Spectral Theory of Regular Sequences

Michael Coons, James Evans, and Neil Mañibo

Received: August 20, 2021
Revised: February 27, 2022

Communicated by Roland Speicher

Abstract. Regular sequences are natural generalisations of fixed points of constant-length substitutions on finite alphabets, that is, of automatic sequences. Using the harmonic analysis of measures associated with substitutions as motivation, we study the limiting asymptotics of regular sequences by constructing a systematic measure-theoretic framework surrounding them. The constructed measures are generalisations of mass distributions supported on attractors of iterated function systems.

2020 Mathematics Subject Classification: 11B85, 42A38, 28A80
Keywords and Phrases: regular sequences, aperiodic order, symbolic dynamics, continuous measures, dilation equations

1 Introduction

A sequence f is called *k-automatic* if there is a deterministic finite automaton that takes in the base-k expansion of a positive integer n and outputs the value $f(n)$. Automatic sequences are ubiquitous in number theory, theoretical computer science and symbolic dynamics, and can be described in many ways, though the one we find most convenient is via the k-kernel,

$$\ker_k(f) := \left\{ (f(k^\ell n + r))_{n\geqslant 0} : \ell \geqslant 0, 0 \leqslant r < k^\ell \right\}.$$

A sequence f is k-automatic if and only if its k-kernel is finite [20, Prop. V.3.3]. It is immediate that an automatic sequence takes only a finite number of values. A natural generalisation to sequences that can be unbounded was given in the early nineties by Allouche and Shallit [2]; a real sequence f is called *k-regular* if the $\mathbb{R}$-vector space

$$\mathcal{V}_k(f) := \langle \ker_k(f) \rangle_{\mathbb{R}}$$

generated by the k-kernel of f is finite-dimensional over $\mathbb{R}$. One nice property of this generalisation is that a sequence taking only a finite number of values is automatic. Additionally, the set of k-regular sequences has algebraic structure; it forms a ring under point-wise addition and (Cauchy) convolution.

The study of automatic sequences is rich from both number-theoretical and dynamical viewpoints. Much of the number-theoretic literature on automatic sequences mirrors that of the rational-transcendental dichotomies of integer power series proved in the first third of the twentieth century, such as those of Fatou [23], Carlson [13] and Szegő [37], and the more recent celebrated result of Adamczewski and Bugeaud [1]. The dynamical literature has focussed on the study of automatic sequences through their related substitution systems—every automatic sequence is a coding of an infinite fixed point of a constant-length substitution on finitely many letters; see Cobham [14]. The long-range order of substitution systems has been well studied; there is an abundance of literature on this still very active area of research going back to the seminal works of Wiener [38] and Mahler [30]. The monographs by Queffélec [34] and Baake and Grimm [6] contain details about both the tiling and symbolic pictures of these systems as well as the associated diffraction and spectral measures—the modern-classical means of examining the long-range order of these systems.

The spectral results concerning substitution systems are not dichotomies, but classifications based on the Lebesgue Decomposition Theorem. In this context, one starts with a substitution and forms a measure—usually the spectral measure or the diffraction measure. The culminating result is then determining the spectral type of the measure. The Thue–Morse dynamical system is a paradigmatic example; the Thue–Morse sequence $\mathfrak{t} := abbabaab\cdots$ is 2-automatic and is the fixed point of the substitution on two letters given by

$$\varrho_{\mathrm{TM}} : \begin{cases} a \mapsto ab \\ b \mapsto ba\,. \end{cases}$$

Given a weight function $w\colon \{a,b\} \to \mathbb{C}$, one can calculate the diffraction measure (or correlation measure) μ_w of the sequence $\mathfrak{t}$ which is determined by its autocorrelation coefficients

$$\widehat{\mu}_w(t) := \lim_{n\to\infty} \frac{1}{n} \sum_{0\leqslant i,\, i-t<n} w(\mathfrak{t}_i)\overline{w(\mathfrak{t}_{i-t})} \tag{1}$$

for $t \in \mathbb{N}$ with $\widehat{\mu}_w(-t) := \overline{\widehat{\mu}_w(t)}$; see [6, 34]. The diffraction measure of the Thue–Morse sequence with weight function $w(a) = 1, w(b) = -1$ is purely singular continuous with respect to Lebesgue measure. This result was proved by Mahler [30]—who was the first author to explicitly record a singular continuous measure—and later in a dynamical setting by Kakutani [25]. For details on diffraction see Baake and Grimm [6, Ch. 9].

Transitioning to regular sequences, the number-theoretic story is much the same as automatic sequences, mirroring that of rational-transcendental dichotomies.

The generalisation of the Cobham–Loxton–van der Poorten Conjecture for regular sequences was proved by Bell, Bugeaud and Coons [8]. But, in contrast to automatic sequences, the study of the long-range order of unbounded regular sequences f, and so also the related spaces $\mathcal{V}_k(f)$, is not so straight-forward; neither diffraction nor spectral measures can be associated with them in a natural way.

As a first step of addressing the long-range order of such objects, Baake and Coons [3] introduced a natural probability measure associated with Stern's diatomic sequence, one simple example of an unbounded regular sequence. Stern's diatomic sequence s is given by $s(0) = 0$, $s(1) = 1$, and for $n \geqslant 1$ by the recurrences $s(2n) = s(n)$ and $s(2n+1) = s(n) + s(n+1)$. Its 2-kernel $\ker_2(s)$ is infinite, but induces a finitely generated vector space;

$$\mathcal{V}_2(s) = \langle \ker_2(s) \rangle_{\mathbb{R}} = \langle \{(s(n))_{n\geqslant 0}, (s(n+1))_{n\geqslant 0}\} \rangle_{\mathbb{R}}.$$

Stern's diatomic sequence has many interesting properties, e.g., the sequence of ratios $(s(n)/s(n+1))_{n\geqslant 0}$ enumerates the non-negative rationals in reduced form and without repeats, and—like the diffraction measure of the Thue–Morse sequence—the associated measure is singular continuous. Baake's and Coons's result rests on the fact that the sequence s satisfies certain self-similar type properties; it has a fundamental region of recursion—between consecutive powers of two—and the sum of s over this fundamental region is linearly recurrent, enabling the use of a volume-averaging process.

In this paper, we generalise the result of Baake and Coons [3]. Here, we provide reasonable assumptions on a k-regular sequence f which guarantee the existence of a natural probability measure μ_f associated to $\mathcal{V}_k(f)$. Indeed, there are regular sequences for which such measures do not exist, and for which the sequence of approximants is not even eventually periodic.

The measures μ_f we construct are analogous to mass distributions supported on attractors of certain iterated function systems—those constructed by repeated subdivision; see Falconer [22] for related definitions and Coons and Evans [15] for a family of extended examples related to generalised Cantor sets. Here, rather than iterating intervals under a finite number of maps, the finite approximants are pure point measures on the unit interval, whose corresponding weights possess the recursive structure. In this way, the existence of a natural measure associated with f and $\mathcal{V}_k(f)$ provides a path to derive interesting dynamical systems from $\mathcal{V}_k(f)$ and opens the possibility of associating (fractal) geometric structures to these sequences and spaces.

To see how this works, we start with a real-valued k-regular sequence f and obtain a basis for $\mathcal{V}_k(f)$, the $\mathbb{R}$-vector space generated by the k-kernel of f. In particular, let $k \geqslant 2$ be an integer, f be a k-regular sequence and let the set of sequences

$$\{f = f_1, f_2, \ldots, f_d\} \subseteq \ker_k(f) \tag{2}$$

be a basis for $\mathcal{V}_k(f) := \langle \ker_k(f) \rangle_{\mathbb{R}}$. Set $\mathbf{f}(m) = (f_1(m), f_2(m), \dots, f_d(m))^T$. For each $a \in \{0, \dots, k-1\}$ let $\mathbf{B}_a$ be a $d \times d$ real matrix such that, for all $m \geqslant 0$,

$$\mathbf{f}(km + a) = \mathbf{B}_a \, \mathbf{f}(m). \tag{3}$$

We refer the reader to the seminal paper of Allouche and Shallit [2] and Nishioka's monograph [32, Ch. 5] for details on existence and the finer definitions. Note that there is $\mathbf{w} \in \mathbb{R}^{d \times 1}$ such that for each $i \in \{1, \dots, d\}$ and $n > 0$, we have

$$f_i(m) = e_i^T \mathbf{B}_{(m)_k} \mathbf{w} = e_i^T \mathbf{B}_{i_0} \mathbf{B}_{i_1} \cdots \mathbf{B}_{i_s} \mathbf{w},$$

where e_i is the ith elementary column vector, $(m)_k = i_s \cdots i_1 i_0$ is the base-k expansion of m and $\mathbf{B}_{(m)_k} := \mathbf{B}_{i_0} \mathbf{B}_{i_1} \cdots \mathbf{B}_{i_s}$. Set

$$\mathbf{B} := \sum_{a=0}^{k-1} \mathbf{B}_a.$$

We note that the matrix $\mathbf{B}$ is analogous to the substitution matrix $\mathbf{M}_\varrho$ of a substitution ϱ on a finite alphabet. In fact, if ϱ is a constant-length substitution of length k on the d letters $0, 1, \dots, d-1$, then the matrices $\mathbf{B}_a$ are the so-called digit (instruction) matrices and $\mathbf{M}_\varrho = \mathbf{B}$; compare [4, 34]. When considering the dynamical properties of substitution systems it is common to assume that the substitution is primitive; that is, the non-negative substitution matrix $\mathbf{M}_\varrho$ is primitive. To continue our analogy with substitutions, we make similar assumptions. At first glance, it is reasonable to restrict ourselves to the assumption that $\mathbf{B}$ is primitive. But if this is the case for the k-regular sequence f, then we can consider f as a k^j-regular sequence with j being the smallest positive integer for which $\mathbf{B}^j$ is positive. Thus, in the context of k-regular sequences the distinction between primitivity and positivity is somewhat blurred. Note that this is also tacitly done for substitutions, where one normally chooses an appropriate power j such that $\mathbf{M}_\varrho^j > 0$, or equivalently, $\varrho^j(a)$ contains all the letters of the alphabet $\mathcal{A}$, for all $a \in \mathcal{A}$. Hence we make the following definition; compare [35, Ch. 2] which deals which primitive morphisms and morphic words.

Definition 1. We call a k-regular sequence f *primitive* provided f only takes non-negative values, is not eventually zero, each of the k digit matrices $\mathbf{B}_a$ is non-negative and the matrix $\mathbf{B}$ is positive.

We are close to being able to state our results—the final ingredients are the introduction of fundamental regions and the definition of the related pure point measures. To this end, we note that, in analogy to the lengths of the iterates of a substitution applied to an initial seed, a k-regular sequence f has a fundamental region, the interval $[k^n, k^{n+1})$. Over these regions, the sum of a regular sequence is linearly recurrent, a property which can be thought of as a sort of self-similarity property for regular sequences. This property provides a structured volume to average by, and using it, we can construct a probability measure. Formally, for each $i \in \{1, \dots, d\}$, set

$$\Sigma_i(n) := \sum_{m=k^n}^{k^{n+1}-1} f_i(m) \tag{4}$$

and define

$$\mu_{n,i} := \frac{1}{\Sigma_i(n)} \sum_{m=0}^{k^{n+1}-k^n-1} f_i(k^n + m)\, \delta_{m/k^n(k-1)}, \tag{5}$$

where δ_x denotes the unit Dirac measure at x. We can view $(\mu_{n,i})_{n\in\mathbb{N}_0}$ as a sequence of probability measures on the 1-torus, the latter written as $\mathbb{T} = [0, 1)$ with addition modulo 1. Here, we have simply re-interpreted the (normalised) values of the sequence $(f_i(m))_{m\geqslant 0}$ between k^n and $k^{n+1} - 1$ as the weights of a pure point probability measure on $\mathbb{T}$ supported on the set $\big\{\frac{m}{k^n(k-1)} : 0 \leqslant m < k^n(k-1)\big\}$. Set

$$\boldsymbol{\mu}_n = \boldsymbol{\mu}_{f,n} := (\mu_{n,1}, \dots, \mu_{n,d})^T. \tag{6}$$

Our first result, Theorem 1 below, provides a unique probability measure on $\mathbb{T}$ associated with the space $\mathcal{V}_k(f)$.

Theorem 1. *Let f be a primitive real-valued k-regular sequence. If $f_1, \dots, f_d$ form the basis of $\mathcal{V}_k(f)$ associated with $\mathbf{B}$, then the vectors $\boldsymbol{\mu}_{f,n}$ of pure point measures converge weakly to a vector of probability measures $\boldsymbol{\mu}_f = (\mu_f, \dots, \mu_f)^T$ on $\mathbb{T}$.*

Of course, one may also wish to consider the measure associated with f apart from the full considerations surrounding that of the space $\mathcal{V}_k(f)$. With a little more specificity one can prove a stronger result. Recall that for a real matrix $\mathbf{B}$, the spectral radius $\rho(\mathbf{B})$ is the largest absolute value of its eigenvalues. The joint spectral radius $\rho^*(\{\mathbf{B}_0, \dots, \mathbf{B}_{k-1}\})$ is the corresponding generalisation for a finite set of matrices; see Eq. (17) below for the formal definition.

Theorem 2. *Let f be a real-valued k-regular sequence. Suppose that $\rho(\mathbf{B})$ is the unique dominant eigenvalue of $\mathbf{B}$, $\rho(\mathbf{B}) > \rho^*(\{\mathbf{B}_0, \dots, \mathbf{B}_{k-1}\})$ and that the asymptotical behaviour of $\Sigma_f(n)$ is determined by $\rho(\mathbf{B})$. If the limit $F_f(x)$ of the sequence $\mu_{f,n}([0, x])$ is a function of bounded variation, then $F_f(x) = \mu_f([0, x])$ is the distribution function of a measure μ_f, which is continuous with respect to Lebesgue measure.*

Note that, in lieu of the non-negativity assumption in Theorem 1, we have assumed a non-degeneracy assumption in Theorem 2 which allows one to cover sequences which take negative values. This more general situation may not result in a probability measure, but possibly, a signed measure.

This article is organised as follows. Section 2 contains the proof of our first theorem, the existence of a measure associated with $\mathcal{V}_k(f)$, and Section 3 contains the proof of our second theorem. In Section 4, we provide witnessing examples which demonstrate why our assumptions are reasonable. Finally, we offer some concluding remarks and open questions in Section 5.

2 A natural probability measure associated with $\mathcal{V}_k(f)$

The aim of this section is to prove Theorem 1. To do this, we attempt to mimic the ideas behind the establishment of spectral measures associated with substitution dynamical systems. Fortunately, the generating power series of k-regular sequences satisfy functional equations that can be thought of as taking the place of a substitution. That is, for each $i \in \{1, \dots, d\}$ setting $F_i(z) := \sum_{m \geqslant 0} f_i(m) z^m$ and $\mathbf{F}(z) = (F_1(z), \dots, F_d(z))^T$, we have [32, p. 153] that $\mathbf{F}(z)$ satisfies the Mahler functional equation

$$\mathbf{F}(z) = \mathbf{B}(z)\,\mathbf{F}(z^k), \tag{7}$$

where $\mathbf{B}(z) := \sum_{a=0}^{k-1} \mathbf{B}_a\, z^a$. The matrix-valued function $\mathbf{B}(z)$ is analogous to the Fourier cocycle in the renormalisation theory of substitution and inflation systems, which carries information about features of the underlying diffraction and spectral measures; see Bufetov and Solomyak [11, 12] and Baake, Gähler and Mañibo [4]. Note that $\mathbf{B}(1) = \mathbf{B}$; this specialisation will be discussed more below. Equation (7) shows that the functions $F_i(z)$ behave well under the Frobenius map $z \mapsto z^k$. The functional equation (7) is analogous to a substitution with repeated application mirroring the iterated composition of a substitution. This property is essentially what allows us to form certain cocycles (e.g., see (14)) that, under the primitivity assumption above, have convergence properties which provide for the existence of the desired limit measures. To achieve our goal, we require a few preliminary results.

For $i \in \{0, \dots, k-1\}$ let $\varSigma_i(n)$ be as defined in (4) and set

$$\mathbf{\Sigma}(n) := (\varSigma_1(n), \varSigma_2(n), \dots, \varSigma_d(n))^T.$$

In the Introduction, we stated that the sequences $\varSigma_i(n)$ are linearly recurrent, which is a well-known result, whose proof we include here for completeness.

Lemma 1. *If $n \geqslant 1$, then $\mathbf{\Sigma}(n) = \mathbf{B} \cdot \mathbf{\Sigma}(n-1)$.*

Proof. Here we are considering the sums of each f_i over all integers in the interval $[k^n, k^{n+1} - 1]$. These are precisely all of the integers that have $n + 1$ digits in their k-ary expansions. Noting that the k-ary expansion of a non-zero integer cannot begin with a zero, we thus have that $\varSigma_i(n)$ satisfies

$$\sum_{m=k^n}^{k^{n+1}-1} f_i(m) = \sum_{m=k^n}^{k^{n+1}-1} e_i^T \mathbf{B}_{(m)_k} \mathbf{w} = e_i^T \mathbf{B}^n \sum_{a=1}^{k-1} \mathbf{B}_a \mathbf{w}.$$

Thus

$$\boldsymbol{\Sigma}(n) = \mathbf{B}^n \sum_{a=1}^{k-1} \mathbf{B}_a \mathbf{w} = \mathbf{B} \cdot \mathbf{B}^{n-1} \sum_{a=1}^{k-1} \mathbf{B}_a \mathbf{w} = \mathbf{B} \cdot \boldsymbol{\Sigma}(n-1),$$

where the second equality follows from the first by writing $n = 1 + (n-1)$. □

Now, define the polynomials $b_{ij}(z)$ by $\mathbf{B}(z) = (b_{ij}(z))_{1 \leqslant i,j \leqslant d}$ and for each $n \geqslant 1$ define the matrix

$$\mathbf{A}_n(z) := \left(\frac{\Sigma_j(n-1)\, b_{ij}(z)}{\Sigma_i(n)} \right)_{1 \leqslant i,j \leqslant d}. \tag{8}$$

Note that since f is primitive, the denominator $\Sigma_i(n)$ cannot vanish. The matrices $\mathbf{A}_n(z)$ are normalised versions of the matrix $\mathbf{B}(z)$. In particular, they allow us to lift the result of Lemma 1 to the level of measures; see Proposition 1 below. Before proving that result, we note the following corollary of Lemma 1.

Corollary 1. *For any $n \geqslant 1$, the matrix $\mathbf{A}_n(1)$ is a Markov matrix, i.e., all of its row sums are equal to* 1.

Recall that for two finite Borel measures μ and ν on $\mathbb{T}$, the convolution $\mu * \nu$ is defined by

$$(\mu * \nu)(g) = \int_{\mathbb{T}\times\mathbb{T}} g(x+y)\, \mathrm{d}\mu(x)\, \mathrm{d}\nu(y),$$

for continuous functions g on $\mathbb{T}$. Consequently, for Dirac measures δ_x and δ_y, we have $\delta_x * \delta_y = \delta_{x+y}$, which is a linear operation. We let $(\delta_x)^r = \delta_{rx}$ be its r-fold convolution product with itself. This allows one to define the value of $p(\delta_x)$, for any polynomial $p \in \mathbb{C}[z]$. Linearity also allows one to define the convolution product $\mathbf{M}_1 * \mathbf{M}_2$ of two matrices whose entries are linear combinations of Dirac measures, where the usual point-wise multiplication is replaced by the convolution product.

Proposition 1. *For $n \geqslant 1$, one has*

$$\boldsymbol{\mu}_n = \mathbf{A}_n\big(\delta_{1/k^n(k-1)}\big) * \boldsymbol{\mu}_{n-1} \tag{9}$$

where $\mathbf{A}_n(z)$ is the matrix-valued function defined in Eq. (8).

Proof. Fix an $i \in \{1, \ldots, d\}$ and consider the ith entry of the vector on the right-hand side of (9). For this entry, we have

$$
\begin{aligned}
e_i^T \mathbf{A}_n\big(\delta_{1/k^n(k-1)}\big) * \boldsymbol{\mu}_{n-1} &= \sum_{j=1}^{d} \frac{\varSigma_j(n-1)\, b_{ij}\big(\delta_{1/k^n(k-1)}\big)}{\varSigma_i(n)} * \mu_{n-1,j} \\
&= \frac{1}{\varSigma_i(n)} \sum_{j=1}^{d} \sum_{\ell=0}^{k^n-k^{n-1}-1} b_{ij}\big(\delta_{1/k^n(k-1)}\big) * f_j(k^{n-1}+\ell)\, \delta_{\ell/k^{n-1}(k-1)} \\
&= \frac{1}{\varSigma_i(n)} \sum_{\ell=0}^{k^n-k^{n-1}-1} \sum_{a=0}^{k-1}\sum_{j=1}^{d} (\mathbf{B}_a)_{ij}\, \big(\delta_{a/k^n(k-1)}\big) * f_j(k^{n-1}+\ell)\, \delta_{\ell/k^{n-1}(k-1)} \\
&= \frac{1}{\varSigma_i(n)} \sum_{\ell=0}^{k^n-k^{n-1}-1} \sum_{a=0}^{k-1} \delta_{(k\ell+a)/k^n(k-1)} \sum_{j=1}^{d} (\mathbf{B}_a)_{ij}\, f_j(k^{n-1}+\ell) \\
&= \frac{1}{\varSigma_i(n)} \sum_{\ell=0}^{k^n-k^{n-1}-1} \sum_{a=0}^{k-1} \delta_{(k\ell+a)/k^n(k-1)}\, f_i(k^n+k\ell+a) \\
&= \frac{1}{\varSigma_i(n)} \sum_{m=0}^{k^{n+1}-k^n-1} f_i(k^n+m)\, \delta_{m/k^n(k-1)} \\
&= \mu_{n,i} = e_i^T \boldsymbol{\mu}_n.
\end{aligned}
$$

Here, the third equality follows using the definition of $b_{ij}(z)$, the fifth equality follows by invoking (3) and the sixth step is just a change of index. □

The following is an immediate consequence of Proposition 1.

Corollary 2. *For $n \geqslant 1$, the Fourier coefficients $\widehat{\mu}_n(t)$ satisfy*

$$
\widehat{\boldsymbol{\mu}_n}(t) = \mathbf{A}_n\Big(\mathrm{e}^{-\frac{2\pi \mathrm{i} t}{k^n(k-1)}}\Big)\, \widehat{\boldsymbol{\mu}_{n-1}}(t). \tag{10}
$$

for all $t \in \mathbb{Z}$.

Remark 1. While the convergents μ_n are pure point probability measures on $\mathbb{T}$, one can also consider the measures $\nu_n = \delta_{\mathbb{Z}} * \mu_n$, which are $\mathbb{Z}$-periodic measures in $\mathbb{R}$. Here, one has $\widehat{\nu}_n \colon \mathbb{R} \to \mathbb{C}$, with $\widehat{\mathbb{R}} = \mathbb{R}$. In the discussion below, we only carry out the analysis for μ_n, but the (vague) convergence of the relevant matrix products also hold for ν_n. ♢

Using Equation (9), one can construct the infinite matrix convolution and hope for the existence of the limit vector

$$
\boldsymbol{\mu} := \left(\mathop{\Huge *}_{n=1}^{\infty} \mathbf{A}_n\big(\delta_{1/k^n(k-1)}\big) \right) * \boldsymbol{\mu}_0. \tag{11}
$$

The existence of the limit vector $\boldsymbol{\mu}$ depends on the convergence of the Fourier coefficients, which boils down to the compact convergence of the analytic matrix product

$$\prod_{n=1}^{\infty} \mathbf{A}_n\big(\mathrm{e}^{-\frac{2\pi i t}{k^n(k-1)}}\big), \tag{12}$$

a property we demonstrate below.

We require the following lemma that the positivity of $\mathbf{B}$ implies the convergence of the quotients $\Sigma_i(n-1)/\Sigma_j(n)$ to a positive value.

Lemma 2. *Let f be a primitive k-regular sequence. Then for each $i,j \in \{1,\dots,d\}$, $\lim_{n\to\infty} \Sigma_j(n-1)/\Sigma_i(n) = c_{ij}$ where $c_{ij} > 0$.*

Proof. This follows from the fact that $\mathbf{B}$ is positive and that f is non-negative (and not trivial), so that given an $i \in \{1,\dots,d\}$ there is a constant $c_i > 0$ such that

$$\Sigma_i(n) \sim c_i \cdot \rho_{\mathrm{PF}}(\mathbf{B})^n, \tag{13}$$

as $n \to \infty$, where the positive real number $\rho_{\mathrm{PF}}(\mathbf{B})$ is the Perron–Frobenius eigenvalue of $\mathbf{B}$. To see this, let $\rho_{\mathrm{PF}}(\mathbf{B})$ be the Perron–Frobenius eigenvalue of $\mathbf{B}$ and let $\mathbf{v}$ be the corresponding positive Perron–Frobenius eigenvector [33]. Let j be a positive integer so that $\mathbf{B}^j\mathbf{w}$ is entry-wise greater than $\mathbf{v}$. Then for $n > j$,

$$\begin{aligned}\Sigma_i(n) = e_i^T\mathbf{B}^n\,\mathbf{w} = e_i^T\mathbf{B}^{n-j}\,\mathbf{B}^j\,\mathbf{w} &= e_i^T\mathbf{B}^{n-j}\,\mathbf{v} + e_i^T\mathbf{B}^{n-j}\,(\mathbf{B}^j\,\mathbf{w} - \mathbf{v})\\ &\geqslant e_i^T\mathbf{B}^{n-j}\,\mathbf{v} = e_i^T\rho_{\mathrm{PF}}(\mathbf{B})^{n-j}\,\mathbf{v} = (e_i^T\mathbf{v}/\rho_{\mathrm{PF}}(\mathbf{B})^j)\cdot\rho_{\mathrm{PF}}(\mathbf{B})^n.\end{aligned}$$

Since $\rho_{\mathrm{PF}}(\mathbf{B})$ is a simple eigenvalue of $\mathbf{B}$ with maximal modulus, comparing the eigenvalue expansion of the linear recurrent sequence $\Sigma_i(n)$ with this inequality proves (13). The lemma follows immediately. □

To prove the next result, we adapt a technique used by Baake and Grimm in determining the intensities of Bragg peaks of Pisot substitutions via the internal Fourier cocycle [7, Thm. 4.6] to prove compact convergence of infinite products of matrices. This is a higher-dimensional extension of a method employed in [17, Thm. 2.1(b)] in dimension one. This convergence and Lévy's continuity theorem are the main ingredients of the proof.

Theorem 3. *Let f be a primitive k-regular sequence. Then the weak limit measure vector $\boldsymbol{\mu}$ in Eq. (11) exists.*

Proof. It follows from the convergence of the quotients $\Sigma_j(n-1)/\Sigma_i(n)$ in Lemma 2 that $\mathbf{A}_n(z) \to \mathbf{A}(z)$ as $n \to \infty$, with $(\mathbf{A}(z))_{ij} = c_{ij}b_{ij}(z)$. In particular, one has $\mathbf{A}_n(1) \to \mathbf{A}(1)$, where every $\mathbf{A}_n(1)$ is a Markov matrix by Corollary 1. Setting $\mathbf{S}_n(t) := \mathbf{A}_n\big(\mathrm{e}^{\frac{-2\pi i t}{(k-1)}}\big)$, one gets $\mathbf{S}_n(t) \to \mathbf{S}(t) := \mathbf{A}\big(\mathrm{e}^{\frac{-2\pi i t}{(k-1)}}\big)$.

For $1 \leqslant r \leqslant \ell$ define the product

$$\mathbf{S}^{(\ell,r)}(t) := \mathbf{S}_\ell\Big(\frac{t}{k^\ell}\Big)\cdots\mathbf{S}_r\Big(\frac{t}{k^r}\Big).$$

These products satisfy the identity $\mathbf{S}^{(\ell,r)}(t) = \mathbf{S}^{(\ell,c+1)}(t)\,\mathbf{S}^{(c,r)}(t)$, for $1 \leqslant r \leqslant c < \ell$. Moreover, $\mathbf{S}^{(\ell,r)}(0)$ is a Markov matrix for any such ℓ, r, being a finite product of Markov matrices. Here we let $\mathbf{S}^{(\ell,0)}(t) := \mathbf{S}^{(\ell)}(t)$. One then has

$$\prod_{n=1}^{m} \mathbf{A}_n\big(\mathrm{e}^{-\frac{2\pi \mathrm{i} t}{k^n(k-1)}}\big) = \mathbf{S}^{(m)}(t) = \mathbf{S}_m\left(\frac{t}{k^m}\right)\mathbf{S}_{m-1}\left(\frac{t}{k^{m-1}}\right)\cdots\,\mathbf{S}_1\left(\frac{t}{k}\right). \quad (14)$$

In order to prove weak convergence of the approximant measures, we need to show that the vector of Fourier coefficients converges pointwise in $\mathbb{Z}$. This entails showing that the $\mathbf{S}^{(m)}(t)$ converges compactly on $\mathbb{Z}$. To this end, we first prove that for a fixed $m \in \mathbb{N}$, $\mathbf{S}^{(m,\ell)}(t)$ is equicontinuous at $t = 0$ for all ℓ satisfying $1 \leqslant \ell \leqslant m$.

Employing a telescoping argument, one obtains the equality

$$\mathbf{S}^{(m,\ell)}(t) - \mathbf{S}^{(m,\ell)}(0) = \sum_{j=\ell-1}^{m-1}\Big(\mathbf{S}^{(m,j+2)}(0)\,\mathbf{S}^{(j+1,\ell)}(t) - \mathbf{S}^{(m,j+1)}(0)\,\mathbf{S}^{(j,\ell)}(t)\Big)$$

This implies the following

$$\begin{aligned}
&\|\mathbf{S}^{(m,\ell)}(t) - \mathbf{S}^{(m,\ell)}(0)\|_\infty \\
&\qquad \leqslant \sum_{j=0}^{m-1} \|\mathbf{S}^{(m,j+2)}(0)\,\mathbf{S}^{(j+1,\ell)}(t) - \mathbf{S}^{(m,j+1)}(0)\,\mathbf{S}^{(j,\ell)}(t)\|_\infty \\
&\qquad \leqslant \sum_{j=0}^{m-1} \|\mathbf{S}^{(m,j+2)}(0)\|_\infty \cdot \|\mathbf{S}^{(j+1,\ell)}(t) - \mathbf{S}_{j+1}(0)\,\mathbf{S}^{(j,\ell)}(t)\|_\infty \\
&\qquad \leqslant \sum_{j=0}^{m-1} \|\mathbf{S}^{(m,j+2)}(0)\|_\infty \cdot \|\mathbf{S}_{j+1}(\tfrac{t}{k^{j+1}}) - \mathbf{S}_{j+1}(0)\|_\infty \cdot \|\mathbf{S}^{(j,\ell)}(t)\|_\infty \\
&\qquad \leqslant \sum_{j=0}^{m-1} \|\mathbf{S}_{j+1}(\tfrac{t}{k^{j+1}}) - \mathbf{S}_{j+1}(0)\|_\infty,
\end{aligned}$$

where the last step follows from the properties $\|\mathbf{S}^{(m,j+2)}(0)\|_\infty = 1$ and

$$\|\mathbf{S}^{(j,\ell)}(t)\|_\infty \leqslant \|\mathbf{S}^{(j,\ell)}(0)\|_\infty = 1,$$

where the inequality holds since the matrices $\mathbf{B}_a$ are non-negative and the equality follows from $\mathbf{S}^{(j,\ell)}(0)$ being Markov.

Now, let $\varepsilon > 0$ be given and choose $\delta > 0$ such that

$$\|\mathbf{S}_{j+1}(\tfrac{t}{k^{j+1}}) - \mathbf{S}_{j+1}(0)\|_\infty < \frac{\varepsilon}{k^{j+1}}$$

holds whenever $|\frac{t}{k^{j+1}}| < \delta$. This yields

$$\begin{aligned}\|\mathbf{S}^{(m,\ell)}(t) - \mathbf{S}^{(m,\ell)}(0)\|_\infty &\leqslant \sum_{j=0}^{m-1} \|\mathbf{S}_{j+1}(\tfrac{t}{k^{j+1}}) - \mathbf{S}_{j+1}(0)\|_\infty \\ &< \sum_{j=0}^{\infty} \frac{\varepsilon}{k^j} = \varepsilon\left(\frac{k}{k-1}\right) \qquad (15)\end{aligned}$$

which proves that $\mathbf{S}^{(m,\ell)}(t)$ is equicontinuous at $t = 0$.

Compact convergence means for any given compact set $K \subset \mathbb{Z}$, $\mathbf{S}^{(m,0)}(t)$ uniformly converges in K, which we show by proving it is uniformly Cauchy in K. Choose p such that $|\frac{t}{k^p}| < \delta$ for all $t \in K$, which implies $|\frac{t}{k^j}| < \delta$ for all $j \geqslant p$ and $t \in K$. One then has

$$\begin{aligned}\|\mathbf{S}^{(p+q+r)}(t) - \mathbf{S}^{(p+q)}(t)\|_\infty &\leqslant \|\mathbf{S}^{(p+q+r,p+1)}(t) - \mathbf{S}^{(p+q,p+1)}(t)\|_\infty \cdot \|\mathbf{S}^{(p)}(t)\|_\infty \\ &\leqslant \|\mathbf{S}^{(p+q+r,p+1)}(t) - \mathbf{S}^{(p+q,p+1)}(t)\|_\infty,\end{aligned}$$

since $\|\mathbf{S}^{(p)}(t)\| \leqslant 1$. Thus, using the triangle inequality, we obtain

$$\begin{aligned}\|\mathbf{S}^{(p+q+r,p+1)}(t) - \mathbf{S}^{(p+q,p+1)}(t)\|_\infty &\leqslant \|\mathbf{S}^{(p+q+r,p+1)}(t) - \mathbf{S}^{(p+q+r,p+1)}(0)\|_\infty \\ + \|\mathbf{S}^{(p+q+r,p+1)}(0) - \mathbf{S}^{(p+q,p+1)}(0)\|_\infty &+ \|\mathbf{S}^{(p+q,p+1)}(t) - \mathbf{S}^{(p+q,p+1)}(0)\|_\infty\end{aligned}$$

where the first and the third summands are strictly less than $\varepsilon \cdot k/(k-1)$ by invoking Eq. (15) for large enough p. The second summand splits further, into

$$\begin{aligned}\|\mathbf{S}^{(p+q+r,p+1)}(0) - \mathbf{S}^{(p+q,p+1)}(0)\|_\infty &\leqslant \|\mathbf{S}^{(p+q+r,p+1)}(0) - \mathbf{A}(1)^{q+r}\|_\infty \\ + \|\mathbf{A}(1)^{q+r} - \mathbf{A}(1)^q\|_\infty &+ \|\mathbf{S}^{(p+q,p+1)}(0) - \mathbf{A}(1)^q\|_\infty, \qquad (16)\end{aligned}$$

Since $\mathbf{S}_\ell(0)$ converges to $\mathbf{A}(1)$, one can choose p' large enough such that

$$\|\mathbf{S}_\ell(0) - \mathbf{A}(1)\|_\infty < \frac{\varepsilon}{k^\ell}$$

for all $\ell \geqslant p'$. One can then invoke a similar argument to show that the first and the third summands in Eq. (16) are strictly less than $\varepsilon \cdot k/(k-1)$ whenever $p \geqslant p'$. Finally, since $\mathbf{A}(1)$ is Markov, $\mathbf{A}(1)^q$ converges to a steady state matrix whence one can choose q large enough such that the second summand is less than ε. Since r is arbitrary, this means for all $m, n \geqslant \max\{p, p'\} + q$ one has that

$$\|\mathbf{S}^{(m)}(t) - \mathbf{S}^{(n)}(t)\|_\infty < \varepsilon\left(\frac{5k-1}{k-1}\right),$$

and thus, $\mathbf{S}^{(n)}(t)$ converges uniformly on K. Since K is arbitrary, we have shown compact convergence on $\mathbb{Z}$.

This convergence, together with Corollary 2, implies that the limit vector

$$\widehat{\boldsymbol{\mu}}(t) := \Big(\lim_{m\to\infty} \mathbf{S}^{(m)}(t)\Big)\, \widehat{\boldsymbol{\mu}_0}(t)$$

of Fourier coefficients is well defined for all $t \in \mathbb{Z}$. The weak convergence to a limit measure vector in Eq. (11) is now guaranteed by Lévy's continuity theorem; see [10, Sec. 26, Cor. 1] or [9, Thm. 3.14]. □

We next prove that the entries of the limit vector of measure in Theorem 3 are all the same measure.

Proposition 2. *If f is a primitive k-regular sequence, then $\boldsymbol{\mu} = (\mu_f, \dots, \mu_f)^T$, where μ_f is a probability measure on $\mathbb{T}$. That is, for each $i \in \{1, \dots, d\}$, the weak limit of $\mu_{n,i}$ is μ_f.*

Proof. Let $\mathbf{L}(t) := \lim_{m\to\infty} \mathbf{S}^{(m)}(t)$. By Lemma 2, $\mathbf{A}_n(1) \to \mathbf{A}(1)$ as $n \to \infty$, where $\mathbf{A}(1)$ is a primitive Markov matrix, where the primitivity of $\mathbf{A}(1)$ follows from the positivity of $\mathbf{B}$. This means $\mathbf{A}(1)^n$ converges to the rank-1 projector $\mathbf{P}_{\mathbf{A}}$ corresponding to the eigenvector $\mathbf{1} = (1, \dots, 1)^T$. One then gets the equality

$$\mathbf{A}(1)\,\mathbf{L}(t) = \mathbf{L}(t) = \mathbf{P}_{\mathbf{A}}\,\mathbf{L}(t),$$

which implies $\widehat{\boldsymbol{\mu}}(t) = \mathbf{P}_{\mathbf{A}}\,\widehat{\boldsymbol{\mu}}(t) = c(t)\,\mathbf{1}$. This means that the limit measures μ_i have the same Fourier coefficients for all t, and hence must correspond to the same measure $\mu = \mu_f$ for all i satisfying $1 \leqslant i \leqslant d$. □

Theorem 1 follows by combining Theorem 3 and Proposition 2.

As stated previously, the non-negativity assumptions on the k matrices $\mathbf{B}_a$ and the positivity (primitivity, initially) assumption on $\mathbf{B}$ are natural, especially if one views it as the corresponding analogue of the substitution matrix $\mathbf{M}_\varrho$ for shift spaces arising from a substitution ϱ. While these assumptions seem less natural in the context of regular sequences, they are satisfied by many of the regular sequences that concern computer scientists and number theorists.

For simplicity, let ϱ be a primitive *constant-length* substitution on a finite alphabet $\mathcal{A}$ (like the Thue–Morse substitution given in the Introduction). Consider the one-sided hull X_ϱ, which can be defined by picking any fixed point x of ϱ, and building its orbit closure under the shift action T, i.e.,

$$X_\varrho = \overline{\{T^n x \mid n \in \mathbb{N}\}},$$

where the closure is seen with respect to the local topology in $\mathcal{A}^{\mathbb{N}}$.

The primitivity of $\mathbf{M}_\varrho$ implies that the hull X_ϱ is strictly ergodic. We then have the following well known result, which is a consequence of this property; see Baake's and Grimm's monograph [6, Ch. 9] for details and definitions. Here, we define the diffraction measure via Eq. (1).

Fact 1. *Let ϱ be a primitive constant-length substitution on a finite alphabet $\mathcal{A}$ and $x \in X_{\varrho}$. Then for any fixed weight function $w\colon \mathcal{A} \to \mathbb{C}$, the diffraction measure μ_w exists and is the same measure for all $x \in X_{\varrho}$.* □

The result we obtain in Proposition 2 can then be seen as an analogous uniqueness result in the context of regular sequences for the basis $\{f_1, \dots, f_d\}$ of $\mathcal{V}_k(f)$.

3 A natural probability measure associated with f

In this section we will prove Theorem 2, via the general situation described at the end of the Introduction. To set up, we recall the following notation, now for only a single k-regular sequence f,

$$\Sigma_f(n) := \sum_{m=k^n}^{k^{n+1}-1} f(m)$$

and

$$\mu_{f,n} := \frac{1}{\Sigma_f(n)} \sum_{m=0}^{k^{n+1}-k^n-1} f(k^n + m)\, \delta_{m/k^n(k-1)},$$

where δ_x denotes the unit Dirac measure at x. Also as previously, let f be defined by the matrices $\mathbf{B}_0, \dots, \mathbf{B}_{k-1}$ and the vector $\mathbf{w} \in \mathbb{R}^{d\times 1}$ such that $f(m) = e_1^T \mathbf{B}_{(m)_k} \mathbf{w}$, where $(m)_k = i_s \cdots i_1 i_0$ is the base-k expansion of m and $\mathbf{B}_{(m)_k} := \mathbf{B}_{i_0}\mathbf{B}_{i_1}\cdots\mathbf{B}_{i_s}$. As before, set $\mathbf{B} := \sum_{a=0}^{k-1} \mathbf{B}_a$. Let $\rho(\mathbf{M})$ denote the spectral radius of the matrix $\mathbf{M}$ and denote the *joint spectral radius* of a finite set of matrices $\{\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_\ell\}$, by the real number

$$\rho^*(\{\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_\ell\}) = \limsup_{n\to\infty} \max_{1\leqslant i_0,\dots,i_{n-1}\leqslant \ell} \left\| \mathbf{M}_{i_0}\mathbf{M}_{i_1}\cdots\mathbf{M}_{i_{n-1}} \right\|^{1/n}, \quad (17)$$

where $\|\cdot\|$ is any (submultiplicative) matrix norm. This quantity was introduced by Rota and Strang [36] and has a wide range of applications. For an extensive treatment, see Jungers's monograph [24].

Unlike Theorem 1, so also unlike the previous section, we do not yet assume that f is non-negative, nor that $\mathbf{B}$ is positive. However, to avoid degeneracies (discussed in the next section), we assume that the spectral radius $\rho(\mathbf{B})$ is the unique dominant eigenvalue of $\mathbf{B}$, that

$$\rho := \rho(\mathbf{B}) > \rho^*(\{\mathbf{B}_0, \dots, \mathbf{B}_{k-1}\}) =: \rho^*,$$

that for n large enough $\Sigma_f(n) \neq 0$ and that the asymptotical behaviour of $\Sigma_f(n)$ is determined by $\rho(\mathbf{B})$. This last assumption will be highlighted and made explicit below.

To exploit the asymptotical behaviour of $\Sigma_f(n)$, we use a result of Dumas [19, Thm. 3] on the asymptotic nature of the partial sums $\sum_{m\leqslant x} f(m)$. Throughout this paper, we have used the convention that $\mathbf{B}_{(m)_k} := \mathbf{B}_{i_0}\mathbf{B}_{i_1}\cdots\mathbf{B}_{i_s}$, where $(m)_k = i_s \cdots i_1 i_0$ is the base-k expansion of m, however, the result of Dumas [19] that we use here requires multiplying in the opposite order. These two representations are related via matrix transposition;

$$f(m) = e_1^T \mathbf{B}_{(m)_k} \mathbf{w} = \mathbf{w}^T \mathbf{B}_{(m)_k}^T e_1.$$

To set up Dumas' result, we require some further notation. Let the column vector $\mathbf{v}_\rho$ be a ρ-eigenvector of $\mathbf{B}^T$ and let the nonzero real number e_ρ be such that $\mathbf{v}_\rho e_\rho$ equals the component of e_1 in the invariant subspace of $\mathbf{B}^T$ associated with ρ. Finally, we define the matrix-valued function $\mathbf{F}_\rho : \mathbb{R} \to \mathbb{R}^{d\times 1}$ by

$$\mathbf{F}_\rho(x) \cdot \rho = \sum_{a=0}^{k-1} \mathbf{B}_a^T \cdot \mathbf{F}_\rho(kx - a), \tag{18}$$

with the boundary conditions

$$\mathbf{F}_\rho(x) = \begin{cases} 0 & \text{for } x \leqslant 0 \\ \mathbf{v}_\rho & \text{for } x \geqslant 1. \end{cases}$$

The function $\mathbf{F}_\rho$ exists and is unique since $\rho > \rho^*$. Moreover, the function $\mathbf{F}_\rho$ is Hölder continuous with exponent α for any $\alpha < \log_k(\rho/\rho^*)$. Functional equations such as (18) are known as *dilation equations* or *two-scale difference equations* in the literature; seminal work on these was done by Daubechies and Lagarias [17, 18]. See also Micchelli and Prautzsch [31]. A point of interest for our context is that both the above-mentioned papers of Daubechies and Lagarias as well as the seminal paper of Allouche and Shallit [2] introducing k-regular sequences were published within the span of one year. Twenty years later, Dumas [19]—extending ideas of Coquet [16]—connected these ideas by showing explicitly how one can use a dilation equation to determine the asymptotic growth of the partial sums of a regular sequence. We record his result here in the special case fit for our purpose.

Theorem 4 (Dumas). *Suppose that the spectral radius $\rho = \rho(\mathbf{B})$ is the unique dominant eigenvalue of $\mathbf{B}$ and that $\rho(\mathbf{B}) > \rho^*(\{\mathbf{B}_0, \dots, \mathbf{B}_{k-1}\})$. Then*

$$\sum_{m \leqslant x} f(m) = \mathbf{w}^T\, \mathbf{E}_\rho(\log_k(x))\, e_\rho + o(\rho^{\log_k(x)}),$$

where

$$\mathbf{E}_\rho(x) := (\mathbf{I}_d - \mathbf{B}_0^T)\, \mathbf{v}_\rho \left(\frac{1-\rho^{\lfloor x\rfloor+1}}{1-\rho}\right) + \rho^{\lfloor x\rfloor+1}\, \mathbf{F}_\rho(k^{\{x\}-1}),$$

where $\mathbf{I}_d$ denotes the $d \times d$ identity matrix.

A point to be made here is that, while Theorem 4 is certainly technical, in the cases we apply it, for integers $x \in [k^n, k^{n+1})$, the integer part $\lfloor x\rfloor = n$ is constant, so only the dependence on the fractional parts $\{x\}$ will need to be dealt with. For a detailed example of how this theorem can be applied to give a distribution function, see Baake and Coons [3, Sec. 3], where they give an account concerning the Stern sequence.

Now, applying Theorem 4 to the complete sums $\Sigma_f(n)$ and using the transposed representation as that theorem requires, gives

$$\Sigma_f(n) = \sum_{m=k^n}^{k^{n+1}-1} f(m) = \rho^{n+1}\mathbf{w}^T\left(\mathbf{F}_\rho\left(\frac{k-1/k^n}{k}\right) - \mathbf{F}_\rho\left(\frac{1}{k}\right)\right)e_\rho + o(\rho^n).$$

As $n \to \infty$, since $\mathbf{F}_\rho$ is Hölder continuous,

$$\mathbf{F}_\rho\left(\frac{k-1/k^n}{k}\right) = \mathbf{v}_\rho + o(1),$$

so that

$$\Sigma_f(n) = \rho^{n+1}\mathbf{w}^T\left(\mathbf{v}_\rho - \mathbf{F}_\rho\left(\frac{1}{k}\right)\right)e_\rho + o(\rho^n). \tag{19}$$

Since we are assuming that for n large enough $\Sigma_f(n) \neq 0$ and that the asymptotical behaviour of $\Sigma_f(n)$ is determined by $\rho(\mathbf{B})$, we have that

$$\mathbf{w}^T\left(\mathbf{v}_\rho - \mathbf{F}_\rho\left(\frac{1}{k}\right)\right) \neq 0. \tag{20}$$

Arguing as in the previous paragraph allows us to prove the following result.

Theorem 5. *Let f be a real-valued k-regular sequence. Suppose that $\rho(\mathbf{B})$ is the unique dominant eigenvalue of $\mathbf{B}$, $\rho(\mathbf{B}) > \rho^*(\{\mathbf{B}_0, \dots, \mathbf{B}_{k-1}\})$ and that the asymptotical behaviour of $\Sigma_f(n)$ is determined by $\rho(\mathbf{B})$. Then, the limit $F_f(x)$ of the sequence $\mu_{f,n}([0,x])$ exists. Moreover, the function $F_f(x)$ is Hölder continuous with exponent α for any $\alpha < \log_k(\rho/\rho^*)$.*

Proof. Let $x \in \mathbb{T}$ and consider the sequence of functions $\mu_{f,n}([0,x])$. Then, applying the argument of the above paragraph with $(k-1/k^n)/k$ replaced by $1+(k-(1+1/k^n))x$, we have

$$\begin{aligned}\Sigma_f(n)\,\mu_{f,n}([0,x]) &= \sum_{m=k^n \leqslant m \leqslant k^n(1+(k-(1+1/k^n))x)} f(m) \\ &= \rho^{n+1}\mathbf{w}^T\left(\mathbf{F}_\rho\left(\frac{1+(k-(1+1/k^n))x}{k}\right) - \mathbf{F}_\rho\left(\frac{1}{k}\right)\right)e_\rho + o(\rho^n).\end{aligned}$$

As before, using the Hölder continuity of $\mathbf{F}_\rho$, we obtain

$$\Sigma_f(n)\mu_{f,n}([0,x]) = \rho^{n+1}\mathbf{w}^T\left(\mathbf{F}_\rho\left(\frac{1+(k-1)x}{k}\right) - \mathbf{F}_\rho\left(\frac{1}{k}\right) + o(1)\right)e_\rho + o(\rho^n).$$

Using the asymptotic (19) yields

$$\mu_{f,n}([0,x]) = \frac{\mathbf{w}^T \left(\mathbf{F}_\rho\big(\frac{1+(k-1)x}{k}\big) - \mathbf{F}_\rho\big(\frac{1}{k}\big) + o(1)\right) + o(1)}{\mathbf{w}^T \left(\mathbf{v}_\rho - \mathbf{F}_\rho\big(\frac{1}{k}\big)\right) + o(1)}.$$

Since the denominator limits to a nonzero value, the point-wise limit of $\mu_{f,n}([0,x])$ exists for all x; explicitly,

$$F_f(x) = \lim_{n\to\infty} \mu_{f,n}([0,x]) = \frac{\mathbf{w}^T \left(\mathbf{F}_\rho\big(\frac{1+(k-1)x}{k}\big) - \mathbf{F}_\rho\big(\frac{1}{k}\big)\right)}{\mathbf{w}^T \left(\mathbf{v}_\rho - \mathbf{F}_\rho\big(\frac{1}{k}\big)\right)}.$$

Finally, we note that the function $F_f(x)$ is Hölder continuous with exponent α for any $\alpha < \log_k(\rho/\rho^*)$, a property it inherits directly from $\mathbf{F}_\rho$. □

The assumptions of Theorem 5 are not strong enough to guarantee the existence of a measure μ_f for which $F_f(x)$ is a distribution function, but the additional assumption that $F_f(x)$ is of bounded variation, as assumed in the statement of Theorem 2, suffices.

Proof of Theorem 2. We start with the function $F_f(x)$ provided by Theorem 5. Assuming that $F_f(x)$ is of bounded variation, we form the (possibly) signed measure μ_f via a Riemann–Stieltjes integral, assigning the value

$$\mu_f((a,b]) := \int_a^b \mathrm{d}\mu_f := F_f(b) - F_f(a)$$

to any interval $(a,b] \subseteq \mathbb{T}$; see Lang [29, Chp. X] for details on Riemann–Stieltjes integration and measure. This is a Borel measure and has distribution function $\mu_f([0,x]) = F_f(x)$. Since the distribution functions $\mu_{f,n}([0,x])$ of the measures $\mu_{f,n}$ are converging point-wise to the continuous distribution function $\mu_f([0,x])$ of the measure μ_f, the measures $\mu_{f,n}$ are converging weakly to μ_f. Moreover, the measure μ_f has no pure points since $\mu_f([0,x]) = F_f(x)$ is continuous, thus μ_f is continuous with respect to Lebesgue measure. □

We have the following corollary.

Theorem 6. *Suppose that f is a positive real-valued k-regular sequence such that $\rho(\mathbf{B})$ is the unique dominant eigenvalue of $\mathbf{B}$, $\rho(\mathbf{B}) > \rho^*(\{\mathbf{B}_0, \dots, \mathbf{B}_{k-1}\})$ and that the asymptotical behaviour of $\Sigma_f(n)$ is determined by $\rho(\mathbf{B})$. Then the measure μ_f exists and is continuous.*

Proof. This follows immediately using the fact that the partial sums $\sum_{m\leqslant y} f(m)$ are increasing with y so that $\mu_f([0,x])$ is increasing and thus of bounded variation. □

Remark 2. The careful reader will have noticed that in the Introduction, we chose the regular sequence f to be a member of the basis of $\mathcal{V}_k(f)$. While this choice is standard in the area, it is even more convenient in our context—it makes it obvious that the measures obtained in Sections 2 and 3 are indeed the same.

4 Some comments on assumptions

Theorem 1 relies on one main assumption, the primitivity of f, and Theorem 2 relies on a two key assumptions. In this section, we consider these assumptions, highlighting certain pathologies which arise for examples which do not satisfy them.

4.1 An example where $\mu_1 \neq \mu_2$ for a basis $\{f_1, f_2\}$

The assumption of primitivity of the k-regular sequence f in Theorem 1 implies that the matrix $\mathbf{B}$ is positive; recall, this positivity was gained by starting with a primitive matrix and considering f as a k^j-regular sequence, where j was the minimal positive integer such that $\mathbf{B}^j$ is positive. Sometimes such a choice is not immediately possible. For example, consider the Josephus sequence J, which is 2-regular and determined by $J(0) = 0$ and the recursions $J(2n) = 2J(n) - 1$ and $J(2n+1) = 2J(n)+1$. These recursions imply that the sequences J and $\mathbf{1}$ (the constant sequence) form a basis for $\mathcal{V}_2(J)$. With this basis, we arrive at the linear representation $J(n) = e_1^T\, \mathbf{B}_{(n)_2}(0,1)^T$, where

$$\mathbf{B}_0 = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B}_1 = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \quad \text{so that} \quad \mathbf{B} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

In this case, $\mathbf{B}$ is neither positive nor primitive, so Theorem 1 cannot be applied. However, according to Theorem 6, one can apply the same limiting measures $\mu_{n,i}$ via the results in Section 3. In the limit, one gets $\mu_J = h(x) \cdot \lambda$ whereas $\mu_{\mathbf{1}} = \lambda$, where λ is normalised Haar measure on $\mathbb{T}$ and $h(x) = 2x$ is the Radon–Nikodym density for the limit measure associated with J.

But, there is hope here—a change of basis allows the use of Theorem 1. We need only note that if we conjugate $\mathbf{B}_0$ and $\mathbf{B}_1$ by the matrix

$$\mathbf{P} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

we arrive at a new linear representation of the sequence J, and a new basis for $\mathcal{V}_2(J)$, where the matrix $\mathbf{B}$ is replaced by

$$\mathbf{P}\mathbf{B}\mathbf{P}^{-1} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix},$$

which is positive. Thus, we can apply Theorem 1 using this new basis resulting in the vector of measures $\boldsymbol{\mu}_J = \mu_J(1,1)^T$. For more information on the Josephus sequence and for a study on measures associated with affine 2-regular sequences, see Evans [21].

4.2 An example where the limit measure does not exist

Shifting now to Theorem 2, consider the assumption that $\rho(\mathbf{B})$ is the unique dominant eigenvalue of $\mathbf{B}$ (which holds if $\mathbf{B}$ were say, primitive). There are plenty of examples of regular sequences where this is not the case. Dumas [19, Ex. 5] gave an interesting example of a 2-regular sequence, which we denote by D, whose associated matrix $\mathbf{B}$ has a negative integer entry, is not primitive and does not have a maximal eigenvalue. Dumas' sequence D is defined by setting

$$\mathbf{B}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B}_1 = \begin{pmatrix} 3 & -3 \\ 3 & 3 \end{pmatrix},$$

and

$$D(m) = e_1^T \, \mathbf{B}_{(m)_2} e_1 = e_1^T \, \mathbf{B}_1^{s_2(m)} e_1, \tag{21}$$

where $s_2(m)$ is the sum of the bits of m. As in the previous section, we consider the sequence of pure point measures

$$\mu_{D,n} := \frac{1}{\Sigma_D(n)} \sum_{m=2^n}^{2^{n+1}-1} D(m) \, \delta_{m/2^{n+1}}, \tag{22}$$

where

$$\Sigma_D(n) := \sum_{m=2^n}^{2^{n+1}-1} D(m) = \frac{3}{2} \cdot \big((1+\mathrm{i})(4+3\mathrm{i})^n + (1-\mathrm{i})(4-3\mathrm{i})^n \big), \tag{23}$$

since in this case, the matrix $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1$ has eigenvalues $4+3\mathrm{i}$ and $4-3\mathrm{i}$, each of modulus 5. To show that the 'limit' μ_D does not exist, it is enough to prove that the sequence $\big(\mu_{D,n}([0,1/2))\big)_{n \geqslant 0}$ does not have a limit. To see why this is enough, note that if the measure μ_D did exist, then so would its distribution function $\mu_D([0,x)) : [0,1] \to \mathbb{R}$, and necessarily we would have that $\lim_{n\to\infty} \mu_{D,n}([0,1/2)) = \mu_D([0,1/2))$.

Proposition 3. *The sequence* $\big(\mu_{D,n}([0,1/2))\big)_{n \geqslant 0}$ *does not converge. Moreover, it is not eventually periodic.*

Proof. To prove this result, we show that there is a subsequence of this sequence which is unbounded. To this end, note that (22) gives

$$\mu_{D,n}([0,1/2)) = \frac{1}{\Sigma_D(n)} \sum_{m=2^n}^{2^n+2^{n-1}-1} D(m).$$

Using (21) we have that

$$\begin{aligned} \sum_{m=2^n}^{2^n+2^{n-1}-1} D(m) &= e_1^T \Bigg(\sum_{\substack{w \in \{0,1\}^* \\ |w| = n-1}} \mathbf{B}_w \Bigg) \mathbf{B}_0 \mathbf{B}_1 e_1 \\ &= e_1^T \left(\mathbf{B}_0 + \mathbf{B}_1 \right)^{n-1} \mathbf{B}_1 e_1 = \sum_{m=2^{n-1}}^{2^n-1} D(m) = \Sigma_D(n-1), \end{aligned}$$

where we have used that $\mathbf{B}_0$ is the identity matrix to obtain the middle equality. Thus

$$\mu_{D,n}([0,1/2)) = \frac{\Sigma_D(n-1)}{\Sigma_D(n)}.$$

By a simple calculation applying (23) to both the numerator and denominator, with some rearrangement, we have

$$\mu_{D,n}([0,1/2)) = \frac{1}{4+3\mathrm{i}} \cdot \frac{1-\mathrm{e}^{-\mathrm{i}(2\vartheta\cdot(n-1)-\pi/2)}}{1-\mathrm{e}^{-\mathrm{i}(2\vartheta\cdot(n)-\pi/2)}},$$

where $\vartheta \approx 0.6435$ is the solution of $\cos\vartheta = 4/5$. Now since ϑ is irrational and not a rational multiple of π, we have that the sequence of fractional parts

$$\big(\{2\vartheta \cdot (n) - \pi/2\}\big)_{n\geqslant 0}$$

is equidistributed in $[0,1)$. In particular, let $M > 0$ be a positive integer. Then there is an $\varepsilon > 0$ satisfying

$$0 < \varepsilon < \frac{\left|1-\mathrm{e}^{-2\mathrm{i}\vartheta}\right|}{5M+1},$$

and there are infinitely many n such that

$$\left|1-\mathrm{e}^{-\mathrm{i}(2\vartheta\cdot(n)-\pi/2)}\right| < \varepsilon.$$

For these infinitely many n,

$$\left|\mu_{D,n}([0,1/2))\right| > \frac{1}{5}\cdot\frac{\left|1-\mathrm{e}^{-2\mathrm{i}\vartheta}\right|-\varepsilon}{\varepsilon} = \frac{1}{5}\cdot\frac{\left|1-\mathrm{e}^{-2\mathrm{i}\vartheta}\right|}{\varepsilon} - \frac{1}{5} > M.$$

Since $M > 0$ can be chosen arbitrarily large, $\big(\mu_{D,n}([0,1/2))\big)_{n\geqslant 0}$ is unbounded, which is the desired result. □

To observe large values of $\big(\mu_{D,n}([0,1/2))\big)_{n\geqslant 0}$ one must be patient. Figure 1 shows values of $\mu_{D,n}([0,1/2))$ for n from 0 to 100000. Modifying the proof of Proposition 3, *mutatis mutandis*, one can show that the sequences $\big(\mu_{D,n}([0,1/2^k))\big)_{n\geqslant 0}$ are unbounded for every $k \geqslant 1$.

COROLLARY 3. *The sequence of the pure point measures $\mu_{D,n}$ does not converge weakly to a finite Borel measure.*

We now consider the dependence of Theorem 2 on the assumption that the asymptotical behaviour of $\Sigma_f(n)$ is determined by $\rho(\mathbf{B})$. While this assumption adds to the technicality of the statement of Theorem 2, seen from the view of some of common situations, the condition is quite natural. When $\mathbf{B}$'s dominant eigenvalue is simple, this assumption is required to avoid certain degenerate

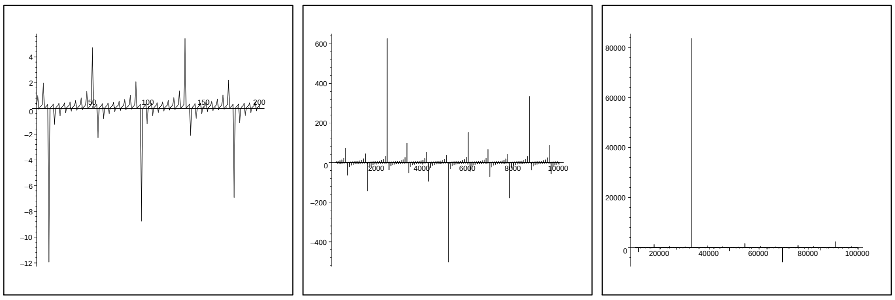

Figure 1: The values of $\mu_{D,n}([0,1/2))$ for (left) $n = 0, \ldots, 200$, (middle) $n = 201, \ldots, 10000$ and (right) $n = 10001, \ldots, 100000$.

situations. For example, consider the positive k-regular sequence f, where $f(m) = e_1^T\, \mathbf{B}_{(m)_k}\, \mathbf{w}$, with

$$\mathbf{B}_0 = \mathbf{B}_1 = \cdots = \mathbf{B}_{k-1} = \begin{pmatrix} * & 0 \\ 0 & c \end{pmatrix}, \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} * \\ 0 \end{pmatrix},$$

where the '$*$' indicates any positive values. For large integers c the dominant eigenvalue of $\mathbf{B}$ will be $k \cdot c$, but none of the matrices $\mathbf{B}_a$ nor $\mathbf{w}$ have any component that interacts with this part of the matrix $\mathbf{B}$, and so this growth is not reflected in the behaviour of $\Sigma_f(n)$. Of course, the linear representation of the sequence is not unique, so this problem is not intrinsic to the sequence. More probable is that one has made a sub-optimal choice of the matrices $\mathbf{B}_a$—as is readily apparent, one can fix this example by deleting the un'$*$' parts.

With these examples in mind, assuming that degenerate situations can be avoided, the sufficient assumptions of Theorem 2 seem quite natural for the existence of such measures. Of course, on a case-by-case basis, the measure could exist without satisfying these conditions, but it is certainly not guaranteed in general.

5 Concluding remarks

In working with examples, curiosities having to do with spectral type are evident, and natural examples with certain properties are elusive. A case in point is the seeming scarcity of regular sequences yielding pure point measures. Of course, Corollary 6 concludes that if f is primitive (and other conditions), then μ_f is continuous. So to find examples, it seems wise to study sequences f which have many zeros. For a trivial example, if one takes $f = \chi_2$ to be the characteristic sequence of the powers of two, then we easily compute that $\mu_{\chi_2} = \delta_0$. Staying with binary sequences for the moment, as soon as the density of ones in the fundamental regions $[2^n, 2^{n+1})$ is bounded below, we are spreading the mass equally to these points, so the mass allocated to any one point in $\mathbb{T}$ must go to zero, and so assuming some regularity on the distribution of the non-zero values of the sequence the measure will be continuous; see Coons and Evans [15] for a family of singular continuous examples related to iterated function systems. Some examples of sequences with pure-point measures exist, see Evans [21]. However, these still occur in a relatively trivial way, where certain terms of the sequence contain a non-zero fraction of its entire mass. One could attempt to produce pure point measures differently, by instead having many small point masses concentrate in a single location, yet we are unable to find a regular sequence whose measure exhibits this property.

For a non-regular example, consider the values of the binomial coefficients, say $f(n) = \binom{2^m}{n-2^m}$ for $2^m \leqslant n < 2^{m+1}$ and follow our process. This sequence is not k-regular for any k since it grows exponentially along a subsequence whereas regular sequences can only grow at most polynomially [2, Thm. 2.10]. More precisely, $f(n_j) \sim 2^{n_j}/\sqrt{\pi \cdot n_j}$ for $n_j = 2^j + 2^{j-1}$. The measure obtained will be $\mu_f = \delta_{1/2}$. This happens because, even though the ratio of every value to the sum is going to zero, the sequence is dominated by the central binomial coefficient and the limiting Gaussian curve, when scaled to the interval $[0, 1)$, is collapsing to its mean.

Our paradigm contrasts starkly with the situation for diffraction measures, where there is a plethora of examples with pure point spectral type, e.g. those arising from model sets (both in the regular and weak sense) and Toeplitz systems; see Keller [26] and Keller and Richard [27]. For example, the diffraction measure of the paperfolding sequence is pure point [6, p. 380], but in our construction the produced measure is Lebesgue. We emphasise here that these two measures, albeit derived from the same object, are not expected to be the same (or even to be of the same spectral type) since they arise from completely different constructions (which capture different aspects of the sequence).

The question of spectral purity also arises. In most of the examples we have produced, the measure μ_f is of pure type. As suggested by Theorem 2, focusing solely on continuous measures, is there a natural example of a regular sequence with continuous measure of mixed spectral type, that is, having both absolutely continuous and singular continuous components in its Lebesgue decomposition? While examples can be constructed—using that the set of k-regular sequences is a group under point-wise addition [2]—we have yet to find a natural non-trivial example or a "good" sufficient criterion for purity. We have some results along this direction and we hope to address this in a future work.

ACKNOWLEDGEMENTS

It is our pleasure to thank Michael Baake for helpful conversations and comments, and for suggesting important references, and the referee for a thorough (and humorous) report containing constructive criticisms and comments. J. Evans acknowledges the support of the Commonwealth of Australia. M. Coons and N. Mañibo acknowledge the support of the German Research Foundation (DFG) via the CRC 1283.

References

[1] Boris Adamczewski and Yann Bugeaud, *On the complexity of algebraic numbers. I. Expansions in integer bases*, Ann. of Math. (2) 165 (2007), no. 2, 547–565.

[2] Jean-Paul Allouche and Jeffrey Shallit, *The ring of k-regular sequences*, Theoret. Comput. Sci. 98 (1992), no. 2, 163–197.

[3] Michael Baake and Michael Coons, *A natural probability measure derived from Stern's diatomic sequence*, Acta Arith. 183 (2018), no. 1, 87–99.

[4] Michael Baake, Franz Gähler, and Neil Mañibo, *Renormalisation of pair correlation measures for primitive inflation rules and absence of absolutely continuous diffraction*, Commun. Math. Phys. 370 (2019), no. 2, 591–635.

[5] Michael Baake, Philipp Gohlke, Marc Kesseböhmer, and Tanja Schindler, *Scaling properties of the Thue–Morse measure*, Discr. Cont. Dynam. Syst. A 39 (2019) no. 7, 4157–4185.

[6] Michael Baake and Uwe Grimm, *Aperiodic order. Vol. 1. A mathematical invitation*, Cambridge University Press, Cambridge, 2013.

[7] Michael Baake and Uwe Grimm, *Fourier transform of Rauzy fractals and point spectrum of 1D Pisot inflation tilings*, Doc. Math. 25 (2020), 2303–2337.

[8] Jason P. Bell, Yann Bugeaud, and Michael Coons, *Diophantine approximation of Mahler numbers*, Proc. Lond. Math. Soc. (3) 110 (2015), no. 5, 1157–1206.

[9] Christian Berg and Gunnar Forst, *Potential theory on locally compact abelian groups*, Springer, Berlin, 1975.

[10] Patrick Billingsley, *Probability and measure*, 3rd ed., John Wiley & Sons, New York, 1995.

[11] Alexander I. Bufetov and Boris Solomyak, *On the modulus of continuity for spectral measures in substitution dynamics*, Adv. Math. 260 (2014), 84–129.

[12] Alexander I. Bufetov and Boris Solomyak, *A spectral cocycle for substitution systems and translation flows*, J. Anal. Math. 141 (2020), no. 1, 165–205.

[13] Fritz Carlson, *Über ganzwertige Funktionen*, Math. Z. 11 (1921), no. 1-2, 1–23.

[14] Alan Cobham, *Uniform tag sequences*, Math. Systems Theory 6 (1972), 164–192.

[15] Michael Coons and James Evans, *A sequential view of self–similar measures; or, what the ghosts of Mahler and Cantor can teach us about dimension*, J. Integer Seq. 24 (2021), no. 2, Art. 21.2.5, 10 pp.

[16] Jean Coquet, *A summation formula related to the binary digits*, Invent. Math. 73 (1983), 107–115.

[17] Ingrid Daubechies and Jeffrey C. Lagarias, *Two-scale difference equations. I. Existence and global regularity of solutions*, SIAM J. Math. Anal. 22 (1991), no. 5, 1388–1410.

[18] Ingrid Daubechies and Jeffrey C. Lagarias, *Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals*, SIAM J. Math. Anal. 23 (1992), no. 4, 1031–1079.

[19] Philippe Dumas, *Joint spectral radius, dilation equations, and asymptotic behavior of radix-rational sequences*, Lin. Alg. Appl. 438 (2013), no. 5, 2107–2126.

[20] Samuel Eilenberg, *Automata, Languages, and Machines, Vol. A*, Academic Press, 1974.

[21] James Evans, *The ghost measures of affine regular sequences*, Houston J. Math. Preprint, 2020, to appear.

[22] Kenneth Falconer, *Fractal geometry*, third ed., Wiley, Chichester, 2014.

[23] Pierre Fatou, *Séries trigonométriques et séries de Taylor*, Acta Math. 30 (1906), no. 1, 335–400.

[24] Raphaël Jungers, *The joint spectral radius: Theory and applications*, Lecture Notes in Control and Information Sciences, 385, Springer, Berlin, 2009.

[25] Shizuo Kakutani, Strictly ergodic symbolic dynamical systems, In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 319–326, Univ. California Press, Berkeley, Calif., 1972.

[26] Gerhard Keller, *Maximal equicontinuous generic factors and weak model sets*, Discrete Contin. Dyn. Syst. 40 (2020), no. 12, 6855–6875.

[27] Gerhard Keller and Christoph Richard, *Dynamics on the graph of the torus parametrisation*, Ergod. Th. & Dynam. Syst. 38 (2018), no. 3, 1048–1085.

[28] Marc Kesseböhmer and Berndt Strattmann, *A multifractal analysis for Stern–Brocot intervals, continued fractions and Diophantine growth rates*, J. Reine Angew. Math. (Crelle) 605 (2007) 133–163.

[29] Serge Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, 142, Springer, New York, 1993.

[30] Kurt Mahler, *The spectrum of an array and its application to the study of the translation properties of a simple class of arithmetical functions. II: On the translation properties of a simple class of arithmetical functions*, J. Math. Phys. (MIT) 6 (1927), 158–163.

[31] Charles A. Micchelli and Hartmut Prautzsch, *Uniform refinement of curves*, Lin. Alg. Appl. 114/115 (1989), 841–870.

[32] Kumiko Nishioka, *Mahler functions and transcendence*, Lecture Notes in Mathematics, 1631, Springer, Berlin, 1996.

[33] Oskar Perron, *Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus*, Math. Ann. 64 (1907), 11–76.

[34] Martine Queffélec, *Substitution dynamical systems—spectral analysis*, second ed., Lecture Notes in Mathematics, 1294, Springer, Berlin, 2010.

[35] Michel Rigo, *Formal languages, automata and numeration systems 1*, John Wiley & Sons, New Jersey, 2014.

[36] Gian-Carlo Rota and Gilbert Strang, *A note on the joint spectral radius*, Indag. Math. 22 (1960), 379–381.

[37] Gábor Szegö, *Tschebyscheffsche Polynome und nichtfortsetzbare Potenzreihen*, Math. Ann. 87 (1922), no. 1-2, 90–111.

[38] Norbert Wiener, *The spectrum of an array and its application to the study of the translation properties of a simple class of arithmetical functions. I: The spectrum of an array*, J. Math. Phys. (MIT) 6 (1927), 145–157.

Michael Coons
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld
Germany
mcoons.math@gmail.com

James Evans
School of Information and Physical Sciences
University of Newcastle
130 University Drive
Callaghan NSW 2308
Australia
james.evans10@uon.edu.au

Neil Mañibo
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld
Germany
cmanibo@math.uni-bielefeld.de