is readily verified that, since $\tau < 1$, the function $Q_0(m,\alpha,\tau)$ increases when m grows, or α tends to zero. In the second subcase, $pm < n/4$, one has $S_{\bar{p}} := k_{l_1} + \cdots + k_{l_{\bar{p}}} \geqslant n/4$. If in addition $\bar{p} = 1$, then

$$|I_\sigma| \leqslant \frac{C_1}{r+1}\frac{C_0}{\kappa(k_{l_1}+2)} \leqslant \frac{C_1}{m^2+1}\frac{C_0}{\kappa(n/4+2)} \leqslant \frac{\alpha}{n+1}$$

if m is large, $m \geqslant Q_3(\alpha,\kappa)$. When $\bar{p} > 1$ the argument is more involved. Consider the product $k_{l_1}k_{l_2}\cdots k_{l_{\bar{p}}}$, where each factor is larger than m and their sum is $S_{\bar{p}}$. If two terms in the product are different, say $m < k_{l_i} < k_{l_j}$, then by replacing k_{l_i} by $k_{l_i} - 1$ and k_{l_j} by $k_{l_j} + 1$ we decrease the product while keeping the sum constant. Therefore, the product is minimal, compatible with the bounds, when $\bar{p} - 1$ of the numbers k_{l_j} are equal to m and the remaining k is such that $(\bar{p}-1)m + k = S_{\bar{p}}$. This proves the first inequality below:

$$\begin{aligned}\frac{\kappa k_{l_1}}{C_0}\cdots\frac{\kappa k_{l_{\bar{p}}}}{C_0} &\geqslant \left(\frac{\kappa m}{C_0}\right)^{\bar{p}-1}\left(\frac{\kappa S_{\bar{p}}}{C_0} - (\bar{p}-1)\frac{\kappa m}{C_0}\right) \\ &\geqslant \left(\frac{\kappa m}{C_0}\right)^{\bar{p}-1}\frac{\kappa S_{\bar{p}}}{C_0\bar{p}} \geqslant \left(\frac{\kappa m}{C_0}\right)\frac{\kappa S_{\bar{p}}}{C_0 2} \geqslant \frac{\kappa m}{C_0}\frac{\kappa n}{8C_0}.\end{aligned}$$

The second inequality follows from $S_{\bar{p}} - (\bar{p}-1)m \geqslant \frac{S_{\bar{p}}}{\bar{p}}$, by the definition of $S_{\bar{p}}$, and the third from the fact that $v^{\bar{p}-1}/\bar{p}$ increases in $\bar{p}$ on the interval $\bar{p} \geqslant 2$ if $v \geqslant 2$: we use this for $v = \kappa m/C_0$, which is larger than two if we require that $m \geqslant 2C_0/\kappa = Q_4(\kappa)$. Finally, $S_{\bar{p}} \geqslant n/4$ yields the last inequality. Hence, in this ($\bar{p} > 1$) case we have (see (8.18))

$$|I_\sigma| \leqslant \frac{C_1}{r+1}\frac{C_0}{\kappa m}\frac{8C_0}{\kappa n} \leqslant \frac{8C_0^2 C_1}{m^3\kappa^2}\frac{1}{n} \leqslant \frac{\alpha}{n+1}$$

if $m \geqslant Q_5(\alpha, \kappa)$.

To conclude: if we let $\bar{m} = \max\{Q_1(\alpha,\tau), Q_2(\alpha), Q_3(\alpha,\kappa), Q_4(\kappa), Q_5(\alpha,\kappa)\}$ and we require that $m \geqslant \bar{m}$, and successively that $n \geqslant \bar{n} = \max\{Q_0(m,\alpha,\tau), m^2\}$, then, for all words $\sigma \in \Sigma^n$ of the form (8.14), we have $|I_\sigma| \leqslant \frac{\alpha}{n+1}$. In other words, once we fix m sufficiently large (depending only on α, τ, κ) the inequality is valid for *all* values of n larger than the threshold $\bar{n}$.

Recall now that (8.14) requires that, being $\sigma \in \Sigma^n$, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, there is an integer r such that $m^2 < r < n - m^2$ for which $\sigma_r \neq \sigma_{r+1}$. Therefore, if *ceteribus paris* $|I_\sigma| > \alpha/(n+1)$, then the word σ must satisfy $\sigma_{m^2+1} = \sigma_{m^2+2} = \cdots = \sigma_{n-m^2}$, and there are at most $(M+1)^{2m^2}$ such σ in Σ^n. Since the cardinality of $L^n(\alpha)$ is clearly bounded for $n \leqslant \bar{n}$, this proves Proposition 8.2. □

Proof of Theorem 8.1. It follows from Proposition 5.1, whose hypotheses are proven by Proposition 8.2, Lemma 8.4 and Remark 5.1. □

Acknowledgements. Studying abstract mathematical models is fun — especially when it provides matter of entertaining discussions with friends: G.M. would like to thank Walter Van Assche and Roberto Artuso.

References

1. G. Alkauskas, 'Integral transforms of the Minkowski question mark function', PhD Thesis, University of Nottingham, Nottingham, 2008, http://etheses.nottingham.ac.uk/641/.

2. G. Alkauskas, 'The Minkowski question mark function: explicit series for the dyadic period function and moments', *Math. Comp.* 79 (2010) 383–418; Addenda and corrigenda, Math. Comp. **80** (2011) 2445–2454.

3. G. Alkauskas, 'Semi-regular continued fractions and an exact formula for the moments of the Minkowski question mark function', *Ramanujan J.* 25 (2011) 359–367.

4. R. Artuso, P. Cvitanović and B. G. Kenny, 'Phase transitions on strange irrational sets', *Phys. Rev. A* 39 (1989) 268–281.

5. M. F. Barnsley and S. G. Demko, 'Iterated function systems and the global construction of fractals', *Proc. Roy. Soc. London* A 399 (1985) 243–275.

6. M. F. Barnsley, S. G. Demko, J. H. Elton and J. S. Geronimo, 'Invariant measures for Markov processes arising from iterated function systems with place dependent probabilities', *Ann. Inst. H. Poincaré Probab. Stat.* 24 (1988) 367–394.

7. D. Bessis and G. Mantica, 'Construction of multifractal measures in dynamical systems from their invariance properties', *Phys. Rev. Lett.* 66 (1991) 2939–2942.

8. D. Bessis and G. Mantica, 'Orthogonal polynomials associated to almost–periodic Schrödinger operators', *J. Comput. Appl. Math.* 48 (1993) 17–32.

9. C. Bonanno and S. Isola, 'Orderings of the rationals and dynamical systems', *Colloq. Math.* 116 (2009) 165–189;

10. A. Brocot, 'Calcul des rouages par approximation, nouvelle méthode', *Revue Chronométrique* 6 (1860) 186–194.

11. M. Degli Esposti, S. Isola and A. Knauf, 'Generalized Farey trees, transfer operators and phase transitions', *Comm. Math. Phys.* 275 (2007) 297–329.

12. A. Denjoy, 'Sur une fonction réelle de Minkowski', *J. Math. Pures Appl.* 17 (1938) 105–151.

13. Z. Dresse and W. Van Assche, 'Orthogonal polynomials for Minkowski's question mark function', *J. Comput. Appl. Math.* 284 (2015) 171–183.

14. P. Erdös and P. Turán, 'On interpolation. III. Interpolatory theory of polynomials', *Ann. of Math.* 41 (1940) 510–553.

15. R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete mathematics* (Addison–Wesley, Reading, MA, 1994).

16. I. Guarneri and G. Mantica, 'On the asymptotic properties quantum mechanics in the presence of a fractal spectrum', *Ann. Inst. H. Poincaré* 61 (1994) 369–379.

17. I. Guarneri and G. Mantica, 'Multifractal energy spectra and their dynamical implications', *Phys. Rev. Lett.* 73 (1994) 3379–3382.

18. M. C. Gutzwiller, 'Bernoulli sequences and trajectories in the anisotropic Kepler problem', *J. Math. Phys.* 18 (1977) 806–823.
19. M. C. Gutzwiller and B. B. Mandelbrot, 'Invariant multifractal measures in chaotic Hamiltonian systems, and related structures', *Phys. Rev. Lett.* 60 (1988) 673–676.
20. J. Hutchinson, 'Fractals and self-similarity', *Indiana J. Math.* 30 (1981) 713–747.
21. T. Jordan and T. Sahlsten, 'Fourier transforms of Gibbs measures for the Gauss map', *Math. Ann.* 364 (2016) 983–1023.
22. M. Kesseböhmer and B. O. Stratmann, 'Stern–Brocot pressure and multifractal spectra in ergodic theory of numbers', *Stoch. Dyn.* 4 (2004) 77–84.
23. M. Kesseböhmer and B. O. Stratmann, 'A multifractal analysis for Stern-Brocot intervals, continued fractions and diophantine growth rates', *J. reine angew. Math.* 605 (2007) 133–163.
24. M. Kesseböhmer and B. O. Stratmann, 'Fractal analysis for sets of non-differentiability of Minkowski's question mark function', *J. Number Theory* 128 (2008) 2663–2686.
25. J. R. Kinney, 'Note on a singular function of Minkowski', *Proc. Amer. Math. Soc.* 11 (1960) 788–794.
26. J. C. Lagarias, 'Number theory and Dynamical Systems', *Proc. Sympos. Appl. Math.* 46 (1992) 35–72.
27. D. S. Lubinsky, 'Singularly continuous measures in Nevai's class', *Proc. Amer. Math. Soc.* 111 (1991) 413–420.
28. G. Mantica, 'Quantum intermittency in almost periodic systems derived from their spectral properties', *Phys. D* 103 (1997) 576–589.
29. G. Mantica, 'Wave propagation in almost-periodic structures', *Phys. D* 109 (1997) 113–127.
30. G. Mantica, 'On Computing Jacobi matrices associated with recurrent and Möbius iterated functions systems', *J. Comput. Appl. Math.* 115 (2000) 419–431.
31. G. Mantica, 'Fourier-Bessel functions of singular continuous measures and their many asymptotics', *Electron. Trans. Numer. Anal.* (Electronic) 25 (2006) 409–430.
32. G. Mantica, 'Minkowski's question mark measure', *J. Approx. Theory* 222 (2017) 74–109.
33. G. Mantica and D. Guzzetti, 'The asymptotic behaviour of the Fourier transform of orthogonal polynomials II: iterated function systems and quantum mechanics', *Ann. Henri Poincaré* 8 (2007) 301–336.
34. A. Máté, P. Nevai and V. Totik, 'Asymptotics for the ratio of leading coefficients of orthonormal polynomials on the unit circle', *Constr. Approx.* 1 (1985) 63–69.
35. F. Mendivil, 'A generalization of IFS with probabilities to infinitely many maps', *Rocky Mountain J. Math.* 28 (1998) 1043–1051.
36. H. Minkowski, 'Zur Geometrie der Zahlen', *Verhandlungen des III* (Internationalen Mathematiker-Kongresses, Heidelberg, 1904) 164–173.
37. K. Okamura, 'Singularity results for functional equations driven by linear fractional transformations', *J. Theoret. Probab.* 27 (2014) 1316–1328.
38. T. Persson, 'On a problem by R. Salem concerning Minkowski's question mark function', Preprint, 2015, arXiv:1501.00876.
39. E. A. Rakhmanov, 'The asymptotic behavior of the ratio of orthogonal polynomials. II', *Mat. Sb. (N.S.)* 118 (1982) 104–117.
40. T. Ransford, *Potential theory in the complex plane* (Cambridge University Press, Cambridge, 1995).
41. E. B. Saff, 'Orthogonal polynomials from a complex perspective', *Orthogonal polynomials* (ed. P. Nevai; Springer, Dordrecht, 1990) 363–393.
42. E. B. Saff, 'Logarithmic potential theory with applications to approximation theory', *Surv. Approx. Theory* 5 (2010) 165–200.
43. R. Salem, 'On some singular monotonic functions which are strictly increasing', *Trans. Amer. Math. Soc.* 53 (1943) 427–439.
44. C. Series, 'The modular surface and continued fractions', *J. Lond. Math. Soc.* 31 (1985) 69–80.
45. H. Stahl and V. Totik, 'n–th root asymptotic behavior of orthogonal polynomials', *Orthogonal polynomials* (ed. P. Nevai; Springer, Dordrecht, 1990) 395–417.
46. H. Stahl and V. Totik, *General Orthogonal Polynomials* (Cambridge University Press, Cambridge, 2010).
47. M. A. Stern, 'Über eine zahlentheoretische Funktion', *J. reine angew. Math.* 55 (1858) 193–220.
48. R. Strichartz, 'Self-similar measures and their Fourier transforms I', *Indiana Univ. Math. J.* 39 (1990) 797–817.
49. R. Strichartz, 'Self-similar measures and their Fourier transforms II', *Trans. Amer. Math. Soc.* 336 (1993) 335–361.
50. R. Strichartz, 'Self similar measures and their Fourier transforms III', *Indiana Univ. Math. J.* 42 (1993) 367–411.
51. J. L. Ullman, 'On the regular behaviour of orthogonal polynomials', *Proc. Lond. Math. Soc.* 24 (1972) 119–148.
52. W. Van Assche and A. P. Magnus, 'Sieved orthogonal polynomials and discrete measures with jumps dense in an interval', *Proc. Amer. Math. Soc.* 106 (1989) 163–173.
53. P. Viader, J. Paradis and L. Bibiloni, 'A new light on Minkowski's $?(x)$ function', *J. Number Theory* 73 (1998) 212–227.
54. S. Yakubovich, 'On some Rajchman measures and equivalent Salem's problem', *Commun. Math. Anal.* 14 (2013) 28–41.
55. S. Yakubovich, 'The affirmative solution to Salem's problem revisited', Preprint, arXiv:1501.00141.

Giorgio Mantica
Center for Nonlinear and Complex Systems
Dipartimento di Scienza ed Alta Tecnologia
Università degli Studi dell' Insubria
Via Valleggio 11, 22100 Como
Italy

and

INFN sezione di Milano
INDAM, GNFM

giorgio.mantica@uninsubria.it

Vilmos Totik
MTA-SZTE Analysis and Stochastics Research Group
Bolyai Institute
University of Szeged
Szeged, Aradi v. tere 1, 6720
Hungary

and

Department of Mathematics and Statistics
University of South Florida
4202 E. Fowler Ave. CMC342
Tampa, FL 33620
U.S.A.

totik@mail.usf.edu