11-19-2018

Markov chains and generalized wavelet multiresolutions

Myung-Sin Song
Southern Illinois University Edwardsville, msong@siue.edu

Palle Jorgensen
The University of Iowa

MARKOV CHAINS AND GENERALIZED WAVELET MULTIRESOLUTIONS

PALLE E.T. JORGENSEN AND MYUNG-SIN SONG

Abstract. We develop some new results for a general class of transfer operators, as they are used in a construction of multi-resolutions. We then proceed to give explicit and concrete applications. We further discuss the need for such a constructive harmonic analysis/dynamical systems approach to fractals.

Contents

1. Introduction and Setting. 1
2. Generalized multi-resolution measures on solenoids. 2
2.1. The case of unbounded w 3
2.2. The bounded case 3
3. A Transfer Operator 6
4. Future Directions 7
4.1. Orthonormal bases generated by Cuntz algebras 8
5. Multi-resolutions and generalized wavelet representations. 14
6. Bibliography 19
References 19

1. Introduction and Setting.

While there are already a number of approaches to harmonic analysis of fractals, "non-smooth" settings, we propose below a focus on a certain family of positive operators. They will serve as transfer operators.

Our paper is divided into two parts: in the first we develop the needed results on transfer operators, and the second part will be concrete applications. There are many justification for the need of a constructive harmonic analysis of fractals; one is the discovery of Jorgensen-Pedersen that certain fractal L^2 spaces admit Fourier bases; while others do not. However the lack of available Fourier bases in many example suggests a need for alternative approaches.

The Cantor fractals are special cases of more general IFS systems. Our present paper will deal with this more general framework. In addition to fractal Fourier analyses (fractals in the large), we shall also study multiresolution and wavelet techniques. In work of Dutkay-Jorgensen, it was shown that the general affine IFS-systems, even if not amenable to Fourier analysis, in fact do admit wavelet bases, and so in particular can be analyzed with the use of multiresolutions; reflecting the inherent self-similarity to the fractal under consideration. But this approach in fact depends on the use of certain transfer operators. The latter in turn ties in with intriguing new work on cascade algorithms, with an analysis of representations of non-commutative generators and relations (especially the Cuntz relations), as well as with certain stochastic processes; and we shall make connections to recent research on Markov processes, and to reproducing kernel theory.

2000 *Mathematics Subject Classification.* Primary 60J20, 42C40, 42C10; Secondary 28A80,
Key words and phrases. Markov chains, multiresolution, fractals, Cuntz relations, optimization.

2. Generalized multi-resolution measures on solenoids.

In section 2 below we introduce a certain multi-resolution approach to problems that arise in analysis of fractals and more generally in stochastic analysis. Examples will be given in section 4, and a wavelet representation approach in section 5.

Definition 2.1. Let X be a compact Hausdorff space. Consider a linear operator

$$R : C(X) \to L^{\infty}(X) \tag{2.1}$$

and assume it is positive. i.e., $f \geq 0 \Rightarrow Rf \geq 0$.

Proof. Riesz. We also consider positive Borel measures λ on X. □

Theorem 2.2. *Let λ be a finite positive Borel measure on X. Then the following are equivalent:*
(1) $\lambda R << \lambda$ with Radon-Nikodym derivative w.
(2) $\int wf d\lambda = \int (Rf) d\lambda$, for all $f \in C(X)$

Proof. This is immediate from the definition of the Radon-Nikodym derivative. To understand it in more detail, it helps to have a generalized Perron-Frobenius theorem. Recall, R satisfies an additional condition thus as follows: In general, it may be difficult, but recall $R : C(X) \to L^{\infty}(X, \lambda)$ also satisfies the following property (i.e., transition operator has the pull-out property.) Let σ be an embedding in the measure space X, and assume that

$$R[(f \circ \sigma)g] = fRg, \quad \forall f, g \in C(X). \tag{2.2}$$

Remark 2.3. The axiom (2.2) is a generalized conditional expectation property. But, in general, as an operator in $L^2(\lambda)$, R may still be unbounded. Nonetheless, if

$$\mu = \lambda \circ R \quad \text{and} \quad \frac{d\mu}{d\lambda} = w \quad \text{Radon-Nikodym derivative.}$$

Note w depends on both R and on λ.

□

We now study domains of the unbounded operators in $L^2(\lambda)$.

Theorem 2.4. *Suppose (2) and (2.2) hold, then*

$$C(X) \subset domain(R^*), \quad \text{the domain of the adjoint operator } R^*; \text{ in general possibly unbounded.} \tag{2.3}$$

Generally, $R^* f = w(f \circ \sigma)$, for all $f \in C(X)$. Moreover, $R : L^2(\lambda) \to L^2(\lambda)$ is bounded if and only if

$$w \in L^{\infty}(\lambda), \quad \text{and then} \quad \|R\|_{2\to 2} = \|R(w)\|_{\infty}^{1/2}. \tag{2.4}$$

But in general, R is an unbounded operator in $L^2(\lambda)$. As noted, we have:
1) $\lambda R << \lambda$
2) $\frac{d\lambda R}{d\lambda} = w$

Lemma 2.5. *Let $C(X) \subset domain(R^*)$, and*

$$R^* f = wf \circ \sigma, \quad \forall f \in C(X) \quad i.e., (R^* f)(x) = w(x) f(\sigma(x)), \quad \forall x \in X. \tag{2.5}$$

implies $RR^* f = fR(w)$.

Proof. (of the Lemma) Assume (2) and (2.2), then

$$\int w(f \circ \sigma) g d\lambda \underset{by (2.2)}{=} \int f(Rg) d\lambda \quad \forall g \in C(X). \tag{2.6}$$

Indeed, the right hand side of (2.6),

$$\text{RHS of (2.6)} = \int fRgd\lambda \underset{by(2.2)}{=} \int R((f\circ g))d\lambda$$
$$\underset{by(2)}{=} \int w(f\circ\sigma)gd\lambda, \quad \text{where we use that } w = \frac{d\mu}{d\lambda} = R^*(\mathbf{1}).$$

Hence,

$$\begin{aligned}\text{RHS of (2.6)} &\underset{Schwarz}{\leq} \int |wf\circ\sigma|^2 d\lambda \int |g|^2 d\lambda \\ &\Rightarrow f\in dom(R^*), \quad \text{and} \\ &R^* f = wf\circ\sigma.\end{aligned}$$

□

this is a weighted composition operator.

2.1. **The case of unbounded w.** Even if $w = \frac{d\mu}{d\lambda}$ is only in $L^1(\lambda)$, then the following two operators are well defined as $L^2(\lambda)\to L^2(\lambda)$ operators; each with $C(X)\subset L^2(\lambda)$ as dense domain, and $R\subset S^*$, $S\subset R^*$; (containments of operators) see (2.9) below.

$$(2.7) \quad L^2(\lambda)\supset \begin{cases} C(X)\ni f \xrightarrow{R} R(f)\text{or} \\ C(X)\ni f \xrightarrow{S} w(f\circ\sigma)\in L^2(\lambda), \quad R: C(X)\to L^\infty(\lambda); \quad Rf\in L^\infty(\lambda), \quad \text{and we have}\end{cases}$$

$$\langle Sf, g\rangle_{L^2(\lambda)} = \langle f, Rg\rangle_{L^2(\lambda)}, \quad \forall f,g\in C(X). \tag{2.8}$$

Proof of (2.8)

$$\int w(f\circ\sigma)gd\lambda = \int fR(g)d\lambda, \tag{2.9}$$

and we verified (2.9) above.

2.2. **The bounded case.** Moreover, assuming $w\in L^\infty(\lambda)$, we get

$$RR^*f = R(wf\circ\sigma) \underset{\text{by (2.9)}}{=} f(Rw)$$

so RR^* is a multiplication operator on $L^2(\lambda)$, i.e., multiplication by the funtion $R(w)\leftarrow L^\infty(\lambda)$. So if $\|RR^*\|_{2\to 2} = \|Rw\|_\infty < \infty$. Recall, by Riesz $\|Rw\|_\infty \leq \|w\|_\infty$. If and only if RR^* is bounded: $L^2(\lambda)\to L^2(\lambda)$. If and only if R is bounded: $L^2(\lambda)\to L^2(\lambda)$. So by the Hilbert space $\|RR^*\|_{2\to 2} = \|R\|^2_{2\to 2} = \|R^*\|^2_{2\to 2}$. (The L^2- operator norms.)

Proof.

$$\int R((f\circ\sigma)g) = \int fRGd\lambda$$
$$\int (f\circ\sigma)gwd\lambda = \int (R^*f)gd\lambda$$
$$R^*f = wf\circ\sigma$$
$$RR^*f = R(wf\circ\sigma) = fRw \quad f: L^2(\lambda)\to L^2(\lambda)$$

$$RR^* \text{ in } L^2(\lambda) \text{ is bounded} \iff Rw\in L^\infty \iff w\in RR^* \iff R \text{ in } L^2(\lambda) \text{ is bounded}$$

□

The converse also holds: If $C(X) \ni f \underset{\longrightarrow}{S} w(f \circ \sigma)$ extends to a bounded operator in $L^2(\lambda)$, then $w \in L^\infty(\lambda)$. Then adjoint operator $R := S^*$ (adjoint with respect to the $L^2(\lambda)-$Hilbert space), then the following holds:

$$R((f \circ \sigma)g) = fR(g), \quad \forall f, g \in C(X).$$

Moreover, if we set $\mu = \lambda R$, i.e., $\int f d\mu = \int Rf d\lambda$, then $\mu << \lambda$, and $w = \frac{d\mu}{d\lambda}$ is the Radon-Nikodym derivative.

Corollary 2.6. *S^*S is a multiplication operator in $L^2(\lambda)$.*

Proof. $S^*Sf = R(wf \circ \sigma) = fR(w)$ so $S^*S = RS$ is a multiplicative operator, $m = R(w)$. Then set $f \longrightarrow mf$ $(\lambda \circ \sigma^{-1})(E) = \lambda(\sigma^{-1}(E))$, $\sigma^{-1}(E) = \{x \in X : \sigma(x) \in E\}$. □

Corollary 2.7. *$\lambda \circ \sigma^{-1} << \lambda$, and $\frac{d\lambda \circ \sigma^{-1}}{d\lambda} = R\left(\frac{1}{w}\right)$.*

Proof.

$$\begin{aligned}\int f d\lambda \circ \sigma^{-1} &= \int f \circ \sigma d\lambda \\ &= \int \frac{1}{w} wf \circ \sigma d\lambda \quad wf \circ \sigma = R^* f \\ &= \int R\left(\frac{1}{w}\right) f d\lambda;\end{aligned}$$

and the corollary follows. □

Assume in addition that $w = \frac{d\mu}{d\lambda}$, $\mu = \lambda \circ R$. Let $w \in L^\infty(\lambda)$, and $\|w\|_\infty \leq 1$, then $\|R\|_{2\to 2} \leq 1$ by the theorem. Then by Hilbert space theorem

$$Rh = h \Rightarrow R^*h = h \tag{2.10}$$

and so

$$\begin{aligned}h &= R^*h = wh \circ \sigma \\ h &= Rh = RR^*h = hR(w), \\ RR^*h(x) &= h(x)R(w)(x),\end{aligned}$$

so

$$h(x) > 0 \Rightarrow R(w)(x) = 1. \tag{2.11}$$

Corollary 2.8. *Suppose $w^{-1} = \frac{1}{w}$ is well defined then X, λ, R, σ especially $R((f \circ \sigma)g) = fRg$, then $\lambda \circ \sigma^{-1} << \lambda$, and $\frac{d\lambda \circ \sigma^{-1}}{d\lambda} = R\left(\frac{1}{w}\right)$; and so λ is $\sigma-$invariant if and only if $R\left(\frac{1}{w}\right) = 1$.*

Proof. Let $f \in C(X)$, then

$$\begin{aligned}\int f d\lambda \circ \sigma^{-1} &= \int (f \circ \sigma) d\lambda \\ &= \int \frac{1}{w} w(f \circ \sigma) d\lambda \quad \text{where } w(f \circ \sigma) = R^* f \\ &= \int R\left(\frac{1}{w}\right) f d\lambda \quad R\left(\frac{1}{w}\right) = \text{Radon-Nikodym} \Rightarrow R\left(\frac{1}{w}\right) d\lambda = \lambda \circ \sigma^{-1}.\end{aligned}$$

the desired conclusion. □

Corollary 2.9.

$$\lambda \circ \sigma^{-1} = \lambda$$
$$R\left(\frac{1}{w}\right) = 1$$

There is a generalized family of multi-resolution measures on solenoids: The solenoid may be defined for any endomorphism $\sigma : X \longrightarrow X$ where X is compact, and σ is assumed to be onto. In addition, we fix a positive operator $R : C(X) \longrightarrow L^\infty(X)$, and $h \geq 0$ function on X such that $Rh = h$. Also, given a finite positive measure λ on X such that $\mu(f) = \int Rf d\mu$ satisfies $\mu << \lambda$.

Theorem 2.10. *Set $w := \frac{d\mu}{d\lambda}$. From this, we define $\mathbb{P}$ on $Sol_\sigma(X)$ such that $\frac{d\mathbb{P}\circ\widetilde{\sigma}}{d\mathbb{P}} = w \circ \pi_0$, where $\widetilde{\sigma}$ is then indeed automorphism on $Sol_\sigma(X)$.*

Remark 2.11. $\widetilde{\sigma}(x_0x_1x_2, \cdots) = (\sigma(x_0)x_0x_1x_2\cdots)$ for all $x \in Sol_\sigma(X)$. R, h, λ or R_i, h_i, $i \in \mathbb{N}$ (properties as stated in the theorem) but it is or item to measure $\mathbb{P}$ on $Sol_\sigma(X)$. Given

$$\begin{cases} R_1, R_2, R_3, \cdots \\ h_1, h_2, h_3, \cdots, \quad h_i \geq 0 \quad \text{Axiom: } R_{i+1}(h_{i+1}) = h_i \end{cases}$$

in sense of that (R_i, h_i) governs the transiton: $\pi_{i-1} \longrightarrow \pi_i$, for all i. But if $R_{i+1}(h_{i+1}) = h_i$ and each R_i satisfy $R_i((f\circ\sigma)g) = fR_i(g)$, for all $f, g \in C(X)$, then there exists a unique $\mathbb{P}_x$ such that we get consisting a cylinder, and so $\mathbb{P}_x$ is well-defined. $\mathbb{P}_x$ on a cylinder function is: Conditions C_n for cylinder functions over n.

Proof. $E_x(cyl^{(n)}) = \int_{\pi_0^{-1}(x)} cyl^{(n)} d\mathbb{P}_x = C_n : f_0(x)R_1(f_1R_2(f_2\cdots R_n(f_nh_n)\cdots))(x)$ The following holds

$$C_{n+1}(f_0, \cdots f_n, \mathbf{1}) = C_n(f_0, \cdots, f_n)$$

holds if $R_{n+1}(h_{n+1}) = h_n$. □

Proof preliminaries: General setting: In detail, recall: $(\pi_i)_{i=0,1,\dots}$, $\pi_i(x) := x_i$ and $Sol_\sigma(X) := \{(x_i)_\sigma^\infty\} \subset \Pi_\sigma^\infty x \supseteq Sol_\sigma(X)$, $\sigma(x_{i+1}) = x_i$.

Definition 2.12. $x = (x_i)_\sigma^\infty$ such that $\sigma(x_{i+1}) = x_i$, for all $i = 0, 1, 2, \cdots$, and we set $\pi_j(x) = x_j$, $j = 0, 1, 2, \cdots$, coordinate functions; $\widetilde{\sigma}(x_0x_1x_2\cdots) = (\sigma(x_0)x_0x_1, x_2\cdots)$; note that $\widetilde{\sigma}$ is an automorphism with inverse $\widetilde{\sigma}^{-1}(x_0x_1x_2\cdots) = (x_1, x_2\cdots)$.

Lemma 2.13. *For all $x \in X$, there exists a unique positive measure $\mathbb{P}_x$ on $\pi_0^{-1}(x) \subset Sol_\sigma(x)$ such that $\int(f_0\pi_0f_1\pi_1\cdots f_n\pi_n)d\mathbb{P}_x = f(x)R(f_1F(f_2\cdots R(f_nh)\cdots))(x)$ or $f_0(x)R_1(f_1R_2(f_2\cdots R_n(f_nh_n)\cdots))(x)$*

General Setting and Assumptions: X compact, $\sigma : X \to X$ endomorphism, onto; λ finite positive measure on X, $R_i : C(X) \to L^\infty(X)$ positive such that

$$R_i((f\circ\sigma)g) = fR_ig, \quad \forall f, g \in C(X) \tag{2.12}$$

or same conditions on R_i, $i \in \mathbb{N}$. Assume there exists

$$h \geq 0, \quad \int hd\lambda = 1, \quad Rh = h, \tag{2.13}$$

More generally $h_1, h_2, \cdots$ $R_{i+1}(h_{i+1}) = h_i$, for all i. Also, $w \in L^\infty(\lambda)$, there exists a finite constant such that

$$\left|\int(Rf)d\lambda\right|^2 \leq const \int |f|^2 d\lambda, \quad \forall f \in C(X). \tag{2.14}$$

Let $\mathbf{1}$ denote constant finite 1 on X. Set $w = R^*\mathbf{1} \in L^2(\lambda)$, by Riesz. i.e.,

$$\int wfd\lambda = \int(Rf)d\lambda \quad \Rightarrow \quad R^*f = wf\circ\sigma. \tag{2.15}$$

Condition (2.14) is really a Radon-Nikodym derivative as follows: Since R is positive, we have:

$$\mu(f) = \lambda(R(f)) = \int_X (Rf)(x) d\lambda(x) = \int f(wd\lambda)$$

is a measure on X by the Riesz theorem; and (2.14) is the case that $d\mu << d\lambda$ (absolute continuous) so the Radon-Nikodym derivative $\frac{d\mu}{d\lambda} = w \in L^1_+(X, \lambda)$ is well defined

$$\int fwd\lambda = \int fd\mu = \int (Rf)d\lambda. \quad \mu = \lambda R \tag{2.16}$$

Conversely, suppose (2.16) holds, then

$$\left| \int fwd\lambda \right|^2 \leq \int |f|^2 d\lambda \int w^2 d\lambda$$

Given $\frac{d\mathbb{P}\circ\widetilde{\sigma}}{d\mathbb{P}}$ where $\mathbb{P} \leftrightarrow (R_1, R_2, R_3, \cdots)$.

Alternative Representation for $\mathbb{P}_x$*:* By Riesz, there exists $\{\mu_x\}_{x \int T}$ such that $\{\pi_i\}$ $(R_i f)(x) = \int_T f(y) d\mu_x^{(i)}(y) = \int f(y) P_i(dy|x)$

$$\begin{aligned} R_1(f_1 R_2(f_2 h))(x) &= \int \int f_1(y_1) f_2(y_2) h_2(y_2) d\mu_{y_1}(y_2) d\mu_x(y_1) \\ &= \int \int f_1(y_1) f_2(y_2) h_2(y_2) P_2(dy_2|y_1) P_1(dy_1|x) \end{aligned}$$

$$R(f_1 R(f_2 \cdots R(f_n h) \cdots)(x) = \int \int \cdots \int f_1(y_1) f_2(y_2) \cdots f_n(y_n) h_n(y_n) P(dy_n|y_{n-1}) \cdots P(dy_1|x)$$

cylinder set $\sigma(E_{i+1}) \subset E_i$,

$$\mathbb{P}_x(cyl^n) = \int_{E_1} \cdots \int_{E_n} h_n(y_n) P_n(dy_n|y_{n-1}) P_{n-1}(dy_{n-1}|y_{n-2}) \cdots P_1(dy_1|x)$$

where $\int_{E_1} \cdots \int_{E_n} h_n(y_n) P_n(dy_n|y_{n-1}) P = \mathbb{P}_x(cyl)$.

3. A Transfer Operator

A popular tool for deciding if a candidate for a wavelet basis is in fact an ONB uses a certain transfer operator. Variants of this operator is used in diverse areas of applied mathematics. It is an operator which involves a weighted average over a finite set of possibilities. Hence it is natural for understanding random walk algorithms. As remarked in for example [15, 16, 17, 10], it was also studied in physics, for example by David Ruelle who used to prove results on phase transition for infinite spin systems in quantum statistical mechanics. In fact the transfer operator has many incarnations (many of them known as Ruelle operators), and all of them based on N-fold branching laws.

In our wavelet application, the Ruelle operator weights in input over the N branch possibilities, and the weighting is assigned by a chosen scalar function w. the and the w-Ruelle operator is denoted R_w. In the wavelet setting there is in addition a low-pass filter function m_0 which in its frequency response formulation is a function on the d-torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$.

Since the scaling matrix A has integer entries A passes to the quotient $\mathbb{R}^d/\mathbb{Z}^d$, and the induced transformation $r_A : \mathbb{T}^d \to \mathbb{T}^d$ is an N-fold cover, where $N = |detA|$, i.e., for every x in $\mathbb{T}^d$ there are N distinct points y in $\mathbb{T}^d$ solving $r_A(y) = x$.

In the wavelet case, the weight function w is $w = |m_0|^2$. Then with this choice of w, the ONB problem for a candidate for a wavelet basis in the Hilbert space $L^2(\mathbb{R}^d)$ as it turns out may be decided by the dimension of a distinguished eigenspace for R_w, by the so called Perron-Frobenius problem.

This has worked well for years for the wavelets which have an especially simple algorithm, the wavelets that are initialized by a single function, called the scaling function. These are called the multiresolution analysis (MRA) wavelets, or for short the MRA-wavelets. But there are instances, for example if a problem must be localized in frequency domain, when the MRA-wavelets do not suffice, where it will by necessity include more than one scaling function. And we are then back to trying to decide if the output from the discrete algorithm, and the $\mathcal{O}_N$ representation is an ONB, or if it has some stability property which will serve the same purpose, in case where asking for an ONB is not feasible.

4. Future Directions

The idea of a scientific analysis by subdividing a fixed picture or object into its finer parts is not unique to wavelets. It works best for structures with an inherent self-similarity; this self-similarity can arise from numerical scaling of distances. But there are more subtle non-linear self-similarities. The Julia sets in the complex plane are a case in point [3, 5, 7, 9, 18, 19]. The simplest Julia set come from a one parameter family of quadratic polynomials $\varphi_c(z) = z^2 + c$, where z is a complex variable and where c is a fixed parameter. The corresponding Julia sets J_c have a surprisingly rich structure. A simple way to understand them is the following: Consider the two brances of the inverse $\beta_{\pm} = z \mapsto \pm\sqrt{z-c}$. Then J_c is the unique minimal non-empty compact subset of $\mathbb{C}$, which is invariant under $\{\beta_{\pm}\}$. (There are alternative ways of presenting J_c but this one fits our purpose. The Julia set J of a holomorphic function, in this case $z \mapsto z^2 + c$, informally consists of those points whose long-time behavior under repeated iteration , or rather iteration of substitutions, can change drastically under arbitrarily small perturbations.) Here “long-time” refers to largen n, where $\varphi^{(n+1)}(z) = \varphi(\varphi^{(n)}(z))$, $n = 0, 1, ...,$ and $\varphi^{(0)}(z) = z$. Please see figures 1 and 2 for examples of Julia set graphs.

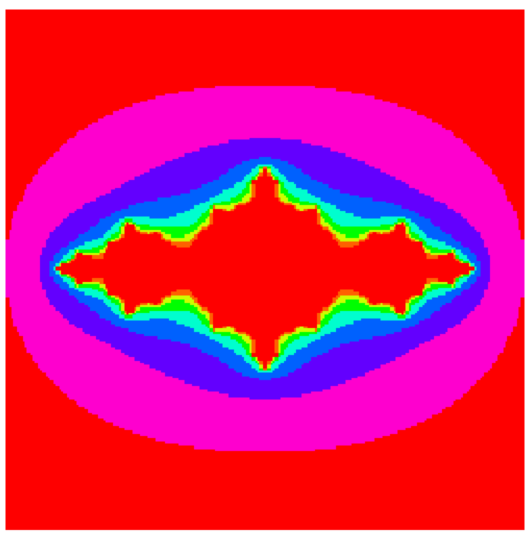

Figure 1. Julia set graphed using Mathematica for $c = -1$. [4], [6], [8]

It would be interesting to adapt and modify the Haar wavelet, and the other wavelet algorithms to the Julia sets. The two papers [11, 12] initiated such a development. Then an attempt to adapt and modify the Haar wavelet to the Julia sets was made, [13] however, there were some limitations in finding the filters. Perhaps trying another fractal set such as tent map or others may work.

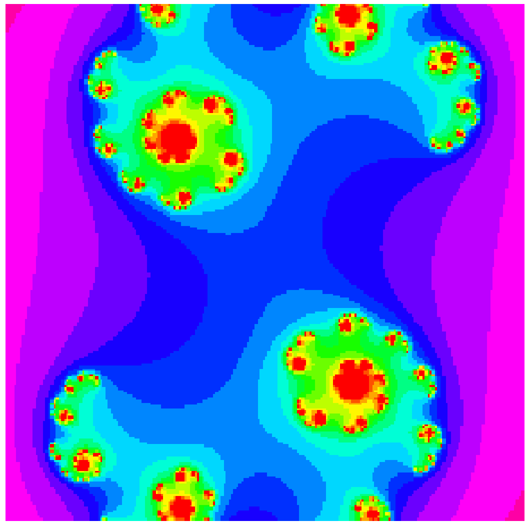

FIGURE 2. Julia set graphed using Mathematica for $c = 0.45 - 0.1428i$. [4], [6], [8]

4.1. **Orthonormal bases generated by Cuntz algebras.** We present new results from [13] by borrowing section 3 and part of section 2 from [13] in the rest of this subsection 4.1. It gives a general criterion for a family generated by the Cuntz isometries to be an orthonormal basis.

Theorem 4.1. [13] *Let $\mathcal{H}$ be a Hilbert space and $(S_i)_{i=0}^{N-1}$ be a representation of the Cuntz algebra $\mathcal{O}_N$. Let $\mathcal{E}$ be an orthonormal set in $\mathcal{H}$ and $f : X \to \mathcal{H}$ a norm continuous function on a topological space X with the following properties:*

(i) $\mathcal{E} = \cup_{i=0}^{N-1} S_i\mathcal{E}$.
(ii) $\overline{\text{span}}\{f(t) : t \in X\} = \mathcal{H}$ *and* $||f(t)||= 1$, *for all* $t \in X$.
(iii) There exist functions $\mathfrak{m}_i : X \to \mathbb{C}$, $g_i : X \to X$, $i = 0, \dots, N-1$ *such that*

$$S_i^* f(t) = \mathfrak{m}_i(t) f(g_i(t)), \quad t \in X. \tag{4.1}$$

(iv) There exist $c_0 \in X$ *such that* $f(c_0) \in \overline{\text{span}}\mathcal{E}$.
(v) The only function $h \in \mathcal{C}(X)$ *with* $h \geq 0$, $h(c) = 1$, $\forall\, c \in \{x \in X : f(x) \in \overline{\text{span}}\mathcal{E}\}$, *and*

$$h(t) = \sum_{i=0}^{N-1} |\mathfrak{m}_i(t)|^2 h(g_i(t)), \quad t \in X \qquad \Longleftrightarrow\ Rh = h \tag{4.2}$$

are the constant functions.

Then $\mathcal{E}$ is an orthonormal basis for $\mathcal{H}$.

Proof. Define

$$h(t) := \sum_{e\in\mathcal{E}} |\langle f(t)\,,\, e\rangle|^2 = ||Pf(t)||^2, \quad t \in X$$

where P is the orthogonal projection onto the closed linear span of $\mathcal{E}$.

Since $t \mapsto f(t)$ is norm continuous we get that h is continuous. Clearly $h \geq 0$. Also, if $f(c) \in \overline{\text{span}}\mathcal{E}$, then $||Pf(c)||= ||f(c)||= 1$ so $h(c) = 1$. In particular, from (ii) and (iv), $h(c_0) = 1$. We check (4.2). Since the sets

$S_i\mathcal{E}$, $i = 0, \dots N-1$ are mutually orthogonal, the union in (i) is disjoint. Therefore for all $t \in X$:

$$h(t) = \sum_{i=0}^{N-1} \sum_{e \in \mathcal{E}} |\langle f(t)\,,\, S_i e\rangle|^2 = \sum_{i=0}^{N-1} \sum_{e \in \mathcal{E}} |\langle S_i^* f(t)\,,\, e\rangle|^2 = \sum_{i=0}^{N-1} |\mathfrak{m}_i(t)|^2 \sum_{e \in \mathcal{E}} |\langle f(g_i(t))\,,\, e\rangle|^2 =$$

$$= \sum_{i=0}^{N-1} |\mathfrak{m}_i(t)|^2 h(g_i(t))$$

By (v), h is constant and, since $h(c_0) = 1$, $h(t) = 1$ for all $t \in X$. Then $||Pf(t)||= 1$ for all $t \in X$. Since $||f(t)||= 1$ it follows that $f(t) \in \text{span}\mathcal{E}$ for all $t \in X$. But the vectors $f(t)$ span $\mathcal{H}$ so $\overline{\text{span}}\mathcal{E} = \mathcal{H}$ and $\mathcal{E}$ is an orthonormal basis.

□

Remark 4.2. [13] The operators of the form

$$Rh(t) = \sum_{i=0}^{N-1} |\mathfrak{m}_i(t)|^2 h(g_i(t)), \quad t \in X, h \in C(X),$$

that appear in (4.2), are sometimes called Ruelle operators or transfer operators, see e.g. [1].

4.1.1. *Piecewise exponential bases on fractals.*

Example 4.3. [13] We consider affine iterated function systems with no overlap. Let R be a $d \times d$ expansive real matrix, i.e., all the eigenvalues of R have absolute value strictly greater than 1.Let $B \subset \mathbb{R}^d$ a finite set such that $N = |B|$. Define the affine iterated function system

$$\tau_b(x) = R^{-1}(x + b) \quad (x \in \mathbb{R}^d,\, b \in B) \tag{4.3}$$

By [14] there exists a unique compact subset X_B of $\mathbb{R}^d$ which satisfies the invariance equation

$$X_B = \cup_{b \in B} \tau_b(X_B) \tag{4.4}$$

X_B is called the attractor of the iterated function system $(\tau_b)_{b \in B}$. Moreover X_B is given by

$$X_B = \left\{ \sum_{k=1}^{\infty} R^{-k} b_k \,:\, b_k \in B \text{ for all } k \geq 1 \right\} \tag{4.5}$$

Also from [14] there is a unique probability measure μ_B on $\mathbb{R}^d$ satisfying the invariance equation

$$\int f d\mu_B = \frac{1}{N} \sum_{b \in B} \int f \circ \tau_b d\mu_B \tag{4.6}$$

for all continuous compactly supported functions f on $\mathbb{R}$. We call μ_B the invariant measure for the iterated function system (IFS) $(\tau_b)_{b \in B}$. By [14], μ_B is supported on the attractor X_B. We say that the IFS has no overlap if $\mu_B(\tau_b(X_B) \cap \tau_b'(X_B)) = \emptyset$ for all $b \neq b'$ in B.

Assume that the IFS $(\tau_b)_{b \in B}$ has no overlap. Define the map $r : X_B \to X_B$

$$r(x) = \tau_b^{-1}(x), \text{ if } x \in \tau_b(X_B) \tag{4.7}$$

Then r is an N-to-1 onto map and μ_B is strongly invariant for r. Note that $r^{-1}(x) = \{\tau_b(x) : b \in B\}$ for μ_B.a.e. $x \in X_B$.

We apply Theorem 4.1 to the setting of Example 4.3, in dimension $d = 1$ for affine iterated function systems, when the set $\frac{1}{R}B$ has a spectrum L [13].

Definition 4.4. [13] Let L in $\mathbb{R}$, $|L| = N$, $R > 1$ such that L is a spectrum for the set $\frac{1}{R}B$. We say that $c \in \mathbb{R}$ is an *extreme cycle point* for (B, L) if there exists $l_0, l_1, \dots, l_{p-1}$ in L such that, if $c_0 = c$, $c_1 = \frac{c_0+l_0}{R}, c_2 = \frac{c_1+l_1}{R} \dots c_{p-1} = \frac{c_{p-2}+l_{p-2}}{R}$ then $\frac{c_{p-1}+l_{p-1}}{R} = c_0$, and $|m_B(c_i)| = 1$ for $i = 0, \dots, p-1$ where

$$m_B(x) = \frac{1}{N} \sum_{b \in B} e^{2\pi i b x} \quad x \in \mathbb{R}.$$

Proposition 4.5. [13] *Let $(m_i)_{i=0}^{N-1}$ be a QMF basis. Define the operators on $L^2(X, \mu)$*

$$S_i(f) = m_i f \circ r, \quad i = 0, \dots, N-1 \tag{4.8}$$

Then the operators S_i are isometries and they form a representation of the Cuntz algebra $\mathcal{O}_N$, i.e.

$$S_i^* S_j = \delta_{ij}, \quad i, j = 0, \dots, N-1, \qquad \sum_{i=0}^{N-1} S_i S_i^* = I \tag{4.9}$$

The adjoint of S_i is given by the formula

$$S_i^*(f)(z) = \frac{1}{N} \sum_{r(w)=z} \overline{m_i}(w) f(w) \tag{4.10}$$

Proof. We compute the adjoint: take f, g in $L^2(X, \mu)$. We use the strong invariance of μ.

$$\langle S_i^* f\,,\, g \rangle = \int f \overline{m_i} \overline{g \circ r}\, d\mu = \int \frac{1}{N} \sum_{r(w)=z} \overline{m_i}(w) f(w) \overline{g}(z) d\mu(z)$$

Then (4.10) follows. The Cuntz relations in (4.9) are then easily checked with Proposition 2.6 in [13]. □

Definition 4.6. [13] We denote by L^* the set of all finite words with digits in L, including the empty word. For $l \in L$ let S_l be given as in (4.8) where m_l is replaced by the exponential e_l. If $w = l_1 l_2 \dots l_n \in L^*$ then by S_w we denote the composition $S_{l_1} S_{l_2} \dots S_{l_n}$.

Theorem 4.7. [13] *Let $B \subset \mathbb{R}$, $0 \in B$, $|B| = N$, $R > 1$ and let μ_B be the invariant measure associated to the IFS $\tau_b(x) = R^{-1}(x + b)$, $b \in B$. Assume that the IFS has no overlap and that the set $\frac{1}{R}B$ has a spectrum $L \subset \mathbb{R}$, $0 \in L$. Then the set*

$$\mathcal{E}(L) = \{S_w e_{-c} : c \text{ is an extreme cycle point for } (B, L), w \in L^*\}$$

is an orthonormal basis in $L^2(\mu_B)$. Some of the vectors in $\mathcal{E}(L)$ are repeated but we count them only once.

Proof. Let c be an extreme cycle point. Then $|m_B(c)| = 1$. Using the fact that we have equality in the triangle inequality $(1 = |m_B(c)| \le \frac{1}{N} \sum_{b \in B} |e^{2\pi i b c}| = 1)$, and since $0 \in B$, we get that $e^{2\pi i b c} = 1$ so $bc \in \mathbb{Z}$ for all $b \in B$. Also there exists another extreme cycle point d and $l \in L$ such that $\frac{d+l}{R} = c$. Then we have: $S_l e_{-c}(x) = e^{2\pi i l x} e^{2\pi i (Rx-b)(-c)}$, if $x \in \tau_b(X_B)$. Since $bc \in \mathbb{Z}$ and $R(-c) + l = -d$, we obtain

$$S_l e_{-c} = e_{-d} \tag{4.11}$$

We use this property to show that the vectors $S_w e_{-c}$, $S_{w'} e_{-c'}$ are either equal or orthogonal for w, w' in L^* and c, c' extreme cycle points for (B, L). Using (4.11), we can append some letters at the end of w and w' suh that the new words have the same length:

$$S_w e_{-c} = S_{w\alpha} e_{-d}, \quad S_{w'} e_{-c'} = S_{w'\beta} e_{-d'}, \quad |w\alpha| = |w'\beta| \quad \text{where } d, d' \text{ are cycle points.}$$

Moreover, repeating the letters for the cycle points d and d' as many times as we want, we can assume that α ends in a repetition of the letters associated to d and similarly for β and d'. But, since $|w\alpha| = |w'\beta|$, the

Cuntz relations imply that $S_{w\alpha}e_{-d} \perp S_{w'\beta}e_{-d'}$ or $w\alpha = w'\beta$. Assume $|w| \leq |w'|$. Then $\alpha = w''\beta$ for some word w''. Then $S_{w\alpha}e_{-d} \perp S_{w'\beta}e_{-d}$ iff $S_\alpha e_{-d} \perp S_{w''\beta}e_{-d'}$. Also, α consists of repetitions of the digits of the cycle associated to d and similarly for d'. So $S_\alpha e_{-d} = e_{-f}$, $S_{w''\beta}e_{-d'} = e_{-f'}$, and all points d, d', f, f', c, c' all belong to the same cycle. So the only case when $S_w e_{-c}$ is not orthogonal to $S_{w'}e_{-c'}$ is when they are equal.

Next we check that the hypotheses of Theorem 4.1 are satisfied. We let $f(t) = e_{-t} \in L^2(\mu_B)$. To check (i) we just to have to see that $e_{-c} \in \cup_{l\in L} S_l\mathcal{E}(L)$. But this follows from (4.11). Requirement (ii) is clear. For (iii) we compute

$$S_l^* e_{-t}(x) = \frac{1}{N}\sum_{b\in B} e^{-2\pi i l\cdot\frac{1}{R}(x+b)}e^{-2\pi i t\cdot\frac{1}{R}(x+b)} = e^{-2\pi x\cdot\frac{1}{R}(t+l)}\frac{1}{N}\sum_{b\in B}e^{-2\pi i b(\frac{t+l}{R})} =$$

$$= \overline{m_B}\left(\frac{t+l}{R}\right)e_{-\frac{t+l}{R}}(x)$$

So (iii) is satisfied with $\mathfrak{m}_l(t) = \overline{m_B}(\frac{t+l}{R})$, $g_l(t) = \frac{t+l}{R}$.

For (iv) take $c_0 = -c$ for any extreme cycle point (0 is always one). For (v), take h continuous on $\mathbb{R}$, $0 \leq h \leq 1$, $h(c) = 1$ for all c with $e_{-c} \in \overline{\text{span}}\mathcal{E}(L)$, and

$$h(t) = \sum_{l\in L}\left|m_B\left(\frac{t+l}{R}\right)\right|^2 h\left(\frac{t+L}{R}\right) := Rh(t)$$

In particular, we have $h(c) = 1$ for every extreme cycle point c. Assume $h \not\equiv 1$. First we will restrict our attention to $t \in I := [a, b]$ with $a \leq \frac{\min L}{R-1}$, $b \geq \frac{\max L}{R-1}$, and note that $g_l(I) \subset I$ for all $l \in L$. Let $m = \min_{t\in I} h(t)$. Then let $h' = h - m$, assume $m < 1$. Then $Rh'(t) = h'(t)$ for all $t \in \mathbb{R}$, h' has a zero in I and $h \geq 0$ on I, $h'(z_0) = 0$. But this implies that $|m_B(g_l(z_0))|^2 h'(g_l(z_0)) = 0$ for all $l \in L$. Since $\sum_{l\in L}|m_B(g_l(z_0))|^2 = 1$, it follows that for one of the $l_0 \in L$ we have $h'(g_{l_0}(z_0)) = 0$. By induction, we can find $z_n = g_{l_{n-1}}\cdots g_{l_0}z_0$ such that $h'(z_n) = 0$. We prove that z_0 is a cycle point. Suppose not. Since m_B has finitely many zeros, for n large enough $g_{\alpha_k}\cdots g_{\alpha_1}z_n$ is not a zero for m_B, for any choice of digits $\alpha_1, \dots, \alpha_k$ in L. But then, by using the same argument as above we get that $h'(g_{\alpha_k}\cdots g_{\alpha_1}z_n) = 0$ for any $\alpha_1, \dots, \alpha_k \in L$. The points $\{g_{\alpha_k}\cdots g_{\alpha_1}z_n : \alpha_1, ...\alpha_k \in L, k \in \mathbb{N}\}$ are dense in the attractor X_L of the IFS $\{g_l\}_{l\in L}$, thus h' is constant 0 on X_L. But the extreme cycle points c are in X_L and since $h(c) = 1$ we have $0 = h'(c) = 1 - m$, so $m = 1$. Thus $h = 1$ on I. Since we can let $a \to -\infty$ and $b \to \infty$ we obtain that $h \equiv 1$. ☐

Remark 4.8. [13] The functions in $\mathcal{E}(L)$ are piecewise exponential. The formula for $S_{l_1\dots l_n}e_{-c}$ is

$$S_{l_1\dots l_n}e_{-c}(x) = e^{\alpha(b,l,c)}\cdot e_{l_1+Rl_2+\dots+R^{n-1}l_{n-1}+R^n(-c)}(x) \tag{4.12}$$

where $\alpha(b, l, c) = -[b_1l_2 + (Rb_1 + b_2)l_3 + \dots + (R^{n-2}b_1 + \dots + b_{n-1})l_n] + (R^{n-1}b_1 + \dots + b_n)\cdot c$ if $x \in \tau_{b_1}\dots\tau_{b_n}X_B$. We have

$$S_{l_1}\dots S_{l_n}e_{-c}(x) = e_{l_1}(x)e_{l_2}(rx)\dots e_{l_n}(r^{n-1}x)e_c(r^n x)$$

If $x \in \tau_{b_1}\dots\tau_{b_n}X_B$ then $rx \in \tau_{b_2}\dots\tau_{b_n}X_B$, $r^{n-1}x \in \tau_{b_n}X_B$. So

$$\begin{aligned}
rx &= Rx - b_1\\
r^2x &= Rrx - b_2 = R^2x - Rb_1 - b_2\\
&\vdots\\
r^{n-1}x &= R^{n-1}x - R^{n-2}b_1 - \dots - Rb_{n-2} - b_{n-1}\\
r^n x &= R^n x - R^{n-1}b_1 - R^{n-2}b_2 - \dots - Rb_{n-1} - b_n.
\end{aligned}$$

The rest follows from a direct computation.

Corollary 4.9. [13] *In the hypothesis of Theorem 4.1, if in addition* $B, L \subset \mathbb{Z}$ *and* $R \in \mathbb{Z}$*, then there exists a set* Λ *such that* $\{e_\lambda : \lambda \in \Lambda\}$ *is an orthonormal basis for* $L^2(\mu_B)$*.*

Proof. If everything is an integer then, it follows from Remark 4.8 that $S_w e_{-c}$ is an exponential function for all w and extreme cycle points c. Note that, as in the proof of Theorem 4.1, $bc \in \mathbb{Z}$ for all $b \in B$.

□

Example 4.10. [13] We consider the IFS that generates the middle third Cantor set: $R = 3$, $B = \{0, 2\}$. The set $\frac{1}{3}\{0, 2\}$ has spectrum $L = \{0, 3/4\}$. We look for the extreme cycle points for (B, L).
We need $|m_B(-c)| = 1$ so $|\frac{1+e^{2\pi i 2c}}{2}| = 1$, therefore $c \in \frac{1}{2}\mathbb{Z}$. Also c has to be a cycle for the IFS $g_0(x) = x/3$, $g_{3/4}(x) = \frac{x+3/4}{3}$ so $0 \leq c \leq \frac{3/4}{3-1} = 3/8$. Thus, the only extreme cycle is $\{0\}$. By Theorem 4.1 $\mathcal{E} = \{S_w 1 : w \in \{0, 3/4\}^*\}$ is an orthonormal basis for $L^2(\mu_B)$. Note also that the numbers $e^{2\pi i \alpha(b,l,c)}$ in formula (4.12) are ± 1 because $2\pi i B \cdot L \subset \pi i \mathbb{Z}$.

4.1.2. *Walsh bases.* In the following, we will focus on the unit interval, which can be regarded as the attractor of a simple IFS and we use step functions for the QMF basis to generate Walsh-type bases for $L^2[0, 1]$ [13].

Example 4.11. [13] The interval $[0, 1]$ is the attractor of the IFS $\tau_0 x = \frac{x}{2}$, $\tau_1 x = \frac{x+1}{2}$, and the invariant measure is the Lebesgue measure on $[0, 1]$. The map r defined in Example 4.3 is $rx = 2x \text{mod} 1$. Let $m_0 = 1$, $m_1 = \chi_{[0,1/2)} - \chi_{[1/2,1)}$. It is easy to see that $\{m_0, m_1\}$ is a QMF basis. Therefore S_0, S_1 defined as in Proposition 4.5 form a representation of the Cuntz algebra $\mathcal{O}_2$.

Proposition 4.12. [13] *The set* $\mathcal{E} := \{S_w 1 : w \in \{0, 1\}^*\}$ *is an orthonormal basis for* $L^2[0, 1]$*, the Walsh basis.*

Proof. We check the conditions in Theorem 4.1. To see that (i) holds note that $S_0 1 = 1$. Define $f(t) = e_t$, $t \in \mathbb{R}$. (ii) is clear. For (iii) we compute

$$S_1^* e_t(x) = \frac{1}{2}(e^{2\pi i t \cdot x/2} + e^{2\pi i t \cdot (x+1)/2}) = e^{2\pi i t \cdot x/2}\frac{1}{2}(1 + e^{2\pi i t/2})$$

$$S_1^* e_t(x) = \frac{1}{2}(e^{2\pi i t \cdot x/2} - e^{2\pi i t \cdot (x+1)/2}) = e^{2\pi i t \cdot x/2}\frac{1}{2}(1 - e^{2\pi i t/2})$$

Thus (iii) holds with $\mathfrak{m}_0(t) = \frac{1}{2}(1 + e^{2\pi i t/2})$, $\mathfrak{m}_1(t) = \frac{1}{2}(1 - e^{2\pi i t/2})$, $g_0(t) = g_1(t) = \frac{t}{2}$. Since $e_0 = 1$ it follows that (iv) holds.

For (v) take h continuous on $\mathbb{R}$, $0 \leq h \leq 1$, $h(c) = 1$ for all $c \in \mathbb{R}$ with $e_t \in \overline{\text{span}}\mathcal{E}$, in particular $h(0) = 1$ and

$$h(t) = \left|\frac{1}{2}(1 + e^{2\pi i t/2})\right|^2 h(t/2) + \left|\frac{1}{2}(1 - e^{2\pi i t/2})\right|^2 h(t/2) = h(t/2)$$

Then $h(t) = h(t/2^n)$ for all $t \in \mathbb{R}$, $n \in \mathbb{N}$. Letting $n \to \infty$ and using the continuity of h, we get $h(t) = h(0) = 1$ for all $t \in \mathbb{R}$. Since all conditions hold, we get that $\mathcal{E}$ is an orthonormal basis. That $\mathcal{E}$ is actually the Walsh basis follows from the following calculations: for $|w| = n$ in $\{0, 1\}^*$ let $n = \sum_i x_i 2^i$ be the base 2 expansion of n. Because $S_0 f = f \circ r$, $S_1 f = m_1 f \circ r$ and $m_0 \equiv 1$ we obtain the following decomposition:

$$S_w 1(x) = m_1(r^{i_1}x) \cdot m_1(r^{i_2}x) \cdots m_1(r^{i_k}x), \quad \text{where } i_1, i_2, \dots i_k \text{ correspond to those } i \text{ with } x_i = 1$$

Also $m_1(r^i x) = m_1(2^i x \text{mod} i)$ are the Rademacher functions and thus we obtain the Walsh basis (see e.g. [20]).

□

The Walsh bases can be easily generalized by replacing the matrix

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

which appears in the definition of the filters m_0, m_1, with an arbitrary unitary matrix A with constant first row and by changing the scale from 2 to N.

Theorem 4.13. [13] *Let $N \in \mathbb{N}$, $N \geq 2$. Let $A = [a_{ij}]$ be an $N \times N$ unitary matrix whose first row is constant $\frac{1}{\sqrt{N}}$. Consider the IFS $\tau_j x = \frac{x+j}{N}$, $x \in \mathbb{R}$, $j = 0, \dots, N-1$ with the attractor $[0,1]$ and invariant measure the Lebesgue measure on $[0,1]$. Define*

$$m_i(x) = \sqrt{N} \sum_{j=0}^{N-1} a_{ij} \chi_{[j/N,(j+1)/N]}(x)$$

Then $\{m_i\}_{i=0}^{N-1}$ is a QMF basis. Consider the associated representation of the Cuntz algebra $\mathcal{O}_N$. Then the set $\mathcal{E} := \{S_w 1 : w \in \{0, ...N-1\}^*\}$ is an orthonormal basis for $L^2[0,1]$.

Proof. We check the conditions in Theorem 4.1. Let $f(t) = e_t$, $t \in \mathbb{R}$.

To check (i) note that $S_0 1 \equiv 1$. (ii) is clear. For (iii) we compute:

$$S_k^* e_t = \frac{1}{N} \sum_{j=0}^{N-1} \overline{m_k}(\tau_j x) e_t(\tau_j x) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \overline{a_{kj}} e^{2\pi i t \cdot (x+j)/N} = e^{2\pi i t \cdot x/N} \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \overline{a_{kj}} e^{2\pi i t \cdot j/N}$$

So (iii) is true with $\mathfrak{m}_k(t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \overline{a_{kj}} e^{2\pi i t \cdot j/N}$ and $g_k(t) = \frac{t}{N}$.

(iv) is true with $c_0 = 0$. For (v) take $h \in \mathcal{C}(\mathbb{R})$, $0 \leq h \leq 1$, $h(c) = 1$ for all $c \in \mathbb{R}$ with $e_c \in \overline{\text{span}} \mathcal{E}$ (in particular $h(0) = 1$), and

$$h(t) = \sum_{k=0}^{N-1} |\mathfrak{m}_k(t)|^2 h(t/N) = h(t/N) \sum_{k=0}^{N-1} \frac{1}{N} |\sum_{j=0}^{N-1} a_{kj} e^{-2\pi i t \cdot j/N}|^2 = h(t/N) \cdot \frac{1}{N} ||Av||^2$$

where $v = (e^{-2\pi i t \cdot j/N})_{j=0}^{N-1}$. Since A is unitary, $||Av||^2 = ||v||^2 = N$. Then $h(t) = h(t/N^n)$. Letting $n \to \infty$ and using the continuity of h we obtain that $h(t) = 1$ for all $t \in \mathbb{R}$. Thus, Theorem 4.1 implies that $\mathcal{E}$ is an orthonormal basis.

□

Remark 4.14. [13] We can read the constants that appear in the step function $S_w 1$ from the tensor of A with itself n times, where n is the length of the word w.

Let A be an $N \times N$ matrix, B an $M \times M$ matrix. Then $A \otimes B$ has entries :

$$(A \otimes B)_{i_1 + M i_2, j_1 + M j_2} = a_{i_1 j_1} b_{i_2 j_2}, \quad i_1, j_1 = 0, \dots, N-1,\ i_2, j_2 = 0, \dots, M-1$$

$$A \otimes B = \begin{pmatrix} Ab_{0,0} & Ab_{0,1} & \cdots & Ab_{0,M-1} \\ Ab_{1,0} & Ab_{1,1} & \cdots & Ab_{1,M-1} \\ \vdots & \vdots & \ddots & \vdots \\ Ab_{M-1,0} & Ab_{M-1,1} & \cdots & Ab_{M-1,M-1} \end{pmatrix}$$

The matrix $A^{\otimes n}$ is obtained by induction, tensoring to the left: $A^{\otimes n} = A \otimes A^{\otimes(n-1)}$.
Thus $A \otimes A \otimes A \otimes \cdots \otimes A$, n times, has entries

$$A^{\otimes n}_{i_0 + N i_1 + N^2 i_2 + \cdots + N^{n-1} i_{n-1}, j_0 + N j_1 + \cdots + N^{n-1} j_{n-1}} = a_{i_0 j_0} a_{i_1 j_1} \dots a_{i_{n-1} j_{n-1}}$$

Now compute for $i_0, \dots i_{n-1} \in \{0, \dots, N-1\}$:

$$S_{i_0 \dots i_{n-1}} 1(x) = m_{i_0}(x) m_{i_1}(rx) \dots m_{i_{n-1}}(r^{n-1} x)$$

Suppose $x \in [\frac{k}{N^n}, \frac{k+1}{N^n})$, $0 \leq k < N^n$ and $k = N^{n-1} j_0 + N^{n-2} j_1 + \cdots + N j_{n-2} + j_{n-1}$, where $0 \leq j_0, \dots, j_{n-1} < N$.

Then $x \in [\frac{j_0}{N}, \frac{j_0+1}{N})$, $rx = (Nx)\text{mod}1 \in [\frac{j_1}{N}, \frac{j_1+1}{N}), \dots, r^{n-1}x = (N^{n-1}x)\text{mod}1 \in [\frac{j_{n-1}}{N}, \frac{j_{n-1}+1}{N})$, so $m_{i_0}(x) = \sqrt{N}a_{i_0j_0}$, $m_{i_1}(rx) = \sqrt{N}a_{i_1j_1}, \dots, m_{i_{n-1}}(r^{n-1}x) = \sqrt{N}a_{i_{n-1}j_{n-1}}$ hence

$$S_{i_0\dots i_{n-1}}1(x) = \sqrt{N^n}a_{i_0j_0}\dots a_{i_{n-1}j_{n-1}} = \sqrt{N^n}A^{\otimes n}_{i_0+Ni_1+N^2i_2+\dots+N^{n-1}i_{n-1},j0+Nj_1+\dots+N^{n-1}j_{n-1}}$$

Example 4.15. [13] The pictures in Figure 3 show the Walsh functions that correspond to the scale $N = 4$ and the matrix

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

for the words of length 2, indicated at the top.

5. Multi-resolutions and generalized wavelet representations.

As is illustrated in [16], and the references given there; as well as in the papers [10, 11, 12, 13], there is a host of problems from analysis of fractals and more generally in stochastic analysis which lend themselves to the present multi-resolution approach. Below we discuss related wavelet representations.

Lemma 5.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probablistic space, and let $\mathcal{A} : \Omega \to X$ be a random variable with values in a fixed measure space $(X, \mathcal{B}_X)$, then $V_\mathcal{A}f := f \circ \mathcal{A}$ defines an isometry. $L^2(X, \mu_\mathcal{A}) \to L^2(\Omega, \mathbb{P})$ where μ_A is the law of A, i.e., $\mu_\mathcal{A}(\Delta) := \mathbb{P}(\mathcal{A}^{-1}(\Delta))$, for all $\Delta \in \mathbb{B}_X$; and $V^*_\mathcal{A}(x) = \mathbb{E}_{\mathcal{A}=x}(\psi|\mathcal{F}_\mathcal{A})$ for all $\psi \in L^2(\Omega, \mathbb{P})$, and all $x \in X$.*

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a solenoid probability span on $\Omega_X = \prod_{n=0}^\infty X$, we shall apply the lemma to the each vertices $\pi_n : \Omega_X \to X$ given by $\pi_n(x_0x_1x_2\cdots) = x_n$, for all $n \in \mathbb{N}_0$, and the isometry can span to π_n will simply be denoted V_n. The sigma-algebra given by π_n will be denoted $\mathcal{F}_n$. Let $(X, \mathbb{B})$ be fixed, let $Rf(x) = \int f(y)\mu(dy|x)$, $f \in \mathcal{F}(X, \mathcal{B})$ and $Rh = h$ i.e., $\mu(\cdots|x)$ is a probability space and $(X, \mathcal{B})$ a.e. $x \in X$. Suppose there exists an X and w such that $\int \mu(B|x)d\lambda(x) = \int_B wd\lambda$, for all $B \in \mathcal{B}_X$ then there exists a probability space $(\Omega, \mathbb{P})$ which is the all paths on $(X, \mathcal{B})$ such that

$$\int_\Omega (f_0 \circ \pi_0)(f_1 \circ \pi_1)\cdots(f_n \circ \pi_n)d\mathbb{P} = \int_X f_0(x)R(f, R(f_2 \cdots R(f_n)\cdots)(x)d\lambda(x)$$

and $\mathbb{P} \circ \pi_1^{-1} = ((W \circ \pi_0)d\mathbb{P}) \circ \pi_0^{-1}$. Moreover,

$$suppt(\mathbb{P}) = Sol_\sigma(X) \iff$$
$$R[(f \circ \sigma)g] = fR(g), \quad \forall f, g \in \mathcal{F}(X, \mathcal{B}).$$

Suppose (R, λ) has the representation $(Rf)(x) = \int_X f(x)\mu(dy, x)$ where $\mu(,x)$ is a measure of $(X, \mathcal{B})$ for all $x \in X$, and each function X $\mu(B, x)$ is measurable for all $B \in \mathcal{B}$. This is only a mild restriction.

Note that a definition by application Riesz if X is locally compact Hausdorff and $\mathcal{B}_X$ is the Borel-sigma algebra. Suppose $R(1) = 1$, then the following representation of $\mathbb{P}$ on $(Sol(X), cylindersets, \mathbb{P})$ are equivalent: The following are equivalent

(i) $\int_{Sol}(f \circ \pi_0)(g \circ \pi_n)d\mathbb{P} = \int_X f(x)(R^n g)(x)d\lambda(x)$
(ii) $Prob_{\text{w.r.t }\mathbb{P}}(\pi_0 = x,\ \pi_1 \in B_1,\ \cdots, \pi_k \in B_k) = \int_{B_1}\int_{B_2}\cdots\int_{B_k}\mu(dy_1|x)\mu(dy_2|y_1)\cdots\mu(dy_k|y_{k-1})$ for all k and for all $B_i \in \mathcal{B}_X$;
(iii) $Prob_{\text{w.r.t }\mathbb{P}}(\pi_0 = x,\ \pi_1 \in B_1,\ \cdots, \pi_k \in B_k) = R(\chi_{B_1}R(\chi_{B_2}R\cdots R(\chi_{B_k})\cdots(x)) = \mathbb{E}_x(\cdots) = \int_{Sol}\cdots d\mathbb{P}_x$

In the case R $R\mathbf{1} = 1$. But if not, then pick h such that $Rh = h$, and set $R'(f) = \frac{R(fh)}{h}$ will $R'(\mathbf{1}) = \mathbf{1}$.

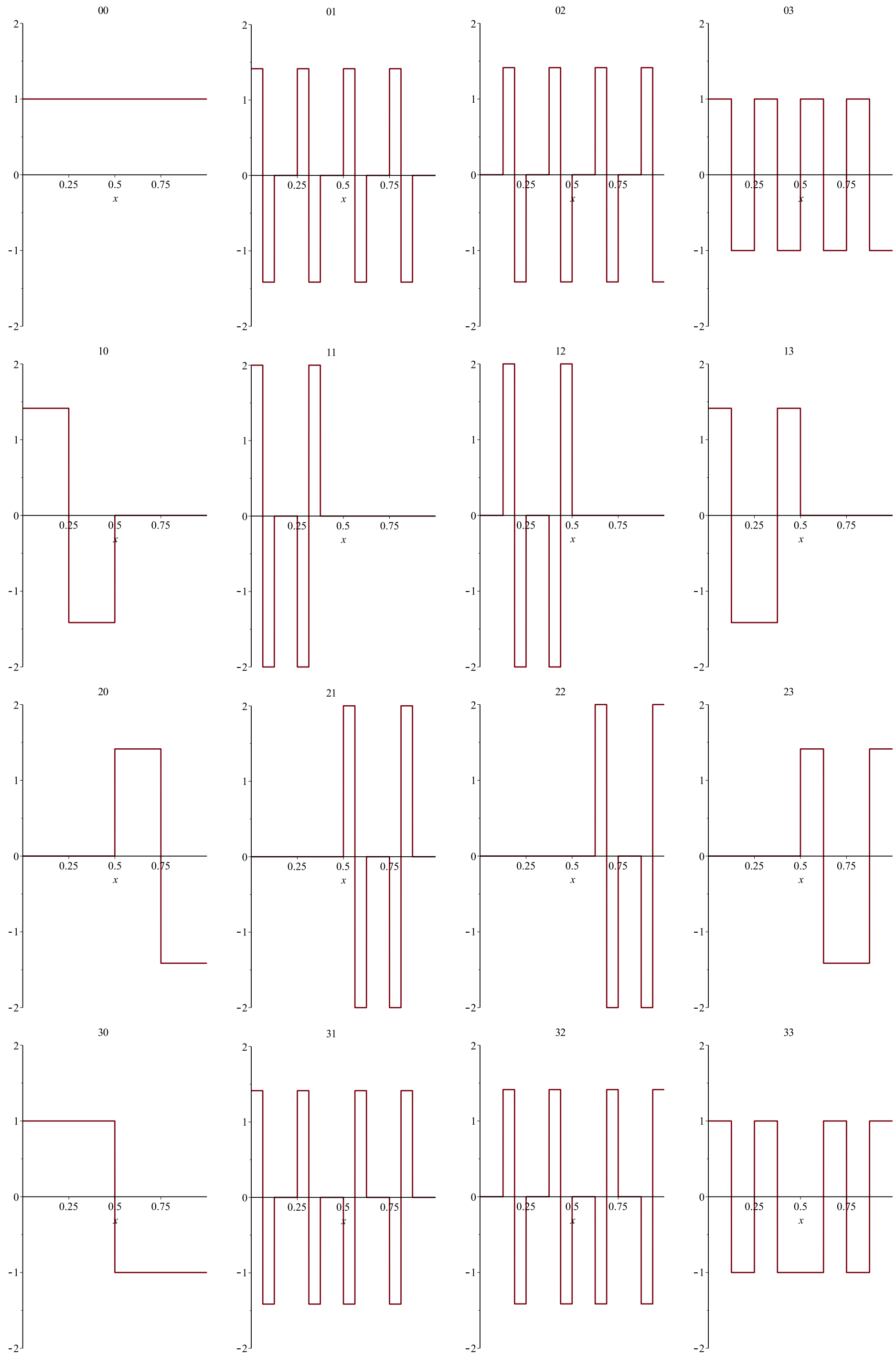

FIGURE 3. Walsh functions graph $S_w 1$ for words w of length 2. [13], [2]

$$Prob(\pi_0 = x, \pi_1 = y_1, \cdots, \pi_k = y_k)$$
$$\frac{1}{N^k} W(y_1)W(y_2)\cdots W(y_k)$$
$$Pr(x \to y_1)Pr(y_1 \to y_2)\cdots Pr(y_{k-1} \to y_k)$$
$$Rf(x) = \int_X f(y)\mu(dy|x) \quad \text{represent R as an....}$$
$$\mu(B|x) := R(\chi_B)(x), \forall B \in \mathcal{B}$$
$$= Prob(\pi_1 \in B|\pi_0 = x).$$

More generally:

$$Prob(\pi_0 = x,\ \pi_1 \in B_1,\ \cdots, \pi_k \in B_k)$$
$$= \int_{B_1}\int_{B_2}\cdots\int_{B_k} \mu(dy_1|x)\mu(dy_2|y_1)\cdots\mu(dy_k|y_{k-1})$$
$$= R(\chi_{B_1}R(\chi_{B_2}R\cdots R(\chi_{B_k})\cdots(x))$$

Same manner prop of $\{\mu(--|x)\}_{x\in X}$.

Lemma 5.2. *If $B \in \mathcal{B}_X$ then*

$$\int_X \mu(B|x)d\lambda(x) = \int_B W(x)d\lambda(x), \quad \textit{where } W = \frac{d\lambda(R)}{d\lambda}.$$

Proof. Let $\{\mu(B|x)\}_{x\in X}$ be a Markov process by $x \in X$ and (X, B) is a fixed measure space and let $\mathbb{P}$ be the corresponding path space measure $\mathbb{P}(\pi_0 = x,\ \pi_1 \in B_1,\ \cdots, \pi_k \in B_k)$. Let

$$\sigma \in End(XB) = \int_{B_1}\int_{B_2}\cdots\int_{B_k} \mu(dy_1|x)\mu(dy_2|y_1)\cdots\mu(dy_k|y_{k-1})$$

then $\mathbb{P}$ is ... as equation $\sigma-$solenoid $Sol_\sigma(X)$ if and only if

$$\mathbb{P}(\pi_{k-1} \in B \cap \sigma^{-1}(A)|\pi_k = x) = \chi_A(x)\mathbb{P}(\pi_k \in B|\pi_k = x)$$

if and only if $supp(\mathbb{P}) \subset Sol_\sigma(X)$. □

Lemma 5.3. *Suppose R has a representation*

$$R(\chi_B)(x) = \mu(B|x), \quad B \in \mathcal{B}_X, \quad x \in X. \tag{5.1}$$

The following are equivalent: as described, i.e., $(Rf)(x) = \int_X f(x)\mu(dy|x)$; then

$$R[(f\circ\sigma)g](x) = f(x)R(g)(x), \quad \forall x, \quad \forall f, g. \tag{5.2}$$

If and only if

$$\mu(\sigma^{-1}(A) \cap B|x) = \chi_A(x)\mu(B|x) \quad \forall A, B \in \mathcal{B} \quad \forall x \in X. \tag{5.3}$$

Notation: Let $\{\mu(\cdot|x)\}_{x\in X}$ be the family of measures on (x, B) define R as in (5.1); and set $\mu_x(\cdot) := \mu(\cdot|x)$ then (5.2) if and only if (conditional measures):

$$\mu_x(\sigma^{-1}(A)|B) = \chi_A(x), \quad \text{where } \mu_x(\cdot|B)$$

denote the conditional probability i.e.,

$$\mu_x(\sigma^{-1}(A)|B) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

$$\mu_x(\sigma^{-1}(A))\mu_x(B|\sigma^{-1}(A)) =$$
$$\mu_x(\sigma^{-1}(A)\cap B) = \chi_A(x)\mu_x(B)$$
$$\mu_x(\sigma^{-1}(A)|B) = \chi_A(x)$$

Proof.

$$\chi_A(x)\mu_x(B|\sigma^{-1}(A)) = \mu(B)\mu_x(\sigma^{-1}(A)|B)$$
$$\chi_A(x)\mu_x(B|\sigma^{-1}(A)) = \chi_A(x)\mu_x(B) = \mu(B)\mu_x(\sigma^{-1}(A)|B)\mu_x(B)$$

□

(i) (LP) $\frac{m_0(0)}{\sqrt{N}} = 1$ Low-Pass. $W = |r_0|$ $(Rf)(x) = \frac{1}{N}\sum_{\sigma(y)=x}(Wf)(y)$
(ii) $R(h) = h$
$\prod_{k=1}^{\infty} \frac{m_0(t/N^k)}{\sqrt{N}} \to |\hat{\varphi}(t)|^2$
$\prod_{k=1}^{\infty} \frac{m_0(t/N^k)}{\sqrt{N}} \in L^2(\mathbb{R})$

Theorem 5.4. $(X, \sigma, R, \lambda) \to Sol_\sigma(X)$. *We have*
(i) $\exists!\ \mathbb{P}$ *such that* $dist(\pi_k) = \mu_k$, $\int_X f d\mu_R = \int_X R^k(f)d\lambda$, *and*
(ii) $\mathbb{P}$ *has the property:* $\frac{d\mathbb{P}\cdot\widehat{\sigma}}{d\mathbb{P}} = w$
Given

$$X \xrightarrow{f} \mathbb{R} \quad \textit{we get}$$
$$X \xrightarrow{f\circ\pi_n} Sol$$
$$V_n f := f\circ\pi_n, \quad \textit{where}$$
$$\mu_n := dist(\pi_n)$$
$$V_n : L^2(X,\mu_n) \xrightarrow{isometry} Sol(X)$$
$$V_n : f \to f\circ\pi_n$$

Here "dist" is short for distribution.

The prove follows from the above discussion.

$\mathcal{H} : (\Omega, \mathcal{F}, \mathbb{P})$. Let λ be on X and $\lambda R << \lambda$ Radon-Nikodym on W. Let U be on $F \in L^2(\Omega,\mathbb{P})$, $R(f)d\lambda = \int fwd\lambda$, $m_0 = \sqrt{W\circ\pi_0}$, $UF = \sqrt{W\circ\pi_0}F\circ\widehat{\sigma}$.

Theorem 5.5. *Let* $L^2(\Omega,\mathbb{P})$, $\mathbf{1}$, U, ρ. *Then there exists a Hilbert space* $\mathcal{H}$, *a representation* ρ *of* $L^2(\Omega,\mathbb{P})$ *on* $\mathcal{H}$, *a unitary operator* U *on* $\mathcal{H}$ *and a vector* φ *in* $\mathcal{H}$ *such that*

$$\rho(f) = m(f\circ\pi_0)$$
$$f\in L^\infty(X) \quad F\to(f\circ\pi_0)F$$

$(\mathcal{H}, \varphi, U, \rho) \to$ *path space measure*
(i) (Covariance)

$$U\rho(f) = \rho(f\circ r)U \quad f\in L^\infty(X). \tag{5.4}$$

(ii) (Scaling equation)

$$F\to(f\circ\pi_0)F\to(f\circ\sigma)\circ\pi_0\sqrt{W\circ\pi_0}F\circ\widehat{\sigma} \tag{5.5}$$

(iii) (Orthogonality)

$$\int_\Omega \rho(f)1d\mathbb{P} = \int(f\circ\pi_0)d\lambda = \int fd\lambda \tag{5.6}$$

(iv)

$$U^{-1}F = \frac{1}{W \circ \pi_1} F \circ \widehat{\sigma}^{-1} \quad \pi_n(w) = x_n \tag{5.7}$$

Example 5.6.

$$UF = \sqrt{W \circ \pi_0} F \circ \widehat{\sigma} \quad \text{where } \sigma : X \to X$$
$$U^* = U^{-1}$$
$$U = \frac{1}{\sqrt{2}} f\left(\frac{x}{2}\right)$$
$$U\widehat{\varphi} = m\widehat{\varphi} \quad \text{where } \widehat{\varphi} \in L^2(\mathbb{R})$$

We are interested in finding the filter function analogous to Dutkay-Jorgensen Haar-Cantor filter. $m_0 = \frac{1+z}{\sqrt{2}}$.

An attempt to find filter functions for Julia set: Let $m_i : X_c \to \mathbb{C}$. For Julia set $r(z) = z^N$ on $\mathbb{T}$. μ is Haar measure on $\mathbb{T}$. Strong invariance of μ with respect to $r(z)$.

$$\int \frac{1}{\#r^{-1}(z)} \sum_{r(w)=z} f(w)dz = \int f(z)d\mu(z)$$

By Brolin's theorem there exists a unique μ strongly invariant for r.

Quadrature mirror filter

$$\frac{1}{\#r^{-1}(z)} \sum_{r(w)=1} |m_0(w)|^2 = 1 \quad z \in J_c.$$

We want to find nice m_0 and m_1.

$w^2 + c = z \Rightarrow w_\pm = \pm\sqrt{z-c}$

$$\frac{1}{2}(|m_0(w_1)|^2 + |m_0(w_2)|^2) = 1$$
$$\frac{1}{2}(|m_1(w_1)|^2 + |m_1(w_2)|^2) = 1$$
$$\frac{1}{2}(m_0(w_1)m_1(w_1) + m_0(w_2)m_1(w_2)) = 1$$

We are interested in solving the following matrix over polynomial

$$\frac{1}{\sqrt{2}} \begin{bmatrix} m_0(w_1) & m_0(w_2) \\ m_1(w_1) & m_1(w_2) \end{bmatrix}$$

where it is unitary and $m_0 = 1$.

Also, we would like to find m_1, high-pass filter

$$\frac{1}{2}(m_1(\sqrt{z-c}) + m_1(-\sqrt{z-c})|) = 0$$
$$\frac{1}{2}(|m_1(\sqrt{z-c})|^2 + |m_1(-\sqrt{z-c})|^2) = 1$$

where $m_1(w) = -m(-w)$.

Acknowledgements. We thank Professors Dorin Dutkay, Gabriel Picioroaga and Judy Packer for helpful discussions.

6. Bibliography

References

[1] Viviane Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[2] Viviane Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[3] M. Braverman and M. Yampolsky. Non-computable Julia sets. *J. Amer. Math. Soc.*, 19(3):551–578 (electronic), 2006.

[4] M. Braverman and M. Yampolsky. Non-computable Julia sets. *J. Amer. Math. Soc.*, 19(3):551–578, 2006.

[5] Mark Braverman. Parabolic Julia sets are polynomial time computable. *Nonlinearity*, 19(6):1383–1401, 2006.

[6] Mark Braverman. Parabolic Julia sets are polynomial time computable. *Nonlinearity*, 19(6):1383–1401, 2006.

[7] Robert L. Devaney and Daniel M. Look. A criterion for Sierpinski curve Julia sets. *Topology Proc.*, 30(1):163–179, 2006. Spring Topology and Dynamical Systems Conference.

[8] Robert L. Devaney and Daniel M. Look. A criterion for Sierpinski curve Julia sets. *Topology Proc.*, 30(1):163–179, 2006. Spring Topology and Dynamical Systems Conference.

[9] Robert L. Devaney, Mónica Moreno Rocha, and Stefan Siegmund. Rational maps with generalized Sierpinski gasket Julia sets. *Topology Appl.*, 154(1):11–27, 2007.

[10] Dorin Ervin Dutkay. The spectrum of the wavelet Galerkin operator. *Integral Equations Operator Theory*, 50(4):477–487, 2004.

[11] Dorin Ervin Dutkay and Palle E. T. Jorgensen. Wavelet constructions in non-linear dynamics. *Electron. Res. Announc. Amer. Math. Soc.*, 11:21–33, 2005.

[12] Dorin Ervin Dutkay and Palle E. T. Jorgensen. Hilbert spaces built on a similarity and on dynamical renormalization. *J. Math. Phys.*, 47(5):053504, 20, 2006.

[13] Dorin Ervin Dutkay, Gabriel Picioroaga, and Myung-Sin Song. Orthonormal bases generated by Cuntz algebras. *J. Math. Anal. Appl.*, 409(2):1128–1139, 2014.

[14] John E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.

[15] Palle E. T. Jorgensen. Matrix factorizations, algorithms, wavelets. *Notices Amer. Math. Soc.*, 50(8):880–894, 2003.

[16] Palle E. T. Jorgensen. *Analysis and probability: wavelets, signals, fractals*, volume 234 of *Graduate Texts in Mathematics*. Springer, New York, 2006.

[17] Palle E. T. Jorgensen. Certain representations of the Cuntz relations, and a question on wavelets decompositions. *Contemp. Math.*, 414:165–188, 2006.

[18] John Milnor. Pasting together Julia sets: a worked out example of mating. *Experiment. Math.*, 13(1):55–92, 2004.

[19] C. L. Petersen and S. Zakeri. On the Julia set of a typical quadratic polynomial with a Siegel disk. *Ann. of Math. (2)*, 159(1):1–52, 2004.

[20] F. Schipp, W. R. Wade, and P. Simon. *Walsh series.* Akademiai Kiado, Adam Hilger Ltd., Bristol, 1990. An introduction to dyadic harmonic analysis.

(Palle E.T. Jorgensen) Department of Mathematics, The University of Iowa, 14 MacLean Hall, Iowa City, IA 52242, USA

E-mail address: palle-jorgensen@uiowa.edu

(Myung-Sin Song) Department of Mathematics and Statistics, Southern Illinois University Edwardsville, Box 1653, Science Building, Edwardsville, IL 62026, USA

E-mail address: msong@siue.edu