On the Beta Transformation - Diary and Notes

Linas Vepstas

December 2017 (Current version: March 2024)*

linasvepstas@gmail.com ORCID: 0000-0002-2557-740X

Abstract

The beta transformation is the iterated map $\beta x \mod 1$. The special case of $\beta = 2$ is known as the Bernoulli map, and is exactly solvable. The Bernoulli map provides a model for pure, unrestrained chaotic (ergodic) behavior: it is the full invariant shift on the Cantor space $\{0,1\}^\omega$. The Cantor space consists of infinite strings of binary digits; it is notable for many properties, including that it can represent the real number line.

The beta transformation defines a subshift: iterated on the unit interval, it singles out a subspace of the Cantor space that is invariant under the action of the left-shift operator. That is, lopping off one bit at a time gives back the same subspace.

The beta transform seems to capture something basic about the multiplication of two real numbers: β and x. It offers insight into the nature of multiplication. Iterating on multiplication, one would get $\beta^n x$ – that is, exponentiation; the mod 1 of the beta transform contorts this in strange ways.

Analyzing the beta transform is difficult. The work presented here is a research diary: a collection of observations and occasional insights. One is that chaos seems to be rooted in how the carry bit behaves during multiplication. Another is that one can surgically insert "islands of stability" into chaotic (ergodic) systems, and have some fair amount of control over how those islands of stability behave. One can have islands with, or without a period-doubling "route to chaos".

The eigenvalues of the transfer operator seem to lie on a circle of radius $1/\beta$ in the complex plane. Given that the transfer operator is purely real, the appearance of such a quasi-unitary spectrum unexpected. The spectrum appears to be the limit of a dense set of quasi-cyclotomic polynomials, the positive real roots of which include the Golden and silver ratios, the Pisot numbers, the n-bonacci (tribonacci, tetranacci, etc.) numbers.

---

1 Introduction

The last three or four decades of mathematical research has seen dramatic advances in the theory of subshifts. This text is mostly not about that, except to point out that this theory has very broad and important impact on many branches of physics and mathematics. From the perspective of the amateur enthusiast, the theory of subshifts finally exposes and makes clear some of the mysterious and intriguing behavior of fractals and of chaotic dynamical systems.

This text focuses almost entirely on just one particular map of the unit interval, the β-transform, defined as the iterated map $\beta x \mod 1$. As such, it is an example of an iterated map on the unit interval of the real number line. Such maps have the form

$$f : [0,1] \to [0,1]$$

and the topic is the exploration of the consequence of iterating the map by composing:

$$f^n(x) = (f \circ f \circ \cdots \circ f)(x) = f(f(\cdots f(x) \cdots))$$

Such one-dimensional iterated maps have been heavily studied, and there is a large body of results, interconnecting many different concepts and results from mathematics, and so having a particularly broad range.

This text attempts to report some brand-new results on the β-transform. This is perhaps surprising, as one might think that the β-transform is sufficiently simple so as to be well-studied and well-understood, it being among the very simplest of iterated one-dimensional maps. This text also attempts to report these results in a naive and unsophisticated fashion, in the hope that this makes the text readable for the interested student and casual enthusiast.

Thus, although the author is personally excited by the advances in the field, this text is neither a survey of known results on the β-transform, nor does it much glance at most of the typical avenues that are available for studying one-dimensional maps. This text does focus extensively on the spectrum of the transfer operator (the "Ruelle Perron Frobenius operator"), and thus it contributes to the general "Koopmania". Little prior knowledge is assumed, and the needed concepts are introduced in a very casual and informal way. This will, no doubt, completely discourage and dismay the formally trained mathematician. The best I can offer is to reiterate: "new results", off the beaten track.

This text begins with some pretty pictures, showing the iterated tent and logistic maps, so as to remind the reader as to why this is an interesting problem to study. The fact is that the β-transformation is far more dry and abstract than the rather sexy logistic map, or its complex cousin, the Mandelbrot set. The hope is that the β-transformation is also simpler, and therefore, perhaps, easier to understand. The reader will soon discover that there is nothing particularly easy about it, and that, at every turn, one bumps into other interesting areas of mathematics that could, perhaps should shed some light, but don't actually seem to do so, in practice.

The most fun for the casual reader might be chapter 5, on the periodic orbits, where the quasi-cyclotomic polynomials appear; these are polynomials of the form $p_n(z) = z^{k+1} - b_0 z^k - b_1 z^{k-1} - \cdots - b_{k-1} z - 1$ for the b_j being binary bits (zero or one). Also

quite fun is the section on the islands of stability, which sheds light on how one can take a purely ergodic (chaotic) system, and surgically insert, as desired, islands of stability. The point here is that the classic logistic map attracted interest precisely because of its interleaving of chaos and stability; it turns out, one can manufacture such systems, at will.

A word about the format of this paper: this is a *de facto* "research diary", not a formal report. This, it contains various unfinished corners, maybe some errors, and notes-to-self.

1.1 Bernoulli shift

The Bernoulli shift, also known as the bit-shift map, the dyadic transform and the full shift, is an iterated map on the unit interval, given by

$$b(x) = \begin{cases} 2x & \text{for } 0 \le x < \frac{1}{2} \\ 2x - 1 & \text{for } \frac{1}{2} \le x \le 1 \end{cases} \tag{1}$$

It can be written much more compactly as $b(x) = 2x \mod 1$. The symbolic dynamics of this map gives the binary digit expansion of x. That is, write

$$b^n(x) = (b \circ b \circ \cdots \circ b)(x) = b(b(\cdots b(x) \cdots))$$

to denote the n-fold iteration of b and let $b^0(x) = x$. The symbolic dynamics is given by the bit-sequence

$$b_n(x) = \begin{cases} 0 & \text{if } 0 \le b^n(x) < \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \le b^n(x) \le 1 \end{cases} \tag{2}$$

Attention: n is a subscript on the left, and a superscript on the right! The left is a sequence, the right is an iteration. Using the letter b one both sides is a convenient abuse of notation. Notation will be abused a lot in this text, except when it isn't. The symbolic dynamics recreates the initial real number:

$$x = \sum_{n=0}^{\infty} b_n(x) 2^{-n-1} \tag{3}$$

All of this is just a fancy way of saying that a real number can be written in terms of it's base-2 binary expansion. That is, the binary digits for x are the $b_n = b_n(x)$, so that

$$x = 0.b_0 b_1 b_2 \cdots$$

is a representation of a real number with a bit-string.

1.2 Bijections

A variety of mathematical objects that can be placed into a bijection with collections of bit-strings, and much of this text is an exploration of what happens when this is done. There will be several recurring themes; these are reviewed here.

The collection of all infinitely-long bit-strings $\{0,1\}^{\omega} = \{0,1\} \times \{0,1\} \times \cdots$ is known as the Cantor space; ω denotes countable infinity, so this is a countable product of repeated copies of two things. Closely related is the Cantor set, which is famously the collection of points $y = \sum_{n=0}^{\infty} b_n(x) 3^{-n-1}$ that results from taking the binary expansion of a real number, and re-expressing it as a base-three expansion. The Cantor set can also be constructed by repeatedly removing the middle-third. If one is careful that the middle-third is always an open set, what remains after a single subtraction is a closed set. What remains after infinite repetition is a "perfect set", and a key theorem is that this perfect set is identical to the collection of points obtained with the sum above. Bouncing between these two distinct constructions requires the definition of the product topology on Cantor space, and thence the Borel sigma algebra, so that one can work in a consistent way with set complements. These ideas will be reviewed as the need arises.

Associated with Cantor space is the infinite binary tree. Any given location in the tree can be specified by giving a sequence of left-right moves, down the tree, starting at the root. Such left-right moves can (of course) be interpreted as bit-strings. After a finite number of moves, one arrives at a node, and under that node extends another infinite binary tree, just like the original. If one has a function $f(b)$ that is defined on every node b of the binary tree, then one can compare this function to $f(Lb)$ and $f(Rb)$ that result from a left move and a right move. If these are equal to each other, or scale in some way when compared to the original, or if there exist two other functions g_L and g_R such that $f(Lb) = g_L(f(b))$ for all b, and likewise, that $f(Rb) = g_R(f(b))$, then one has fractal self-similarity. More explicitly, whenever one has a pair of commuting diagrams, $f \circ L = g_L \circ f$ and , $f \circ R = g_R \circ f$, then one has a dyadic monoid self-symmetry. This is the symmetry of a large class of fractals.

1.2.1 Formalities

The last paragraph is a bit glib, and so some formalities and examples are in order. These are all very straightforward and conventional, almost trivial, belaboring the obvious. Despite the seeming triteness of the next handful of paragraphs, these formal definitions will be needed, so as to avoid future ambiguities and confusions.

Let $\mathbb{M} = \{L,R\}^{<\omega}$ denote the collection of finite-length binary strings. These can be graded by the length ν of the string, so that

$$\mathbb{M} = \{L,R\}^{<\omega} = \bigcup_{\nu=0}^{\infty} \{L,R\}^{\nu} = \varepsilon \cup \{L,R\} \cup (\{L,R\} \times \{L,R\}) \cup \cdots$$

with ε denoting the empty (zero-length) string. This can be turned into a monoid by defining multiplication as string concatenation: given $\gamma \in \{L,R\}^n$ some sequence of L,R moves of length n, and $\gamma' \in \{L,R\}^m$ some other sequence of length m, then $\gamma\gamma' \in \{L,R\}^{n+m}$ is some other string of length $m+n$.

This set is in bijection to the integers in a straight-forward way. This bijection, written as $\kappa : \{L,R\}^{<\omega} \to \mathbb{N}$, can be defined recursively, by a simple commuting diagram. Define $\kappa(\varepsilon) = 1$ and ask that $\kappa(L\gamma) = 2\kappa(\gamma)$ and that $\kappa(R\gamma) = 2\kappa(\gamma) + 1$ for every $\gamma \in \{L,R\}^n$. Thus, $\kappa(L) = 2$ and $\kappa(R) = 3$ and LL, LR, RL, RR map to $4,5,6,7$ respectively.

This bijection commutes with the canonical L, R moves on the natural numbers. Write these as a pair of functions $L : \mathbb{N} \to \mathbb{N}$ and $R : \mathbb{N} \to \mathbb{N}$, defined as $L : m \mapsto 2m$ and $R : m \mapsto 2m + 1$. Interpreting these as strings, one promptly has a pair of commuting diagrams $\kappa \circ L = L \circ \kappa$ and $\kappa \circ R = R \circ \kappa$. True formality would have required writing $\kappa \circ L_s = L_\mathbb{N} \circ \kappa$ or even that $\kappa(L_s \gamma) = L_\mathbb{N}(\kappa(\gamma))$ in order to remind us that L_s is a string (of length one) concatenated onto some other string γ, while $L_\mathbb{N}$ is a map of the natural numbers. It is convenient to drop these labels, as they mostly serve to clutter the text. The intended meaning is always clear from context.

Associated with the set $\{L, R\}^{<\omega}$ is a binary tree $\mathbb{B}$. It can be defined as a graph of vertexes v_j and edges e_{ij} connecting vertex v_i to vertex v_j. Formally, it is the graph $\mathbb{B} = \{v_i, e_{ij} : i \in \mathbb{N}, j \in \{2i, 2i + 1\}\}$. Every vertex $v_i \in \mathbb{B}$ can be given an integer label: it is just the integer i itself. To formalize this, there is a map $\eta : \mathbb{N} \to \mathbb{B}$ that provides this labeling. The canonical labeling gives the root node a label of 1, the left and right sub-nodes 2,3, and so on.

The canonical moves on this binary tree are $L : \mathbb{B} \to \mathbb{B}$ and $R : \mathbb{B} \to \mathbb{B}$ defined by $L : v_i \mapsto v_{2i}$ and $L : v_i \mapsto v_{2i+1}$. Just as above, these commute with the left and right moves on the integers. So, $\eta \circ L = L \circ \eta$ and $\eta \circ R = R \circ \eta$.

The pattern repeats. Consider the set $\mathbb{D}$ of dyadic rationals between zero and one. These are fractions that can be written as $(2n + 1)/2^m$ for some non-negative integers m, n. These are in one-to-one correspondence with the integers: there is a canonical bijection is $\delta : \mathbb{D} \to \mathbb{N}$ given by $\delta : (2n + 1)/2^m \mapsto 2^{m-1} + n$. There are obvious left and right moves, given by $\delta \circ L = L \circ \delta$ and $\delta \circ R = R \circ \delta$. So, for example, $L(1/2) = 1/4$ and $R(1/2) = 3/4$. Viewed as a tree, this places $1/2$ at the root of the tree, and $1/4$ and $3/4$ as the nodes to the left and right.

All of these maps were bijections: they are all invertable. They place elements of all four objects $\mathbb{M}, \mathbb{N}, \mathbb{B}, \mathbb{D}$ in one-to-one correspondence with each other. The commutation of the L, R moves guarantees that the multiplication on $\mathbb{M}$, i.e. string concatenation, allows elements $\gamma \in \mathbb{M}$ to act on $\mathbb{N}, \mathbb{B}, \mathbb{D}$ in the obvious, intended way.

1.2.2   Example: Julia Sets

As a practical example of the above machinery, consider the Julia set of the Mandelbrot map. Recall the Mandelbrot map is an iterated map on the complex plane, given by $z \mapsto z^2 + c$. The Julia set is the set of points of "where things came from" in the Mandelbrot map; it is the inverse "map" $z \mapsto \pm\sqrt{z - c}$. The word "map" is in scare quotes, as the plus-minus in front of the square root indicate that each z maps to either one of two distinct predecessors. The choice of the plus-minus signs can be interpreted as left-right moves, and so the Julia set can be interpreted as a representation of the binary tree, with it's elements labeled by integers, or dyadic fractions, or nodes in the binary tree, or strings of ones and zeros.

We take a moment to make this explicit. Fix a point $c \in \mathbb{C}$ in the complex plane. Define a function $j_c : \mathbb{N} \to \mathbb{C}$ recursively, by writing $j_c(1) = 0$ and then $j_c(2m) = -\sqrt{j_c(m) - c}$ and $j_c(2m + 1) = +\sqrt{j_c(m) - c}$. Equivalently, there are a pair of moves on the complex plane, $L_c : \mathbb{C} \to \mathbb{C}$ and $R_c : \mathbb{C} \to \mathbb{C}$ given by $L_c : z \to -\sqrt{z - c}$ and $R_c : z \to +\sqrt{z - c}$ which commute with the Julia map: $j_c \circ L = L_c \circ j_c$ and $j_c \circ R = R_c \circ j_c$.

The skeleton of the Julia set is a set of points in the complex plane:

$$\mathbb{J}_c = \{z \in \mathbb{C} : z = j_c(n), n \in \mathbb{N}\}$$

The full Julia set is the closure $\overline{\mathbb{J}}_c$ which includes all of the limit points of $\mathbb{J}_c$. The set $\mathbb{J}_c$ is countable; the closure $\overline{\mathbb{J}}_c$ is uncountable.

1.2.3 Example: de Rham Curves

The above construction generalizes. Given any pair of functions $f : X \to X$ and $g : X \to X$ and a point $x_0 \in X$ for any space X, the dyadic monoid induces a map $j : \mathbb{N} \to X$ given recursively as $j(1) = x_0$ and $j(2m) = f(j(m))$ and $j(2m+1) = g(j(m))$. As a commuting diagram, one has that $j \circ L = f \circ j$ and $j \circ R = g \circ j$. This in turn induces a set $\mathbb{J} = \{x \in X : x = j(n), n \in \mathbb{N}\}$. If the space X is a topological space, so that limits can be meaningfully taken, then one also has the closure $\overline{\mathbb{J}}$.

In 1957, Georges de Rham notes that if X is a topological space, so that continuity can be defined, and if there exist two points $x_a, x_b \in X$ such that $g(x_a) = f(x_b)$, then the set $\overline{\mathbb{J}}$ is a continuous curve.[1] The original proof also requires that X be a metric space, and that the two functions f, g be suitably contracting, so that the set $\mathbb{J}$ remains compact. More precisely, so that the Banach fixed-point theorem can be applied, giving two fixed points $x_a = f(x_a)$ and $x_b = g(x_b)$. Without compactness, the curve wanders off to infinity, where conceptions of continuity break down. It is no longer a curve, "out there".

The continuity condition $g(x_a) = f(x_b)$ and the fixed points have a direct interpretation from the viewpoint of the binary tree $\mathbb{B}$. Pick a point $x \in X$, any point at all, and make nothing but left moves: the infinite string $LLL\cdots$. The map converts this to the iteration $f(f(\cdots(x)))$ which converges onto the fixed point $x_a = f(x_a)$. Likewise, all right-moves converge onto $x_b = g(x_b)$. In between, there all other branches in $\mathbb{B}$; but there are also the "gaps" in between the branches.

Consider the two paths $RLLL\cdots$ and $LRRR\cdots$ down the tree. Both start at the root, but end up at different places. Yet, they are immediate neighbors: there are no other branches "in between" these two. Such immediate neighbors always lie at either end of a "gap". Each gap is headed up by the root that sits immediately above them, so that each gap can be labeled by the node from which these two distinct branches diverged. The continuity condition asks that these gaps be closed up: the requirement that $g(x_a) = f(x_b)$ is the requirement that the two sides of the central gap converge to the same point. The curve becomes continuous at this point. By self-similarity, each gap in the tree closes up; the curve is continuous at all such gaps.

As a specific example, consider $X = \mathbb{R}$ and $f(x) = x/2$ and $g(x) = (x+1)/2$. The fixed points are $f(0) = 0$ and $g(1) = 1$ and the continuity condition is satisfied: $g(0) = f(1)$. Iteration produces a curve that is just all of the real numbers of the unit interval. This curve is just the standard mapping of the Cantor space to the unit interval: it is one-to-one for all points that are not dyadic rationals, and it is two-to-one at the dyadic rationals, as the continuity condition explicitly forces the two-to-one mapping.

Note that Julia sets are not de Rham curves: they don't satisfy the continuity criterion.

1.2.4 Shifts

Adjoint to the left and right moves is a shift τ that undoes what the L and R moves do. It cancels them out, so that $\tau \circ L = \tau \circ R = e$ with e the identity function.

Given a string $\gamma \in \mathbb{M}$ of length ν, consisting of letters $a_0 a_1 \cdots a_{\nu-1}$ so that each $a_k \in \{L, R\}$ is a single letter, define the shift $\tau : \mathbb{M} \to \mathbb{M}$ as the function that lops off a single letter from the front, so that $\tau : a_0 a_1 \cdots a_{\nu-1} \mapsto a_1 a_2 \cdots a_{\nu-1}$ is a string that is one letter shorter. This shift is adjoint to the moves L, R, which prepend either L or R to the string. That is, $L : \mathbb{M} \to \mathbb{M}$ which acts on strings as $L : \gamma \mapsto L\gamma$, and likewise, $R : \gamma \mapsto R\gamma$. Then, taken as functions, $\tau \circ L = \tau \circ R = e$ with e being the identity function on $\mathbb{M}$, the function that does nothing. The shift τ is only an adjoint, not an inverse, since there is no way to reattach what was lopped off, at least, not without knowing what it was in the first place. Thus $L \circ \tau \neq e \neq R \circ \tau$. The maps L, R were one-to-one but not onto; the map τ is onto but not one-to-one.

The shift can be composed with either κ or with δ, to have the obvious effect. It's handy to introduce a new letter and a new function $T : \mathbb{N} \to \mathbb{N}$ so that one has $T \circ \kappa = \kappa \circ \tau$ acting as $T(2m) = T(2m+1) = m$. Since τ applied to the empty string returns the empty string, so also $T(1) = 1$. Recycling the same letter $T : \mathbb{D} \to \mathbb{D}$ and defining it so that $\delta \circ T = T \circ \delta$ one can infer that $Tx = 2x \mod 1$ for any $x \in \mathbb{D}$. So, for example, $T(1/4) = T(3/4) = 1/2$.

On the binary tree $\mathbb{B}$, the shift moves back up the tree, from either the left or the right side. That is, given a vertex $v \in \mathbb{B}$, it is the map $\tau : v_{2m} \mapsto v_m$ and likewise $\tau : v_{2m+1} \mapsto v_m$.

For the Julia set example, it has a meaningful form: $\tau : z = j_c(n) \mapsto z^2 + c$. It re-does what the two Julia set maps undid. It is onto: it maps $\mathbb{J}_c$ into all of $\mathbb{J}_c$, and likewise $\overline{\mathbb{J}}_c$ onto $\overline{\mathbb{J}}_c$. For the de Rham curve example, it maps the curve back onto itself. In all three examples, these sets are fixed points of τ. Taking $\mathbb{J}_c \subset \mathbb{C}$ as a subset of the complex plane, it is invariant under the action of τ, so that one has $\tau(\mathbb{J}_c) = \mathbb{J}_c$ and likewise $\tau(\overline{\mathbb{J}}_c) = \overline{\mathbb{J}}_c$. Likewise, the de Rham curve stays fixed in X. These are all examples of invariant subspaces.

1.2.5 Completions

The previous section defined a binary tree $\mathbb{B}$, but this tree is not the "infinite binary tree" alluded to in the opening paragraphs. It is incomplete, in that it does not go "all the way down" to it's leaves. It is not compact, in the same sense that the dyadic fractions $\mathbb{D}$ are not compact: the limit points are absent. The Cantor tree $\overline{\overline{\mathbb{B}}}$ is the closure or completion ore compactification of $\mathbb{B}$; it contains all infinitely-long branches, all the way down to the "leaves" of the tree. The Cantor tree $\overline{\overline{\mathbb{B}}}$ is in one-to-one correspondence with the Cantor space $\{0,1\}^\omega$, and both can be mapped down to the reals on the unit interval, using eqn 3. None of this is particularly deep, but a few paragraphs to articulate these ideas will help avoid later confusion and imprecision.

To convert letter strings to binary strings, define a function $b : \mathbb{M} \to \{0,1\}$ such that it returns 0 if the first letter of a string is L, and otherwise it returns 1. If the string is of zero length, then one has a choice: one can take b to be undefined, or let it be 0, or 1, or introduce a wild-card character $* = 0 \vee 1$ denoting "either zero or one". For the

present, any of these choices is satisfactory. The wildcard is appealing when working with the product topology; but, at the moment, we have no topologies at play.

To extract the n'th letter from a string, define $b_n : \mathbb{M} \to \{0,1\}$ as $b_n = b \circ \tau^n$. Thus, given $\gamma \in \mathbb{M}$ of length ν, one can create a bitstring $b_0 b_1 \cdots b_{\nu-1}$. It can be assigned the obvious numerical value:

$$\left[\delta^{-1} \circ \kappa \right] (\gamma) = \sum_{n=0}^{\nu-1} b_n (\gamma) \, 2^{-n-1}$$

Comparing this to eqn 3, the completion is obvious: $\mathbb{M} = \{L,R\}^{<\omega}$ is completed as $\overline{\mathbb{M}} = \{L,R\}^\omega$ so that it contains all strings of infinite length. This is consistent with the completion of the dyadic rationals $\mathbb{D}$ being the entire real unit interval: $\overline{\mathbb{D}} = [0,1]$. There is no completion $\overline{\mathbb{N}}$ of the countable numbers, at least, not unless one wishes to say that it is the uncountable infinity. This could be done, but then the games gets even more circular, as this completion is just the Cantor set, and we already have that. It seems best to leave $\overline{\mathbb{N}}$ undefined, to avoid circular confusions. The completion $\overline{\mathbb{B}}$ engenders similar confusion. In the original definition, $\mathbb{B}$ was defined as a graph with a countable number of vertexes, each labeled with an integer. This labeling must be abandoned: $\overline{\mathbb{B}}$ is a graph with an uncountable number of vertexes, each labeled by an element from $\overline{\mathbb{M}}$.

The distinction between $\overline{\mathbb{M}}$, $\overline{\mathbb{B}}$ and $\{0,1\}^\omega$ becomes hard to maintain at this point: they are all isomorphic. The distinction between $\overline{\mathbb{M}}$ and $\{0,1\}^\omega$ is particularly strained: they are both collections of strings in two symbols. The primary purpose of trying to maintain this distinction is to remind that $\overline{\mathbb{M}}$ should be though of as a collection of actions that can be applied to sets, while $\{0,1\}^\omega$ is a set, a collection of points that sometimes act as labels for things. This distinction is useful for avoiding off-by-one mistakes during calculations; it is a notational convenience. This is a variant of common practice in textbooks: after showing that two things are isomorphic, only rarely is the notation collapsed into one big tangle. One maintains a Rosetta Stone of different ways of writing the same thing. And so here: a distinction without a difference.

1.3 Shift space

The shift τ was defined above as an operator that takes a sequence of letters, and lops off the left-most symbol, returning a new sequence that is the remainder of the string. A shift space S is any subset of the set of all infinite strings that remains invariant under the shift: $\tau S = S$.

In general settings, one considers a vocabulary of N letters, and the set of infinite sequences N^ω, so that a shift space $S \subseteq N^\omega$ is a subset of the "full shift" N^ω (which is trivially invariant under τ). Shifts that are proper subsets of a full shift will be called subshifts. For the Bernoulli shift, there are $N = 2$ letters, and the Bernoulli shift is by definition the full shift $2^\omega = \{0,1\}^\omega$. A trivial example of a subshift that is not a full shift is the set $S = \{0^\omega, 1^\omega\}$; it has two elements, and is obviously invariant; both 0^ω and 1^ω are fixed points of τ. Another example is $S = \{(01)^\omega, (10)^\omega\}$, where

$(01)^{\omega} = 0101\cdots$ is a repeating periodic string. Any collection of such periodic strings forms a subshift.

Clearly, the union of two subshifts is a subshift, and so, to classify subshifts, one wants to find all of the indecomposable pieces, and characterize those. Factors include periodic strings of fixed period; but not all of these are unique: so, the period-4 string $(0101)^{\omega}$ is really just the period-two string in disguise.

Subshifts consisting entirely of periodic strings can be characterized in terms of Lyndon words. Lyndon words are fixed length strings that are not decomposable into shorter sequences. Thus, each one characterizes a periodic subshift. Cyclic permutations of a Lyndon word give the same subshift; for example, both $(01)^{\omega}$ and $(10)^{\omega}$ belong to the same subshift. The number of distinct, unique subshifts of length v is given by Moreau's necklace counting function: it counts the number of distinct sequences of a given length, modulo cyclic permutations thereof.

Characterizing the subshifts that do not consist of periodic orbits is considerably harder. For example, consider the string $s = 010^{\omega}$. It has an orbit: $\tau s = 10^{\omega}$ and $\tau^2 s = 0^{\omega}$ and so one can write down a set $\langle s \rangle = \{010^{\omega}, 10^{\omega}, 0^{\omega}\}$ which has the property that $\tau \langle s \rangle \subset \langle s \rangle$. However, it is not a subshift, because $\tau \langle s \rangle \neq \langle s \rangle$. The first two points "wander away" under the application of the shift; they are part of the "wandering set". What remains is the fixed point $\tau 0^{\omega} = 0^{\omega}$. The ergodic decomposition theorem states that all such sets X having the property that $\tau X \subset X$ can be decomposed into two pieces: $X = S \cup W$ with S a subshift, $\tau S = S$ and W the wandering set or dissipative set, that eventually dissipates into the empty set: $\lim_{n \to \infty} \tau^n W = \varnothing$. Subshifts are fixed-points; everything else disappears.

Given some real number $x \in [0,1]$, and it's binary expansion $x = 0.b_0 b_1 \cdots$, defined in eqn 2, what is the nature of $\langle s \rangle$ for $s = b_0 b_1 \cdots$? That is, defining

$$\langle s \rangle = \{\gamma = \tau^n s : n \in \mathbb{N}\}$$

what portion of $\langle s \rangle$ is wandering, and which part is a subshift? If $x = p/q$ is a rational number, the answer is easy: rational numbers have binary expansions that are eventually periodic. They consist of some finite-length prefix of non-repeating digits, followed by an infinite-length cyclic orbit. The finite-length prefix is the wandering set; the cyclic part is a subshift. If the period of the cyclic part is v, then the subshift contains precisely v elements.

For the Bernoulli shift $\tau x = 2x \mod 1$, for the real numbers, the answer is provided by the ergodic theorem. For all real numbers $x \in [0,1] \setminus \mathbb{Q}$, that is, the unit interval excluding the rationals, the orbit of x is ergodic: given any real number $\varepsilon > 0$ and any $y \in [0,1]$ there exists some $n \in \mathbb{N}$ such that $|y - \tau^n x| < \varepsilon$. Iteration takes x arbitrarily close to any location on the unit interval. In terms of symbolic dynamics, the binary expansion of y and the binary expansion of $\tau^n x$ will have $m = \lfloor \log_2 \varepsilon \rfloor$ digits in common. The number m can be made arbitrarily large; the subsequence will occur somewhere in the expansion. Put differently, every finite-length string $\gamma \in \{0,1\}^{<\omega}$ occurs as a prefix of (uncountably many) of the strings in $\langle x \rangle$.

In essence, that takes care of that, for the Bernoulli shift, at least, if one is looking at it from the point of view of point dynamics. As long as one thinks with the mind-set of points and their orbits, there is not much more to be said. The above is, in a sense,

a complete description of the Bernoulli shift. But the introduction to this text gave lie to this claim. If one instead looks at the shift in terms of it's transfer operator acting on distributions, then much more can be said. The spectrum of the transfer operator is non-trivial, and the eigenfunctions are fractal, in general. This will be examined more carefully, later; but for now, the topic of point dynamics in the Bernoulli shift is exhausted. This is the end of the line.

The Bernoulli shift is not the only shift on Cantor space. And so, onward ho.

1.4 Beta shift

The beta shift is similar to the Bernoulli shift, replacing the number 2 by a constant real-number value $1 < \beta \leq 2$. It can be defined as

$$T_\beta(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases} \tag{4}$$

This map, together with similar maps, is illustrated in figure 6 below.

Just as the Bernoulli shift generates a sequence of digits, so does the beta shift: write

$$k_n = \begin{cases} 0 & \text{if } 0 \leq T_\beta^n(x) < \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leq T_\beta^n(x) \leq 1 \end{cases} \tag{5}$$

Given the symbolic dynamics, one can reconstruct the original value whenever $1 < \beta$ as

$$x = \frac{k_0}{2} + \frac{1}{\beta}\left(\frac{k_1}{2} + \frac{1}{\beta}\left(\frac{k_2}{2} + \frac{1}{\beta}\left(\frac{k_3}{2} + \frac{1}{\beta}(\cdots)\right)\right)\right)$$

Written this way, the $T_\beta(x)$ clearly acts as a shift on this sequence:

$$T_\beta(x) = \frac{k_1}{2} + \frac{1}{\beta}\left(\frac{k_2}{2} + \frac{1}{\beta}\left(\frac{k_3}{2} + \frac{1}{\beta}\left(\frac{k_4}{2} + \frac{1}{\beta}(\cdots)\right)\right)\right)$$

In this sense, this shift is "exactly solvable": the above provides a closed-form solution for iterating and un-iterating the sequence.

Multiplying out the above sequence, one obtains the so-called "β-expansion" of a real number x, namely the series

$$x = \frac{1}{2}\sum_{n=0}^{\infty} \frac{k_n}{\beta^n} \tag{6}$$

The bit-sequence that was extracted by iteration can be used to reconstruct the original real number. Setting $\beta = 2$ in eqn 5 gives the Bernoulli shift: $T_2(x) = b(x)$.

Unlike the Bernoulli shift, not every possible bit-sequence occurs in this system. It is a subshift of the full shift: it is a subset of $\{0,1\}^\omega$ that is invariant under the action of T_β. The structure of this shift is explored in detail in a later section.

1.5 Associated polynomial

The iterated shift can also be written as a finite sum. A later section will be devoted entirely to the properties of this sum. Observe that

$$T_\beta(x) = \beta\left(x - \frac{k_0}{2}\right)$$

and that

$$T_\beta^2(x) = \beta^2 x - \frac{\beta}{2}(\beta k_0 + k_1)$$

and that

$$T_\beta^3(x) = \beta^3 x - \frac{\beta}{2}\left(\beta^2 k_0 + \beta k_1 + k_2\right)$$

The general form is then:

$$T_\beta^p(x) = \beta^p x - \frac{\beta}{2}\sum_{m=0}^{p-1} k_m \beta^{p-m-1} \tag{7}$$

Since the k_m depend on both β and on x, and are only piece-wise continuous functions, this is not a true polynomial. It does provide a polynomial-like representation with a range of interesting properties.

1.6 Density Visualizations

The long-term dynamics of the β-shift can be visualized by means of a bifurcation diagram. The idea of a bifurcation diagram gained fame with the Feigenbaum map (shown in figure 5). The same idea is applied here: track orbits over long periods of time, and see where they go. This forms a density, which can be numerically explored by histogramming. This is shown in figure 2.

When this is done for the β-shift, one thing becomes immediately apparent: there are no actual "bifurcations", no "islands of stability", no extended period-doubling regions, none of the stuff normally associated with the Feigenbaum map. Although there are periodic orbits, these form a set of measure zero: the iteration produces purely chaotic motion for almost all values of x and all values of $\beta > 1$. In this sense, the beta transform provides a clean form of "pure chaos",[1] without the pesky "islands of stability" popping up intermittently.

The visualization of the long-term dynamics is done by generating a histogram and then taking the limit. The unit interval is divided into a fixed sequence of equal-sized bins; say, a total of N bins, so that each is $1/N$ in width. Pick a starting x, and then iterate: if, at the n'th iteration, one has that $j/N \le b_\beta^n(x) < (j+1)/N$, then increment the count for the j'th bin. After a total of M iterations, let $c(j;M)$ be the count in the

---

[1]Formal mathematics distinguishes between many different kinds of chaotic number sequences: those that are ergodic, those that are weakly or strongly Bernoulli, weakly or strongly mixing. The beta transform is known to be ergodic,[2] weakly mixing[3] and weakly Bernoulli.[4]

j'th bin. This count is the histogram. In the limit of a large number of iterations, as well as small bin sizes, one obtains a distribution:

$$\rho(y;x) = \lim_{N \to \infty} \lim_{M \to \infty} \frac{c(j;M)}{M} \text{ for } \frac{j}{N} \leq y < \frac{j+1}{N}$$

This distribution depends on the initial value *x* chosen for the point to be iterated; a "nice" distribution results when one averages over all starting points:

$$\rho(y) = \int_0^1 \rho(y;x)\,dx$$

Numerically, this integration can be achieved by randomly sampling a large number of starting points. By definition, $\rho(y)$ is a probability distribution:

$$1 = \int_0^1 \rho(x)\,dx$$

Probability distributions are the same thing as measures; they assign a density to each point on the unit interval. It can be shown that this particular distribution is invariant under iteration, and thus is often called the invariant measure, or sometimes the Haar measure.

For each fixed β, one obtains a distinct distribution $\rho_\beta(y)$. The figure 1 illustrates some of these distributions for a selection of fixed β. Note that, for $\beta < 1$, the distribution is given by $\rho_\beta(y) = \delta(y)$, a Dirac delta function, located at $y = 0$.

The general trend of the distributions, as a function of β, can be visualized with a Feigenbaum-style "bifurcation diagram", shown in figure 2. This color-codes each distribution $\rho_\beta(y)$ and arranges them in a stack; a horizontal slice through the diagram corresponds to $\rho_\beta(y)$ for a fixed value of β. A related visualization is in 3, which highlights the discontinuities in 2. Periodic orbits appear wherever the traceries in this image intersect. A characterization of these orbits occupies a large portion of this text.

1.7   Tent Map

The tent map is a closely related iterated map, given by iteration of the function

$$v_\beta(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} \\ \beta(1-x) & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases}$$
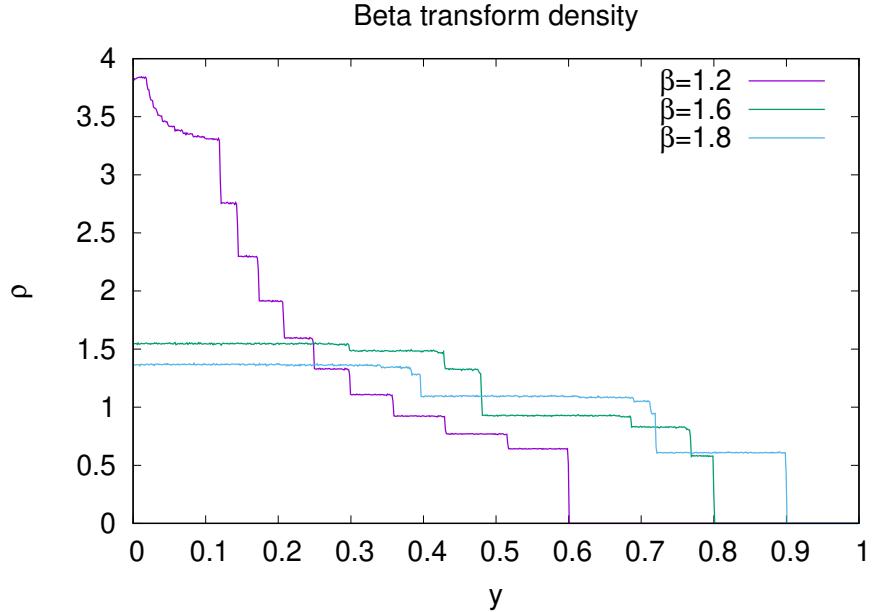
Its similar to the beta shift, except that the second arm is reflected backwards, forming a tent. The bifurcation diagram is shown in figure 4. Its is worth contemplating the similarities between this, and the corresponding beta shift diagram. Clearly, there are a number of shared features.

1.8   Logistic Map

The logistic map is related to the tent map, and is given by iteration of the function
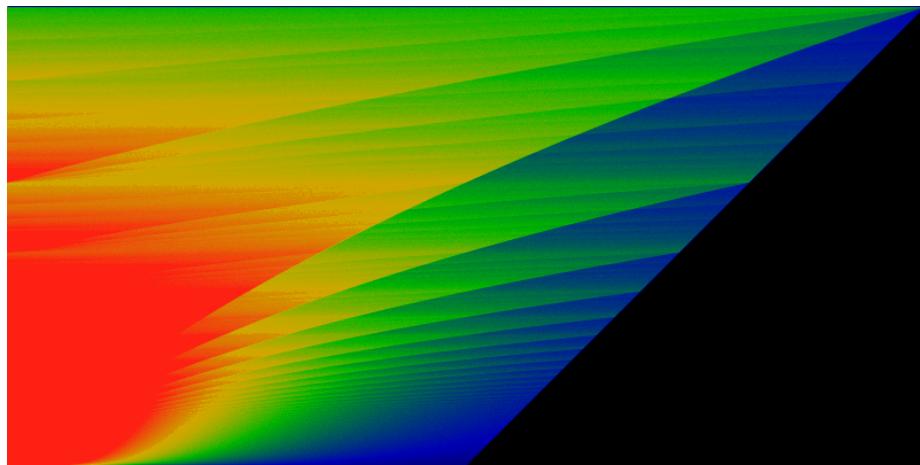
$$f_\beta(x) = 2\beta x(1-x)$$
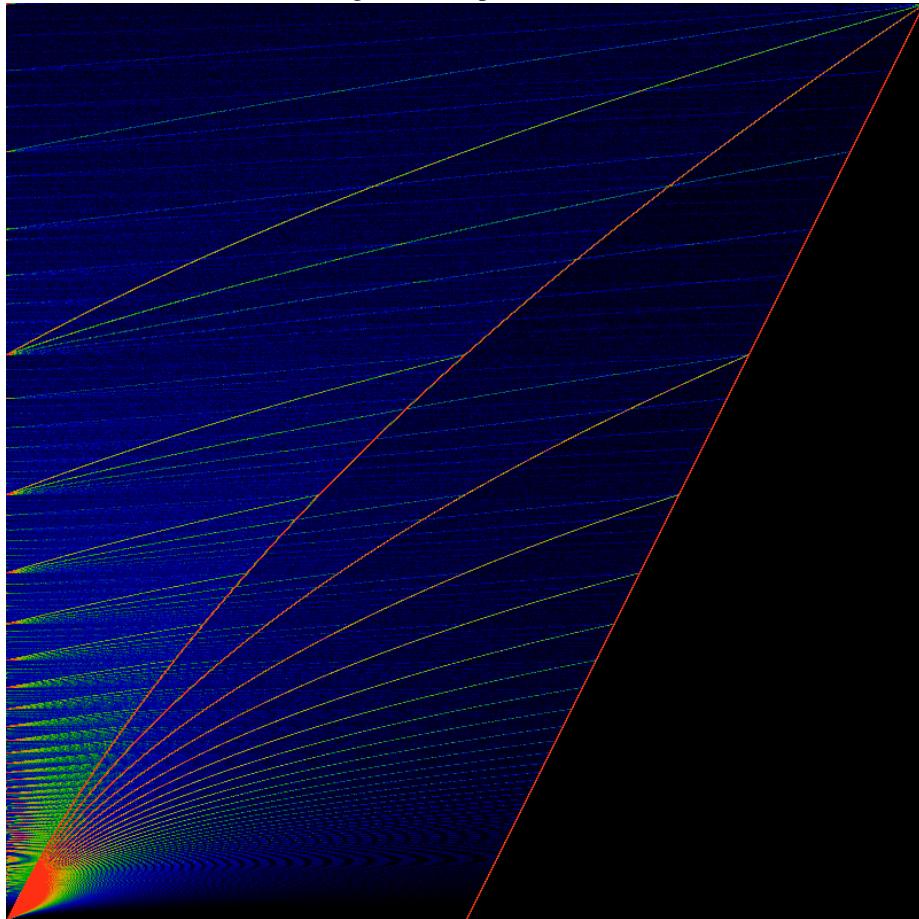
12

Figure 1: Beta-shift Density Distribution

The above figure shows three different density distributions, for $\rho_{1.2}(y)$, $\rho_{1.6}(y)$ and $\rho_{1.8}(y)$, calculated numerically. These are obtained by histogramming a large number of point trajectories, as described in the text. The small quantities of jitter are due to a finite number of samples. To generate this figure, a total of $M = 4000$ iterations were performed, using randomly generated arbitrary-precision floats (using the Gnu GMP package), partitioned into $N = 800$ bins, and sampled 24000 times (or 30 times per bin) to perform the averaging integral. It will later be seen that the discontinuities in this graph occur at the "iterated midpoints" $m_p = T_\beta^p(\beta/2)$. The flat plateaus are not quite flat, but are filled with microscopic steps. There is a discontinuous step at every p; these are ergodically distributed, *i.e.* dense in the interval, so that there are steps everywhere. This is the general case; for special cases, when the midpoint has a finite orbit, then there are a finite number of perfectly flat plateaus. The first such example occurs at $\beta = \left(1 + \sqrt{5}\right)/2 = \varphi$ the Golden Ratio. In this case, there are only two such plateaus.

Figure 2: Beta-shift Bifurcation Diagram
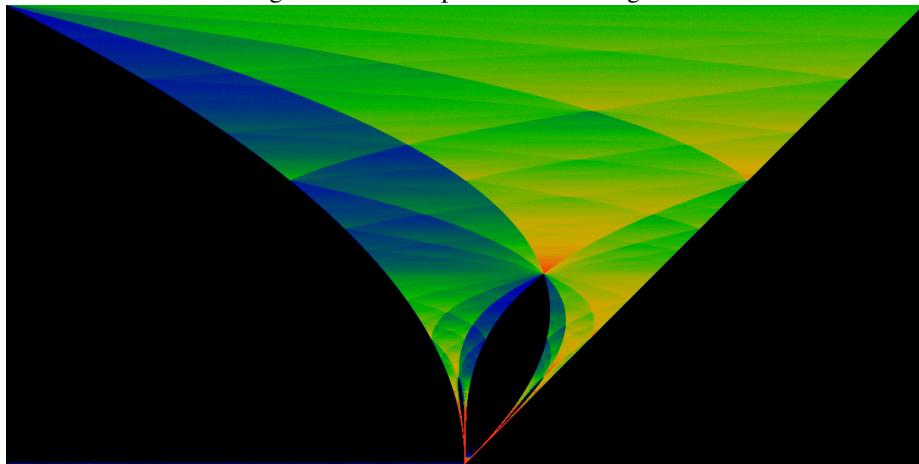
This figure shows the density $\rho_\beta(y)$, rendered in color. The constant β is varied from 1 at the bottom to 2 at the top; whereas y runs from 0 on the left to 1 on the right. Thus, a fixed value of β corresponds to a horizontal slice through the diagram. The color green represents values of $\rho_\beta(y) \approx 0.5$, while red represents $\rho_\beta(y) \gtrsim 1$ and blue-to-black represents $\rho_\beta(y) \lesssim 0.25$. The lines forming the fan shape are not actually straight, they only seem to be; in fact, they have a slight curve. The form will be provided in a later section, it is a variant of the polynomail in eqn 7. The discontinuities in this figure are more clearly highlighted in the next figure, 3.

Figure 3: Midpoint Trace

Traces of midpoint iteration. Each horizontal line corresponds to a fixed β, with β running fom 1 at the bottom, to 2 at the top. At each fixed β, the midpoint $x = 1/2$ is iterated to generate $T_\beta^n(1/2)$. At each such location (from left to right, of 0 to 1), the corresponding pixel is given a color assignment, fading from red, through a rainbow, to black, as n increases. This is a variant of 2, highlighting the edges. A formal analysis of the traceries begins with figure 28
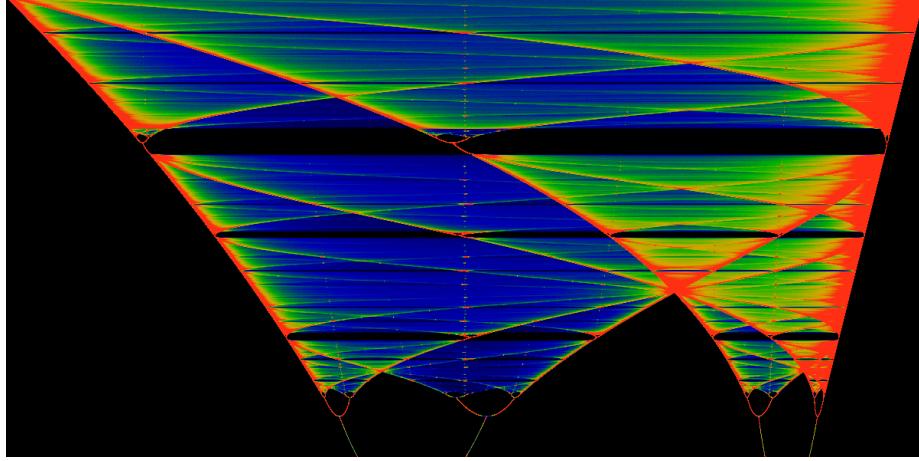
Figure 4: Tent Map Bifurcation Diagram

The bifurcation diagram for the tent map. The value of β runs from 1 at the bottom of the image, to 2 at the top. The color scheme is adjusted so that green represents the average value of the distribution, red represents areas of more than double the average value, while blue shows those values that are about half the average value. Note that this is a different color scheme than that used in figure 2; that scheme would obliterate the lower half of this figure in red.

The black areas represent parts of the iterated range that are visited at most a finite number of times. To the right, the straight boundary indicates that after one iteration, points in the domain $\beta/2 \leq x \leq 1$ are never visited. To the left, points in the domain $0 \leq x \leq \beta\left(1 - \beta/2\right)$ are never visited more than a finite number of times.

Figure 5: Logistic Map Bifurcation Diagram

The logistic map bifurcation diagram. The value of β runs from 1.75 at the bottom of the image, to 2 at the top. The color scheme is adjusted so that green represents the average value of the distribution, red represents areas of more than double the average value, while blue shows those values that are about half the average value. Clearly, the orbits of the iterated points spend much of their time near the edges of the diagram.

This is a very widely reproduced diagram. The goal here is not to waste space reproducing it yet again, but to draw attention to the similarities between this diagram, and the the corresponding diagram for the beta shift.

It essentially replaces the triangle forming the tent map with a parabola of the same height. That is, the function is defined here so that the the same value of β corresponds to the same height for all three maps. Although the heights of the iterators have been aligned so that they match, each exhibits rather dramatically different dynamics. The β-transform has a single fixed point for $\beta < 1$, and then explodes into a fully chaotic regime above that. By contrast, the logistic map maintains a single fixed point up to $\beta = 3/2$, where it famously starts a series of period-doubling bifurcations. The onset of chaos is where the bifurcations come to a limit, at $\beta = 3.56995/2 = 1.784975$. Within this chaotic region are "islands of stability", which do not appear in either the β-transform, or in the tent map. The tent map does show a period-doubling regime, but in this region, there are no fixed points: rather, the motion is chaotic, but confined to multiple arms. At any rate, the period doubling occurs at different values of β than for the logistic map.

The bifurcation diagram is shown in figure 5. Again, it is worth closely examining the similarities between this, and the corresponding tent-map diagram, as well as the β-transform diagram. Naively, it would seem that the general structure of the chaotic regions are shared by all three maps. Thus, in order to understand chaos in the logistic map, it is perhaps easier to study it in the β-transform.

The general visual similarity between the figures 2, 4 and 5 should be apparent,

and one can pick out and find visually similar regions among these three illustrations. Formalizing this similarity is a bit harder, but it can be done: all three of these maps are topologically conjugate to one-another. This is perhaps surprising, but is based on the observation that the "islands of stability" in the logistic map are countable, and are in one-to-one correspondence with certain "trouble points" in the iterated beta transformation. These are in turn in one-to-one correspondence with rational numbers. With a slight distortion of the beta transformation, the "trouble points" can be mapped to the islands of stability, in essentially the same way that phase locking regions (Arnold tongues) appear in the circle map. This is examined in a later section; it is mentioned here only to whet the appetite.

1.9   Beta Transformation

After exactly one iteration of the beta shift, all initial points $\beta/2 \le x \le 1$ are swept up into the domain $0 \le x < \beta/2$, and never leave. Likewise, the range of the iterated beta-shift is $0 \le x < \beta/2$. Thus, an alternative representation of the beta shift, filling the entire unit square, can be obtained by dividing both the domain and range by $\beta/2$ to obtain the function

$$t_\beta(u) = \begin{cases} \beta u & \text{for } 0 \le u < \frac{1}{\beta} \\ \beta u - 1 & \text{for } \frac{1}{\beta} \le u \le 1 \end{cases} \tag{8}$$

This can be written more compactly as $t_\beta(x) = \beta x \mod 1$. In this form, the function is named "the beta-transform", written as the β-transformation, presenting a typesetting challenge to search engines when used in titles of papers. The orbit of a point x in the beta-shift is identical to the orbit of a point $u = 2x/\beta$ in the beta-transformation. Explicitly comparing to the beta-shift of eqn 4:

$$T_\beta^n(x) = \frac{\beta}{2} t_\beta^n \left(\frac{2x}{\beta} \right) \tag{9}$$

The beta-shift and the β-transformation are essentially the same function; this text works almost exclusively with the beta-shift, and is thus idiosyncratic, as it flouts the more common convention of working with the β-transformation. The primary reason for this is a historical quirk, as this text was started before the author became aware of the β-transformation.

After a single iteration of the tent map, a similar situation applies. After one iteration, all initial points $\beta/2 \le x \le 1$ are swept up into the domain $0 \le x < \beta/2$. After a finite number of iterations, all points $0 < x \le \beta(1 - \beta/2)$ are swept up, so that the remaining iteration takes place on the domain $\beta(1 - \beta/2) < x < \beta/2$. It is worth defining a "sidetent" function, which corresponds to the that part of the tent map in which iteration is confined. It is nothing more than a rescaling of the tent map, ignoring those parts outside of the above domain that wander away. The sidetent is given by

$$s_\beta(u) = \begin{cases} \beta(u-1)+2 & \text{for } 0 \le u < \frac{\beta-1}{\beta} \\ \beta(1-u) & \text{for } \frac{\beta-1}{\beta} \le u \le 1 \end{cases}$$

Performing a left-right flip on the side-tent brings it closer in form to the beta-transformation. The flipped version, replacing $u \to 1 - u$ is

$$f_\beta(u) = \begin{cases} \beta u & \text{for } 0 \le u < \frac{1}{\beta} \\ 2 - \beta u & \text{for } \frac{1}{\beta} \le u \le 1 \end{cases}$$

The tent map (and the flipped tent) exhibits fixed points (periodic orbits; mode-locking) for the smaller values of β. These can be eliminated by shifting part of the tent downwards, so that the diagonal is never intersected. This suggests the "sidetarp":

$$a_\beta(u) = \begin{cases} \beta u & \text{for } 0 \le u < \frac{1}{\beta} \\ \beta(1 - u) & \text{for } \frac{1}{\beta} \le u \le 1 \end{cases}$$

The six different maps under consideration here are depicted in figure 6. It is interesting to compare three of the bifurcation diagrams, side-by-side. These are shown in figure 7.
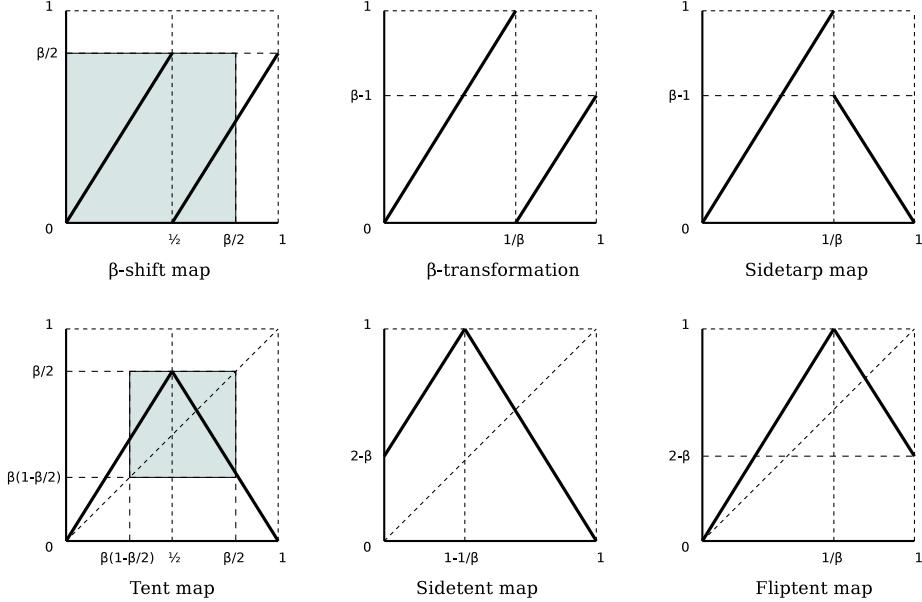
1.10 Dynamical Systems

A brief review of dynamical systems is in order, as it provides a coherent language with which to talk about and think about the beta-shift. The technical reason for this is that a subshift $S \subset \{0,1\}^\omega$ provides a more natural setting for the theory, and that a lot of the confusion about what happens on the unit interval is intimately entangled with the homomorphism 3 (or 6 as the case may be). Disentangling the subshift from the homomorphism provides a clearer insight into what phenomena are due to which component.

The review of dynamical systems here is more-or-less textbook-standard material; it is included here only to provide a firm grounding for later discussion.

The Cantor space $\{0,1\}^\omega$ can be given a topology, the product topology. The open sets of this topology are called "cylinder sets". These are the infinite strings in three symbols: a finite number of 0 and 1 symbols, and an infinite number of * symbols, the latter meaning "don't care". Set union is defined location-by-location, with $0 \cup * = 1 \cup * = *$ and set intersection as $0 \cap * = 0$ and $1 \cap * = 1$. Set complement exchanges 0 and 1 and leaves * alone: $\overline{0} = 1$, $\overline{1} = 0$ and $\overline{*} = *$. The topology is then the collection of all cylinder sets. Note that the intersection of any finite number of cylinder sets is still a cylinder set, as is the union of an infinite number of them. The product topology does *not* contain any "points": strings consisting solely of just 0 and 1 are *not* allowed in the topology. By definition, topologies only allow finite intersections, and thus don't provide any way of constructing "points". Of course, points can always be added "by hand", but doing so tends to generate a topology (the "box topology") that is "too fine"; in particular, the common-sense notions of a continuous function are ruined by fine topologies. The product topology is "coarse".

The Borel algebra, or sigma-algebra, takes the topology and also allows set complement. This effectively changes nothing, as the open sets are still the cylinder sets, although now they are "clopen", as they are both closed and open.
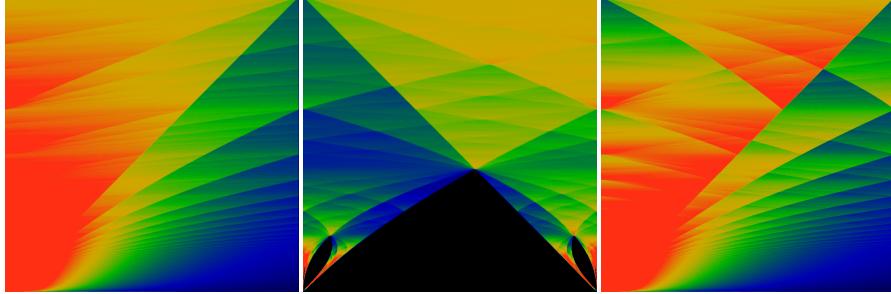
Figure 6: Iterated piece-wise linear maps

The beta shift map, shown in the upper left, generates orbits that spend all of their time in the shaded area: a box of size $\frac{\beta}{2} \times \frac{\beta}{2}$. Enlarging this box to the unit square gives the β-transformation. The tent map resembles the beta shift, except that one arm is flipped to make a tent-shape. After a finite number of iterations, orbits move entirely in the shaded region; enlarging this region to be the unit square gives the sidetent map. Flipping it left-right gives the fliptent map. Although it is not trivially obvious, the fliptent map and the sidetent map have the same orbits, and thus the same bifurcation diagram.

The bottom three maps all have prominent fixed points and periodic orbits, essentially because the diagonal intersects the map. The top three maps have periodic orbits, but these occur only for a countable number of β values. General orbits are purely chaotic, essentially because the diagonal does not intersect them. Note that the slopes and the geometric proportions of all six maps are identical; they are merely rearrangents of the same basic elements.

Figure 7: Beta transform and Side-tent

The left figure shows the bifurcation diagram for the β-transform, as it is normally defined as the $\beta x \mod 1$ map. It is the same map as the beta shift, just rescaled to occupy the entire unit square. In all other respects, it is identical to 2.

The middle figure is a similarly-rescaled tent map, given the name "side tent" in the main text. It is essentially identical to 4, with the middle parts expanded and the sides removed. In both figures, β runs from 1 at the bottom to 2 at the top. The right-hand-side figure is the "sidetarp", clearly its an oddly-folded variant of the beta transform.

Denote the Borel algebra by $\mathscr{B}$. A shift is now a map $T : \mathscr{B} \to \mathscr{B}$ that lops off the leading symbol of a given cylinder set. This is provides strong theoretical advantages over working with "point dynamics": confusions about counting points and orbits and defining densities go away. This is done by recasting discussion in terms of functions $f : \mathscr{B} \to \mathbb{R}$ from Borel sets to the reals (or the complex numbers $\mathbb{C}$ or other fields, when this is interesting). An important class of such functions are the measures. These are functions $\mu : \mathscr{B} \to \mathbb{R}$ that are positive, and are "compatible" with the sigma algebra, in that $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A \cap B = \varnothing$ and (for product-space measures) that $\mu(A \cap B) = \mu(A)\mu(B)$ for all $A, B \in \mathscr{B}$. The measure of the total space $\Omega = \{0,1\}^{\omega}$ is by convention unity: $\mu(\Omega) = 1$.

The prototypical example of a measure is the Bernoulli measure, which assigns probability p to any string containing a single 0 and the rest all *'s. By complement, a string containing a single 1 and the rest all *'s has probability $1 - p$. The rest follows from the sigma algebra: a cylinder set consisting of m zeros and n ones has measure $p^m(1-p)^n$. It is usually convenient to take $p = 1/2$, the "fair coin"; the Bernoulli process is a sequence of coin tosses.

The map given in equation 3 is a homomorphism from the Cantor space to the unit interval. It extends naturally to a map from the Borel algebra $\mathscr{B}$ to the algebra of intervals on the unit interval. It is not an isomorphism: cylinder sets are both open and closed, whereas intervals on the real number line are either open, or closed (or half-open). It is convenient to take the map as a map to closed intervals, so that it's a surjection onto the reals, although usually, this detail does not matter. What does matter is if one takes $p = 1/2$, then the Bernoulli measure is preserved: it is mapped onto the conventional measure on the real-number line. Thus, the cylinder set $0***\cdots$

21

is mapped to the interval $[0,1/2]$ and $1***\cdots$ is mapped to $[1/2,1]$ and both have a measure of $1/2$ and this extends likewise to all intersections and unions. Points have a measure of zero. That is, the homomorphism 3 preserves the fair-coin Bernoulli measure.

Much of what is said above still holds for subshifts. Recall, a subshift S is a subspace $S \subset \{0,1\}^\omega$ that is invariant under the shift T, so that $TS = S$. The space S inherits a topology from $\{0,1\}^\omega$; this is the subspace topology. The Borel algebra $\mathscr{B}$ is similarly defined, as are measures. One can now (finally!) give a precise definition for an invariant measure: it is a measure μ such that $\mu \circ T^{-1} = \mu$, or more precisely, for which $\mu\left(T^{-1}(\sigma)\right) = \mu(\sigma)$ for almost all cylinder sets $\sigma \in S$. This is what shift invariance looks like. Note carefully that T^{-1} and not T is used in the definition. This is because T^{-1} is a surjection while T is not: every cylinder set σ in the subshift "came from somewhere"; we want to define invariance for all σ and not just for some of them.

The T^{-1} is technically called a "pushforward", and it defines a linear operator $\mathscr{L}_T$ on the space $\mathscr{F}$ of all functions $f : \mathscr{B} \to \mathbb{R}$. It is defined as $\mathscr{L}_T : \mathscr{F} \to \mathscr{F}$ by setting $\mathscr{L}_T : f \mapsto f \circ T^{-1}$. It is obviously linear, in that $\mathscr{L}_T(af+bg) = a\mathscr{L}_T(f) + b\mathscr{L}_T(g)$. This pushforward is canonically called the "transfer operator" or the "Ruelle-Frobenius-Perron operator". Like any linear operator, it has a spectrum. The precise spectrum depends on the space $\mathscr{F}$.

The canonical example is again the Bernoulli shift. For this, we invoke the inverse of the mapping of eqn 3 so that $f : [0,1] \to \mathbb{R}$ is a function defined on the unit interval, instead of $f : \mathscr{B} \to \mathbb{R}$. When $\mathscr{F}$ is the space of real-analytic functions on the unit interval, that is, the closure of the space of all polynomials in $x \in [0,1]$, then the spectrum of $\mathscr{L}_T$ is discrete. It consists of the Bernoulli polynomials $B_n(x)$ corresponding to an eigenvalue of 2^{-n}. That is, $\mathscr{L}_T B_n = 2^{-n} B_n$. Note that $B_0(x) = 1$ is the invariant measure on the full shift. For the space of square-integrable functions $f : [0,1] \to \mathbb{R}$, the spectrum of $\mathscr{L}_T$ is continuous, and consists of the unit disk in the complex plane; the corresponding eigenfunctions are fractal. Even more interesting constructions are possible; the Minkowski question mark function provides an example of a measure on $\{0,1\}^\omega$ that is invariant under the shift defined by the Gauss map $h(x) \mapsto \frac{1}{x} - \lfloor \frac{1}{x} \rfloor$. That is, as a measure, it solves $\mathscr{L}_T ?' = ?'$ with ? the Minkowski question mark function, and $?'$ it's derivative; note that the derivative is "continuous nowhere". This rather confusing idea (of something being "continuous nowhere") can be completely dispelled by observing that it is well-defined on all cylinder sets in $\mathscr{B}$ and is finite on all of them – not only finite, but less than one, as any good measure must obey.

These last examples are mentioned so as to reinforce the idea that working with $\mathscr{B}$ instead of the unit interval $[0,1]$ really does offer some strong conceptual advantages. They also reinforce the idea that the Bernoulli shift is not the only "full shift". In the following text, we will be working with subshifts, primarily the beta-shift, but will draw on ideas from the above so as to make rigorous statements about measurability and invariance, without having to descend into either *ad hoc* hand-waving or provide painfully difficult (and confusing) reasoning about subsets of the real-number line.

1.11 Beta Transformation Literature Review and References

The β-transformation, in the form of $t_\beta(x) = \beta x \mod 1$ has been well-studied over the decades. The beta-expansion 5 was introduced by A. Renyi[2] in 1957, who demonstrates the existence of the invariant measure. The ergodic properties of the transform were proven by W. Parry[3] in 1960, who also shows that the system is weakly mixing.

An explicit expression for the invariant measure was obtained independently by A.O. Gel'fond[5] in 1959, and by W. Parry[3], as a summation of step functions

$$v_\beta(y) = \frac{1}{F} \sum_{n=0}^{\infty} \frac{\varepsilon_n(y)}{\beta^n} \tag{10}$$

where ε_n is the digit sequence

$$\varepsilon_n(y) = \begin{cases} 1 & \text{if } y < t_\beta^n(1) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

and F is a normalization constant. By integrating $\varepsilon_n(y)$ under the sum, the normalization is given by

$$F = \sum_{n=0}^{\infty} \frac{t_\beta^n(1)}{\beta^n}$$

Analogous to the way in which a dyadic rational $p/2^n$ has two different binary expansions, one ending in all-zeros, and a second ending in all-ones, so one may also ask if and when a real number x might have more than one β-expansion (for fixed β). In general, it can; N. Sidorov shows that almost every number has a continuum of such expansions![6] This signals that the beta shift behaves rather differently from the Cantor set in it's embedding.

Conversely, the "univoke numbers" are those values of β for which there is only one, unique expansion for $x = 1$. These are studied by De Vries.[7]

The β-transformation has been shown to have the same ergodicity properties as the Bernoulli shift.[4] The fact that the beta shift, and its subshifts are all ergodic is established by Climenhaga and Thompson.[8]

An alternative to the notion of ergodicity is the notion of universality: a β-expansion is universal if, for any given finite string of bits, that finite string occurs somewhere in the expansion. This variant of universality was introduced by Erdös and Komornik[9]. Its is shown by N. Sidorov that almost every β-expansion is universal.[10] Conversely, there are some values of β for which rational numbers have purely periodic β-expansions;[11] all such numbers are Pisot numbers.[12]

The symbolic dynamics of the beta-transformation was analyzed by F. Blanchard[13]. A characterization of the periodic points are given by Bruno Maia[14]. A discussion of various open problems with respect to the beta expansion is given by Akiyama.[15]

When the beta expansion is expanded to the entire real-number line, one effectively has a representation of reals in a non-integer base. One may ask about arithmetic properties, such as the behavior of addition and multiplication, in this base - for example, the sum or product of two β-integers may have a fractional part! Bounds on the lengths of these fractional parts, and related topics, are explored by multiple authors.[16, 17, 18]

Certain values of β – generally, the Pisot numbers, generate fractal tilings,[19, 20, 21, 11, 15] which are generalizations of the Rauzy fractal. An overview, with common terminology and definitions is provided by Akiyama.[22] The tilings, sometimes called (generalized) Rauzy fractals, can be thought of as living in a direct product of Euclidean and p-adic spaces.[23]

The set of finite beta-expansions constitutes a language, in the formal sense of model theory and computer science. This language is recursive (that is, decidable by a Turing machine), if and only if β is a computable real number.[24]

The zeta function, and a lap-counting function, are given by Lagarias[25]. The Hausdorff dimension, the topological entropy and general notions of topological pressure arising from conditional variational principles is given by Daniel Thompson[26]. A proper background on this topic is given by Barreira and Saussol[27].

None of the topics or results cited above are made use of, or further expanded on, or even touched on in the following. This is not intentional, but rather a by-product of different goals.

1.12 Glossary

A collection of basic definitions. Almost none of these are actually needed or used in the text that follows. However, they do provide context for the overall project. This are cribbed from multiple sources, including [28]

1.12.1 Dynamical systems

Basic definitions.

- A map $T : X \to X$ is *strongly transitive* if for every open set $U \subset X$, $\bigcup_{n=1}^{\infty} T^n U = X$.

1.12.2 Subshifts of finite type

Basic definitions.

- Let $\tau : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$ be a shift, where Σ a set of N distinct elements (the alphabet). A *subshift* or *shift space* is a subset $\Lambda \subset \Sigma^{\mathbb{Z}}$ that is compact and invariant under the action of τ. Here, compact means that it is closed under the product topology.

- A finite string of symbols $d_1 d_2 \cdots d_k \in \Sigma^*$ is called a *block*. Let $\mathscr{F}$ be a set of blocks. Define $\Lambda_{\mathscr{F}} \subset N^{\mathbb{Z}}$ as the set of strings that do not contain any block in $\mathscr{F}$ at any location. Then $\Lambda_{\mathscr{F}}$ is a subshift. A subset $\Lambda \subset N^{\mathbb{Z}}$ is a subshift if and only if there is some set $\mathscr{F}$ such that $\Lambda = \Lambda_{\mathscr{F}}$. The elements of $\mathscr{F}$ are called *forbidden blocks*.

- A subshift is of finite type (SFT) if $\mathscr{F}$ is finite. A subshift is *k-step* if $\mathscr{F} \subset \Sigma^{k+1}$ and k is the smallest integer that gives $\Lambda = \Lambda_{\mathscr{F}}$. A 1-step SFT is a *topological Markov chain* (TMC). All SFT's (of any k) can be recoded to a TMC; that is, they are conjugate to a TMC, with "recoding" and "conjugate" defined below.

- Given a shift space Λ, define the ball $B_k(\Lambda)$ as the set of all blocks of length k in Λ. The *language* of Λ is $B(\Lambda) = \bigcup_{k=1}^{\infty} B_k(\Lambda)$.

- A *finite block code* is a map $\Phi : B_k(\Lambda) \to \Sigma'$ mapping blocks of length k to some other alphabet Σ'. If $k = 1$, it is called a *1-block map*.

- A *sliding block code* is is a continuous map $\phi : \Lambda \to \Lambda'$ between shift spaces Λ, Λ', built from alphabets Σ, Σ', that commutes with the shift, so that $\phi \circ \tau = \tau \circ \phi$.

- The Curtis–Lyndon–Hedlund theorem states that *all* sliding block codes are generated by finite block codes. That is, given ϕ there exists an integer k and a finite block code Φ such that $[\phi(x)]_i = \Phi(x_i x_{i+1} \cdots x_k)$ where x_m is the m'th digit in $x \in \Lambda$ and $[\phi(x)]_i$ is the i'th digit in $\phi(x)$. See [28]

- If a sliding block code is a bijection, then it is called a *conjugacy*. If it is a surjection, then it is called a *factor map* or a *factor code*.

- Combining the above gives an alternate definition for an SFT. A subshift is of finite type if there exists directed graph G having N vertices, such that the only transitions between successive digits in the subshift correspond to edges in the graph. That is, for $x \in \Lambda$ having a digit expansion $\cdots d_{k-1} d_k d_{k+1} \cdots$ the only digit transitions $d_k d_{k+1}$ are those that correspond to a (directed) edge in the graph. Every such graph has a corresponding $N \times N$ adjacency matrix A with entries $A_{ij} = 1$ if there is an edge $E : i \mapsto j$ and zero otherwise. The adjacency matrix can be thought of as a Markov chain, and so subshifts of finite type earn the name of "topological Markov chain".

- Let $G = (V, E)$ be an undirected graph. Let $A \subset V$. The boundary of A is $\partial A = \{v \in V \setminus A : (a, v) \in E, a \in A\}$. For directed graphs, we have the incoming and outgoing boundaries. Boundaries define nearest-neighbors, and are used to define interaction potentials for constructing Gibbs measures.

1.12.3 Sofic systems

Basic definitions

- A sofic shift is a factor map of an SFT. By recoding to a TMC, all sofic shifts are 1-block factor maps of TMC's. The graph definition of TMC's above implicitly assumed each vertex label was distinct. A sofic shift maps some of these labels in a way that they are no longer distinct.

- Sofic shifts correspond to regular languages. This follows immediately from the definition in terms of a graph.

- Given a word $w \in B(\Lambda)$, the *follower set* of w is $F(w) = \{u \in B(\Lambda) : wu \in B(\Lambda)\}$. The set of follower sets is $F(\Lambda) = \{F(w) : w \in B(\Lambda)\}$. A shift space is sofic if and only if $F(\Lambda)$ is finite. The predecessor set $P(w)$ is defined analogously; a shift space is sofic iff $P(\Lambda)$ is finite.

- If Λ is sofic, then Λ has dense periodic points. If Λ has dense periodic points, then Λ is non-wandering. See Chandgotia.[28]

1.12.4 Topological Markov fields

Topological Markov fields (TMF) are explored in Chandgotia.[28] TMF's are interesting because they sit "in between" SFT's and sofic systems: every TMF is sofic; not all sofic systems are TMF. Also, there are TMF's that are not SFT's.

- A *topological Markov field* (TMF) is a shift space Λ for which, given two letters $a, b \in \Sigma$ and two words u, v of the same length $|u| = |v|$, then for two strings $aub \in B(\Lambda)$ and $xavby \in B(\Lambda)$ then $xauby \in B(\Lambda)$. This is a 1-step TMF; a k-step TMF replaces a, b by length-k strings.

- A shift space Λ is a TMF if and only if, for all $x, y \in \Lambda$ and all finite cylinders $C \subset \mathbb{Z}$ such that $x = y$ on ∂C, the point $z \in \Sigma^{\mathbb{Z}}$ given by

$$z = \begin{cases} x & \text{on } C \cup \partial C \\ y & \text{on } (C \cup \partial C)^c \end{cases}$$

  is also $z \in \Lambda$. Here, ∂C is the set of integers in $\mathbb{Z} \setminus C$ that are immediately adjacent to elements of C.

1.12.5 Measures

A measure is a function $\mu : \Lambda \to \mathbb{R}^+$ satisfying the usual sigma-additivity axioms on the Borel set. Some additional definitions follow.

- A *cylinder set* is a finite subset of the integers $C \subset \mathbb{Z}$ and a map $c : C \to \Sigma$. The cylinder sets provide the the base of open sets for the topology; together with complementation, this gives the Borel sets.

- A measure is *stationary* if it is shift-invariant, i.e. if $\mu \circ \tau^{-1} = \mu$.

- A shift space Λ is *non-wandering* if $\forall u \in B(\Lambda) \ \exists v \in B(\Lambda)$ s.t. $uvu \in B(\Lambda)$.

- The *support of a measure* is the (closed, compact) subspace of Λ from which all cylinder sets of zero measure have been removed.

- The support of any stationary measure is non-wandering.

1.12.6 Markov measures

Markov measures are stationary; general Markov random fields are not. Again Chandgotia[28] provides gruesome details. Stationary MRF's are Markov chains. Markov chains are Gibbs measures, where Gibbs measures are given by nearest-neighbor interaction potentials. The equivalence is more-or-less due to the boundary conditions, which are nearest-neighbor conditions.

- A *conditional measure* takes the form $\mu(C|D)$ where C, D are cylinder sets. That is, the measure of C is conditionally dependent on D.

- A Markov chain is a conditional measure of the form $\mu(d_1 d_2 \cdots d_n | d_0)$ where $d_1 d_2 \cdots d_n$ are the digits of a contiguous cylinder set. That is, the measure for a Markov chain depends only on the previous digit (and the cylinder that follows) (and there is no other dependency.)

- The support of a Markov chain is exactly a TMC.

- A Markov random field (MRF) is a conditional measure of the form $\mu(d_1 d_2 \cdots d_n | d_0, d_{n+1})$ and so the measure depends on both the preceding and the following digit. A kind of two-sided Markov chain.

- The support of a stationary MRF is exactly a TMF.

1.12.7 Entropy

Basic definitions.

- A finite partition $\mathscr{P} = \{P_1, \cdots, P_n\}$ of a space X is a collection of subsets $P_k \subset X$ such that they are pairwise disjoint, up to measure zero: $\mu(P_i \cap P_j) = 0$ for $i \neq j$, and that $X = \bigcup_k P_k \mod \mu$. That is, $\mu(\bigcup_k P_k) = 1$.

- The entropy of a partition $\mathscr{P}$ is $H(\mathscr{P}) = -\sum_k \mu(P_k) \ln \mu(P_k)$.

- The measure-theoretic entropy $h_\mu(\tau, \mathscr{P})$ of a shift τ w.r.t. $\mathscr{P}$ is

$$h_\mu(\tau, \mathscr{P}) = \lim_{n \to \infty} \frac{1}{n} H\left(\mathscr{P} \vee \cdots \vee \tau^{-n+1} \mathscr{P}\right)$$

where $\mathscr{P} \vee \cdots \vee \tau^{-n+1} \mathscr{P}$ is a partition of X generated by τ by shifting $\mathscr{P}$.

- The metric (or measure-theoretic) entropy $h_\mu(\tau)$ of a shift τ is the supremum of $h_\mu(\tau, \mathscr{P})$ over all finite measurable partitions $\mathscr{P}$.

1.12.8 Equilibria

Equilibrium measures.

- Given a continuous function $\phi : \Lambda \to \mathbb{R}$, the "equilibrium measure" (or "equilibrium state") is defined as the shift-invariant measure μ that maximizes $E = h_\tau(\mu) + \int_\Lambda \phi d\mu$. (Is this is the Gibbs free energy? The sign seems wrong. The Gibbs measure?) Here $h_\tau(\mu)$ is the measure-theoretic entropy of τ w.r.t. μ.

1.13 Other factoids

Other random factoids and definitions that are interesting or useful.

- The topological entropy of of the β-shift is $\log \beta$.

- Every strongly transitive piece-wise monotonic map on the unit interval is topologically conjugate to a β-transfomation.

- A β-shift is of finite type if and only if the β-expansion of 1 is finite.

- A β-shift is sofic if and only if the β-expansion of 1 is eventually periodic.

- If β is a Pisot–Vijayaragavhan number, then the β-shift is sofic. A Pisot–Vijayaragavhan number is an algebraic integer all of whose conjugates have modulus less than 1.

- If a β-shift is sofic then β is a Perron number. (We already see this numerically.)

Anything more?

2 Symbolic Dynamics

A brief review of the Cantor set is in order.

The Bernoulli shift corresponds to the sequence of binary digits of a real number. Such sequences can be imagined to belong to the space of all possible sequences of binary digits, the Cartesian product of infinitely many copies of the set containing two elements:

$$\{0,1\} \times \{0,1\} \times \{0,1\} \times \cdots = \{0,1\}^{\omega} = 2^{\omega}$$

This set is the Cantor set. It has a natural topology, the product topology, whose open sets are finite lengths of digits, possibly interspersed with some "don't care" markers ("don't care" meaning "either 0 or 1"), and terminated by an infinite tail of "don't care" markers. Equation 3 provides a homomorphism from the Cantor set to the unit interval. There are other homomorphisms into a broad variety of fractals. Essentially all of the strange phenomena of fractals and of iterated functions follows from the product topology on this sequence.

The self-similarity of many kinds of fractals can be described with the Cantor set. This arises from the self-similarity of the product space under the action of a shift: specifically, the left-shift, which discards the left-most digit, and shifts the rest over by one. The shift operator itself is that operator that performs this shift; self-similar fractals can be seen to be eigenstates of the shift operator.

The infinitely-deep binary tree is another manifestation of the Cantor set. The set $\{0,1\}$ of the product space can be interpreted as the set $\{L,R\}$ of left-right moves. At each point in a binary sequence, one can make a choice of one of two things: to move left or right. This naturally suggests a binary decision tree.

A byproduct is the presence of some implicit, ambient hyperbolic space. The infinite binary tree, when drawn on flat two-dimensional space, simply "runs out of room", as each subsequent branching pushes closer together. The infinite binary tree can be embedded in the simplest hyperbolic space, the Poincaré disk or upper-half-plane, in such a way that the distance, the spacing between two neighboring nodes is always the same. Visually, this takes the form of some prototypical M.C. Escher drawing, of a repeated fractal form moving out to the edge of a disk. This makes the self-similar shape of the infinite binary tree manifest: as one moves from one location to another, one always sees "the same thing" in all directions: the space is homogeneous.

The rational numbers play a very special role in the infinite binary tree. Dyadic rationals, of the form $(2p+1)/2^n$ for integers p and n correspond to bit sequences (eqn 2) that terminate in all-zeros after a finite number of moves. That is, after an initial "chaotic" sequence, they settle down to a fixed point of period one. General rational numbers p/q behave similarly, in that after an initial "chaotic" sequence, they settle down to periodic orbits of some fixed period. The bit-sequence becomes cyclic. This cyclic behavior implies that most of classical number theory can be dragged into the proceedings. Any particular statement that classical number theory makes with regard to rational numbers, or even modular forms, can be promptly ported over to a statement about the bit-sequences and the orbits of the Bernoulli shift, usually taking on a strange and unrecognizable form.

All of these things go together, like hand in glove: whenever one is manifest and

visible, the others are lurking right nearby, in the unseen directions. All of these things can be given a formal and precise definition, and their explicit interrelationships articulated. This has been done by a wide variety of authors over the last four decades; a proper bibliography would be overwhelming. I have written extensively on all of thee topics, trying to present them in the simplest, most jargon-free way that I can, in a dozen different texts available wherever you found this one. The ideas will not be repeated here; they are not immediately useful to the current proceedings. Nonetheless, the general interplay between all of these concepts is extremely important to understand, and burbles constantly under the surface of the current proceedings. In essence, shifts and subshifts are interesting precisely because they touch on so any different topics; and, conversely, so many different areas of mathematics can inform the subshift.

2.1 Symbolic Dynamics

Given that iteration can generate strings of binary digits, and that these can be reassembled back into real numbers, it is interesting to ask what those mappings look like. To firm up the notation, let $B = (b_n) = (b_0, b_{1,}, \cdots)$ denote a sequence of bits (or symbols) and write

$$x_\beta(B) = \frac{1}{2} \sum_{n=0}^{\infty} \frac{b_n}{\beta^n} \tag{12}$$

as the real number generated from that sequence. Conversely, given a real number x, let $K_\beta(x) = (k_{n;\beta}(x))$ denote the sequence of bits obtained by iterating the beta shift on x with constant β; that is, the sequence generated by eqn. 5. The bit sequence for $K_2(x)$ is just the bit sequence $(b_n(x))$ generated by eqn 2. The transformations between symbol sequences and real numbers make sense only when $1 < \beta \leq 2$.

Two interesting functions can be considered. One is the compressor

$$\mathrm{cpr}_\beta(y) = x_2\left(K_\beta(y)\right)$$
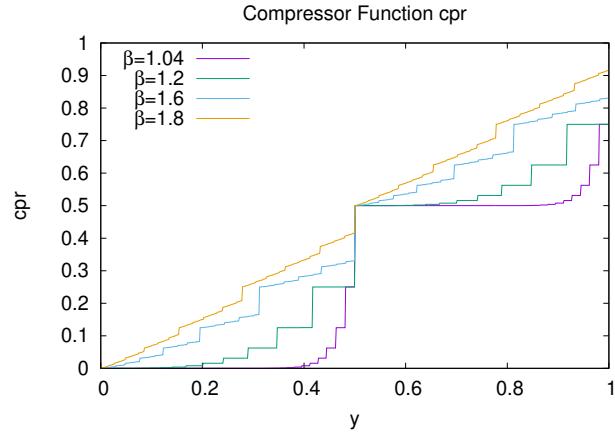
and the other is the expander

$$\mathrm{pdr}_\beta(y) = x_\beta\left(K_2(y)\right) \tag{13}$$

The terms "compressor" and "expander" are being invented here to indicate negative and positive Lyapunov exponents associated with the two functions. For almost all y, the compressor function is pushing nearby points closer together; the total measure of the range of the compressor function is less than one. Likewise, for almost all y, the expander function is pushing nearby points apart. These two functions are illustrated in figures 8 and 9.
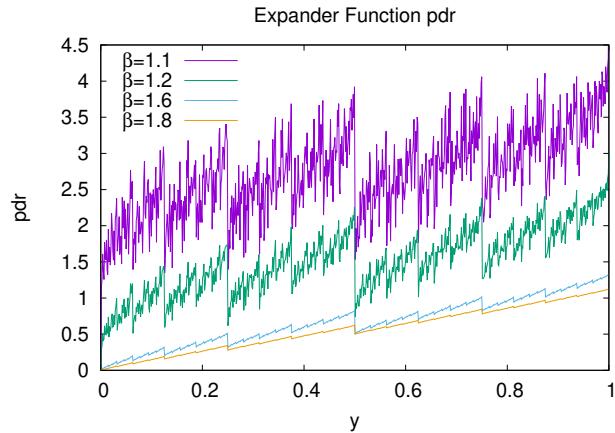
The two functions are adjoint; specifically, one has that $\mathrm{pdr}_\beta\left(\mathrm{cpr}_\beta(y)\right) = y$ for almost all y but that $\mathrm{cpr}_\beta\left(\mathrm{pdr}_\beta(y)\right) \neq y$. The former relation is equivalent to eqn. 6. Not all possible sequences of bit strings appear in the beta shift sequence $K_\beta(x)$; that is, this function is not a surjection onto $\{0,1\}^\omega$. This manifests itself as the gaps in the range of the compressor function, clearly visible in figure 8. If a sequence of bits is

Figure 8: Compressor Function

Compressor Function cpr

This illustrates the compressor function for various values of β. Some caution is advised in interpreting the curves at $\beta = 1.2$ and less. These curves appear to have plateaus; this is a visual artifact only. They are sloped, much as for larger betas, just at a microscopic scale. Each curve can be thought of as a remapping of the Cantor set: each discontinuity is a "middle third" step. The image consists of an uncountable collection of disjoint perfect sets.

Figure 9: Expander Function

Expander Function pdr

This illustrates the expander function for various values of β. As should be clear, almost all neighboring input values are mapped to wildly different output values.

31

viewed as a sequence of left-right moves walking down a binary tree, this implies that some branches of the tree are never taken, and can be pruned. Only branches on the right are ever pruned: That is, there can be arbitrarily long sequences of zeros in the expansion, but the longest possible sequence of 1's is always bounded. The longest run of 1's possible is the largest value of n that satisfies

$$2 \geq \frac{1 + \beta + \beta^2 + \cdots \beta^{n-1}}{\beta^{n-1}}$$

Solving, the bound is

$$n = 1 + \left\lfloor \frac{-\log(2-\beta)}{\log \beta} \right\rfloor \tag{14}$$

That is, every n'th right branch is pruned from the binary tree. For example, a run of three 1's in a row is possible only for $\beta \geq \left(1 + \sqrt{5}\right)/2 = 1.618034\cdots$ the Golden Ratio. The range of $\mathrm{cpr}_\beta(y)$ is most of, but not all of the Cantor set. The figure 10 visualizes the range of the compressor as a function of β.

2.2 Shifts with holes

Viewed as a shift space, as opposed to a cut-down binary tree, the trimming can be thought of as a punching of holes into the full shift. This requires a bit of mental gymnastics. Let (a,c) be an (open) interval on the real number line: $(a,c) = \{x \,|\, a < x < c\}$. Given the Bernoulli shift $b(x) = T_2(x)$ from eqns 1 or 4, consider the set

$$\mathscr{I}(a,c) = \{x \,|\, b^n(x) \notin (a,c) \text{ for any } n \geq 0\}$$

That is, as one iterates on some fixed x, one requests that no iterate $b^n(x)$ ever lands in the interval (a,c). In essence, one has punched a hole in the unit interval; this corresponds to a "hole" in the full Bernoulli shift. The set $\mathscr{I}(a,c)$ is what remains after punching such a hole.

How can this be visualized? Considering the case $n = 0$, its clear that $\mathscr{I}(a,c)$ cannot contain (a,c). That is, $\mathscr{I}(a,c) \cap (a,c) = \emptyset$. For $n = 1$, the interval (a,c) can come from one of two places: either from $\left(\frac{a}{2}, \frac{c}{2}\right)$ or from $\left(\frac{a+1}{2}, \frac{c+1}{2}\right)$, and so neither of these can be in $\mathscr{I}(a,c)$. Continuing, for $n = 2$, the intervals $\left(\frac{a}{4}, \frac{c}{4}\right)$, $\left(\frac{a+1}{4}, \frac{c+1}{4}\right)$, $\left(\frac{a+2}{4}, \frac{c+2}{4}\right)$ and $\left(\frac{a+3}{4}, \frac{c+3}{4}\right)$ must also be gone. Continuing in this fashion, one proceeds with an infinite hole-punch: to obtain $\mathscr{I}(a,c)$, one just cuts out (a,c) and everything that iterates to (a,c). For the holes, write
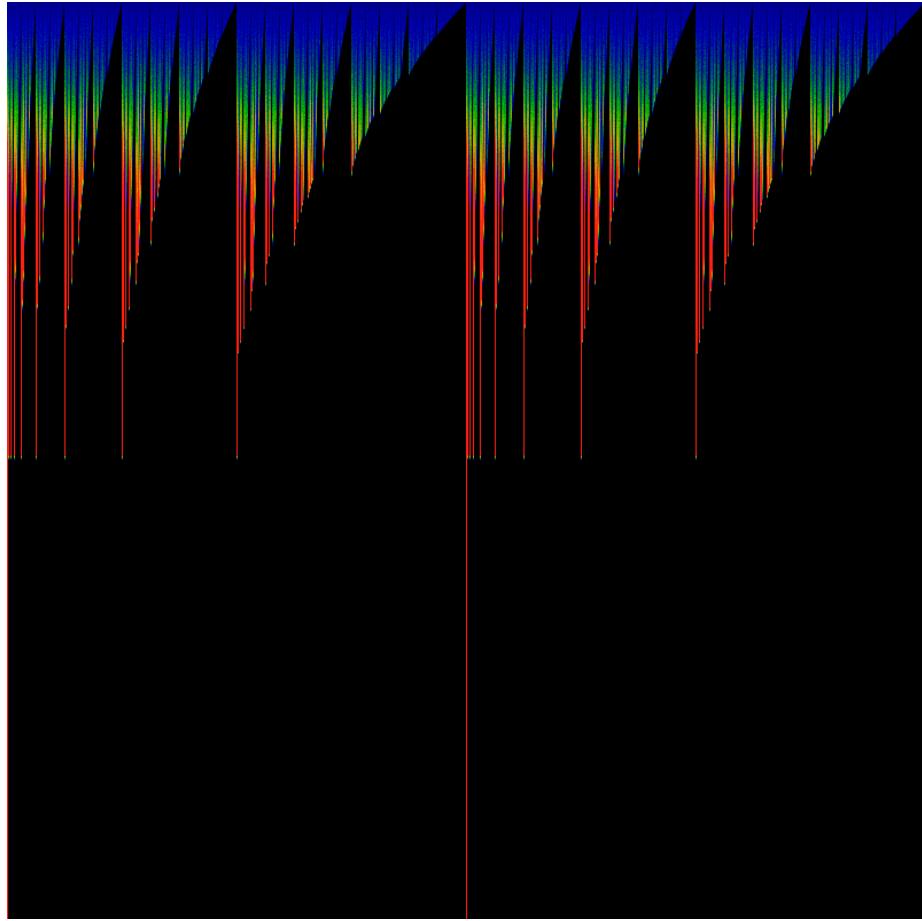
$$\mathscr{H}(a,c) = \bigcup_{n=0}^{\infty} \bigcup_{k=0}^{2^n-1} \left(\frac{a+k}{2^n}, \frac{c+k}{2^n}\right)$$

and for the interval with the holes punched out:

$$\mathscr{I}(a,c) = [0,1] \setminus \mathscr{H}(a,c)$$

where $\bigcup$ denotes set-union and $\setminus$ denotes set subtraction. It is not hard to see that, in the end, this forms a contorted Cantor set, using the standard midpoint-subtraction

Figure 10: Range of the compressor

This figure illustrates a color coded visualization of the range of the compressor function. As before β varies from 0 at the bottom to 2 at the top, and y varies from 0 on the left to 1 on the right. In general, the compressor function maps intervals of the real number line to single points; the color corresponds to the size (the measure) of the intervals that were mapped to that particular point. Blue corresponds to a compression of the measure by about 1, green to a compression of about 2-3, and yellow-red to a compression greater than that.

algorithm, but with different endpoints. The canonical Cantor set is built by taking $(a,c) = \left(\frac{1}{3}, \frac{2}{3}\right)$.

Note that both $\mathscr{H}(a,c)$ and $\mathscr{I}(a,c)$ are subshifts: applying the left-shift to them just returns the same set again. Bot are invariant under the action of the shift operator. In formulas,

$$b\mathscr{H}(a,c) = \mathscr{H}(a,c)$$

and

$$b\mathscr{I}(a,c) = \mathscr{I}(a,c)$$

where, for notational simplicity, the parenthesis are not written, so that for the set S, write $bS = b(S)$. As shifts, its more appropriate to view both as sets of bit-sequences, so that the proper relationship between one and the other should have been written as

$$\mathscr{I}(a,c) = \{0,1\}^{\omega} \setminus \mathscr{H}(a,c)$$

How should these subshifts be visualized as strings? Let $(b_n(x))$ be the bit sequence generated by x, for some $a < x < c$. The cut operation states that such strings can never occur anywhere in $\mathscr{I}(a,c)$. Explicitly, $\mathscr{I}(a,c)$ never contains sequences of the form $d_0 d_1 d_2 \cdots d_k b_0(x) b_1(x) b_2(x) \cdots$ for any arbitrary leading bits $d_0 d_1 d_2 \cdots d_k$.

How should these subshifts be visualized as binary trees? The simplest case to visualize is to take $a = m/2^n$ and $c = (m+1)/2^n$ being dyadic rationals, for some integers m, n. In this case, one takes the bit-expansion for both have the same n leading bits: one starts at the root of the tree, and walks down the binary tree, making left-right moves in accordance with this sequence, and after n moves, arrives at a node above a subtree. Just cut out this subtree, in it's entirety. That's the first cut. Now repeat the process, for the left and right subtrees, from off the root, *ad infinitum*. For a and c not dyadic rationals, the process is more complicated. If a and c are ordinary rationals, thus having a repeating bit-sequence, one performs in the same way, but cyclically walking down the side branches of subtrees. For a and c irrational, the algorithm is considerably more complicated, and is left as an exercise for the reader :-).

A general classification of shifts with holes, for the beta transform, was performed by Lyndsey Clark[29].

2.3 Generalized compressors and expanders

The range of the compressor function is a shift with a hole. Specifically, for a given β, the range of cpr_β is $\mathscr{I}\left(\frac{\beta}{2}, \frac{1}{2}\right)$. The construction for shifts with holes can then be applied to construct generalized compressor and expander functions. One way, which is really rather cheesy, but it works, is to define the function

$$\text{dcpr}_{\beta,\gamma}(a;x) = \sum_{n=0}^{\infty} \left[\frac{1}{\gamma^{n+1}} \sum_{k=0}^{2^n-1} \delta\left(x - \frac{a+k}{\beta^n}\right) \right]$$

and then define the generalized compressor as

$$\text{cpr}(a;x) = \int_0^x \text{dcpr}(a;y)\, dy$$

34

That is, as one walks along the unit interval, from left to right, one picks up points with weights on them, obtaining a generalized Devil's staircase (Cantor-Vitali) function. This generalization does not seem to be terribly useful here, and is left to rot.

2.4 Self-similarity

Subshifts are, by definition, self-similar. If S is a subshift, and T is the shift operator, then $TS = S$ is the key relation obeyed by a subshift. It is fun to see how this actually manifests itself on the unit interval.

So, the two functions cpr and pdr are self-similar. The pdr function demonstrates classic period doubling self-similarity: namely, under $g(x) = x/2$, it behaves as

$$\left(\text{pdr}_\beta \circ g \right)(x) = \text{pdr}_\beta \left(\frac{x}{2} \right) = \frac{1}{\beta} \text{pdr}_\beta (x)$$

while under reflection $r(x) = 1 - x$, it behaves as

$$\left(\text{pdr}_\beta \circ r \right)(x) = \text{pdr}_\beta (1 - x) = \frac{\beta}{2(\beta - 1)} - \text{pdr}_\beta (x)$$

Note that

$$\lim_{x \to 1} \text{pdr}_\beta (x) = \frac{\beta}{2(\beta - 1)}$$

The full dyadic monoid is generated by the generators g and r. It consists of all finite-length strings of the form

$$\gamma = g^{a_1} r g^{a_2} r g^{a_3} r \cdots r g^{a_m}$$

for a finite-length list of positive integers $(a_1, a_2, a_3, \cdots, a_m)$. Note that this list can be interpreted as a continued fraction; this establishes an isomorphism between the continued fractions and the dyadic monoid. This isomorphism is conventionally called the Minkowski question-mark function. We can be a bit more explicit. Write

$$g_\beta (x) = \frac{x}{\beta}$$

and

$$r_\beta (x) = \frac{\beta}{2(\beta - 1)} - x$$

Note that $r_\beta \circ r_\beta = 1$ so that it is a reflection, and that $r_2 (x) = 1 - x$ is the reflection of the unit interval. For strings of binary digits, the reflection performs the exchange of $0 \leftrightarrow 1$; its a mirroring symmetry. The above relationships then become

$$g_\beta \circ \text{pdr}_\beta = \text{pdr}_\beta \circ g_2$$

and

$$r_\beta \circ \text{pdr}_\beta = \text{pdr}_\beta \circ r_2$$

35

The symmetry under the dyadic monoid is simply the statement that

$$\gamma_\beta \circ \text{pdr}_\beta = \text{pdr}_\beta \circ \gamma_2$$

for any γ of the above form.

The function cpr also exhibits self-similarity, although it alters (expands) what happens on the x axis. Several self-similarities are apparent. First, for $0 \le x \le 1$, one has

$$\text{cpr}_\beta \left(\frac{x}{2} \right) = \frac{1}{2} \text{cpr}_\beta \left(\frac{\beta x}{2} \right)$$

Equivalently, for $0 \le y \le \beta/2$ one can trivially restate the above as

$$\text{cpr}_\beta \left(\frac{y}{\beta} \right) = \frac{1}{2} \text{cpr}_\beta (y) \tag{15}$$

Although this follows trivially, this restatement helps avoid certain confusions later in this text. This can also be written as

$$g_2 \circ \text{cpr}_\beta = \text{cpr}_\beta \circ g_\beta$$

There is obviously no corresponding reflection symmetry. The left and right halves are identical to one-another, but offset:

$$\text{cpr}_\beta \left(\frac{1}{2} + \frac{x}{2} \right) = \frac{1}{2} + \text{cpr}_\beta \left(\frac{x}{2} \right)$$

It follows that

$$\text{cpr}_\beta \left(\frac{1}{2} + \frac{y}{\beta} \right) = \frac{1}{2} + \frac{1}{2} \text{cpr}_\beta (y)$$

Combining the above results into one, one has that

$$\text{cpr}_\beta \left(\frac{y}{\beta} \right) + \text{cpr}_\beta \left(\frac{1}{2} + \frac{y}{\beta} \right) = \frac{1}{2} + \text{cpr}_\beta (y)$$
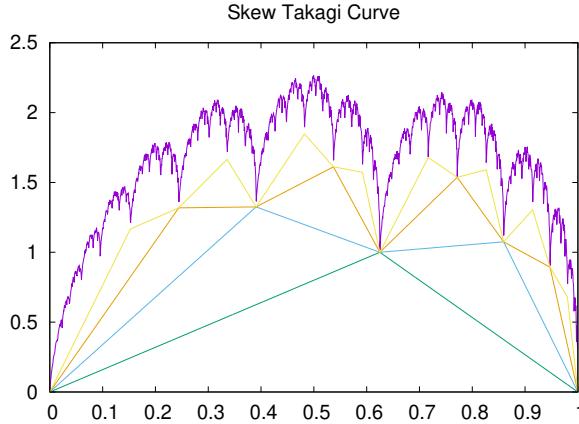
This last form is interesting, as it makes an appearance in relation to the transfer operator, defined below.

2.5 Other things with similar symmetry

The cpr curve is just one that belongs to a class of such curves. As an example, one may construct a Takagi (blancmange) curve by iterating triangles whose peak is located at $1/\beta$. The Takagi curve is an example of a curve transforming under a 3-dimensional representation of the dyadic monoid; the cpr curves transforms under a two-dimensional representation. See my paper on the Takagi curve for details. Figure 11 shows such a curve. Denote by $\text{tak}_{\beta;w}(x)$ a curve constructed in this fashion. The transformation properties of this curve include self-similarity on the left, as

$$\text{tak}_{\beta;w} \left(\frac{x}{\beta} \right) = x + w \, \text{tak}_{\beta;w} (x)$$

36

Figure 11: Skew Takagi Curve

Skew Takagi Curve

This figure shows a skew Takagi curve, and the first four steps of its construction. The initial triangle is of height 1; the apex is located at $1/\beta$, for $\beta = 1.6$ in this figure. Subsequent triangles obtain a height of w=0.7 above the apex point, and are similarly skew.

for $0 \le x \le 1$ and self-similarity on the right, as

$$\text{tak}_{\beta;w}\left(\frac{1}{\beta} + x\left(1 - \frac{1}{\beta}\right)\right) = 1 - x + w\,\text{tak}_{\beta;w}(x)$$

Both of these properties follow directly from the construction of the curve; they can be taken as the defining equations for the curve. That is, the curve can be taken as that function which satisfies these two recursion relations.

The derivative of the skew Takagi curve is shown in figure 12, and, for lack of a better name, could be called the skew Haar fractal wavelet. It can be defined as the formal derivative
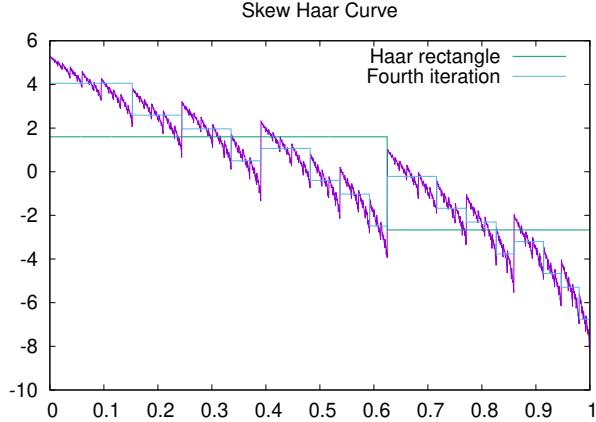
$$\text{har}_{\beta;w}(x) = \frac{d}{dx}\,\text{tak}_{\beta;w}(x)$$

This formal derivative is well-defined, as the skew Takagi is smooth and piecewise-linear almost everywhere; the places where it has corners is a dense set of measure zero. That is, the derivative is defined everywhere, except on a set of measure zero, which happens to be dense in the unit interval.

Note that the Haar fractal wavelet is piece-wise constant everywhere. It is constructed from a "mother wavelet" given by

$$h_{\beta}(x) = \begin{cases} \beta & \text{for } 0 \le x < \frac{1}{\beta} \\ \frac{-\beta}{\beta-1} & \text{for } \frac{1}{\beta} \le x \le 1 \end{cases} \tag{16}$$

37

Figure 12: Skew Haar Wavelet

This figure shows the derivative of the skew Takagi curve. Note that it is piece-wise constant everywhere. The mother wavelet is shown, as well as the fourth iteration. The specific values graphed are $\beta = 1.6$ and $w = 0.7$.

which is then iterated on to form the fractal curve $\mathrm{har}_{\beta;w}(x)$. The self symmetries are

$$\mathrm{har}_{\beta;w}\left(\frac{x}{\beta}\right) = \beta + w\,\mathrm{har}_{\beta;w}(x)$$

and

$$\mathrm{har}_{\beta;w}\left(\frac{1}{\beta} + x\left(1 - \frac{1}{\beta}\right)\right) = -\frac{\beta}{\beta-1} + w\,\mathrm{har}_{\beta;w}(x)$$

2.6   Fixed Points; Periodic Orbits

The Bernoulli shift, given by eqn 2, generates every possible bit-sequence. As was observed in a previous section, not every possible bit-sequence occurs in the beta shift. The longest sequence of all-ones possible was given by eqn 14. Arbitrary finite lengths of zeros do appear; but are there fixed points, i.e. sequences that terminate in all-zeros? Clearly, $x = 1/2\beta^n$ is such a fixed point: after $n+1$ iterations of eqn 4, x goes to zero, and stays there. Is this the only such fixed point? The answer depends on β. If β can be written in the form of $\beta^n = 2m+1$ for some integers n and m, then the values of x which can iterate down to zero in $n+1$ steps are dense in the interval $[0, \beta/2]$. Curiously, such values β are dense in the interval $[1, 2)$. A later chapter performs explores periodic orbits in great detail.

38

3   Transfer operators

Given any map from a space to itself, mapping points to points, the pushforward maps distributions to distributions. The pushforward is a linear operator, called the transfer operator or the Ruelle–Frobenius–Perron operator. The spectrum of this operator, broken down into eigenfunctions and eigenvalues, can be used to understand the time evolution of a given density distribution. The invariant measure is an eigenstate of this operator, it is the eigenstate with eigenvalue one. There are other eigenstates; these are explored in this section.

Restricting to the unit interval, given an iterated map $f : [0, 1] \to [0, 1]$, the transfer operator acting on a distribution $\rho : [0, 1] \to \mathbb{R}$ is defined as

$$\left[\mathscr{L}_f \rho \right] (y) = \sum_{x = f^{-1}(y)} \frac{\rho(x)}{|f'(x)|}$$

The next subsection gives an explicit expression for this, when f is the β-transform. After that, a subsection reviewing the invariant measure, and then a discussion of some other eigenfunctions.

3.1   Motivation

There are two broad approaches for studying iterated functions. One is to examine the point dynamics and orbits: "where does the point x go, when iterated?". The other is in terms of distributions: "how does a scattered dust of points evolve over time?". Within the context of physics, these give two broad views of reality. The first is of microscopic, time-reversible systems whose future is deterministic and known with infinite precision. The second is of macroscopic, time-irreversible thermodynamics, where time can only go forward, and the future is unknown and unknowable. Of these two approaches, the first is commonplace and inescapable; the second remains obscure, poorly-recognized and opaque. Thus, a large part of this text is devoted to this second approach.

If iterating a map $x \mapsto f(x) \mapsto f(f(x)) \mapsto \cdots$ pulls a point x through f, through time, then the action of the map f on a distribution ρ is a pushforward:

$$\rho(A) \mapsto \rho\left(f^{-1}(A)\right) \mapsto \rho\left(f^{-1}\left(f^{-1}(A)\right)\right) \mapsto \cdots$$

The proper definition of a pushforward requires a significant development of the concepts of measurable spaces and Borel sigma algebras, topics that will be gently reviewed a bit further in this introductory section. For the present, it is enough to take $\rho : [0, 1] \to \mathbb{R}$ to be some function defined on the unit interval. In the above, $A \subset [0, 1]$ is a subset of the unit interval, so that $\rho(A) = \int_A \rho(x) \, dx$ is an ordinary integral, the "size" of the set A with respect to the distribution.

The challenge is to find an explicit expression for the pushforward $\rho\left(f^{-1}(A)\right)$. This can be obtained as a change of variable $y = f(x)$ under integration. Start with any function h; it's integral over the set A is as above: $h(A) = \int_A h(y) \, dy$. Under the change of variable, this becomes

$$h(A) = \int_A h(y)\,dy = \int_{f^{-1}(A)} h(f(x))\left|f'(x)\right|dx$$

Writing the integrand as $\rho(x) = h(f(x))\left|f'(x)\right|$ and working backwards, one recovers

$$h(y) = \frac{\rho\left(f^{-1}(y)\right)}{\left|f'\left(f^{-1}(y)\right)\right|}$$

Plugging this back through gives the identity

$$\int_A \frac{\rho\left(f^{-1}(y)\right)}{\left|f'\left(f^{-1}(y)\right)\right|}\,dy = \int_{f^{-1}(A)} \rho(x)\,dx = \rho\left(f^{-1}(A)\right)$$

The right-hand-side is the desired pushforward; the left-hand side is an explicit expression for it. There was a minor sleight-of-hand in the above derivation: the map $y = f(x)$ may not be one-to-one. Thus, there may be several distinct points $x = f^{-1}(y)$. In this case, the above needs to be amended as

$$h(y) = \sum_{x \in f^{-1}(y)} \frac{\rho(x)}{\left|f'(x)\right|}$$

As h depends only on ρ and f, the sum construction on the right-hand side can be thought of as an operation $\mathscr{L}$, defined by f, acting on ρ; in short-hand, $h = \mathscr{L}_f \rho$.

The symbol $\mathscr{L}$ is used to remind that this is a linear operator: $\mathscr{L}(a\rho + b\sigma) = a\mathscr{L}\rho + b\mathscr{L}\sigma$ for any pair of real numbers a, b and any functions ρ, σ. The pushforward sequence now becomes

$$\rho(A) \mapsto \left[\mathscr{L}_f \rho\right](A) \mapsto \left[\mathscr{L}_f \mathscr{L}_f \rho\right](A) \mapsto \cdots$$

Thus, we've defined a linear operator $\mathscr{L}_f$ that depends only on the iterated function f, and has the property of mapping distributions to other distributions as it is iterated. It is the result of commuting with function composition: $\mathscr{L}_f \circ \rho = \rho \circ f^{-1}$; it's a kind of a trick with function composition. Indeed, one can define an analogous operator $\mathscr{K}_f \circ \rho = \rho \circ f$, the "composition operator" or "Koopman operator", that acts as a kind of (one-sided) inverse to $\mathscr{L}_f$.

Formally, the pushforward $\mathscr{L}_f$ is called the "transfer operator" or the "Ruelle–Frobenius–Perron operator". As a linear operator, the full force of operator theory comes into play. The primary task is to describe it's spectrum (it's eigenfunctions and eigenvalues). Two aspects of this spectrum are interesting. The first is the so-called "invariant measure", the distribution $\mu : [0, 1] \to \mathbb{R}$ that defines a density on the unit interval that is invariant under the application of the pushforward: $\mathscr{L}_f \mu = \mu$. An informal example of such an invariant measure are the rings of Saturn: an accumulation of dust and gravel, orbiting Saturn, coupled by gravitation to both Saturn and orbiting moons, yet in a stable dynamical distribution. This is the physical meaning and importance of the invariant measure; more generally, it appears as the "ground state" or "thermodynamic equilibrium state" in a vast variety of dynamical systems.

Aside from the invariant measure, there is also the question of the rest of the spectrum. These are described by the eigenfunctions ρ satisfying $\mathscr{L}\rho = \lambda\rho$. By the theorem of Frobenius–Perron, all these other solutions have eigenvalue $|\lambda| < 1$. In physics, these correspond to the decaying modes, to the distributions that disappear over time. For the example of Saturn, these are anything not orbiting in the plane of the rings: tidal forces and perturbations from the moons will force such orbits either into the ring, or crash into the planet, or possibly fly away to infinity. The other orbits are not stable. Thus, a characterization of the decaying spectrum is of general interest.

A much stronger conception is that the decaying spectrum has something to do with the irreversibility of time. In the macroscopic world, this is plainly obvious. In the microscopic world, the laws of physics are manifestly time-reversible. Somehow, complex dynamical systems pass through a region of chaotic and turbulent motion, culminating in thermodynamic equilibrium. The decaying spectrum provides a conceptual framework in which one can ponder this transition.

This is where the fun begins. The spectrum is not a "fixed thing", but depends strongly on the space of functions on which $\mathscr{L}_f$ is allowed to act. If one limits oneself to ρ drawn from the space of piece-wise continuous and smooth functions, *i.e.* polynomials, then $\mathscr{L}_f$ will in general have a discrete spectrum. If instead, $\rho \in L^2[0,1]$ the space of functions that are square-integrable on the unit interval, then the spectrum will often be continuous, and perhaps may have a large kernel. Larger spaces exhibit even wilder behavior: if one asks only that ρ be L^1-integrable (not square-integrable), then it is possible for continuous-nowhere functions to appear as eigenfunctions of $\mathscr{L}_f$. An explicit example of the latter is the Minkowski measure for the transfer operator of the Minkowski Question Mark function: it vanishes on the rationals, but can be integrated just fine; it's integral is the Question Mark function. In short, a rich variety can often be found. In the present case, it seems, nothing quite this rich, but getting there.

3.2 The β-shift Transfer Operator

This text works primarily with the β-shift, instead of the more common β-transform. These two are more-or-less the same thing, differing only by scale factors, as given in eqn. 9. The transfer operators are likewise only superficially different, being just rescalings of one-another; both are given below.

The transfer operator the beta-shift map $T_\beta(x)$ is

$$[\mathscr{L}_\beta f](y) = \begin{cases} \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\right] & \text{for } 0 \le y \le \beta/2 \\ 0 & \text{for } \beta/2 < y \le 1 \end{cases}$$

or, written more compactly

$$[\mathscr{L}_\beta f](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\right]\Theta\left(\frac{\beta}{2} - y\right) \tag{17}$$

where Θ is the Heaviside step function. The transfer operator for the beta-transform map $t_\beta(x)$ is

$$[\mathscr{M}_\beta f](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{\beta}\right)\Theta(\beta - 1 - y)\right]$$

The density distributions graphed in figure 1 are those functions satisfying

$$\left[\mathscr{L}_\beta \mu\right](y) = \mu(y) \tag{18}$$

That is, the $\mu(y)$ satisfies

$$\mu(y) = \frac{1}{\beta}\left[\mu\left(\frac{y}{\beta}\right) + \mu\left(\frac{y}{\beta} + \frac{1}{2}\right)\right]\Theta\left(\frac{\beta}{2} - y\right) \tag{19}$$

Likewise, the Gelfond-Parry measure of eqn 10 satisfies

$$\left[\mathscr{M}_\beta \nu_\beta\right](y) = \nu_\beta(y)$$

Recall that $\mu(y) = \frac{2}{\beta}\nu_\beta\left(\frac{2y}{\beta}\right)\Theta\left(\frac{\beta}{2} - y\right)$; the two invariant measures are just scaled copies of one-another. Both are normalized so that $\int_0^1 \mu(y)\,dy = \int_0^1 \nu_\beta(y)\,dy = 1$.

Both of these invariant measures are the Ruelle-Frobenius-Perron (RFP) eigenfunctions of the corresponding operators, as they correspond to the largest eigenvalues of the transfer operators, in each case being the eigenvalue one.

More generally, one is interested in characterizing the spectrum

$$\left[\mathscr{L}_\beta \rho\right](y) = \lambda \rho(y)$$

for eigenvalues $|\lambda| \leq 1$ and eigenfunctions $\rho(y)$. Solving this equation requires choosing a space of functions in which to work. Natural choices include piece-wise continuous smooth functions (piece-wise polynomial functions), any of the Banach spaces, and in particular, the space of square-integrable functions. In general, the spectrum will be complex-valued: eigenvalues will be complex numbers.

If a distribution $\rho(y)$ is nonzero on the interval $[\beta/2, 1]$, the operator $\mathscr{L}_\beta$ will map it to one that is zero on this interval. Thus, it makes sense to restrict oneself to densities that are non-zero only on $[0, \beta/2]$. When this is done, eqn 17 has the slightly more convenient form

$$\left[\mathscr{L}_\beta f\right](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right)\Theta(m_0 - y) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta(m_1 - y)\right]$$

with $m_0 = \beta/2$ and $m_1 = \beta(\beta - 1)/2 = T_\beta(m_0)$. It is always the case that $m_1 < m_0$ and so the second term above vanished on the interval $[m_1, m_0]$. This can be gainfully employed in a variety of settings; typically to write $\mathscr{L}_\beta f$ on $[m_1, m_0]$ as a simple rescaling of $\mathscr{L}_\beta f$ on $[0, m_1]$.

This equation can be treated as a recurrence relation; setting $\mathscr{L}_\beta f = f$ gives the $\lambda = 1$ eigenstate. Performing this recursion gives exactly the densities shown in figure 1. Computationally, these are much cheaper to compute than trying to track a scattered cloud of points; the result is free of stochastic sampling noise. This density is the Ruelle–Frobenius–Perron eigenstate; an explicit expression was given by Gelfond and by Parry, as described in the next section.

3.3 (No) Real-analytic solutions

What happens if we start with the assumption that there might be a real-analytic solution to the eigenequation? It is reasonable to expect that this hope is doomed, but the process is instructive anyway. Start by positing

$$f(y) = \sum_{n=0}^{\infty} a_n y^n$$

Ignoring, for a moment, the step function, and plugging this into the above gives

$$\frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\right] = \sum_{m=0}^{\infty} \frac{y^m}{\beta^{m+1}}\left[a_m + \sum_{j=0}^{\infty}\binom{m+j}{m}2^{-j}a_{m+j}\right]$$

Matching, term by term, this gives the eigenequation $\lambda a_m = \sum_k L_{mk}a_k$ with L_{mk} an upper-triangular matrix, having matrix entries

$$L_{mk} = \frac{1}{\beta^{m+1}}\left[1 + \frac{1}{2^{k-m}}\binom{k}{m}\right]$$

for $k \geq m$ and $L_{mk} = 0$ otherwise. Although this is upper-triangular, and thus solvable, the omission of the step function was critical: the eigenfunction has to vanish for $y > \beta/2$ and there is no way to do this with a real-analytic function. Inserting a step function into the mix ruins the term-by-term comparison. That is, positing

$$f(y) = \Theta\left(\frac{\beta}{2} - y\right)\sum_{n=0}^{\infty} a_n y^n$$

prevents the term-by-term comparison from passing through.

Perhaps working with the beta-shift here is a bad idea. Let's try again with the beta-transform. Without ignoring the step function, and plugging through as before gives:

$$\frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{\beta}\right)\Theta(\beta - 1 - y)\right] = \sum_{m=0}^{\infty}\frac{y^m}{\beta^{m+1}}\left[a_m + \sum_{j=0}^{\infty}\binom{m+j}{m}\beta^{-j}a_{m+j}\Theta(\beta - 1 - y)\right]$$

The step function prevents term-by-term comparison. If it is conveniently ignored or forgotten, then a pseudo-solution is easy: the term-by-term solution gives the eigenequation $\lambda a_m = \sum_k M_{mk}a_k$ with M_{mk} an upper-triangular matrix, having matrix entries

$$M_{mk} = \frac{1}{\beta^{m+1}}\left[1 + \frac{1}{\beta^{k-m}}\binom{k}{m}\right]$$

for $k \geq m$ and $M_{mk} = 0$ otherwise. This matrix is solvable. Its solutions are effectively the Bernoulli polynomials. By ignoring the step function, all that has happened is that we've rediscovered the real-analytic solutions to the Bernoulli process; indeed, the term-by-term comparison does work for the special case of $\beta = 2$, when the step function really does disappear.

3.4 Almost-solutions

If one ignores the Heaviside step function in the definition 17, one easily finds a number of "almost solutions" to the transfer operator. These are most easily discussed by defining the operator

$$\left[\mathscr{P}_\beta f\right](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\right]$$

Solving this operator is relatively straight-forward. Consider, for example, the monomial $f(y) = y^n$. Clearly, $\left[\mathscr{P}_\beta y^n\right]$ is a polynomial of degree n and that therefore, the space of polynomials is closed under the action of $\mathscr{P}_\beta$. But this result is even stronger: the monomials provide a basis in which $\mathscr{P}_\beta$ is upper-triangular, *i.e.* solvable. It's eigensolutions in this basis are polynomials. The eigenspectrum is clearly discrete, and is given by $(\beta)^{-n-1}$ for integers n corresponding to the degree of the polynomial eigensolution.

This all goes horribly wrong if one instead considers $\mathscr{L}_\beta$ and the almost-monomials $f(y) = y^n \Theta\left(\frac{\beta}{2} - y\right)$. This does not provide a basis that is closed under the action of $\mathscr{L}_\beta$. Attempting to find the closure by iterating on $\mathscr{L}_\beta$ generates a splatter of step functions. This case is examined more closely in the next chapter.

Attempting some guess-work, the self-similarity of the cpr function suggests an opening. Specifically, let $\mathrm{ei}_\beta(x) = \mathrm{cpr}_\beta(x) - 1/2$. The one finds that

$$\begin{aligned}
\left[\mathscr{P}_\beta \mathrm{ei}_\beta\right](y) &= \frac{1}{\beta}\left[\mathrm{ei}_\beta\left(\frac{y}{\beta}\right) + \mathrm{ei}_\beta\left(\frac{y}{\beta} + \frac{1}{2}\right)\right] \\
&= \frac{\mathrm{ei}_\beta(y)}{\beta}
\end{aligned}$$

This is a non-polynomial, fractal eigenfunction of $\mathscr{P}_\beta$, and, with a bit of elbow-grease, one can find many more. This includes the Takagi functions, and their higher-order analogs, which are, roughly speaking, Takagi functions constructed from polynomials. These all have interesting self-similarity properties under the dyadic monoid.

Unfortunately, one has that $\mathrm{ei}_\beta(x) \neq 0$ when $\beta < 2x$; it won't do as an eigenfunction of $\mathscr{L}_\beta$. There is no obvious, simple modification of $\mathrm{ei}_\beta(x)$ that would cause it to be a valid eigensolution of $\mathscr{L}_\beta$. Manually installing a factor of $\Theta\left(\frac{\beta}{2} - y\right)$ and then iterating to find the closure leads to the same splatter of step functions as in the case of the polynomials.

Another interesting case arises if one attempts a Fourier-inspired basis. Define

$$\mathrm{e}_{\beta;n;l}(x) = \exp i2\pi(2l+1)\beta^n x$$

for integer l. One then obtains a shift sequence

$$\left[\mathscr{P}_\beta \mathrm{e}_{\beta;n;l}\right](x) = \frac{1}{\beta}\mathrm{e}_{\beta;n-1;l}(x)\left(1 + \mathrm{e}_{\beta;n;l}\left(\frac{1}{2}\right)\right)$$

This is not a viable candidate for $\mathscr{L}_\beta$, as it is again beset by the step function. As a shift sequence, it can be used to construct coherent states that are eigenfunctions of $\mathscr{P}_\beta$,

having any eigenvalue within the unit disk. Specifically, observe that $e_{\beta;0;l}(1/2) = \exp i\pi(2l+1) = -1$ so that $\left[\mathscr{P}_{\beta}e_{\beta;0;l}\right](x) = 0$ and so the shift sequence terminates after finite iteration. Given a complex value z, construct the coherent state as

$$\phi_{l;z}(x) = \sum_{n=0}^{\infty} z^n e_{\beta;n;l}(x)$$

The shift is then

$$\left[\mathscr{P}_{\beta}\phi_{l;z}\right](x) = \frac{z}{\beta} \sum_{n=0}^{\infty} z^n \left(1 + e_{\beta;n+1;l}\left(\frac{1}{2}\right)\right) e_{\beta;n;l}(x)$$

This is not particularly useful, until one notices that for for certain values of β, this contains nilpotent sub-series.

Specifically, fix a value of $n = N$ and consider those values of β for which $e_{\beta;N;l}(1/2) = -1$. This holds whenever β^N is an odd integer, that is, whenever $\beta = (2m+1)^{1/N}$ (and, as always, $\beta \leq 2$). For these special values of β, one has that $\left[\mathscr{P}_{\beta}e_{\beta;N;l}\right](x) = 0$ and so the functions

$$\phi_{l;z;N}(x) = \sum_{n=0}^{N} z^n e_{\beta;n;l}(x)$$

vanish after N iterations of $\mathscr{P}_{\beta}$. That is, these can be used to form a a basis in which $\mathscr{P}_{\beta}$ is nilpotent. Conversely, letting m and N be free, the values for which $\beta = (2m+1)^{1/N}$ are dense in the interval $[1,2)$ and so any β is arbitrarily close to one with a nilpotent function space. These values of β are exactly the same values for which the bit sequences given by eqn 5 eventually terminate in all zeros; i.e. become periodic fixed points with period 1.

The existence of a dense set of fixed points is dual to the the existence of nilpotent densities. That is, one "causes" or "forces" the other to happen. This idea should be further elaborated, as it establishes a duality between countable and uncountable sets, which has an element of curiosity to it.

Presumably, there are special values of β which allow a periodic orbits to originate from a dense set. Such values of β, and such periodic orbits, should then correspond to specific self-similarities of the $\phi_{l;z}(x)$ function, specifically manifesting as cyclic behavior in $\left(1 + e_{\beta;n+1;l}\left(\frac{1}{2}\right)\right)^p$ for some period p. Whether there is some similar manifestation for $\mathscr{L}_{\beta}$ is wholly unclear; however, the examination of the periodic orbits of the beta shift, undertaken in a later chapter, will provide a strong clue.

3.5 Complex eigenvalues

Since the operator $\mathscr{L}_{\beta}$ is purely real, then if it has a complex spectrum, the eigenvalues and eigenfunctions must come in complex-conjugate pairs. This can make numerical searches and numerical convergence behave in unexpected ways, so some brief commentary is in order.

Assume that there exists some complex-valued eigenfunction $\rho_{\lambda}(w)$ for fixed, complex eigenvalue λ. Write it's real and complex components as

$$\rho_{\lambda}(w) = \sigma(w) + i\tau(w)$$

while also writing $\lambda = a + ib$. Then

$$
\begin{aligned}
\left[\mathscr{L}_\beta \sigma \right](w) &= \frac{1}{2} \left[\mathscr{L}_\beta \left(\rho_\lambda + \bar{\rho}_\lambda \right) \right](w) \\
&= \frac{1}{2} \left(\lambda \rho_\lambda(w) + \bar{\lambda} \bar{\rho}_\lambda(w) \right) \\
&= a\sigma(w) - b\tau(w)
\end{aligned}
$$

Both left and right sides of the above are real. If one had somehow stumbled upon $\sigma(w)$ numerically, as an eigenvector-candidate, then the above admixing of the imaginary component would quickly throw one off the hunt. Thus, a numeric search for complex-valued eigenfunctions must necessarily take into account eigenfunction pairs, with real and imaginary components that mix together as above.

Consistency requires that

$$
\begin{aligned}
\left[\mathscr{L}_\beta \tau \right](w) &= \frac{1}{2i} \left[\mathscr{L}_\beta \left(\rho_\lambda - \bar{\rho}_\lambda \right) \right](w) \\
&= \frac{1}{2i} \left(\lambda \rho_\lambda(w) - \bar{\lambda} \bar{\rho}_\lambda(w) \right) \\
&= a\tau(w) + b\sigma(w)
\end{aligned}
$$

3.6   The Gelfond–Parry measure

An explicit expression for the solution to $\mathscr{M}_\beta \nu = \nu$ was given by Gelfond[5] and by Parry[3]. It is the expression given by eqn 10. Unfortunately, I find the Russian original of Gelfond's article unreadable, and Parry's work, stemming from his PhD thesis, is not available online. Therefore, it is of some interest to provide a proof suitable for the current text. A generalization of this proof, stated in terms of a Borel algebra, is used in the subsequent section to construct general eigenfunctions.

There are two routes: either a direct verification that eqn 10 is correct, or a derivation of eqn 10 from geometric intuition. The direct verification is useful for practical purposes; the geometric construction, as a stretch-cut-stack map, provides insight. Both are given.

The Gelfond–Parry measure includes a normalization factor. It will be of recurring interest, and so a graph f it is presented in figure 13.
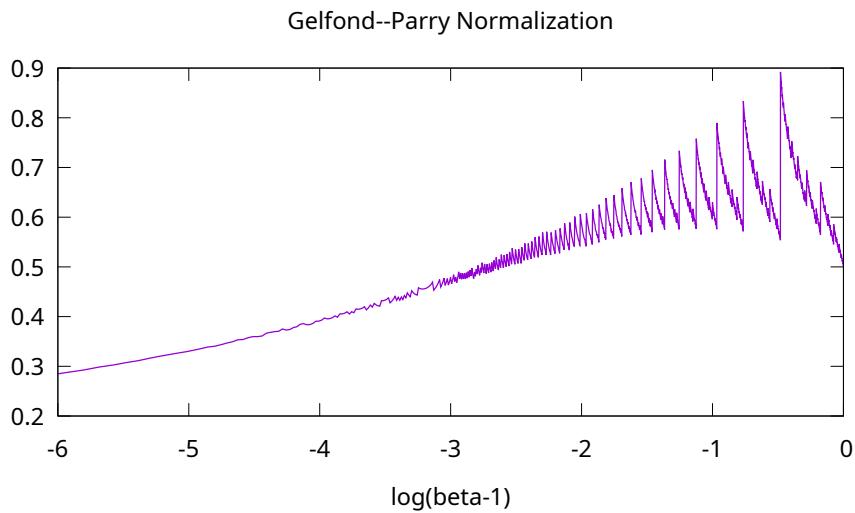
3.6.1   Direct verification

A direct verification of correctness is done below, explicitly showing all steps in laborious detail. It's not at all difficult; just a bit hard on the eyes.

As before, let $t(x) \equiv t_\beta(x) = \beta x \mod 1$ be the β-transformation of eqn 8, and $t^n(x)$ the iterated transformation. Let $\Theta(x)$ be the Heaviside step function as always, and to keep notation brief, let $t_n \equiv t^n(1)$. The Gelfond–Parry measure is then

$$
\nu(y) = \frac{1}{F} \sum_{n=0}^{\infty} \frac{\Theta(t_n - y)}{\beta^n}
$$

Figure 13: Gelfond–Parry Normalization

Gelfond--Parry Normalization

The above figure shows $1/F$ for the normalization constant $F = \sum_{n=0}^{\infty} t_n \beta^{-n}$ as a function of β. The horizonal axis is stretched out using $\log(\beta - 1)$ so as to amplify the behavior as $\beta \to 1$. One has that $F \to \infty$ in this limit (so $1/F \to 0$); the curve suggests just how catastrophic that limit is. A graph of F vs. β, without the rescaling of the horizontal axis, is shown in figure 30.

where the normalization F is given by

$$F = \sum_{n=0}^{\infty} \frac{t_n}{\beta^n}$$

The transfer operator $\mathscr{M}$ for the beta-transformation is slightly more convenient to work with than $\mathscr{L}$ for this particular case. It is given by

$$[\mathscr{M}f](y) = \frac{1}{\beta} \left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right) \Theta(\beta - 1 - y) \right]$$

and we wish to verify that $\mathscr{M}v = v$. Plugging in directly,

$$P = v\left(\frac{y}{\beta}\right) + v\left(\frac{y+1}{\beta}\right) \Theta(\beta - 1 - y) =$$

$$= \frac{1}{F} \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta\left(t_n - \frac{y}{\beta}\right) + \Theta\left(t_n - \frac{y+1}{\beta}\right) \Theta(\beta - 1 - y) \right)$$

$$= \frac{1}{F} \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta(\beta t_n - y) + \Theta(\beta t_n - 1 - y) \right)$$

The second line follows from the first, since $\Theta(ax) = \Theta(x)$ for all constants a, and the $\Theta(\beta - 1 - y)$ can be safely dropped, since $t_n \leq 1$ for all n. These terms simplify, depending on whether t_n is small or large. Explicitly, one has

$$\Theta(\beta t_n - 1 - y) = 0 \text{ if } \beta t_n - 1 < 0$$
$$\Theta(\beta t_n - y) = 1 \text{ if } \beta t_n - 1 > 0$$

and so

$$P = \frac{1}{F} \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta(1 - \beta t_n) \Theta(\beta t_n - y) + \Theta(\beta t_n - 1)(1 + \Theta(\beta t_n - 1 - y)) \right)$$

These can be collapsed by noting that

$$\beta t_n = t_{n+1} \text{ if } \beta t_n - 1 < 0$$
$$\beta t_n - 1 = t_{n+1} \text{ if } \beta t_n - 1 > 0$$

and so

$$P = \frac{1}{F} \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta(1 - \beta t_n) \Theta(t_{n+1} - y) + \Theta(\beta t_n - 1)(1 + \Theta(t_{n+1} - y)) \right)$$

$$= \frac{1}{F} \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta(t_{n+1} - y)[\Theta(1 - \beta t_n) + \Theta(\beta t_n - 1)] + \Theta(\beta t_n - 1) \right)$$

$$= \frac{1}{F} \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta(t_{n+1} - y) + \Theta(\beta t_n - 1) \right)$$

$$= \beta v(y) - \frac{\beta}{F} + \frac{1}{F} \sum_{n=0}^{\infty} \frac{\Theta(\beta t_n - 1)}{\beta^n}$$

$$= \beta v(y)$$

The last sum on the right is just the β-expansion for 1. That is, the β-expansion of x is

$$x = \sum_{n=0}^{\infty} \frac{\Theta(\beta t^n(x) - 1)}{\beta^{n+1}}$$

This is just eqn 6 written in a different way (making use of the equivalence 9). Thus $P = \beta v(y)$ and so $\mathscr{M}v = v$ as claimed.

3.6.2 Stretch–Cut–Stack Map

The measure can be geometrically constructed and intuitively understood as the result of the repeated application of a bakers-map-style stretch and squash operation. The idea is to iterate $v_{n+1} = \mathscr{M}_\beta v_n$, starting with $v_0 = 1$ and then take the limit $n \to \infty$; the result is $v_n \to v$ in the limit, given as above. The basic stretch-cut-stack operation is given below, providing the intuition. The proof that it converges as desired is nearly identical to the (shorter) proof above.

Begin with a first approximation that v_0 is constant on the interval $[0,1]$, so that $v_0(y) = \Theta(t_0 - y)$ with endpoint $t_0 = 1$. The operation of $\mathscr{M}$ acting on this is to stretch it out to the interval $[0,\beta]$, chop off the $[1,\beta]$ part, move it to $[0,\beta-1]$, stack it on top, doubling the density in this region. The doubling, though, is partly counteracted by the stretching, which thins out the density to $1/\beta$ uniformly over the entire interval $[0,\beta]$. This operation preserves the grand-total measure on the unit interval. Writing $\beta - 1 = t_1 = t(t_0) = t(1)$ for the first iterate of the endpoint $t_0 = 1$, this stretch, cut and stack operation should result in

$$v_1(y) = \begin{cases} \frac{2}{\beta} & \text{for } y \in [0, t_1] \\ \frac{1}{\beta} & \text{for } y \in [t_1, t_0] \end{cases}$$

The stretch–cut–stack operation is the intuitive, geometrical explanation for obtaining the first iterate. The same result is obtained algebraically, by writing $v_1 = \mathscr{M}v_0$ and then plugging and chugging:

$$v_1(y) = \frac{1}{\beta}\left[\Theta(t_0 - y) + \Theta(t_1 - y)\right]$$

$$= \frac{v_0(y)}{\beta} + \frac{\Theta(t_1 - y)}{\beta}$$

A recursive formula for v_n is needed. Rearrange the above as

$$v_1(y) = \frac{1}{\beta}\left[1 + \Theta(t_1 - y)\right]$$

$$= \Theta(t_0 - y) + \frac{\Theta(t_1 - y)}{\beta} - \left(1 - \frac{1}{\beta}\right)$$

$$= \sum_{k=0}^{1} \frac{\Theta(t_k - y)}{\beta^k} - c_1 v_0$$

$$= \rho_1 - c_1 v_0$$

49

The appropriate general form will use ρ_n defined as

$$\rho_n = \sum_{k=0}^{n} \frac{1}{\beta^k} \Theta(t_k - y)$$

and c_n defined as

$$c_n = 1 - \sum_{k=0}^{n-1} \frac{\Theta(\beta t_k - 1)}{\beta^{k+1}} \tag{20}$$

It will be shown that the recursive form is

$$v_n = \rho_n - \sum_{k=0}^{n-1} c_{n-k} v_k$$

Iteration is performed directly, so that $v_{n+1} = \mathcal{M} v_n$.

It is convenient to isolate the action of $\mathcal{M}$ on ρ_n. Plugging through,

$$\beta \mathcal{M} \rho_n = P_n = \rho_n\left(\frac{y}{\beta}\right) + \rho_n\left(\frac{y+1}{\beta}\right) \Theta(t_1 - y)$$

$$= \sum_{k=0}^{n} \frac{1}{\beta^k} \left(\Theta\left(t_k - \frac{y}{\beta}\right) + \Theta\left(t_k - \frac{y+1}{\beta}\right) \Theta(t_1 - y) \right)$$

$$= \sum_{k=0}^{n} \frac{1}{\beta^k} \left(\Theta(\beta t_k - y) + \Theta(\beta t_k - 1 - y) \right)$$

The $\Theta(t_1 - y)$ can be safely dropped, because $t_k \leq 1$ for all k. That is, $\Theta(t_1 - y) = \Theta(\beta - 1 - y)$ and since $\beta t_k - 1 - y \leq \beta - 1 - y$ the extra factor of $\Theta(t_1 - y)$ has no effect.

These terms simplify, depending on whether t_k is small or large. Explicitly, one has

$$\Theta(\beta t_k - 1 - y) = 0 \text{ if } \beta t_k - 1 < 0$$
$$\Theta(\beta t_k - y) = 1 \text{ if } \beta t_k - 1 > 0$$

and so

$$P_n = \sum_{k=0}^{n} \frac{1}{\beta^k} \left(\Theta(1 - \beta t_k) \Theta(\beta t_k - y) + \Theta(\beta t_k - 1)(1 + \Theta(\beta t_k - 1 - y)) \right)$$

These can be collapsed by noting that

$$\beta t_k = t_{k+1} \text{ if } \beta t_k - 1 < 0$$
$$\beta t_k - 1 = t_{k+1} \text{ if } \beta t_k - 1 > 0$$

and so

$$P_n = \sum_{k=0}^{n} \frac{1}{\beta^k} \left(\Theta(1 - \beta t_k) \Theta(t_{k+1} - y) + \Theta(\beta t_k - 1)(1 + \Theta(t_{k+1} - y)) \right)$$

$$= \sum_{k=0}^{n} \frac{1}{\beta^k} \left(\Theta(t_{k+1} - y) \left[\Theta(1 - \beta t_k) + \Theta(\beta t_k - 1) \right] + \Theta(\beta t_k - 1) \right)$$

$$= \sum_{k=0}^{n} \frac{1}{\beta^k} \left(\Theta(t_{k+1} - y) + \Theta(\beta t_k - 1) \right)$$

$$= \beta \sum_{k=0}^{n+1} \frac{1}{\beta^k} \Theta(t_k - y) - \beta + \beta \sum_{k=0}^{n} \frac{\Theta(\beta t_k - 1)}{\beta^{k+1}}$$

$$= \beta \rho_{n+1} - \beta \left(1 - \sum_{k=0}^{n} \frac{\Theta(\beta t_k - 1)}{\beta^{k+1}} \right)$$

$$= \beta \left(\rho_{n+1} - c_{n+1} \right)$$

Returning to the Ansatz for v_n in terms of ρ_n and plugging through,

$$v_{n+1} = \mathcal{M} v_n = \frac{1}{\beta} P_n - \mathcal{M} \sum_{k=0}^{n-1} c_{n-k} v_k$$

$$= \rho_{n+1} - c_{n+1} - \sum_{k=0}^{n-1} c_{n-k} v_{k+1}$$

$$= \rho_{n+1} - \sum_{k=0}^{n} c_{n+1-k} v_k$$

and so the hypothesized recursive form is preserved.

The measure is preserved with each iteration, by construction. That is, $\int_0^1 v_n(y)\, dy = 1$ for all n. Verifying this:

$$1 = \int_0^1 v_n(y)\, dy = \sum_{k=0}^{n} \frac{1}{\beta^k} \int_0^1 \Theta(t_k - y)\, dy - \sum_{k=0}^{n-1} c_{n-k} \int_0^1 v_k$$

$$= \sum_{k=0}^{n} \frac{t_k}{\beta^k} - \sum_{k=1}^{n} c_k$$

This identity can be verified by plugging through, exchanging sums and then comparing terms; the underlying identity is

$$t_n = \beta^n c_n$$

which follows, as each $\Theta(\beta t_k - 1)$ records a decision to decrement, or not, the product βt_k occurring in the iteration $t_{k+1} = \beta t_k \mod 1$.

A curious identity arises in the verification of the above. After rearranging terms and swapping the summation order, one gets

$$n+1 = \sum_{k=0}^{n} \beta^{-k} \left(t_k + (n-k+1) \Theta \left(\beta t_{k-1} - 1 \right) \right)$$

where, to keep the sum well-defined, set $t_{-1} = 0$.

The sum in c_n is a partial β-expansion for 1. Each term can be recognized as a bit from the bit-expansion:

$$\Theta \left(\beta t_k - 1 \right) = d_n \left(\frac{1}{2} \right) = k_n \left(\frac{\beta}{2} \right) = \varepsilon_n \left(\frac{1}{\beta} \right)$$

with k_n as defined in eqn 5, ε_n defined as in eqn 11 and d_n as in eqn 23 (apologies for the variety of notations; each is "natural" in a specific context.) The β-expansion of a real number x is as given in eqn 6. In the present context, this is the expansion for $x = \beta/2$, or, after rescaling, the expansion for $x = 1$:

$$1 = \sum_{k=0}^{\infty} \frac{\Theta \left(\beta t_k - 1 \right)}{\beta^{k+1}}$$

Thus, $|c_n| < \beta^{-n+1}$ and $c_n \to 0$ as $n \to \infty$.

The Gelfond–Parry normalization is the $n \to \infty$ limit

$$F = \sum_{k=0}^{\infty} \frac{t_k}{\beta^k}$$

The sequence of v_n constructed above give a density that is the result of n repeated stretch-cut-stack operations. They preserve the total density. The $n \to \infty$ limit is the invariant measure v.

Only the first few steps were overtly geometric; the recursive step required a bit of algebraic grinding. The piece that is being cut and stacked is $\Theta \left(\beta t_k - 1 - y \right)$ but it gets a bit lost in the shuffle. Its hard to see the pancake-stacking that is happening in the algebra, at least, not without drawing careful pictures. The stretch operation moves many of the βt_k past 1, where they are restacked, while many other βt_k do not cross past the cut at 1; these are the two possibilities. Drawing a pancake diagram clarifies the ultimately geometric aspect of this recursive operation.

3.6.3 Stacking generic functions

The above proof can be extended to other shapes. For example, consider iterating $v_0(y) = y - 1/2$. Based on experience with the Bernoulli shift, one might expect this to generate an eigenfunction with eigenvalue $1/\beta$. Similarly, the Bernoulli polynomials $v_0(y) = B_n(y) = y^n + \mathscr{O}\left(y^{n-1} \right)$ might be expected to generate eigenfunctions with eigenvalue $1/\beta^n$. Thus, reproducing the above proof, with a general initial distribution v_0 promises generalized results.

The rest of this section is filled with false starts, motivating arguments, repeated calculations, double and triple checks. To get the final result, while skipping the detailed proof/derivation, hop to the next subsection, title "Stacking generic functions – conclusion".

To discover the general form, perform the first few iterations by blunt force. Mirror the earlier proof as closely as possible; it provides a working template. Start with

$$v_1(y) = \frac{1}{\beta}\left[v_0\left(\frac{y}{\beta}\right) + v_0\left(\frac{y}{\beta} + \frac{1}{\beta}\right)\Theta(t_1 - y)\right]$$

$$= v_0(y)\Theta(t_0 - y) + \frac{1}{\beta}v_0\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y) - \left[v_0(y) - \frac{1}{\beta}v_0\left(\frac{y}{\beta}\right)\right]$$

$$= \rho_1 - c_1$$

where

$$\rho_1 = v_0(y) + \frac{1}{\beta}v_0\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

and

$$c_1 = v_0(y) - \frac{1}{\beta}v_0\left(\frac{y}{\beta}\right)$$

As before, contemplate the application of $\mathcal{M}$ to each of these pieces, individually. Examining ρ, the critical structure is the step function, which is the primary object to track. Separate the step function into it's own factor, so that it can be explicitly tracked and manipulated. Thus, define f and g such that

$$\rho_1 = f(y) + g(y)\Theta(t_1 - y)$$

Then

$$\beta\mathcal{M}\rho_1 = f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

$$+ g\left(\frac{y}{\beta}\right)\Theta(\beta t_1 - y) + g\left(\frac{y+1}{\beta}\right)\Theta(\beta t_1 - y - 1)\Theta(t_1 - y)$$

The g terms are susceptible to

$$\Theta(\beta t_k - 1 - y) = 0 \text{ if } \beta t_k - 1 < 0$$
$$\Theta(\beta t_k - y) = 1 \text{ if } \beta t_k - 1 > 0$$

Together with the identity

$$1 = \Theta(\beta t_1 - 1) + \Theta(1 - \beta t_1)$$

53

This allows

$$\beta \mathscr{M} \rho_1 = f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

$$+ g\left(\frac{y}{\beta}\right)\Theta(\beta t_1 - y)\left[\Theta(\beta t_1 - 1) + \Theta(1 - \beta t_1)\right]$$

$$+ g\left(\frac{y+1}{\beta}\right)\Theta(\beta t_1 - y - 1)\Theta(t_1 - y)\left[\Theta(\beta t_1 - 1) + \Theta(1 - \beta t_1)\right]$$

$$= f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

$$+ g\left(\frac{y}{\beta}\right)\left[\Theta(\beta t_1 - 1) + \Theta(\beta t_1 - y)\Theta(1 - \beta t_1)\right]$$

$$+ g\left(\frac{y+1}{\beta}\right)\Theta(\beta t_1 - y - 1)\Theta(\beta t_1 - 1)$$

As before, make use of

$$\beta t_k = t_{k+1} \text{ if } \beta t_k - 1 < 0$$
$$\beta t_k - 1 = t_{k+1} \text{ if } \beta t_k - 1 > 0$$

to write

$$\beta \mathscr{M} \rho_1 = f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

$$+ g\left(\frac{y}{\beta}\right)\left[\Theta(\beta t_1 - 1) + \Theta(t_2 - y)\Theta(1 - \beta t_1)\right]$$

$$+ g\left(\frac{y+1}{\beta}\right)\Theta(t_2 - y)\Theta(\beta t_1 - 1)$$

$$= f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

$$+ g\left(\frac{y}{\beta}\right)\Theta(\beta t_1 - 1)$$

$$+ \left[g\left(\frac{y}{\beta}\right)\Theta(1 - \beta t_1) + g\left(\frac{y+1}{\beta}\right)\Theta(\beta t_1 - 1)\right]\Theta(t_2 - y)$$

$$= f\left(\frac{y}{\beta}\right) + f\left(\frac{y+1}{\beta}\right)\Theta(t_1 - y)$$

$$+ g\left(\frac{y}{\beta}\right)\Theta(\beta t_1 - 1)$$

$$+ g\left(\frac{y + \Theta(\beta t_1 - 1)}{\beta}\right)\Theta(t_2 - y)$$

54

Rearranging,

$$\beta \mathcal{M} \rho_1 = f\left(\frac{y}{\beta}\right) + g\left(\frac{y}{\beta}\right) \Theta(\beta t_1 - 1)$$
$$+ f\left(\frac{y+1}{\beta}\right) \Theta(t_1 - y)$$
$$+ g\left(\frac{y + \Theta(\beta t_1 - 1)}{\beta}\right) \Theta(t_2 - y)$$

which clearly has parts that have a Heaviside $\Theta(t_k - y)$ and other parts that don't.
Now take a look at the action of $\mathcal{M}$ on c_1. Recall that

$$c_1 = v_0(y) - \frac{1}{\beta} v_0\left(\frac{y}{\beta}\right)$$

Then plugging through,

$$\beta \mathcal{M} c_1 = \beta v_1 - \frac{1}{\beta}\left[v_0\left(\frac{y}{\beta^2}\right) + v_0\left(\frac{y+1}{\beta^2}\right) \Theta(t_1 - y)\right]$$

which has a term with $\Theta(t_k - y)$ and the rest without it. Putting these together,

$$v_2 = \mathcal{M} v_1 = \mathcal{M} \rho_1 - \mathcal{M} c_1$$

The next goal is to repeat the iteration, recursively.

General form The above suggests that the correct Ansatz is to split the repeated iteration so that

$$v_n = \rho_n - d_n \quad \text{and} \quad v_{n+1} = \mathcal{M} v_n = \mathcal{M} \rho_n - \mathcal{M} d_n$$

with

$$\rho_n(y) = \sum_{k=0}^{n} g_{nk}(y) \Theta(t_k - y)$$

and

$$d_n(y) = \sum_{k=0}^{n-1} a_{nk} v_k(y) - \sum_{k=0}^{n-1} f_{nk}(y) \Theta(\beta t_k - 1)$$

By definition, $g_{n0} = v_0$ for all n.
For $n = 0$, consistency requires that $\rho_0 = g_{00} = v_0$ and $d_0 = 0$.
For $n = 1$, we started with

$$g_{11}(y) = \frac{1}{\beta} v_0\left(\frac{y+1}{\beta}\right)$$

and

$$a_{10} = 1 \quad \text{and} \quad f_{10}(y) = \frac{1}{\beta} v_0\left(\frac{y}{\beta}\right)$$

After one iteration, it was deduced that

$$g_{21}(y) = \frac{1}{\beta} v_0 \left(\frac{y+1}{\beta} \right) + \frac{1}{\beta} f_{10} \left(\frac{y+1}{\beta} \right)$$

$$g_{22}(y) = \frac{1}{\beta} g_{11} \left(\frac{y + \Theta(\beta t_1 - 1)}{\beta} \right)$$

and

$$f_{20}(y) = \frac{1}{\beta} v_0 \left(\frac{y}{\beta} \right) + \frac{1}{\beta} f_{10} \left(\frac{y}{\beta} \right)$$

$$f_{21}(y) = \frac{1}{\beta} g_{11} \left(\frac{y}{\beta} \right)$$

and

$$a_{20} = a_{21} = 1$$

Above looks very regular and entirely manageable. Unrolling the loop obscures some of this structure, but the regularity does peak through.

$$g_{21}(y) = \frac{1}{\beta} v_0 \left(\frac{y+1}{\beta} \right) + \frac{1}{\beta^2} v_0 \left(\frac{y+1}{\beta^2} \right)$$

$$g_{22}(y) = \frac{1}{\beta^2} v_0 \left(\frac{y + \beta + \Theta(\beta t_1 - 1)}{\beta^2} \right)$$

while

$$f_{20}(y) = \frac{1}{\beta} v_0 \left(\frac{y}{\beta} \right) + \frac{1}{\beta^2} v_0 \left(\frac{y}{\beta^2} \right)$$

$$f_{21}(y) = \frac{1}{\beta^2} v_0 \left(\frac{y + \beta}{\beta^2} \right)$$

Iteration step   Armed with a hypothesis for the general form, iteration will reveal how it transforms. Starting with

$$\rho_n(y) = \sum_{k=0}^{n} g_{nk}(y) \Theta(t_k - y)$$

one gets

$$\beta \mathcal{M} \rho_n = \sum_{k=0}^{n} g_{nk} \left(\frac{y}{\beta} \right) \Theta(\beta t_k - y) + \sum_{k=0}^{n} g_{nk} \left(\frac{y+1}{\beta} \right) \Theta(\beta t_k - y - 1) \Theta(t_1 - y)$$

Making use of

$$\Theta(\beta t_k - 1 - y) = 0 \text{ if } \beta t_k - 1 < 0$$

$$\Theta(\beta t_k - y) = 1 \text{ if } \beta t_k - 1 > 0$$

together with the identity

$$1 = \Theta(\beta t_k - 1) + \Theta(1 - \beta t_k)$$

allows

$$
\begin{aligned}
\beta \mathcal{M} \rho_n &= \sum_{k=0}^{n} g_{nk}\left(\frac{y}{\beta}\right) \Theta(\beta t_k - y)\left[\Theta(\beta t_k - 1) + \Theta(1 - \beta t_k)\right] \\
&\quad + \sum_{k=0}^{n} g_{nk}\left(\frac{y+1}{\beta}\right) \Theta(\beta t_k - y - 1)\,\Theta(t_1 - y)\left[\Theta(\beta t_k - 1) + \Theta(1 - \beta t_k)\right] \\
&= \sum_{k=0}^{n} g_{nk}\left(\frac{y}{\beta}\right) \left[\Theta(\beta t_k - 1) + \Theta(\beta t_k - y)\,\Theta(1 - \beta t_k)\right] \\
&\quad + \sum_{k=0}^{n} g_{nk}\left(\frac{y+1}{\beta}\right) \Theta(\beta t_k - y - 1)\,\Theta(\beta t_k - 1)
\end{aligned}
$$

As before, make use of

$$
\begin{aligned}
\beta t_k &= t_{k+1} \text{ if } \beta t_k - 1 < 0 \\
\beta t_k - 1 &= t_{k+1} \text{ if } \beta t_k - 1 > 0
\end{aligned}
$$

to write

$$
\begin{aligned}
\beta \mathcal{M} \rho_n &= \sum_{k=0}^{n} g_{nk}\left(\frac{y}{\beta}\right) \left[\Theta(\beta t_k - 1) + \Theta(t_{k+1} - y)\,\Theta(1 - \beta t_k)\right] \\
&\quad + \sum_{k=0}^{n} g_{nk}\left(\frac{y+1}{\beta}\right) \Theta(t_{k+1} - y)\,\Theta(\beta t_k - 1) \\
&= \sum_{k=0}^{n} g_{nk}\left(\frac{y}{\beta}\right) \Theta(\beta t_k - 1) \\
&\quad + \sum_{k=0}^{n} \left[g_{nk}\left(\frac{y}{\beta}\right) \Theta(1 - \beta t_k) + g_{nk}\left(\frac{y+1}{\beta}\right) \Theta(\beta t_k - 1)\right] \Theta(t_{k+1} - y) \\
&= \sum_{k=0}^{n} g_{nk}\left(\frac{y}{\beta}\right) \Theta(\beta t_k - 1) \\
&\quad + \sum_{k=0}^{n} g_{nk}\left(\frac{y + \Theta(\beta t_k - 1)}{\beta}\right) \Theta(t_{k+1} - y) \\
&= \sum_{k=0}^{n} g_{nk}\left(\frac{y}{\beta}\right) \Theta(\beta t_k - 1) + \sum_{k=1}^{n+1} g_{n,k-1}\left(\frac{y + \Theta(\beta t_{k-1} - 1)}{\beta}\right) \Theta(t_k - y)
\end{aligned}
$$

This clearly preserves the hypothesized general form: aspects of both ρ_{n+1} and d_{n+1} are visible in the above. The remaining pieces require a look at $\mathcal{M} d_n$. Given

$$d_n(y) = \sum_{k=0}^{n-1} a_{nk} v_k(y) - \sum_{k=0}^{n-1} f_{nk}(y)\,\Theta(\beta t_k - 1)$$

57

This is

$$\beta \mathscr{M} d_n = \beta \sum_{k=1}^{n} a_{n,k-1} v_k(y)$$

$$- \sum_{k=0}^{n-1} f_{nk}\left(\frac{y}{\beta}\right) \Theta(\beta t_k - 1)$$

$$- \sum_{k=0}^{n-1} f_{nk}\left(\frac{y+1}{\beta}\right) \Theta(\beta t_k - 1)\Theta(t_1 - y)$$

We've arrived at the final form:

$$v_{n+1} = \mathscr{M} v_n = \mathscr{M} \rho_n - \mathscr{M} d_n$$

Performing term by term compares,

$$\beta g_{n+1,k} = g_{n,k-1}\left(\frac{y + \Theta(\beta t_{k-1} - 1)}{\beta}\right) \qquad \text{for } k > 1$$

$$\beta g_{n+1,1} = g_{n0}\left(\frac{y+1}{\beta}\right) + \sum_{k=0}^{n-1} f_{nk}\left(\frac{y+1}{\beta}\right) \Theta(\beta t_k - 1)$$

$$g_{n+1,0} = v_0$$

$$a_{n+1,k} = a_{n,k-1} = 1$$

$$\beta f_{n+1,k} = f_{nk}\left(\frac{y}{\beta}\right) + g_{nk}\left(\frac{y}{\beta}\right)$$

The above can be checked numerically, and it correctly reproduces the Gelfond–Parry measure, when $v_0 = 1$. The result that $a_{n,k} = 1$ is mildly unexpected, given the earlier calculations. If would appear that there is some re-summation that is possible when $v_0 = 1$ but is unavailable when working with the general case.

The above also correctly generates the decaying eigenfunctions associated with the periodic orbits, when $v_0(y) = y - 1/2$. A few of the low-ranked periodic orbits were checked. In conclusion, the above is correct.

Double-check This general result can be compared to the earlier special-case results for $n = 0, 1, 2$. They agree. They are reposted here, for verification.

For $n = 0$, $g_{00} = v_0$ and $f_{00} = 0$.

The general formula above gives the same $n = 1$ result as previously obtained by direct iteration:

$$g_{11}(y) = \frac{1}{\beta} v_0\left(\frac{y+1}{\beta}\right)$$

and

$$f_{10}(y) = \frac{1}{\beta} v_0\left(\frac{y}{\beta}\right)$$

Deduce from general formulas that, for $n = 2$, that

$$g_{22}(y) = \frac{1}{\beta} g_{11} \left(\frac{y + \Theta(\beta t_1 - 1)}{\beta} \right)$$

$$= \frac{1}{\beta^2} v_0 \left(\frac{y + \beta + \Theta(\beta t_1 - 1)}{\beta} \right)$$

$$g_{21}(y) = \frac{1}{\beta} v_0 \left(\frac{y + 1}{\beta} \right) + \frac{1}{\beta} f_{10} \left(\frac{y + 1}{\beta} \right)$$

$$= \frac{1}{\beta} v_0 \left(\frac{y + 1}{\beta} \right) + \frac{1}{\beta^2} v_0 \left(\frac{y + 1}{\beta^2} \right)$$

and

$$f_{21}(y) = \frac{1}{\beta} g_{11} \left(\frac{y}{\beta} \right)$$

$$= \frac{1}{\beta^2} v_0 \left(\frac{y + \beta}{\beta^2} \right)$$

$$f_{20}(y) = \frac{1}{\beta} f_{10} \left(\frac{y}{\beta} \right) + \frac{1}{\beta} g_{10} \left(\frac{y}{\beta} \right)$$

$$= \frac{1}{\beta} v_0 \left(\frac{y}{\beta} \right) + \frac{1}{\beta^2} v_0 \left(\frac{y}{\beta^2} \right)$$

These also agree with earlier special-case results.

Loop unroll   Given the general form for v_n given above, together with the recursive formula for g_{nk} and f_{nk}, it is desirable to unroll the sequence of recursive steps into a series summation. This is the basic goal of "analytic algorithmics": to convert recursive formulas into series summations. This is similar to the idea of "analytic combinatorics", except that each Θ occuring in the recursive equation can be thought of as an if-then statement, albeit a very primitive one: $\Theta(x)$ is equivalent to "if $x > 0$ then 1 else 0". Recursive expressions are conventional in functional programming; loops are conventional in imperative programming. Loops and recursion are dual to one-another, explicitly so when tail recursion is possible. The recursive step for g_{nk} is tail-recursive, so fairly easy to express. The recursive step for f_{nk} generates a summation.

By definition, $g_{n0} = v_0$ for all n. It seems that $a_{nk} = 1$ always, so we'll ignore that from now on. Then $f_{nn} = 0$ and generally, $f_{nk} = 0$ for $k \geq n$. Likewise, $g_{nk} = 0$ for $k > n$.

It is convenient to introduce a short-hand for the bit-sequence generated by the iterated mid-point. The shorthand is

$$b_n = \Theta(\beta t_k - 1) = d_n \left(\frac{1}{2} \right) = \varepsilon_n \left(\frac{1}{\beta} \right)$$

with ε_n defined as in eqn 11 and d_n as in eqn 23.

The recursion relations to unroll are

$$\beta g_{n,k}(y) = g_{n-1,k-1}\left(\frac{y+b_{k-1}}{\beta}\right) \qquad \text{for } k > 1$$

$$\beta g_{n,1}(y) = v_0\left(\frac{y+1}{\beta}\right) + \sum_{k=0}^{n-2} b_k f_{n-1,k}\left(\frac{y+1}{\beta}\right)$$

$$g_{n,0}(y) = v_0$$

$$\beta f_{n,k}(y) = f_{n-1,k}\left(\frac{y}{\beta}\right) + g_{n-1,k}\left(\frac{y}{\beta}\right)$$

Recursing on g_{nk} for $k > 1$ is straight-forward, as only the arguement is iterated. This gives

$$g_{nk}(y) = \frac{1}{\beta^{k-1}} g_{n-k+1,1}\left(\frac{y}{\beta^{k-1}} + \sum_{j=1}^{k-1} b_j \beta^{-j}\right)$$

Recursing on f_{nk} terminates because $f_{nk} = 0$ for $k \geq n$, leaving only g_{nk} behind. This gives

$$f_{n,k}(y) = \sum_{j=1}^{n-k} \beta^{-j} g_{n-j,k}\left(\frac{y}{\beta^j}\right)$$

Thus,

$$g_{n,1}(y) = \frac{1}{\beta} v_0\left(\frac{y+1}{\beta}\right) + \sum_{k=0}^{n-2} b_k \sum_{j=1}^{n-k-1} \beta^{-j-1} g_{n-j-1,k}\left(\frac{y+1}{\beta^{j+1}}\right)$$

$$= \frac{1}{\beta} v_0\left(\frac{y+1}{\beta}\right) + \sum_{k=0}^{n-2} b_k \sum_{j=0}^{n-k-2} \beta^{-j-2} g_{n-j-2,k}\left(\frac{y+1}{\beta^{j+2}}\right)$$

$$= \frac{1}{\beta} v_0\left(\frac{y+1}{\beta}\right) + \sum_{j=0}^{n-2} \beta^{-j-2} v_0\left(\frac{y+1}{\beta^{j+2}}\right)$$

$$+ \sum_{k=1}^{n-2} b_k \sum_{j=0}^{n-k-2} \beta^{-j-2} g_{n-j-2,k}\left(\frac{y+1}{\beta^{j+2}}\right)$$

$$= \sum_{j=0}^{n-1} \beta^{-j-1} v_0\left(\frac{y+1}{\beta^{j+1}}\right) + \sum_{k=1}^{n-2} b_k \sum_{j=0}^{n-k-2} \beta^{-j-2} g_{n-j-2,k}\left(\frac{y+1}{\beta^{j+2}}\right)$$

Exchanging the order of summation doesn't seem to do anything interesting.

$$g_{n,1}(y) = \sum_{j=0}^{n-1} \beta^{-j-1} v_0\left(\frac{y+1}{\beta^{j+1}}\right) + \sum_{j=0}^{n-2} \beta^{-j-2} \sum_{k=1}^{n-j-2} b_k g_{n-j-2,k}\left(\frac{y+1}{\beta^{j+2}}\right)$$

$$= \sum_{j=0}^{n-1} \beta^{-j-1} v_0\left(\frac{y+1}{\beta^{j+1}}\right)$$

$$+ \sum_{j=0}^{n-2} \frac{1}{\beta^{j+2}} \sum_{k=1}^{n-j-2} b_k \frac{1}{\beta^{k-1}} g_{n-k-j-1,1}\left(\frac{y+1}{\beta^{k+j+1}} + \sum_{i=1}^{k-1} b_i \beta^{-i}\right)$$

It does not seem reasonable to try to unroll this any further.

60

Verification   The above gives a more-or-less complete unrolled version of the recursion relations. It can be double-checked by setting $v_0 = 1$ and looking to see how the Gelfond–Parry measure arises.

The $g_{nk}(y)$ must be independent of y, and so

$$g_{n0} = 1$$

$$g_{n1} = \sum_{j=0}^{n-1} \frac{1}{\beta^{j+1}} + \sum_{j=0}^{n-2} \frac{1}{\beta^{j+2}} \sum_{k=1}^{n-j-2} b_k \frac{1}{\beta^{k-1}} g_{n-k-j-1,1}$$

$$g_{nk} = \frac{1}{\beta^{k-1}} g_{n-k+1,1}$$

$$= \frac{1}{\beta^k} \left[\sum_{j=0}^{n-k} \frac{1}{\beta^j} + \sum_{j=0}^{n-k-1} \frac{1}{\beta^{j+1}} \sum_{k=1}^{n-k-j-1} b_k \frac{1}{\beta^{k-1}} g_{n-2k-j,1} \right]$$

$$f_{nk} = \sum_{j=1}^{n-k} \beta^{-j} g_{n-j,k}$$

Returning to the form $v_n = \rho_n - d_n$, the above are to be inserted into

$$\rho_n(y) = \sum_{k=0}^{n} g_{nk} \Theta(t_k - y)$$

$$= 1 + \sum_{k=1}^{n} \frac{1}{\beta^{k-1}} g_{n-k+1,1} \Theta(t_k - y)$$

and

$$d_n(y) = \sum_{k=0}^{n-1} v_k - \sum_{k=0}^{n-1} b_k f_{nk}$$

$$= \sum_{k=0}^{n-1} v_k - \sum_{k=0}^{n-1} b_k \sum_{j=1}^{n-k} \beta^{-j} g_{n-j,k}$$

This is opaque, and does not have the structure of earlier results. Lets write out the first few.

$$\rho_1 = 1 + \frac{1}{\beta} \Theta(t_1 - y)$$

$$\rho_2 = 1 + \frac{\beta+1}{\beta^2} \Theta(t_1 - y) + \frac{1}{\beta^2} \Theta(t_2 - y)$$

while

$$d_1 = 1 - \frac{1}{\beta}$$

$$d_2 = 1 + v_1 - \frac{1}{\beta} - \frac{1}{\beta^2} - b_1 \frac{1}{\beta^2}$$

61

So

$$v_1 = \rho_1 - d_1 = \frac{1}{\beta} \left[1 + \Theta\left(t_1 - y\right) \right]$$

$$v_2 = \rho_2 - d_2 = \frac{1}{\beta^2} \left[1 + b_1 + \Theta\left(t_1 - y\right) + \Theta\left(t_2 - y\right) \right]$$

Bleh. The most obvious mistake that we made was a failure to group the $\sum_{k=0}^{n-1} v_k$ with the g_{nk} where they should have gone. Lets try to amend that mistake now.

Restructure   Lets undo the above mistake. Write

$$v_n = \sigma_n + c_n$$

and hypothesize that

$$\sigma_n\left(y\right) = \sum_{k=0}^{n} h_{nk}\left(y\right) \Theta\left(t_k - y\right)$$

$$c_n\left(y\right) = \sum_{k=0}^{n-1} e_{nk}\left(y\right) \Theta\left(\beta t_k - 1\right)$$

Iterating gives

$$\beta \mathcal{M} \sigma_n = \sum_{k=0}^{n} h_{nk}\left(\frac{y}{\beta}\right) \Theta\left(\beta t_k - 1\right) + \sum_{k=1}^{n+1} h_{n,k-1}\left(\frac{y + \Theta\left(\beta t_{k-1} - 1\right)}{\beta}\right) \Theta\left(t_k - y\right)$$

and

$$\beta \mathcal{M} c_n = \sum_{k=0}^{n-1} e_{nk}\left(\frac{y}{\beta}\right) \Theta\left(\beta t_k - 1\right) + \sum_{k=0}^{n-1} e_{nk}\left(\frac{y+1}{\beta}\right) \Theta\left(\beta t_k - 1\right) \Theta\left(t_1 - y\right)$$

Comparing term by term gives the recursion relations

$$\beta h_{n+1,k} = h_{n,k-1}\left(\frac{y + \Theta\left(\beta t_{k-1} - 1\right)}{\beta}\right) \qquad \text{for } k > 1$$

$$\beta h_{n+1,1} = h_{n0}\left(\frac{y+1}{\beta}\right) + \sum_{k=0}^{n-1} e_{nk}\left(\frac{y+1}{\beta}\right) \Theta\left(\beta t_k - 1\right)$$

$$h_{n+1,0} = 0$$

$$h_{00} = v_0$$

$$\beta e_{n+1,k} = e_{nk}\left(\frac{y}{\beta}\right) + h_{nk}\left(\frac{y}{\beta}\right) \qquad \text{for } k < n$$

$$\beta e_{n+1,n} = h_{nn}\left(\frac{y}{\beta}\right)$$

The above converges correctly for $v_0 = 1$ as well as $v_0 = y - 1/2$ for assorted low-order periodic orbits. So this checks out numerically.

Unroll   Lets unroll the recursion. Easier to rewrite above to eliminate off-by-one:

$$h_{n,k} = \frac{1}{\beta} h_{n-1,k-1}\left(\frac{y+b_{k-1}}{\beta}\right) \quad \text{for } k > 1$$

$$h_{n,1} = \frac{1}{\beta}\sum_{k=0}^{n-2} b_k e_{n-1,k}\left(\frac{y+1}{\beta}\right) \quad \text{for } n > 1$$

$$h_{1,1} = \frac{1}{\beta} v_0 \left(\frac{y+1}{\beta}\right)$$

$$h_{n+1,0} = 0$$

$$h_{00} = v_0$$

$$e_{n,k} = \frac{1}{\beta}\left[e_{n-1,k}\left(\frac{y}{\beta}\right) + h_{n-1,k}\left(\frac{y}{\beta}\right)\right]$$

$$e_{n,n} = 0$$

Recursing on h_{nk} for $k > 1$ is straight-forward, as only the arguement is iterated. This gives

$$h_{nk}(y) = \frac{1}{\beta^{k-1}} h_{n-k+1,1}\left(\frac{y}{\beta^{k-1}} + \sum_{j=1}^{k-1} b_j \beta^{-j}\right)$$

Recursing on e_{nk} terminates because $e_{nk} = 0$ for $k \geq n$, leaving only h_{nk} behind. This gives

$$e_{n,k}(y) = \sum_{j=1}^{n-k} \beta^{-j} h_{n-j,k}\left(\frac{y}{\beta^j}\right)$$

The sum $h_{n,1}$ can be left as-is, or expanded; there does not seem to be any benefit to doing the below, but it can be contemplated:

$$h_{n,1} = \frac{1}{\beta}\sum_{k=0}^{n-2} b_k \sum_{j=1}^{n-k-1} \beta^{-j} h_{n-j-1,k}\left(\frac{y+1}{\beta^{j+1}}\right)$$

$$= \sum_{k=0}^{n-2} b_k \sum_{j=2}^{n-k} \frac{1}{\beta^j} h_{n-j,k}\left(\frac{y+1}{\beta^j}\right)$$

The above pass numerical checks.

Basic checks   Check it for $n = 1$.

$$\beta h_{1,1} = h_{00}\left(\frac{y+1}{\beta}\right)$$

$$h_{1,0} = 0$$

$$h_{00} = v_0$$

$$\beta e_{1,0} = h_{00}\left(\frac{y}{\beta}\right)$$

so

$$v_1 = \sigma_1 + c_1$$
$$= h_{11}(y)\,\Theta(t_1 - y) + e_{10}(y)$$
$$= \frac{1}{\beta} v_0 \left(\frac{y+1}{\beta} \right) \Theta(t_1 - y) + \frac{1}{\beta} v_0 \left(\frac{y}{\beta} \right)$$

which is correct & verified.

Check for $n = 2$.

$$c_2(y) = e_{20}(y) + b_1 e_{21}(y)$$
$$e_{20} = \frac{1}{\beta^2} h_{00} \left(\frac{y}{\beta^2} \right)$$

and

$$h_{21} = \frac{1}{\beta} e_{10} \left(\frac{y+1}{\beta} \right)$$
$$= \frac{1}{\beta} h_{1,1} \left(\frac{y}{\beta} \right)$$
$$= \frac{1}{\beta^2} v_0 \left(\frac{y}{\beta^2} + \frac{1}{\beta} \right)$$

More unrolling   Consolidating the results so far into a short section:

$$h_{n,k} = \frac{1}{\beta^{k-1}} h_{n-k+1,1} \left(\frac{y}{\beta^{k-1}} + \sum_{j=1}^{k-1} b_j \beta^{-j} \right) \quad \text{for } k > 1$$
$$h_{n,1} = \frac{1}{\beta} \sum_{k=0}^{n-2} b_k e_{n-1,k} \left(\frac{y+1}{\beta} \right) \quad \text{for } n > 1$$
$$h_{1,1} = \frac{1}{\beta} v_0 \left(\frac{y+1}{\beta} \right) \quad \text{for } y < t_1$$
$$h_{n+1,0} = 0$$
$$h_{00} = v_0$$
$$e_{n,k} = \sum_{j=1}^{n-k} \beta^{-j} h_{n-j,k} \left(\frac{y}{\beta^j} \right) \quad \text{for } n > k$$
$$e_{n,k} = 0 \quad \text{for } k \geq n$$

The second row simplifies after noting the earlier definition

$$c_n(y) = \sum_{k=0}^{n-1} b_k e_{nk}(y)$$

and so

$$h_{n,1}(y) = \frac{1}{\beta} c_{n-1} \left(\frac{y+1}{\beta} \right) \quad \text{for } n > 1$$

which enables

$$h_{n,k} = \frac{1}{\beta^k} c_{n-k} \left(\frac{y+1}{\beta^k} + \sum_{j=1}^{k-1} b_j \beta^{-j} \right) \quad \text{for } n > k, k > 0$$

The h_{nk} will only ever be used with $y < t_k$ so there is no expectation that the argument is valid for larger y. This can be verified using the identity

$$t_{k+1} = \beta t_k - b_k$$

$$= \beta^{k+1} - \sum_{j=0}^{k} b_j \beta^{k-j}$$

$$\sum_{j=1}^{k-1} b_j \beta^{-j} = t_1 - \frac{t_k}{\beta^{k-1}}$$

So, working backwards,

$$\frac{y+1}{\beta^k} + \sum_{j=1}^{k-1} b_j \beta^{-j} = t_1 + \frac{1}{\beta^k} (y - \beta t_k + 1)$$

$$\leq 1 \quad \text{when } y \leq t_k$$

Thus, it is appropriate to write

$$h_{n,k} = \frac{1}{\beta^k} c_{n-k} \left(t_1 + \frac{1}{\beta^k} (y - \beta t_k + 1) \right) \quad \text{for } n > k, k > 0$$

The $n = k$ case needs special attention. It is given by

$$h_{k,k} = \frac{1}{\beta^{k-1}} h_{1,1} \left(\frac{y}{\beta^{k-1}} + \sum_{j=1}^{k-1} b_j \beta^{-j} \right)$$

$$= \frac{1}{\beta^k} v_0 \left(\frac{y}{\beta^k} + \frac{1}{\beta} + \frac{1}{\beta} \sum_{j=1}^{k-1} b_j \beta^{-j} \right)$$

$$= \frac{1}{\beta^k} v_0 \left(1 - \frac{t_k - y}{\beta^k} \right)$$

As before, this is defined only for $y \leq t_k$.

This can then be plugged into e_{nk} to write

$$e_{nk} = \sum_{j=1}^{n-k} \beta^{-j} h_{n-j,k} \left(\frac{y}{\beta^j} \right) \quad \text{for } n > k, k > 0$$

$$= \frac{\beta^k}{\beta^n} h_{kk} \left(\frac{\beta^k y}{\beta^n} \right) + \sum_{j=1}^{n-k-1} \beta^{-j} h_{n-j,k} \left(\frac{y}{\beta^j} \right)$$

$$= \frac{1}{\beta^n} v_0 \left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n} \right) + \sum_{j=1}^{n-k-1} \beta^{-j-k} c_{n-k-j} \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{y}{\beta^{k+j}} \right)$$

65

Use $m = n - k - j$ or $j = n - k - m$ to write, for $n > k$ and $y < t_k$ and $k > 0$,

$$e_{nk} = \frac{1}{\beta^n} \left[v_0 \left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n} \right) + \sum_{m=1}^{n-k-1} \beta^m c_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n} \right) \right]$$

while, for $k = 0$

$$e_{n0} = \frac{1}{\beta^n} v_0 \left(\frac{y}{\beta^n} \right)$$

This allows for the replacement of e_{nk} in the expression for c_n to provide a recursion relation for c_n only, and no other "external" factors.

Renaming trick There's two more tricks we can play to simplify things. So far, we've concluded that

$$h_{nk}(y) = \frac{1}{\beta^k} c_{n-k} \left(t_1 + \frac{1}{\beta^k} (y - \beta t_k + 1) \right) \qquad \text{for } n > k, k > 0$$

$$h_{k,k}(y) = \frac{1}{\beta^k} v_0 \left(1 - \frac{t_k - y}{\beta^k} \right)$$

$$h_{n,0} = 0 \quad \text{for } n > 0$$

But if we define $c_0 = v_0$ then the special case for h_{kk} is not needed. This also means that we can now use $h_{n0} = 0$ always, and so the special casing there goes away, too.

Validation Lets see what happens if $v_0 = 1$. Then

$$e_{nk} = \frac{1}{\beta^n} \left[1 + \sum_{m=1}^{n-k-1} \beta^m c_m \right] \qquad \text{for } k > 0$$

$$e_{n0} = \frac{1}{\beta^n}$$

$$c_n = \sum_{k=0}^{n-1} b_k e_{nk}$$

$$h_{nk} = \frac{c_{n-k}}{\beta^k}$$

Rearranging,

$$c_n = \frac{1}{\beta^n} \left[1 + \sum_{k=1}^{n-1} b_k \left[1 + \sum_{m=1}^{n-k-1} \beta^m c_m \right] \right]$$

so that

$$v_n(y) = c_n + \sum_{k=1}^{n} \frac{c_{n-k}}{\beta^k} \Theta(t_k - y)$$

Comparing to the Gelfond-Parry result, we conclude that $c_n \to 1/F$ as $n \to \infty$, with F the normalization constant as before:

$$1 = \frac{1}{F} \left[\sum_{k=0}^{\infty} \frac{t_k}{\beta^k} \right]$$

66

Fibonacci The sequence for c_n is takes the form of a generalized Fibonacci sequence. First, rearrange

$$c_n = \frac{1}{\beta^n} \left[1 + \sum_{k=1}^{n-1} b_k \sum_{m=0}^{n-k-1} \beta^m c_m \right]$$

Write $\alpha_n = \beta^n c_n$; this gives the beguiling form

$$\alpha_n = 1 + \sum_{k=1}^{n-1} b_k \sum_{m=0}^{n-k-1} \alpha_m$$

Denote the partial sum as $f_i = \sum_{m=0}^{i} \alpha_m$. This implies $\alpha_{i+1} + f_i = f_{i+1}$

$$\alpha_n = 1 + b_1 f_{n-2} + b_2 f_{n-3} + \cdots$$
$$f_n = 1 + f_{n-1} + b_1 f_{n-2} + b_2 f_{n-3} + \cdots$$

For the finite bitsequence $b_1 = 1$ and $b_k = 0$ for $k > 1$ this is just the Fibonacci sequence, off-by-one:

$$(f_n + 1) = (f_{n-1} + 1) + (f_{n-2} + 1)$$

3.6.4   Stacking Generic Functions – Conclusion

To summarize: a recursive series has been obtained for $v_n = \mathscr{M}^n v_0$ for an arbitrary function $v_0(y)$ and $\mathscr{M}$ the transfer operator.

The proof provided in the stretch–cut–stack subsection goes through with only minor modifications needed, to obtain this general form. The trick is to track two distinct travellers: one that travels with $\Theta(t_k - y)$ and another that travels with the bitsequence $b_k = \Theta(\beta t_k - 1)$. The general solution has the form

$$v_n(y) = c_n(y) + \sum_{k=1}^{n} h_{nk}(y) \Theta(t_k - y)$$

$$c_n(y) = \sum_{k=0}^{n-1} b_k e_{nk}(y)$$

The functions h_{nk} and e_{nk} can be solved for recursively; the recursive relations are simple enough that they can be rolled up as series summations. These are given by

$$h_{nk}(y) = \frac{1}{\beta^k} c_{n-k}\left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{y}{\beta^k} \right) \qquad \text{for } n \geq k$$

which express h_{nk} in terms of the function c_n. This is given as a recursive series, obtained by iterating

$$e_{nk}(y) = \frac{1}{\beta^n} \left[v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n} \right) + \Theta(k) \sum_{m=1}^{n-k-1} \beta^m c_m\left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n} \right) \right]$$

The starting point for iteration is $c_0 = v_0$.

The function c_n is polynomial if v_0 is, and of the same degree; c_n is analytic, if v_0 is, and so on. The function h_{nk} is likewise, on the domain $0 \le y \le t_k$; the discontinuities in v_n are entirely due to the $\Theta(t_k - y)$ term.

The proper calculation of the $n \to \infty$ limit of c_n remains a mystery. This is the primary obstacle to constructing general eigenfunctions from this series.

Verification   When $v_0 = 1$, all three functions c_n, e_{nk} and h_{nk} become constants. In this case, $c_n \to 1/F$ as $n \to \infty$, with F the normalization constant from before. Defining $\alpha_n = \beta^n c_n$, the recursion relation takes the curious form

$$\alpha_n = 1 + \sum_{k=1}^{n-1} b_k \sum_{m=0}^{n-k-1} \alpha_m$$

This is an integer sequence: each α_n is an integer, starting with $\alpha_0 = \alpha_1 = 1$. This has the form of a generalized Fibonacci sequence. Denote the partial sum as $s_i = \sum_{m=0}^{i} \alpha_m$. This implies $\alpha_i = s_i - s_{i-1}$ and so

$$\alpha_n = 1 + b_1 s_{n-2} + b_2 s_{n-3} + \cdots$$
$$s_n = 1 + s_{n-1} + b_1 s_{n-2} + b_2 s_{n-3} + \cdots$$

If the orbit is of finite length p, so that $b_k = 0$ for $k > p$, then this can be recast directly into generalized Fibonacci form, by defining $f_n = s_n + \sum_{k=1}^{p} b_k$. The recursion relation is then

$$f_n = f_{n-1} + b_1 f_{n-2} + \cdots + b_p f_{n-p}$$

For example, the orbit generated by $\beta = \varphi = 1.618\cdots$ the golden mean has $p = 1, b_1 = 1$ and the sequence is $\alpha_m = 1, 1, 2, 3, 5, 8, 13, \cdots$. Finite orbits and generalized Fibonacci sequences will be treated at length in the next chapter. One of the interesting properties is that $\beta = \lim_{n \to \infty} f_n/f_{n-1}$ holds in the general case, and not just for the golden mean. The bitsequences are self-describing; this is byproduct from the identity $\beta = \sum_{k=0}^{\infty} b_k \beta^{-k}$.

Unfinished aspirations   The unfulfilled hope of the above was to obtain an enlightening form in the $n \to \infty$ limit. The sum over piece-wise continuous terms in h_{nk} is not an issue, since the h_{nk} can be calculated. The stumbling block is the behavior of $c_n(y)$ in the $n \to \infty$ limit. It is clear that if v_0 is holomorphic, then so is $c_n(y)$; it has no discontinuities.

The next section determines that

$$c_n(y) \to \text{const}$$

in the $n \to \infty$ limit, explicitly for $v_0(y) = y$. The const is rather opaque. Numerical work suggests that this holds in general, at least for polynomial $v_0(y)$.

3.7 Linear form

Given the above form, apply some blunt technique, and see what happens. The primary result is that the recursion relation can be specialized to $v_0(y) = y$ and an explicit form can be given. The linear part results in generalized Fibonacci sequences being generated, while the constant term is messy and opaque (but not hard to compute).

3.7.1 Finite orbits

Lets take a look at above, for finite orbits. These are associated with finite sequences b_k that terminate after a certain point.

As above, define $\alpha_n(y) = \beta^n c_n(y)$ and

$$s_{nk}(y) = \sum_{m=1}^{n-k-1} \alpha_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n} \right)$$

Then

$$\phi_{nk}(y) = \beta^n e_{nk}(y) = v_0 \left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n} \right) + \Theta(k) s_{nk}(y) \qquad \text{for } k > 0$$

and

$$\alpha_n(y) = \sum_{k=0}^{n-1} b_k \phi_{nk}(y)$$

Well, after all the burbling above about Fibonacci, this doesn't look so promising. Mash it into one large line

$$\alpha_n(y) = v_0 \left(\frac{y}{\beta^n} \right) + \sum_{k=1}^{n-1} b_k \left[v_0 \left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n} \right) + \sum_{m=1}^{n-k-1} \alpha_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n} \right) \right]$$

3.7.2 Linear generator

Plug in $v_0(y) = y$ into above. Then

$$\sum_{k=1}^{n-1} b_k v_0 \left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n} \right) = \left(1 + \frac{y}{\beta^n} \right) \sum_{k=1}^{n-1} b_k - \sum_{k=1}^{n-1} b_k \frac{t_k}{\beta^k}$$

$$= \gamma_n + \frac{\delta_n}{\beta^n} y$$

The two constants γ_n, δ_n are both bounded: $\gamma_n < n+1$ and $\delta_n < n$. Note that δ_n is an integer! The α_n will have a similar form:

$$\alpha_n(y) = \varepsilon_n + \frac{\zeta_n}{\beta^n} y$$

and so plugging through:

$$\alpha_n(y) = \frac{y}{\beta^n} + \gamma_n + \frac{\delta_n}{\beta^n}y + \sum_{k=1}^{n-1} b_k \sum_{m=1}^{n-k-1} \left[\varepsilon_m + \frac{\zeta_m}{\beta^m} \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n} \right) \right]$$

$$= \gamma_n + \frac{y}{\beta^n} \left[1 + \delta_n + \sum_{k=1}^{n-1} b_k \sum_{m=1}^{n-k-1} \zeta_m \right]$$

$$+ \sum_{k=1}^{n-1} b_k \sum_{m=1}^{n-k-1} \varepsilon_m$$

$$+ \sum_{k=1}^{n-1} b_k \left(t_1 - \frac{\beta t_k - 1}{\beta^k} \right) \sum_{m=1}^{n-k-1} \frac{\zeta_m}{\beta^m}$$

The constant terms are quite the mess, but the linear terms are not so bad; not only is δ_n and integer, but so is ζ_n:

$$\delta_n = \sum_{k=1}^{n-1} b_k$$

$$\zeta_n = 1 + \delta_n + \sum_{k=1}^{n-1} b_k \sum_{m=1}^{n-k-1} \zeta_m$$

The ζ_n sequence is identical to the "generalized Fibonacci", given in the prior section. This is not immediately apparent, but can be verified numerically. In particular, this implies that

$$\zeta_n = 1 + \sum_{k=1}^{n-1} b_k \sum_{m=0}^{n-k-1} \zeta_m$$

The index on the last sum runs from $m = 0$ instead of $m = 1$. In the asymptotic limit, the ratio of successive terms converges $\zeta_n / \zeta_{n-1} \to \beta$. This is the generalized form of the standard Fibonacci series converging to the golden mean.

Lets take a look at the constant terms, now. If the bitstring b_k is of finite length, then both of these stop changing, for n greater than the bitstring length (since $b_k = 0$ from that point on).

$$\gamma_n = \delta_n - \sum_{k=1}^{n-1} b_k \frac{t_k}{\beta^k}$$

$$\varepsilon_n = \gamma_n + \sum_{k=1}^{n-1} b_k \left[\left(\sum_{m=1}^{n-k-1} \varepsilon_m \right) + \left(t_1 - \frac{\beta t_k - 1}{\beta^k} \right) \sum_{m=1}^{n-k-1} \frac{\zeta_m}{\beta^m} \right]$$

But, by definition,

$$t_k = \beta t_{k-1} - b_{k-1}$$

$$= \beta^k \left[1 - \sum_{m=0}^{k-1} \beta^{-m-1} b_m \right]$$

70

Inside the square brackets is a portion the identity $1 = \sum_{m=0}^{\infty} b_m \beta^{-m-1}$. The δ_n cancels, and so

$$\gamma_n = \sum_{k=1}^{n-1} b_k \sum_{m=0}^{k-1} \beta^{-m-1} b_m$$

The largest γ_n and ε_n occur when $\beta \to 2$. In these cases, there is a string of $b_k = 1$ up to $k = p$ and then $b_k = 0$ after that. For these cases, the asymptotic limits appear to be

$$\gamma_n = p - 1 + \mathcal{O}\left(\beta^{-p+2} \log p\right)$$
$$\varepsilon_n = 3\beta^{p-2}\left[1 + \mathcal{O}\left(\beta^{-p+2} \log p\right)\right]$$

where, in both cases, the extra $\log p$ is a wild guess based on numerics. As before, this is for $n > p$, since, for finite bitsequnces, these stop changing once $n > p$.

Time to wrap it up. The goal of this exercise was to understand the $n \to \infty$ limit of $c_n(y)$. We have achieved that, at least for the case of $v_0(y) = y$. The result is that
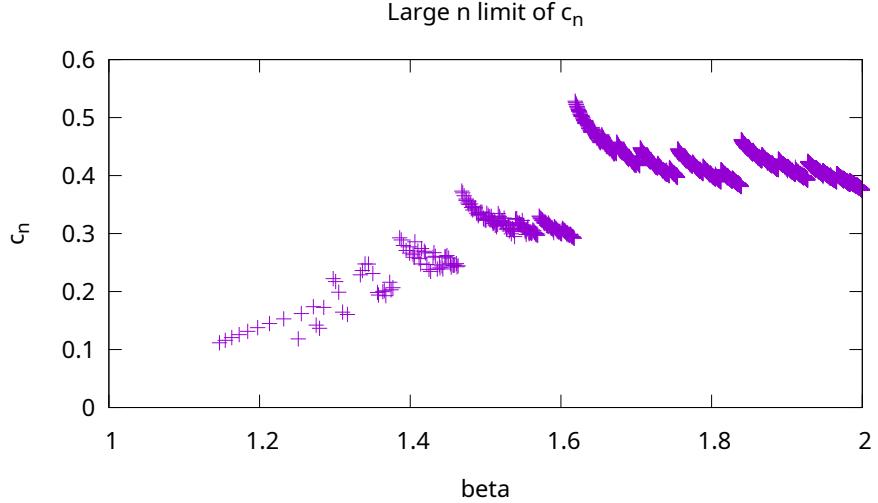
$$c_n(y) = \beta^{-n} \alpha_n(y)$$
$$= \frac{1}{\beta^n}\left[\varepsilon_n + \frac{\zeta_n}{\beta^n} y\right]$$

In the worst case, ε_n grows as β^n so the constant term is of order 1, but more commonly small. In the limit, the ratio of succesive generalized Fibonacci terms converge to their generator, so $\zeta_n/\zeta_{n-1} \to \beta$, and thus, $\zeta_n \beta^{-n} \to 1$. As this is suppressed by an additional term of β^{-n}, the final result is that

$$c_n(y) \to \text{const}$$

for a somewhat impenetrable const that seems to always be less than one. Based on numerical work with polynomial $v_0(y)$, this seems to hold in general: only the constant survives, and the polynomial parts are clobbered by a factor of β^{-n}.

The obvious next step is to make a graph of c_n in the large-n limit and see if it looks like anything interesting. This is shown in the image immediately below.

Large n limit of c_n

A clear, repeating pattern is visible. Asterisk: the above holds, assuming all above equations were derived correctly, and implemented correctly. This has not been double-checked. However, the same general magnitude of c_n was already seen in unrelated, independent explorations, so this seems to probably be okay. It's plausible given other experience.

3.7.3 Golden Mean

Lets try $\beta = \varphi$ so that $b_1 = 1$ and all others vanish. Then

$$
\begin{aligned}
\alpha_n(y) &= \phi_{n0}(y) + \phi_{n1}(y) \\
&= v_0\left(\frac{y}{\beta^n}\right) + v_0\left(1 - \frac{t_1}{\beta} + \frac{y}{\beta^n}\right) + s_{n1}(y) \\
&= v_0\left(\frac{y}{\beta^n}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^n}\right) + \sum_{m=1}^{n-2} \alpha_m\left(\frac{1}{\beta} + \frac{\beta^m y}{\beta^n}\right)
\end{aligned}
$$

Lets write some of these out.

$$\alpha_1(y) = \phi_{10}(y) = v_0\left(\frac{y}{\beta}\right)$$

$$\alpha_2(y) = v_0\left(\frac{y}{\beta^2}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^2}\right)$$

$$\alpha_3(y) = v_0\left(\frac{y}{\beta^3}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^3}\right) + \alpha_1\left(\frac{1}{\beta} + \frac{y}{\beta^2}\right)$$

$$= v_0\left(\frac{y}{\beta^3}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^3}\right) + v_0\left(\frac{1}{\beta^2} + \frac{y}{\beta^3}\right)$$

$$\alpha_4(y) = v_0\left(\frac{y}{\beta^4}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^4}\right) + \alpha_1\left(\frac{1}{\beta} + \frac{y}{\beta^3}\right) + \alpha_2\left(\frac{1}{\beta} + \frac{y}{\beta^2}\right)$$

$$= v_0\left(\frac{y}{\beta^4}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^4}\right) + v_0\left(\frac{1}{\beta^2} + \frac{y}{\beta^4}\right) + v_0\left(\frac{1}{\beta^3} + \frac{y}{\beta^4}\right) + v_0\left(\frac{1}{\beta} + \frac{1}{\beta^3} + \frac{y}{\beta^4}\right)$$

Lots of heat, very little light. The above does make clear that if v_0 is a polynomial of degree p then so are all of the α_n. Likewise, if v_0 is holomorphic/meromorphic, then so are the α_n. There are no discontinuities in the α_n; the piece-wise parts are all in the h_{nk} sum.

Try again. Assume an analytic series, or at least a polynomial.

$$v_0(x) = \sum_{i=0}^{\infty} w_i x^i \qquad \text{and} \qquad \alpha_n(x) = \sum_{i=0}^{\infty} u_{ni} x^i$$

Then

$$v_0(x+y) = \sum_{i=0}^{\infty} w_i (x+y)^i$$

$$= \sum_{i=0}^{\infty} w_i \sum_{j=0}^{i} \binom{i}{j} x^j y^{i-j}$$

So

$$\alpha_n(x) = v_0\left(\frac{y}{\beta^n}\right) + v_0\left(\frac{1}{\beta} + \frac{y}{\beta^n}\right) + \sum_{m=1}^{n-2}\sum_{i=0}^{\infty} u_{mi}\left(\frac{1}{\beta} + \frac{\beta^m y}{\beta^n}\right)^i$$

$$\sum_{i=0}^{\infty} u_{ni} y^i = \sum_{i=0}^{\infty} w_i \left(\frac{y}{\beta^n}\right)^i + \sum_{i=0}^{\infty} w_i \sum_{j=0}^{i} \binom{i}{j}\left(\frac{y}{\beta^n}\right)^j \frac{1}{\beta^{i-j}} + \sum_{m=1}^{n-2}\sum_{i=0}^{\infty} u_{mi} \sum_{j=0}^{i} \binom{i}{j}\left(\frac{\beta^m y}{\beta^n}\right)^j \frac{1}{\beta^{i-j}}$$

Quite the mess. Equating term by term,

$$u_{np} = w_p \frac{1}{\beta^{pn}} + \sum_{i=p}^{\infty} w_i \binom{i}{p}\frac{1}{\beta^{pn}}\frac{1}{\beta^{i-p}} + \sum_{m=1}^{n-2}\sum_{i=p}^{\infty} u_{mi}\binom{i}{p}\frac{\beta^{mp}}{\beta^{np}}\frac{1}{\beta^{i-p}}$$

$$= \frac{1}{\beta^{pn}}\left[w_p + \sum_{i=p}^{\infty}\binom{i}{p}\frac{1}{\beta^{i-p}}\left[w_i + \sum_{m=1}^{n-2} u_{mi}\beta^{mp}\right]\right]$$

73

Yikes! Even if we assume that quintic terms and higher are zero, *i.e.* that $w_p = 0$ for $p > 4$, this does not simplify in any meaningful way. Mind you, the above is for the special case of $\beta = \varphi$. The general case is slightly more complex, but not much more.

Lets try $v_0(y) = y$ so that $w_1 = 1$ and all other $w_i = 0$. Given $\alpha_1(y) = v_0(y/\beta)$ conclude $u_{10} = 0$ and $u_{11} = 1$ and all other $u_{1p} = 0$. Then

$$u_{2,p} = \frac{1}{\beta^{2p}} \left[w_p + \sum_{i=p}^{\infty} \binom{i}{p} \frac{1}{\beta^{i-p}} w_i \right]$$

$$u_{2,0} = \frac{w_1}{\beta^i}$$

$$u_{2,1} = \frac{2w_1}{\beta^2}$$

and all other $u_{2,p} = 0$. Next,

$$u_{3,p} = \frac{1}{\beta^{3p}} \left[w_p + \sum_{i=p}^{\infty} \binom{i}{p} \frac{1}{\beta^{i-p}} [w_i + u_{1i}\beta^p] \right]$$

$$= \frac{1}{\beta^{3p}} \left[w_p + \binom{1}{p} \frac{1}{\beta^{1-p}} [w_1 + u_{11}\beta^p] \right] \quad \text{for } p \le 1$$

$$u_{3,0} = \frac{1}{\beta} [w_1 + u_{11}]$$

$$u_{3,1} = \frac{1}{\beta^3} [2w_1 + u_{11}\beta]$$

and $u_{3,p} = 0$ for all other p. Recursively, $u_{np} = 0$ for all $p > 1$, but we knew this already. Thus, all we really have is

$$u_{n0} = \sum_{m=1}^{n-2} u_{m0} + \frac{1}{\beta} \sum_{m=1}^{n-2} u_{m1}$$

$$u_{n1} = \frac{1}{\beta^n} \left[2w_1 + \sum_{m=1}^{n-2} u_{m1}\beta^m \right]$$

Oof. It's doable, but icky.

3.7.4   Integral constraints

Compute the integral $\int c_n$, doe it reveal anything?

$$\int_0^1 c_n(y)\,dy = \sum_{k=0}^{n-1} b_k \int_0^1 e_{nk}(y)\,dy$$

where

$$\int_0^1 e_{nk}(y)\,dy = \frac{1}{\beta^n} \left[\int_0^1 v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n}\right) + \Theta(k) \sum_{m=1}^{n-k-1} \beta^m \int_0^1 c_m\left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n}\right) \right]$$

$$= \int_p^q v_0(x)\,dx + \Theta(k) \sum_{m=1}^{n-k-1} \beta^m \int_r^s c_m(x)\,dx$$

74

where $p = 1 - t_k \beta^{-k}$ and $q = p + \beta^{-n}$ and $r = t_1 - (\beta t_k - 1)\beta^{-k}$ and $s = r + \beta^{m-n}$. Plugging in (for example) $v_0(y) = y$ does not look promising.

3.8 Coherent Disk

Using the same trick as further below (sorry for out-of-order presentation), we can add a complex z to the mix. Cut-n-pasting from up above, a sequence of functions satisfying $\mathscr{M}\mu_n = \lambda\mu_{n+1}$ are given by

$$\mu_n(y) = c'_n(y) + \sum_{k=1}^{n} h'_{nk}(y)\Theta(t_k - y)$$

$$c'_n(y) = v_0(y) + \frac{1}{(\lambda\beta)^n}\sum_{k=1}^{n-1} b_k \left[v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n}\right) + \sum_{m=1}^{n-k-1}(\lambda\beta)^m c'_m\left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n}\right)\right]$$

which is identical to before but prime is used to indicate the presence of the λ in the eqns. The functions h'_{nk} are given by

$$h'_{nk}(y) = \frac{1}{(\lambda\beta)^k} c'_{n-k}\left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{y}{\beta^k}\right) \quad \text{for } n \geq k$$

and iteration starts with $c'_0 = v_0$.

3.8.1 Verify

For $v_0 = 1$ we should get old results. This implies there no y dependence in c'_n and so

$$c'_n = v_0 + \frac{1}{(\lambda\beta)^n}\sum_{k=1}^{n-1} b_k \left[v_0 + \sum_{m=1}^{n-k-1}(\lambda\beta)^m c'_m\right]$$

and

$$\mu_n(y) = c'_n + \sum_{k=1}^{n} \frac{1}{(\lambda\beta)^k}c'_{n-k}\Theta(t_k - y)$$

The Parry–Gelfond result is obtained by setting $\lambda = 1$ so that

$$\lim_{n\to\infty} c'_n + \sum_{k=1}^{n}\frac{1}{\beta^k}c'_{n-k}\Theta(t_k - y) = \frac{1}{F}\left[v_0 + \sum_{k=1}^{\infty}\frac{1}{\beta^k}\Theta(t_k - y)\right]$$

Comparing term by term, this is possible only if $c'_n \to 1/F$. A quick numeric check confirms this is the case. So everything looks OK, here.

3.8.2 Remarks

There is room for confusion, here. Lets clear the air. The above generates a sequence μ_n for any value of β and λ, satisfying $\mathscr{M}\mu_n = \lambda\mu_{n+1}$ by construction. What happens if $\lambda \neq 1$? Well, selecting $v_0 = 1$ it appears that $\lambda\mu_{n+1} \approx \mu_n$ so that $\lambda^n\mu_n \to v$ in the

75

limit. That is, each iterate gets closer and closer to Gelfond–Parry while also blowing up geometrically. There is no free lunch.

This result follows, because $\mathcal{M}$ preserves volume. That is, $\int \mu_n = \int \mathcal{M}\mu_n = \lambda \int \mu_{n+1}$ so if $\int \mu_n \neq 0$ then its no surprise that there is a geometric increase with each application. This is a rediscovery that the decaying eigenfuncs must be orthogonal to the invariant measure.

The next step is very interesting. A report of numeric observations. Fix $\beta = 1.6$. Then:

- If $\mu_0(x) = x - 1/2$, then $\lambda^n \mu_n$ converges to $\approx -0.0869229\,\nu$ with ν the Gelfond–Parry invariant measure. This is independent of λ. This is not a surprise; the result should be independent of λ, as it is just a rescaling.

- If $\mu_0(x) = x - 1/2 + 0.0869229$ and $\lambda = 1/\beta$, then the μ_n appear to bounce around ergodically, maintaining a bounded norm. Three behaviors are apparent: $\int \mu_n = \text{const} \approx 0.2767$. The L_1 and L_2 norms bounce around but remain bounded: $0.3 < \int |\mu_n| < 0.8$ and $0.1 < \int |\mu_n|^2 < 1.0$, and each iterate is approximately orthogonal to the prior one: $\int \mu_n \mu_{n+1} \approx 0$. Here, $\int f = \int_0^1 f(x)\,dx$ is just short-hand notation.

- If $\mu_0(x) = x - 1/2 + 0.0869229$ and $\lambda = e^{i\phi}/\beta$ and n is held fixed while the phase ϕ real is varied indicates each μ_n at different ϕ is just a scaled copy of the $\phi = 0$ version.

- If $\mu_0(x) = x - 1/2 + 0.0869229$ and $\lambda \neq 1/\beta$ then $\lambda^n \beta^n \mu_n$ are exactly equal to above. This is independent of the choice of λ.

The λ-independence just says that λ is a red herring, having nothing to do with anything. It was supposed to be a calculational convenience, but serves mostly to confuse the issue. The code base rescales by β anyway, and that is all that is needed. Yellow flag: continue use of λ only with caution.

This leaves two questions: what is this magic constant 0.0869229? What is this ergodic bouncing? The second question has a relatively simple (but interesting) answer: μ_n has converged to a superposition of multiple complex eigenfunctions, all having an eigenvalue with the same norm.

Is it a single complex eigenvalue? No. Assume a single complex eigenvalue. Assume $\mathcal{M}(\psi + i\chi) = \lambda e^{i\phi}(\psi + i\chi)$ for unknown ψ, χ, ϕ. This implies $\mathcal{M}\psi = \lambda(\psi\cos\phi - \chi\sin\phi)$. Assume $\int \psi\chi = 0$. Then $\int \psi \mathcal{M}\psi = \lambda\cos\phi \int \psi^2$. Assume $\mu_n \to \psi$. Then $\int \mu_n \mathcal{M}\mu_n = \int \mu_n \mu_{n+1} = \lambda\cos\phi \int \mu_{n+1}^2$. We already know that $\lambda = 1/\beta$ so estimate $\cos\phi \approx \beta \int \mu_n \mu_{n+1} / \int \mu_{n+1}^2$. But this is not supported by numerics.

Did we make a mistake with assuming $\mu_n \to \psi$? Assume instead that $\mu_n \to K(\psi\cos\theta + \chi\sin\theta)$ for some unknown normalization K and mixing angle θ. Then

$$\mu_{n+1} = \lambda K(\psi\cos(\theta - \phi) + \chi\sin(\theta - \phi))$$

and $\int \mu_n \mu_{n+1} = K^2 \lambda \cos\phi$ after assuming that $\int \psi^2 = \int \chi^2 = 1$ and $\int \psi\chi = 0$. So the corrected hypothesis does not rescue the conclusion that this is not due to a single complex eigenvalue.

Is it a mixture? Yes. Assume there are multiple eigenfunctions $\mathcal{M}(\psi_n + i\chi_n) = \frac{1}{\beta} e^{i\phi_n}(\psi_n + i\chi_n)$ so that all have $|\lambda| = 1/\beta$ but differing in phase ϕ_n. Assume that $\mu_n \to \sum_k a_k \psi_k$. Then further iteration will just tangle the phase into each of these into a mess. That is,

$$\mathcal{M}\mu_n = \mu_{n+1} = \frac{1}{\beta} \sum_k a_k (\psi_k \cos\phi_k - \chi_k \sin\phi_k)$$

and so $\int \mu_n \mathcal{M}\mu_n = \int \mu_n \mu_{n+1} = \frac{1}{\beta} \sum_k a_k^2 \cos\phi_k$. Hmm. Well, that is a constant, independent of n but the numerics are not converging to a constant here. Are the numerics suffering from rounding errors? Possibly. Probably?! Since cancellation of Fibonacci-scale terms is required, this destroys precision after not very many (a dozen) iterations.

Conclusion: the result of iteration is a mixture. The mixture looks like an ergodic bouncing around. The phase angle ϕ_k can be called the "energy" of the eigenvalue, so that, up to rescaling, the time evolution is unitary. The "ergodic" sequence is just a sum of unitarily-evolving eigenfunctions. Huh. Wow!

3.8.3   General

Then $\mu_n \to \mu$ with $\mathcal{M}\mu = \lambda\mu$ if $|\mu_{n+1} - \mu_n| \to 0$. Write this out.

$$\mu_{n+1} - \mu_n = c'_{n+1} - c'_n + h'_{n+1,n+1}\Theta(t_{n+1} - y) + \sum_{k=1}^{n} \left[h'_{n+1,k} - h'_{nk}\right]\Theta(t_k - y)$$

We need several parts, here.

$$
\begin{aligned}
c'_{n+1} - c'_n &= \frac{1}{(\lambda\beta)^{n+1}} \sum_{k=1}^{n} b_k \left[v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^{n+1}}\right) + \sum_{m=1}^{n-k} (\lambda\beta)^m c'_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^{n+1}}\right) \right] \\
&\quad - \frac{1}{(\lambda\beta)^n} \sum_{k=1}^{n-1} b_k \left[v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n}\right) + \sum_{m=1}^{n-k-1} (\lambda\beta)^m c'_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n}\right) \right] \\
&= \frac{1}{(\lambda\beta)^{n+1}} b_n v_0 \left(1 + \frac{y - \beta t_n}{\beta^{n+1}}\right) \\
&\quad + \frac{1}{(\lambda\beta)^n} \sum_{k=1}^{n-1} b_k \left[\frac{1}{\lambda\beta} v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^{n+1}}\right) - v_0\left(1 - \frac{t_k}{\beta^k} + \frac{y}{\beta^n}\right) \right] \\
&\quad + \frac{1}{(\lambda\beta)^n} \sum_{k=1}^{n-1} b_k \left[\frac{1}{\lambda\beta} \sum_{m=1}^{n-k} (\lambda\beta)^m c'_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^{n+1}}\right) \right] \\
&\quad - \frac{1}{(\lambda\beta)^n} \sum_{k=1}^{n-1} b_k \left[\sum_{m=1}^{n-k-1} (\lambda\beta)^m c'_m \left(t_1 - \frac{\beta t_k - 1}{\beta^k} + \frac{\beta^m y}{\beta^n}\right) \right]
\end{aligned}
$$

while

$$h'_{n+1,n+1}(y) = \frac{1}{(\lambda\beta)^{n+1}} v_0 \left(t_1 + \frac{1}{\beta^{n+1}} + \frac{y - \beta t_{n+1}}{\beta^{n+1}}\right)$$

Blech.

3.9 Inverse Problem

The results above and below suggest that the correct object of study is $\mathcal{N} = \mathcal{M} - I$ so that $\mathcal{N}v = 0$ is in the kernel. Note that $\mathcal{N}$ is no more singular than $\mathcal{M}$ since $\mathcal{M}$ already has a large and extensive kernel. Thus $\mathcal{N}$ just has some more.

Two questions now arise: first, is $\mathcal{N}$ the pushforward of something? What is that something? Second question: what happens if this process is repeated? We wish to arrive at an $\mathcal{N}$ where the only only eigenfunctions were those that were on the $1/\beta$ ring. In other words, $\mathcal{N} \to U$ for some unitary U.

This engenders a third set of questions: is such a U the pushforward of something? If so, what? Then, what is the spectrum of U? Interpreting the phases as "energies", what is the corresponding Hamiltonian H satisfying $U = \exp iH$?

3.10 Analytic Gelfond-Parry function

The technique above can be repeated verbatim for a "rotated" or "coherent" function

$$v_{\beta;z}(y) = \sum_{n=0}^{\infty} z^n \frac{\Theta(t_n - y)}{\beta^n} \tag{21}$$

for a given complex-valued z. No changes are required, and the result can be read off directly:

$$\left[\mathcal{M} v_{\beta;z}\right](y) = \frac{v_{\beta;z}(y)}{z} - \frac{1}{z} + \sum_{n=0}^{\infty} z^n \frac{\Theta(\beta t_n - 1)}{\beta^{n+1}}$$

$$= \frac{v_{\beta;z}(y)}{z} + C(\beta;z)$$

with $C(\beta;z)$ being a constant independent of y. If there are values of β and/or z at which $C(\beta;z) = 0$, then this becomes the eigenequation for $\mathcal{M}$.

The eigenfunction for $\mathcal{L}$ is the same, up to rescaling of $y \mapsto \beta x/2$. Recycling notation slightly, write

$$v_{\beta;z}(x) = \sum_{n=0}^{\infty} \frac{d_n(x)}{\beta^n} z^n \tag{22}$$

where the $d_n(x)$ are exactly the same digits as defined by Parry, just rescaled for the beta-shift. That is,

$$d_n(x) = \varepsilon_n\left(\frac{2x}{\beta}\right) = \Theta\left(\frac{\beta}{2} t_n - x\right) = \Theta\left(T^n\left(\frac{\beta}{2}\right) - x\right) = \begin{cases} 1 & \text{if } x < T^n\left(\frac{\beta}{2}\right) \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

where T the beta-shift map of eqn 4 and eqn 9 is used. The iterated end-point becomes the iterated midpoint:

$$t^n(1) = \frac{2}{\beta} T^n\left(\frac{\beta}{2}\right)$$

Holding both β and x fixed, the summation is clearly convergent (and holomorphic in z) for complex numbers z within the disk $|z| < \beta$. The eigenequation has the same form:

$$\left[\mathscr{L}_\beta v_{\beta;z}\right](x) = \frac{1}{z}v_{\beta;z}(x) + C(\beta;z)$$

where $C(\beta;z)$ is a constant independent of x. Numeric verification reveals we were a bit glib: $C(\beta;z)$ is a constant for $x \leq \beta/2$ and is zero otherwise! (This is normal; $\mathscr{L}$ was defined in such a way that it is always exactly zero for $x > \beta/2$.)

The interesting limit is where $|z| \to \beta$ and so its convenient to re-express C in terms of $\zeta = z/\beta$, so that everything is mapped to the unit disk. With some rearrangements, one obtains

$$E(\beta;\zeta) \equiv \zeta\beta C(\beta;\zeta\beta) = -1 + \zeta\sum_{n=0}^{\infty}\zeta^n d_n\left(\frac{1}{2}\right) \qquad (24)$$

Given that $d_n(1/2) = \Theta(\beta t_n - 1)$, the above can be recognized as the rotated/coherent form of eqn 20 in the $n \to \infty$ limit. The primary task is to characterize the zeros of $E(\beta;\zeta)$. This is a straight-forward sum to examine numerically; results will be presented in the section after the next.

3.11 Analytic ergodics

This section proposes that the entire β-subshift can be tied together with a single holomorphic equation. The holomorphic equation effectively provides a continuum (*i.e.* uncountable number) of distinct relationships between different parts of the subshift. This can be interpreted either as a form of interaction across the subshift, or as a kind of mixing. Given the nature of the relationship, the moniker of "fundamental theorem of analytic ergodics" is an amusing name to assign to the result.

The constant term can be independently derived through a different set of manipulations. Explicitly plugging in eqn 22 into the transfer operator yields

$$\begin{aligned}C(\beta;z) &= \left[\mathscr{L}_\beta v_{\beta;z}\right](y) - \frac{v_{\beta;z}(y)}{z}\\ &= \sum_{n=0}^{\infty}\frac{z^n}{\beta^n}\left[\frac{1}{\beta}\left[d_n\left(\frac{y}{\beta}\right) + d_n\left(\frac{y}{\beta} + \frac{1}{2}\right)\right]\Theta\left(\frac{\beta}{2} - y\right) - \frac{1}{z}d_n(y)\right]\end{aligned}$$

Replacing z by $\zeta = z/\beta$ gives

$$\begin{aligned}E(\beta;\zeta) &= \zeta\beta C(\beta;\zeta\beta)\\ &= \sum_{n=0}^{\infty}\zeta^n\left[\zeta\left[d_n\left(\frac{y}{\beta}\right) + d_n\left(\frac{y}{\beta} + \frac{1}{2}\right)\right]\Theta\left(\frac{\beta}{2} - y\right) - d_n(y)\right]\end{aligned}$$

This is holomorphic on the unit disk $|\zeta| < 1$, as each individual d_n is either zero or one; there won't be any poles inside the unit disk. Note that $d_n(y) = 0$ for all $y > \beta/2$, and

so one may pull out the step function to write

$$E(\beta;\zeta) = \sum_{n=0}^{\infty} \zeta^n \left[\zeta d_n \left(\frac{y}{\beta} \right) + \zeta d_n \left(\frac{y}{\beta} + \frac{1}{2} \right) - d_n(y) \right] \Theta \left(\frac{\beta}{2} - y \right)$$

confirming the earlier observation that $E(\beta;\zeta)$ vanishes for all $y > \beta/2$.

The bottom equation holds without assuming that $E(\beta;\zeta)$ is independent of y. However, we've already proven that it is; and so a simplified expression can be given simply by picking a specific y. Setting $y = 0$, noting that $d_n(0) = 1$ and canceling terms, one obtains eqn. 24 again.

Staring at the right-hand side of the sum above, it is hardly obvious that it should be independent of y. In a certain sense, this is not "one equation", this holds for a continuum of y, for all $0 \leq y \leq 1$. It is an analytic equation tying together the entire subshift. For each distinct y, it singles out three completely different bit-sequences out of the subshift, and ties them together. It is a form of mixing. Alternately, a form of interaction: the bit-sequences are not independent of one-another; they interact. This section attempts to make these notions more precise.

The tying-together of seemingly unrelated sequences seems somehow terribly important. It is amusing to suggest that this is a kind of "fundamental theorem of analytic ergodics".

For such a claim, it is worth discussing the meaning at length, taking the effort to be exceptionally precise and verbose, perhaps a bit repetitive. Equation 4 defined a map, the β-shift. Equation 5 defined a bit-sequence, the β-expansion of a real number $0 \leq x \leq 1$, where equation 6 is the definition of the β-expansion. The set of all such bit-sequences defines the shift. To emphasize this point, its best to compare side-by-side. Copying equation 5, one bitsequence records the orbit of x relative to the midpoint:

$$k_n(x) = \begin{cases} 0 & \text{if } 0 \leq T_\beta^n(x) < \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leq T_\beta^n(x) \leq 1 \end{cases}$$

while a different bitsequence records the orbit of the midpoint, relative to x:

$$d_n(x) = \begin{cases} 1 & \text{if } x < T^n \left(\frac{\beta}{2} \right) = T^{n+1} \left(\frac{1}{2} \right) \\ 0 & \text{otherwise} \end{cases}$$

The iterations are running in opposite directions; this is as appropriate, since the the transfer operator was a pushforward.

It is useful to return to the language of sigma algebras and cylinder sets, as opposed to point dynamics. Recall that the Borel algebra $\mathscr{B}$ was defined as the sigma algebra, the collection of all cylinder sets in the product topology of $\{0,1\}^\omega$. A subshift is a subset $\mathscr{S} \subset \mathscr{B}$ together with a map $T : \mathscr{S} \to \mathscr{S}$ that lops off the leading symbol of a given cylinder set but otherwise preserves the subshift: $T\mathscr{S} = \mathscr{S}$. The inverted map T^{-1} is a pushforward, in that it defines the transfer operator, a linear operator $\mathscr{L}_T : \mathscr{F} \to \mathscr{F}$ on the space $\mathscr{F}$ of all functions $f : \mathscr{S} \to \mathbb{R}$; explicitly, it is given by $\mathscr{L}_T : f \mapsto f \circ T^{-1}$. Insofar as the d_n arose in the exploration of the transfer operator, it is not surprising that the shift is acting "backwards".

The problem with the language of point dynamics is that one cannot meaningfully write $T^{-n}(x)$ for a real number, a point x, at least, not without severe contortions that lead back to the Borel algebra. Not for lack of trying; the $T^{-n}(x)$ is called the "Julia set" (to order n) of x: it is the preimage, the set of all points that, when iterated, converge onto x.

Can the analytic relation be restated in terms of cylinder sets? Yes, and it follows in a fairly natural way. The first step is to extend d_n to a map $d_n : \mathscr{S} \to [0,1]$. Let $\mu : \mathscr{B} \to [0,1]$ be the Bernoulli measure. Using the Bernoulli mapping 3, the interval $[0, T^n(\beta/2)]$ maps to some cylinder; call it Δ_n. Then, given some cylinder $A \in \mathscr{S}$, define

$$d_n(A) = \mu(A \cap \Delta_n)$$

The rotated (pre-)measure is extended likewise:

$$v(A) = \sum_{n=0}^{\infty} \zeta^n d_n(A)$$

with $\zeta = z/\beta$ as before, recovering the Parry measure by setting $z = 1$. The Parry measure should be invariant under the action of $T^{-1} : \mathscr{S} \to \mathscr{S}$, and otherwise yield eqn 24.

Let's check. The proof will mirror the one of the previous section. Here, it is convenient to use the β-transform t instead of the β-shift T. This is primarily a conceptual convenience; the subshift is more easily visualized in terms of the mod 1 map. Otherwise, the same notation is used, but rescaled, so that Δ_n is the cylinder corresponding to the interval $[0, t^n(1)]$.

Recall that for every $A \in \mathscr{B}$ and every $y \in A$, one will find that $y/\beta \in t^{-1}(A)$ and, whenever $y \leq \beta - 1$ that also $(y+1)/\beta \in t^{-1}(A)$. Thus, $t^{-1}(A)$ naturally splits into two parts: the cylinder that maps to $[0, 1/\beta]$, call it D and the complement $\overline{D}$.

The pushforward action is then

$$v(t^{-1}(A)) = \sum_{n=0}^{\infty} \zeta^n \mu(\Delta_n \cap t^{-1}(A))$$

$$= \sum_{n=0}^{\infty} \zeta^n \left[\mu(\Delta_n \cap D \cap t^{-1}(A)) + \mu(\Delta_n \cap \overline{D} \cap t^{-1}(A))\right]$$

Two distinct cases emerge. When $t^n(1) < 1/\beta$ then one has that $\Delta_n \cap \overline{D} = \varnothing$. Thus, the second term can be written as

$$\mu(\Delta_n \cap \overline{D} \cap t^{-1}(A)) = \Theta\left(t_n - \frac{1}{\beta}\right)\mu(\Delta_n \cap \overline{D} \cap t^{-1}(A))$$

$$= \Theta\left(t_n - \frac{1}{\beta}\right)\frac{1}{\beta}\mu(\Delta_{n+1} \cap A)$$

where the second line follows from the first by linearity, and that $\overline{D}$ selected out one of the two branches of $t^{-1}(A)$. Meanwhile, when $t^n(1) > 1/\beta$, then $D \subset \Delta_n$ so that

$D \cap \Delta_n = D$. Thus, the first term splits into two:

$$\Theta\left(t_n - \frac{1}{\beta}\right) \mu\left(\Delta_n \cap D \cap t^{-1}(A)\right) = \Theta\left(t_n - \frac{1}{\beta}\right) \mu\left(D \cap t^{-1}(A)\right)$$

$$= \Theta\left(t_n - \frac{1}{\beta}\right) \frac{1}{\beta} \mu(A)$$

while

$$\Theta\left(\frac{1}{\beta} - t_n\right) \mu\left(\Delta_n \cap D \cap t^{-1}(A)\right) = \Theta\left(\frac{1}{\beta} - t_n\right) \frac{1}{\beta} \mu\left(\Delta_{n+1} \cap A\right)$$

Reassembling these pieces and making use of $\Delta_0 \cap A = A$ one gets

$$\nu\left(t^{-1}(A)\right) = \sum_{n=0}^{\infty} \frac{\zeta^n}{\beta} \left[\mu(A)\Theta\left(t_n - \frac{1}{\beta}\right) + \mu\left(\Delta_{n+1} \cap A\right)\right]$$

$$= \frac{1}{z}\nu(A) - \frac{\mu(A)}{z} + \mu(A)\sum_{n=0}^{\infty} \frac{\zeta^n}{\beta}\Theta\left(t_n - \frac{1}{\beta}\right)$$

$$= \frac{1}{z}\nu(A) + \frac{\mu(A)}{z}E(\beta;z)$$

with the constant term as before, in eqn 24:

$$E(\beta;z) = -1 + \zeta \sum_{n=0}^{\infty} \zeta^n \Theta\left(t_n - \frac{1}{\beta}\right)$$

$$= -1 + \zeta \sum_{n=0}^{\infty} \zeta^n d_n\left(\frac{1}{2}\right)$$
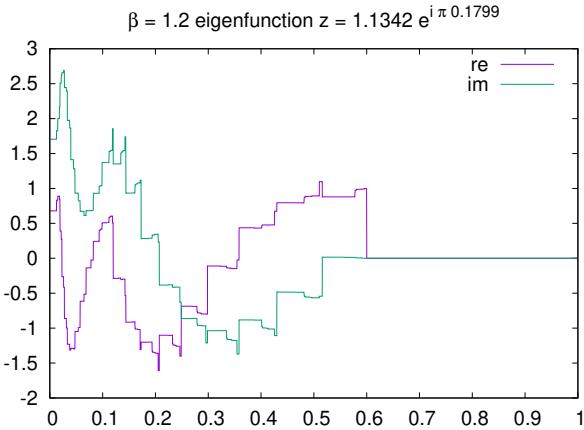
As before, one has for $z = 1$ that $E(\beta;1) = 0$ and so $\nu \circ t^{-1} = \nu$ is indeed the measure invariant under t^{-1}. Other eigenvalues can be found for those values of z for which $E(\beta;z) = 0$. The task at hand is then to characterize $E(\beta;z)$.

3.12 Exploring $E(\beta;z)$

The function is easily explored numerically. It is clearly convergent in the unit disk $|\zeta| < 1$ and has no poles in the disk. For almost all β, there seem to be a countable number of zeros within the disk, accumulating uniformly on the boundary as $|\zeta| \to 1$. An example is shown in figure 15. The notion of "uniformly" will be made slightly more precise in the next section, where it is observed that, for certain special values of β, the bit-sequence $d_n\left(\frac{1}{2}\right)$ is periodic, and thus is a polynomial. When it is polynomial, there are a finite number of zeros (obviously; the degree of the polynomial), which are distributed approximately uniformly near the circle $|\zeta| = 1$. As the degree of the polynomial increases, so do the number of zeros; but they remain distributed approximately evenly. In this sense, the limit of infinite degree seems to continue to hold.

A handful of selected zeros are listed in the table below. The numbers are accurate to about the last decimal place.

Figure 14: Typical Eigenfunction

β = 1.2 eigenfunction z = 1.1342 e^{i π 0.1799}

A typical eigenfunction $v_{\beta;z}(x)$ solving $\mathscr{L}_\beta v = \lambda v$ for $\beta = 1.2$. This is the eigen-function is as defined in eqn 22. This eigenfunction corresponds to the zero $z = 1.1342 \exp i\pi 0.1799$ of $E(\beta;z)$, alternately of $C(\beta;z)$, as defined in eqn 24. The eigen-value is $\lambda = 1/z$. Since the eigenvalue is complex, so is the eigenfunction. The real and imaginary parts are paired in a way that vaguely resembles sine and cosine; such phased offsets are generic.
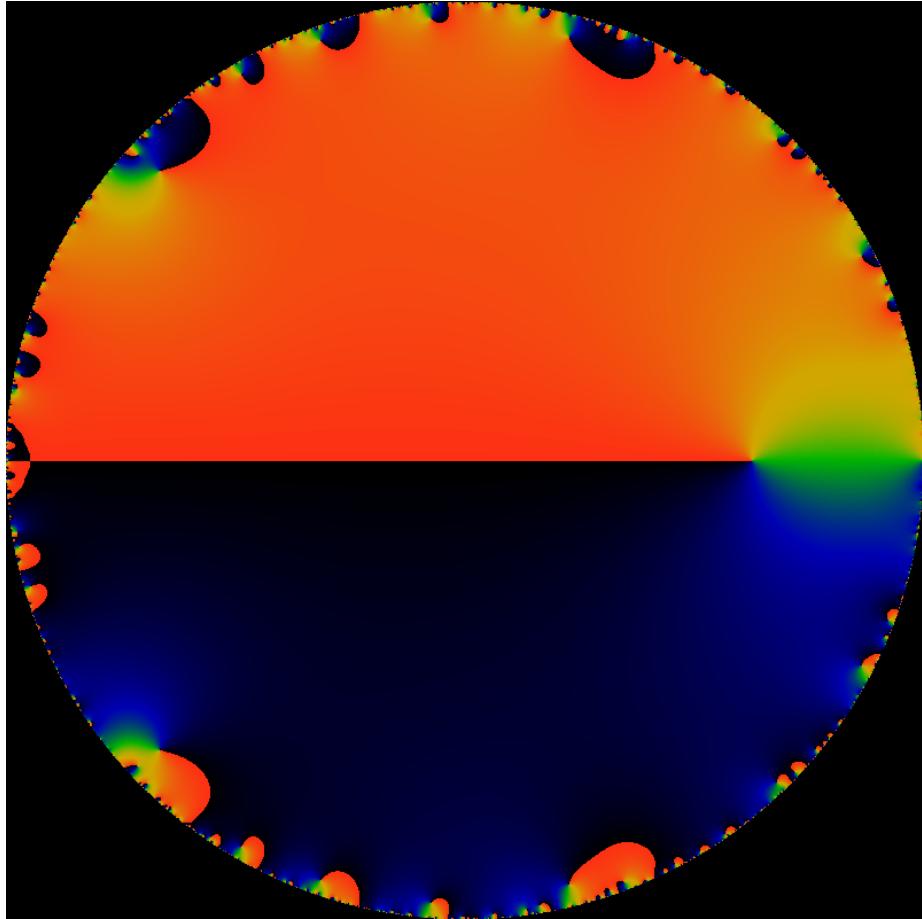
β	z	$|z|$	$1/z$
1.8	-1.591567859	1.59156785	-0.6283112558
1.8	-1.1962384 +i 1.21602223	1.70578321	-0.41112138 - i 0.41792066
1.8	0.99191473 +i 1.44609298	1.75359053	0.32256553 -i 0.47026194
1.6	-1.06365138 +i 1.00895989	1.46606764	-0.49487018 -i 0.46942464
1.4	0.55083643 +i 1.17817108	1.30057982	0.32564816 -i 0.69652119
1.2	0.95788456 +i 0.60733011	1.13419253	0.74462841 -i -0.47211874

These are not particularly meaningful numbers; they just give a flavor for some locations of eigenvalues. Given a zero, the corresponding eigenfunction is also very easily computed. A typical eigenfunction is shown in figure 14; this is for the zero listed in the last row of the table above. Although it is unlike the figure 1, in that it is not strictly decreasing, it does have the same general plateau-like regions. Note that all such eigenfunctions are bounded and generally, differentiable-nowhere.

As the zeros accumulate onto the circle $|\zeta| \to 1$, there appears to be no way to holomorphically continue the function $E(\beta;z)$ outside of the unit circle. This indicates that there is a lower bound on the possible eigenvalues that can be reached via the eigenfunctions of eqn 22.

The apparent reason for this is that the coherent state was constructed as a deformed version of the repeated iteration of $v_0 = 1$ a constant, as described in earlier sections. Iterating on $v_0(y)$ a polynomial appears to generate eigenfunctions with eigenvalue

Figure 15: Disk of $E(\beta;z)$

The above shows a visualization of $E(\beta;z)$ for $\beta = 1.6$ in the complex $\zeta = z/\beta$ plane. The disk consists of all ζ values with $|\zeta| \leq 1$. The plot is a phase plot, showing the argument $\arg E(\beta;z) \in [-\pi, \pi]$. The color-coding of blue-black puts the phase just above $-\pi$, green near 0, and red just below $+\pi$. Locations where the phase wraps around counter-clockwise (right-handed) are zeros of $E(\beta;z)$; this follows from Cauchy's principle. The most prominent zero at the right-hand side of the image corresponds to $z = 1$, located at $\zeta = 1/\beta$ so well inside the circle. This corresponds to the Gelfond–Parry invariant measure. Other zeros are seen at the end of black-red whiskers, wrapping through green. These accumulate at the $\zeta = 1$ boundary; in general, there are a countable number of such zeros. The corresponding eigenvalues are located at $\lambda = 1/z = 1/\beta\zeta$, so the accumulation ring of zeros occurs at $|\lambda| = 1/\beta$. A different presentation of the same information can be found in figure 16.

bounded by $1/\beta^p$ for a polynomial of degree p. One explicit example, generated by a parabola, is given in a much later section. A simple, direct construction remains elusive.

3.12.1 Convergence

The analytic structure of the ergodic disk

$$E(\beta;z) = -1 + \zeta \sum_{n=0}^{\infty} b_n \zeta^n$$

is daunting. Clearly, when the bitsequence b_n is not finite, then there is a pole at $\zeta = 1$. There is also a pole at $\zeta = -1$ unless the number of even and odd bits asymptotically cancel out, and then the series is conditionally convergent.

More generally, the series will be conditionally convergent for $\zeta = e^{i\phi}$ only when the bitsequence b_n is ultimately periodic. In this case, $\zeta = e^{i\phi}$ that are roots of unity can be arranged so that they precisely cancel for each period.

For ergodic bitsequences b_n, it is not clear how to formulate a proper conjecture for the sum.

3.12.2 Jentzsch's theorem

Relevant to the present case is Jentzch's theorem (1914), which states that if

$$g(z) = 1 + a_1 z + \cdots + a_n z^n + \cdots$$

has a radius of convergence of one, then every point on the unit circle is a cluster-point of zeros of partial sums

$$s_n(z) = 1 + a_1 z + \cdots + a_n z^n$$

A proof is given in [30], which generalizes to polynomial sequences that converge on arbitrary regions of the complex plane (having positive capacitance).

Perhaps a simpler statement is that this follows from Rouché's theorem, in that $1 - z^n$ will have n roots, located at the roots of unity, and all the other terms only perturb the location of the roots.

3.13 Unitary Component

Do the zeros merely accumulate onto the circle $|\zeta| \to 1$, or are any of them exactly on $|\zeta| = 1$? If they are exactly on $|\zeta| = 1$, then $\mathscr{M} = \mathscr{N} \oplus U$ for some unitary $U = \exp iH$.

Lets plug through and see what happens. We had

$$E(\beta;z) = -1 + \zeta \sum_{n=0}^{\infty} b_n \zeta^n$$

with $b_n = d_n\left(\frac{1}{2}\right)$ due to notational confusion. Write $z = e^{i2\pi\phi}/\beta$ so that $\zeta = e^{i2\pi\phi}$. Then the task is to find solutions

$$1 = e^{i2\pi\phi} \sum_{n=0}^{\infty} b_n e^{i2n\pi\phi} \tag{25}$$

for real ϕ, if any. From general principles, such a ϕ must be irrational, as otherwise otherwise, it would be a root of unity, and the sum would diverge whenever the sequence b_n is infinite. For finite-length sequences, the above corresponds to the usual polynomials, $p_n \left(e^{-i2\pi\phi}/\beta \right) = 0$ and we already know that these do not have any solutions for real ϕ.

Periodic sequences also cannot have solutions; these reduce to the case of the finite polynomials. For example, $b_{2n} = 1$ and $b_{2n+1} = 0$ reduces the above to $\zeta^{-2} - \zeta^{-1} - 1 = 0$ which is just the usual $\zeta = 1/\beta$ which has no solutions for $|\zeta| = 1$. All periodic sequences work this way. What about the ultimately-periodic sequences, which have a chaotic initial segment, followed by a periodic tail? Those won't work either. They are described by a polynomial in ζ and in order for that polynomial to have a solution $|\zeta| = 1$, it would have to have a factor of $\zeta^k - 1$ for some integer k. But we already know that the ultimately periodic sequences do not have such a factor.

That leaves the possibility that only the chaotic bit-sequences might have such solutions. The prospects look dim. The periodic sequences are dense in β and the reals are seperable, and so behavior is dominated by the periodic sequences; this is a generic theorem; what is it's name? The theorem is that when something is dense in a seperable space, then the thing that is dense governs the rest. But this theorem can be violated; clearly, the Minkowski Question Mark violates this theorem. Hmm. Confusing. I've got forgotten theorems mixed up.

3.13.1 Existence

How do we go about finding solutions to the above?

Well, the bitsequence b_n must belong to the infinite comb (below). What's more is that this is a Fourier transform of the infinite comb. What's more is that the periodic orbits are "almost solutions", in that they do accumulate onto the circle $1/\beta$ without actually lying on the circle. So the limit is going in the correct direction for such solutions to exist. We must now prove that, for a valid infinite sequence b_n, the truncated finite sequence $\check{b}_n^k$ has roots $r_k \to e^{i2\pi\phi}/\beta$ for a fixed ϕ. Hmm. But it is clear that there are countably many such roots. As the bitsequence lengthens, the number of roots increases by one; each root approaches the circle and so approaches a limit point. These limit poaints are dense on the unit circle. The limit points are the desired ϕ. So this is the proof sketch. QED.

Here, the truncated finite sequence is $\check{b}_n^k = b_n$ for $k < n$ and $\check{b}_k^k = 1$ for $k = n$ and $\check{b}_n^k = 0$ for $k > n$. The condition that $\check{b}_k^k = 1$ for $k = n$ arises from the bracketing condition on the polynomials: the final bit must always be one.

As a result, we now have a countable number of ϕ satsifying eqn 25 and these are dense on the unit circle. Based on general experience, these will be uniformly distributed. We know that $\phi = 0$ is not in the set. We don't have an explicit equation giving valid ϕ other than the vague presecription of "find roots numerically and take limits numerically" which is unsatisfying. In particular, we don't have a Hamiltonian, yet; we just have a proof that it exists. What is the Hamiltonian?

3.13.2 Estimates

Given a sequence b_n truncated to length k, there will be a convergent bracket $\ell \mapsto m \Leftarrow \rho$ lead by the polynomial p_m such that

$$2m+1 = b_0 b_1 b_2 \cdots b_k \tag{26}$$

$$p_m(x) = x^{k+1} - \sum_{n=0}^{k} b_n x^{k-n}$$

$$E_k(\zeta) = -1 + \zeta \sum_{n=0}^{k} b_n \zeta^n = -\zeta^{k+1} p_m\left(\frac{1}{\zeta}\right)$$

Suppose that ζ_k is a complex root, so that $E_k(\zeta_k) = 0$. We now want the estimate $\zeta_{k+1} = \zeta_k(1 + \varepsilon_k)$ where it is assumed that ε_k is small. Well, $E_{k+1}(\zeta) = -\zeta^{k+2} p_{m'}$ where $m' = 2m$ if $b_{k+1} = 0$ else $m' = 2m+1$. This follows from eqn 31 which takes the form

$$p_{m'}(x) = \begin{cases} x(p_m(x) + 1) - 1 & \text{for } b_{k+1} = 1 \\ x p_m(x) - 1 & \text{for } b_{k+1} = 0 \end{cases}$$

Thus

$$0 = p_{m'}\left(\frac{1}{\zeta_{k+1}}\right)$$

$$= \frac{1}{\zeta_{k+1}}\left(p_m\left(\frac{1}{\zeta_{k+1}}\right) + b_{k+1}\right) - 1$$

$$\approx \frac{1}{\zeta_k}(1 - \varepsilon_k)\left(p_m\left(\frac{1}{\zeta_k}(1 - \varepsilon_k)\right) + b_{k+1}\right) - 1$$

$$\approx \frac{1}{\zeta_k}(1 - \varepsilon_k)\left(-\varepsilon_k p_m'\left(\frac{1}{\zeta_k}\right) + b_{k+1}\right) - 1$$

$$\approx -\varepsilon_k\left(\frac{1}{\zeta_k} p_m'\left(\frac{1}{\zeta_k}\right) + b_{k+1}\right) + \frac{1}{\zeta_k} b_{k+1} - 1$$

so that

$$\varepsilon_k \approx \frac{b_{k+1} - \zeta_k}{p_m'\left(\frac{1}{\zeta_k}\right) + \zeta_k b_{k+1}}$$

where p'_m denotes the derivative:

$$p'_m(x) = \frac{d}{dx}\left(x^{k+1} - \sum_{n=0}^{k} \check{b}_n x^{k-n}\right)$$

$$= (k+1)x^k - \sum_{n=0}^{k-1} \check{b}_n (k-n) x^{k-n-1}$$

$$= k\left(x^k - \sum_{n=0}^{k-1} \check{b}_n x^{k-n-1}\right) + x^k + \sum_{n=1}^{k-1} \check{b}_n n x^{k-n-1}$$

$$= \frac{k}{x}\left(x^{k+1} - \sum_{n=0}^{k} \check{b}_n x^{k-n} + \check{b}_k\right) + x^k\left(1 + \sum_{n=1}^{k-1} \check{b}_n n x^{-n-1}\right)$$

$$= \frac{k}{x}\left(p_m(x) + 1\right) + x^k\left(1 + \sum_{n=1}^{k-1} b_n n x^{-n-1}\right)$$

Above $\check{b}$ is used instead of b because $\check{b}$ is appropriate for the convergent. Note that $\check{b}_n = b_n$ for $n < k$ while $\check{b}_k = 1$ so that the series is correctly terminated to get a proper, valid polynomial.

Thus

$$p'_m\left(\frac{1}{\zeta_k}\right) = k\zeta_k + \frac{1}{\zeta_k^k}\left(1 + \sum_{n=1}^{k-1} b_n n \zeta_k^{n+1}\right)$$

Keep in mind that $|\zeta_k| \approx 1$, so it is almost pure phase. So indeed, ε_k is small, but it is hard to tell which way it is pointing. We want it to be pointing mostly radially outward.

So this isn't going anywhere.

The other issue is there's no insight into where new zeros are sprouting. Time for numerics.

3.13.3 Graphs
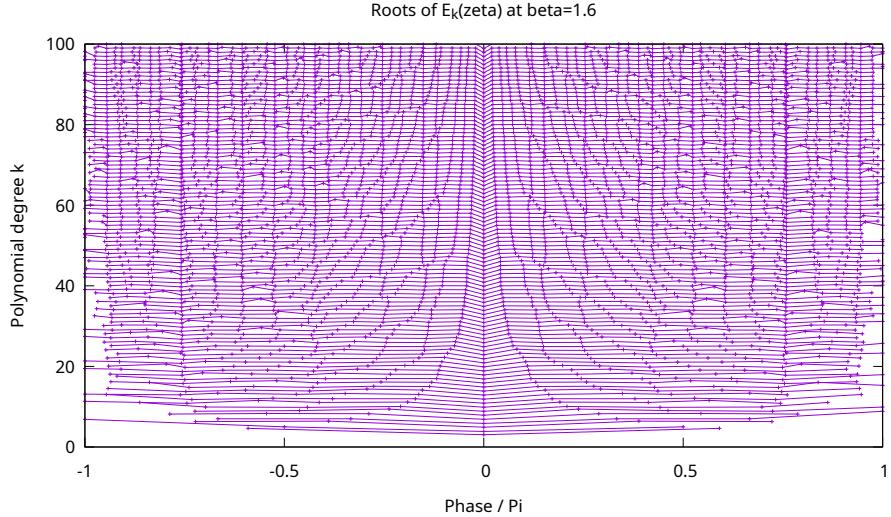
A "typical" example of the complex roots of $E_k(\zeta)$ is shown in figure 16. The value of β is held fixed at $\beta = 1.6$; this generates an infinite sequence of b_n that can be truncated at any point, to get a polynomial $E_k(\zeta)$ of finite degree $k+1$.

There is a curiosity in this figure worth discussing. The locations of the roots appear to follow continuous trajectories, as $k \to \infty$. Now, formally, this is "incorrect", in that each root is discrete, and each k is an integer. Yet, it would appear that it should somehow be possible to interpolate to fractional k. How would this work? It seems that we fall off the map of conventional mathematics by asking this question.

Another notable aspect to this figure is that the traceries of the roots forma a weaving pattern, resembling ripples in lakeshore sand. Such rippling is not uncommon in analytic functions derived from arithmetic series commonly studied in number theory. A similar rippling can be seen in the polylogarithm. Where do the ripples come from?

Oh, there is a simple answer for the rippling. To first order, the location of the zeros are given by Jentzsch's theorem, which are roots of unity, in this case: $1 - \zeta^k = 0$ because $b_k = 1$. They are perturbed by b_{k-1} but if this is zero, nothing happens. Next

Figure 16: Complex Roots of $E_k(\zeta)$ for $\beta = 1.6$

Roots of E$_k$(zeta) at beta=1.6

This figure shows the location of all roots to the polynomial $E_k(\zeta)$ for $3 \le k \le 100$. The roots are plotted in the form $re^{i\phi}$ with the phase ϕ running from $-\pi$ to $+\pi$ on the horizontal axis. The vertical axis shows $r + k$, so that for each degree k the roots are vertically offset. Each root is marked with a cross; horizontal line segments connect successive roots coming from the same polynomial. The roots accumulate to fixed locations as $E_k \to E$, with the limit $E(\beta; z)$ shown in figure 15.

---

is b_{k-2} and so on. The first one that is not vanishing bumps, and creates that ripple. This seems plausible, except that the ripples do not quite align.

Lets pursue the wild idea above, as far as we can. It would seem that there is some function of two variables, $F(\eta, \zeta)$ such that for for integer $\eta = n$, it interpolates the polynomials $E_n(\zeta)$, so that $F(n, \zeta) = E_n(\zeta)$. We want this interpolation to be smooth in some way, so that, for a given root ζ_α of $E_n(\zeta_\alpha) = 0$ it is possible to track the location of ζ_α as a smooth function of η. That is, $\zeta_\alpha(\eta)$ should be smooth, or as close to being smooth as we can make it.
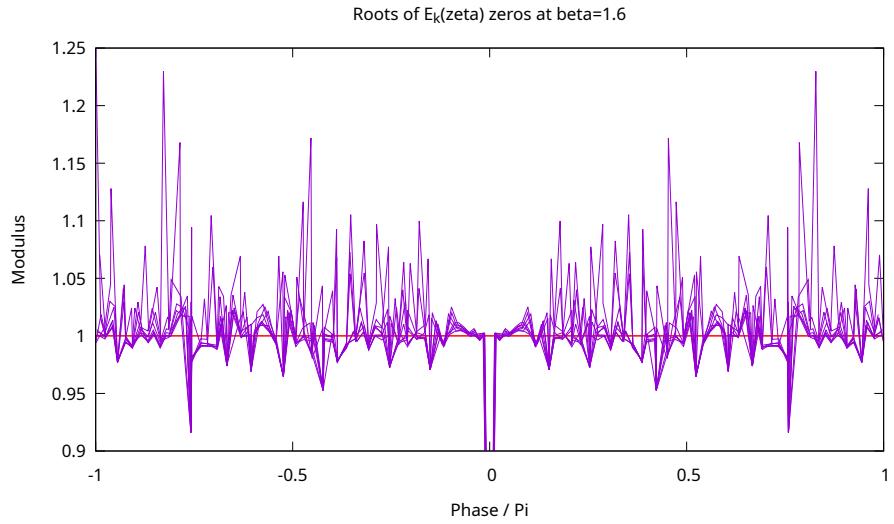
Should $\zeta_\alpha(\eta)$ be real-analytic in η? What about $F(\eta, \zeta)$? One conjecture would be to write

$$F(\eta, \zeta) = \sum_{j=0}^{\infty} b_j(\eta)\zeta^j$$

with $b_j(\eta)$ being functions of η, having all the correct limits for $\eta = n$. But this seems strange, because for $\eta = n$ the function $F(\eta, \zeta)$ is a polynomial, having $n + 1$ complex roots ζ_α, while, for $\eta \neq n$, it is holomorphic in ζ and thus has a countably infinite number of roots in ζ. So these come and go.

For $\eta = n$, the $b_j(\eta)$ are of course integers, and not just integers, but bits, zero or

Figure 17: Complex Roots of $E_k(\zeta)$ for $k \geq 120$

This figure shows the location of all roots to the polynomial $E_k(\zeta)$ for $k = 120 + 5j$ and $0 \leq j \leq 7$. The roots are plotted in the form $re^{i\phi}$ with the phase ϕ running from $-\pi$ to $+\pi$ on the horizontal axis and the modulus r on the vertical axis. Line segments are used to join each successive root, for fixed k. Thus, each root lies at the tip of a sharp spike. There is a suggestive accumulation, but it remains difficult to describe it. This is a large-k version of figure 16, with the offset removed.

one. Is it necessary to assume that $b_j(\eta)$ are smooth functions of η? Or can they be discontinuous, yet resulting in $F(\eta, \zeta)$ that is smooth in η?

Is it reasonable to assume that $b_j(\eta)$ are real-analytic? I guess they would have to be, if we want $b_j(\eta)$ to depend smoothly on η. Thus

$$b_j(\eta) = \sum_{i=0}^{\infty} b_{ji}\eta^i$$

Of couse, the b_j are also functions of β.

Consider $\zeta_\alpha(\eta)$. Since $\zeta_\alpha(\eta) \to \text{const}$ in the limit of $\eta \to \infty$ it seems that the correct expansion is either that

$$\zeta_\alpha(\eta) = \text{const} + \frac{a}{\eta} + \frac{b}{\eta^2} + \cdots$$

or that

$$\zeta_\alpha(\eta) = \text{const} + e^{-a\eta^2}\left(b + c\eta + d\eta^2 + \cdots\right)$$

Are either of these reasonable? The second form is particularly beguiling, as it is reminiscent of the simple harmonic oscillator.

But how would this even work? For $\eta = n$, we have that

$$E_n(\zeta) = A \prod_{\alpha=0}^{n} (\zeta - \zeta_\alpha)$$

and the constant $A = 1$ since $b_n = 1$. For $\eta \neq n$, this has to be of the form

$$E_\eta(\zeta) = A_\eta(\zeta) \prod_{\alpha=0}^{\lfloor \eta \rfloor} (\zeta - \zeta_\alpha)$$

such that

$$A_n(\zeta) = 1$$

while $A_\eta(\zeta) \to (\zeta - \zeta_n)$ as $\eta \to n$ from below. So, for $n - 1 \leq \eta < n$ there is a linear interpolation

$$A_\eta(\zeta) = (1 + \eta - n)(\zeta - \zeta_n) + n - \eta$$

or perhaps a (smooth) interpolating function $f(x)$ with $f(0) = 0$ and $f(1) = 1$ and

$$A_\eta(\zeta) = f(1 + \eta - n)(\zeta - \zeta_n) + f(n - \eta)$$

So we have that A_η is this sawtooth, as a function of η and it is constantly sprouting new monomials, and each sprouted monomial... well, its a blue-sky catastrophe; first there is not any root ζ_n and then all of sudden, there is. So we want some Morse-theory-like thing to make these roots appear out of thin air. How? Is there any existing theoretical framework for this?

3.14 Convolutions

Given any polynomial $p(x) = \sum_{i=0}^{N} a_k x^i$ with roots $p(x) = 0$, the result of inserting a root into the series $E(\zeta) = -1 + \sum_{n=0}^{\infty} b_n \zeta^{n+1}$ can be understood as a convolution, in that $x^N = \frac{1}{a_N} \sum_{i=0}^{N-1} a_i x^i$ can be plugged in, replacing any given power ζ^{N+k} with a sequence of lower powers. This can also be done for any x^i for $1 \leq i \leq N$. Write $E *_i p$ for the result of this substitution, performed once, one every ζ^i. That is, substitute $\zeta^{i+k} \mapsto a_i^{-1} \zeta^k \sum_{j \neq i} a_j \zeta^j$. This results in a formal series $E *_i p$ with the property that $[E *_i p](x) = E(x)$ whenever x satisfies $p(x) = 0$. As this can be done repeatedly and recursively, one obtains a large equivalence class of formal series having this property. This can be thought of as a series resummation technique, but one that is valid only when $p(x) = 0$.

The question is then: are any of these rewritten series absolutely convergent (holding $p(x)$ fixed)? Naively, one might guess this is the empty set, but this is not obvious. Alternately, are there any $p(x)$ for which one the rewritten series becomes absolutely convergent?

3.14.1 Fourier series

To make the above more concrete, define as above

$$E_k(\zeta) = -1 + \zeta \sum_{n=0}^{k} \check{b}_n \zeta^n$$

Just as before, the above is defined with $\check{b}$ is used instead of b, because $\check{b}$ is appropriate for the convergent. Note that $\check{b}_n = b_n$ for $n < k$ while $\check{b}_k = 1$ so that the series is correctly terminated to get a proper, valid polynomial.

Consider the sequence $E_k(1)$ as well as $E_k(\exp i2\pi/m)$ or perhaps just generally $E_k(\exp i2\pi x)$.

3.15 Function field arithmetic

In algebraic number theory, one studies function fields over fields of characteristic p for p a prime number. Multiplication of a field element by an integer n can be taken as the beta transform, with $\beta = n$ integer, acting on the field. Here of course we are working with a field of characteristic zero (the reals), or more properly, the sigma algebra of the Cantor set. None-the-less, what sort of inspiration can be taken?

This is spurred by the observation that the mapping $\mathscr{L}_T : \mathscr{F} \to \mathscr{F}$ given by $\mathscr{L}_T f = f \circ T^{-1}$ is richer than just just a linear operator. That is, not only does one have the obvious identity $\mathscr{L}_T(af + bg) = a\mathscr{L}_T f + b\mathscr{L}_T g$ for functions f, g and constants a, b, but one also has $\mathscr{L}_T(fg) = \mathscr{L}_T f \cdot \mathscr{L}_T g$ and $\mathscr{L}_T(f/g) = \mathscr{L}_T f / \mathscr{L}_T g$ so that if $\mathscr{F}$ is a function field, then $\mathscr{L}_T$ preserves that structure.

Additional algebraic identities follow from the sigma algebra of the product topology. The product topology implies that one can work with function spaces, e.g. measures, for which set intersection and set union behave in a field-like way.

If e belongs to the discrete or continuous spectrum of $\mathscr{L}_T$, in the sense that $\mathscr{L}_T e = \lambda e$, then it also belongs to the spectrum of any iterate of $\mathscr{L}_T$, in the sense that $\mathscr{L}_T^n e =$

$\lambda^n e$, and so the $\mathscr{L}_T^n$ form a basis for a vector space. By linearity, one can arrange for operators of arbitrary spectrum, in that one has (at least formally) that

$$\left(\sum_{n=0}^{\infty} a_n \mathscr{L}_T^n \right) e = \sum_{n=0}^{\infty} a_n \lambda^n e$$

for some arbitrary series of numbers a_n (presumably convergent, in order to get reasonable results).

3.16 Iterated transfer operator

To understand the nature of the steady-state solution (the Frobenius-Perron eigenstate), its is worth iterating on the recurrence relation for it, by hand, the first few times. To do this, it is convenient to write it in the form

$$\left[\mathscr{L}_\beta f \right] (y) = \frac{\Theta(y)}{\beta} \left[f(\alpha(y)) + f(\omega(y)) \right]$$

where $\Theta(y) = 1$ if $y \le \beta/2$ else zero; this is a step function to denote the vanishing for the operator for $2y > \beta$. (This differs from the use of Θ as the Heaviside step function in earlier sections; the intent is the same, but the goal is to have a briefer notation here. Which is which should be clear from context.) The functions $\alpha(y) = y/\beta$ and $\omega(y) = \frac{1}{2} + \alpha(y)$ are convenient shorthands for symbolic iteration.

Iterating once gives

$$\left[\mathscr{L}^2 f \right] (y) = \quad \frac{\Theta(y)}{\beta^2} \quad \left[\Theta(\alpha(y)) \left[f(\alpha^2(y)) + f((\omega \circ \alpha)(y)) \right] + \right.$$
$$\left. \Theta(\omega(y)) \left[f((\alpha \circ \omega)(y)) + f(\omega^2(y)) \right] \right]$$
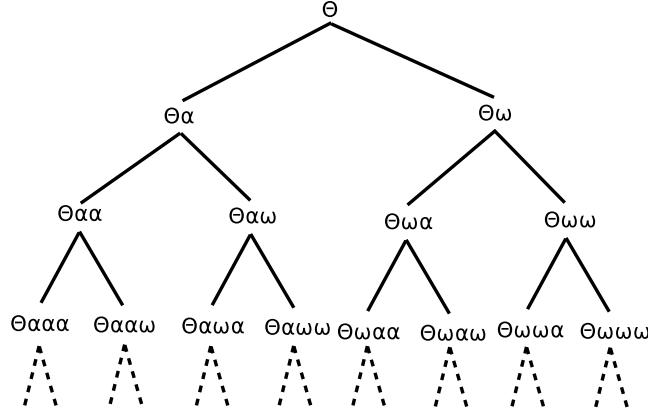
Using a simplified notation $g(y) = f(\alpha(y)) + f(\omega(y))$ allows this to be iterated a third time:

$$\left[\mathscr{L}^3 f \right] (y) = \quad \frac{\Theta(y)}{\beta^3} \quad \left[\Theta(\alpha(y)) \left[\Theta(\alpha^2(y)) g(\alpha^2(y)) + \Theta(\omega\alpha(y)) g(\omega\alpha(y)) \right] + \right.$$
$$\left. \Theta(\omega(y)) \left[\Theta(\alpha\omega(y)) g(\alpha\omega(y)) + \Theta(\omega^2(y)) g(\omega^2(y)) \right] \right]$$

and a fourth time, this time omitting the argument, and the various nesting parenthesis.

$$\left[\mathscr{L}^4 f \right] (y) = \quad \frac{\Theta(y)}{\beta^4} \quad \left[\Theta\alpha\Theta\alpha^2 \left[\Theta\alpha^3 g\alpha^3 + \Theta\omega\alpha^2 g\omega\alpha^2 \right] + \right.$$
$$\Theta\alpha\Theta\omega\alpha \left[\Theta\alpha\omega\alpha g\alpha\omega\alpha + \Theta\omega^2\alpha g\omega^2\alpha \right]$$
$$\Theta\omega\Theta\alpha\omega \left[\Theta\alpha^2\omega g\alpha^2\omega + \Theta\omega\alpha\omega g\omega\alpha\omega \right]$$
$$\left. \Theta\omega\Theta\omega^2 \left[\Theta\alpha\omega^2 g\alpha\omega^2 + \Theta\omega^3 g\omega^3 \right] \right]$$

Notice that the primary structure is given by a product of step functions. This is more conveniently visualized as a tree:

For any given iteration, the result is the sum of the vertexes at a given level, while the product of step functions is the product of the step functions in the tree, following the path to each node. Because any particular step function might be zero, it effectively acts to cut off the tree at that location. It is therefore interesting to understand general products of the α and β functions.

It is convenient to define

$$\gamma_x(y) = \frac{x}{2} + \frac{y}{\beta}$$

so that $\alpha(y) = \gamma_0(y)$ and $\omega(y) = \gamma_1(y)$, so that a general iterated sequence of inter-mixed α's and ω's can be written uniformly in terms of γ. Given a sequence of bits $b_0 b_1 b_2 \cdots b_n$ with each b_k being either zero or one, the iterated sequence of functions can be written as

$$(\gamma_{b_0} \gamma_{b_1} \gamma_{b_2} \cdots \gamma_{b_n})(y) = \frac{1}{2}\left[b_0 + \frac{b_1}{\beta} + \frac{b_2}{\beta^2} + \cdots + \frac{b_n}{\beta^n} \right] + \frac{y}{\beta^{n+1}} \tag{27}$$

So, for example:

$$\alpha^n(y) = \frac{y}{\beta^n}$$

while
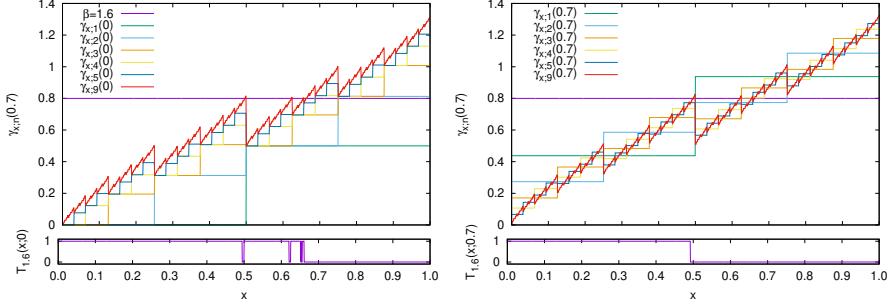
$$\omega^2(y) = \frac{1}{2} + \frac{1}{\beta}\left(\frac{1}{2} + \frac{y}{\beta} \right)$$

and, in general, that

$$\omega^n(y) = \frac{1}{2}\left[1 + \frac{1}{\beta} + \frac{1}{\beta^2} + \cdots \frac{1}{\beta^{n-1}} \right] + \frac{y}{\beta^n}$$

Iterated sequences of pairs of functions, of the form $\gamma_{b_0} \gamma_{b_1} \gamma_{b_2} \cdots \gamma_{b_n}$ are reminiscent of de Rham curves, which generalize Cesaro curves and the Koch snowflake. The proper definition of a de Rham curve assumes the sequence is of infinite length, and applies a certain continuity condition, and is generally carried out on the complex plane, so that a continuous, nowhere-differentiable curve results. Here, the curve is distinctly not continuous: eqn 27 is a finite-length form of the shift series 6 which can be visualized as the expander function pdr 13, as shown in figure 9.

Figure 18: Gamma functions

Examples of "typical" gamma functions. Both figures show gamma functions for $\beta = 1.6$; the one on the left shows them for $y = 0$, while the one on the right shows them for $y = 0.7$. Every gamma function is a sequence of plateaus; the zig-zag line is a high-order gamma, essentially showing the limiting case. The tree function is unity whenever all of these curves are below $\beta/2$, and is zero when above. So, for example, for the left figure, the tree function is unity, for all values of x less than about 0.4952; it drops to zero, then returns to unity above $x = 0.5$, until about 0.6221, when it briefly plunges and rises again. Then, another dip, before finally settling to zero near 0.6541. For the right figure, a high-order zig-zag rises above 0.8 somewhere near 0.4914; then $\gamma_{x;1}(0.7)$ rises above 0.8 and stays there, driving the tree function to zero, rendering all other orders irrelevant.

3.17 The Tree Function

Given a bit sequence (b_k) and value for y, define the tree function as
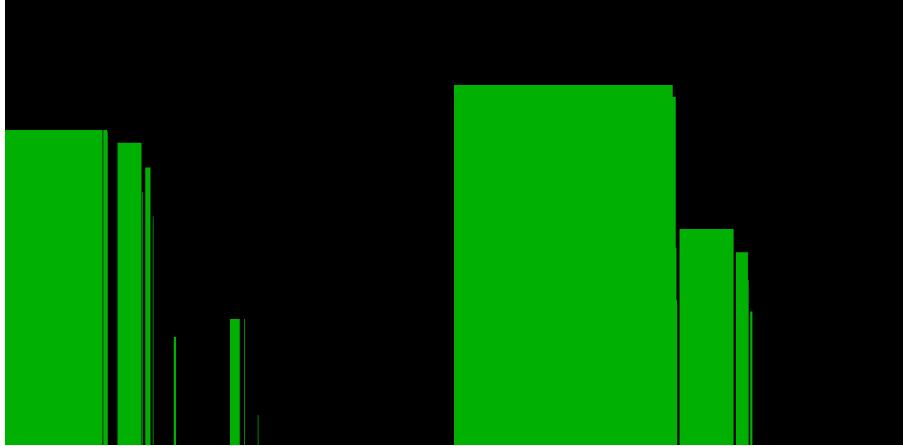
$$T_\beta\left((b_k);y\right) = \Theta(y) \prod_{n=0}^{\infty} \Theta\left(\gamma_{b_0}\gamma_{b_1}\gamma_{b_2}\cdots\gamma_{b_n}(y)\right)$$

For any given fixed sequence of bits and value of y, this function is either zero or one. One way to understand this function is to ask how it varies for fixed β and y, but with the bit sequence coming from the Bernoulli shift of eqn 2, so that $b_n = b_n(x)$. This simplifies notation, so that one can write

$$T_\beta(x;y) = T_\beta\left((b_k(x));y\right) = \Theta(y) \prod_{n=0}^{\infty} \Theta\left(\gamma_{x;n}(y)\right)$$

with $\gamma_{x;n}(y) = \gamma_{b_0}\gamma_{b_1}\gamma_{b_2}\cdots\gamma_{b_n}(y)$. Its clear that the tree function has maximum support when $y = 0$. Figure 18 shows several gamma functions, and the corresponding tree function that results. Figure 19 shows the x vs. y behavior of the tree functions. Figure 20 shows the β vs. x behavior of the functions. Figure 21 shows a unified visualization of the three preceding charts.

95

Figure 19: Tree functions

The above illustrate the y vs. x dependence of the tree functions; the left image is for $\beta = 1.4$, the right is for $\beta = 1.6$. Green indicates the regions where the tree function is unity, and black where it is zero. To be clear, this shows $T_\beta(x; y)$ with x and y plotted along the x and y axes. The tree functions shown in figure 18 are just two horizontal slices taken from the right image: a slice along the bottom, and a slice a bit above the middle.

3.18 Haar Basis Matrix Elements

The symmetric Haar wavelets are built from the mother wavelet

$$h(x) = \begin{cases} 1 & \text{for } 0 \le x < 1/2 \\ -1 & \text{for } 1/2 \le x < 1 \end{cases}$$

and has individual wavelets given by

$$h_{nj}(x) = 2^{n/2} h(2^n x - j) \ \text{ for } 0 \le j \le 2^n - 1$$

The matrix elements of the transfer operator are

$$\langle mi | \mathscr{L}_\beta | nj \rangle = \int_0^1 h_{mi}(x) \left[\mathscr{L}_\beta h_{nj} \right](x)\, dx$$
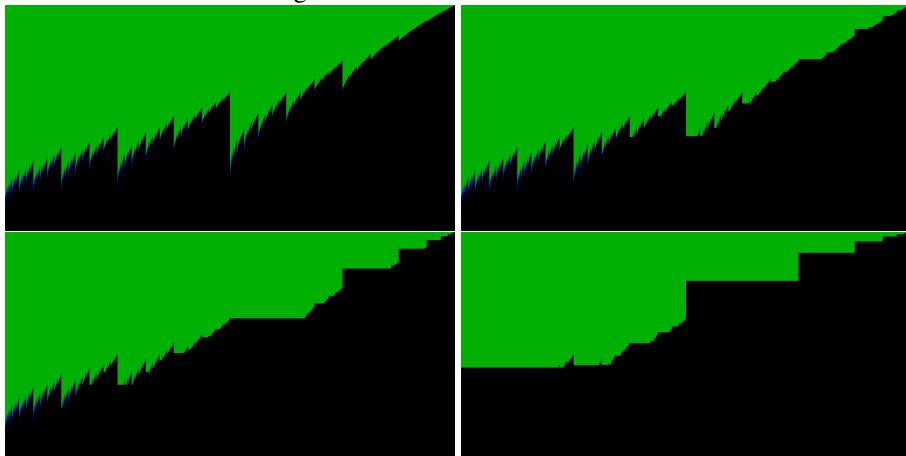
where the operator $\mathscr{L}_\beta$ is given by eqn 17. Computing these by hand promptly pushes into a big mess. One can obtain explicit expressions, just that they are tedious to obtain. Some preliminary observations include that

$$\langle mi | \mathscr{L}_\beta | nj \rangle = 0 \ \text{ if } \beta \le i/2^{m-1}$$

because the transfer operator vanishes above $\beta/2$. In the same vein, matrix elements vanish unless
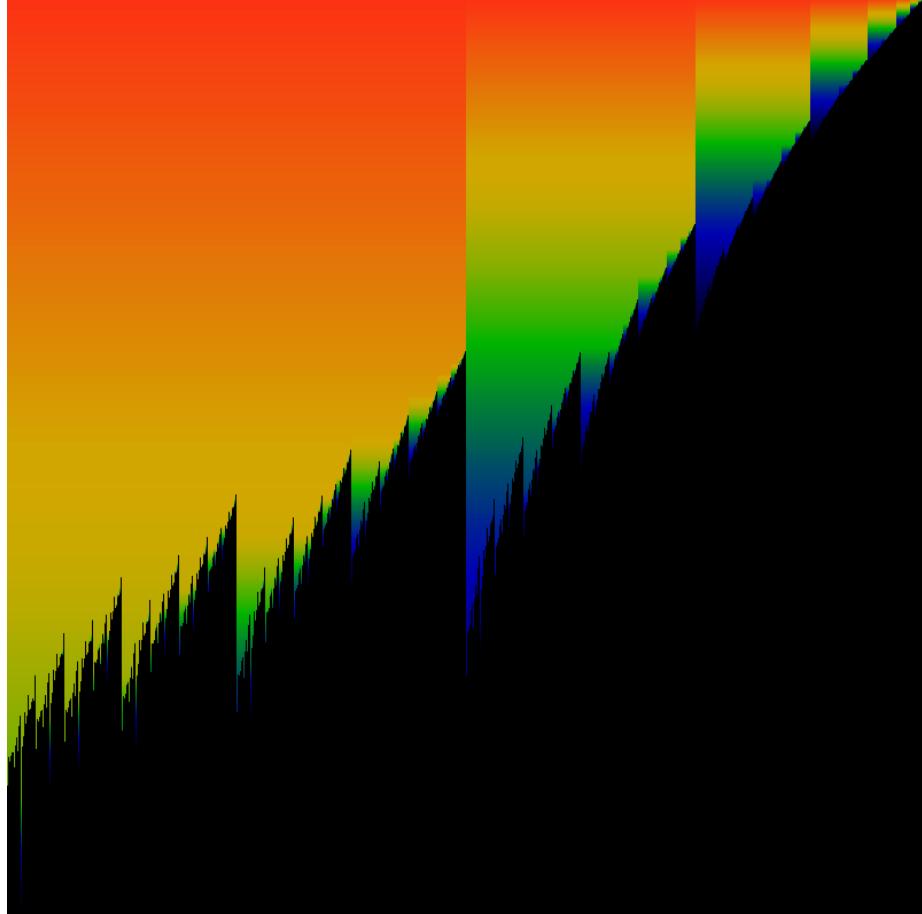
$$\left[\frac{i}{2^m}, \frac{i+1}{2^m} \right] \cap \left[\frac{\beta j}{2^n}, \frac{\beta(j+1)}{2^n} \right] \ne \emptyset$$

Figure 20: Tree function variations

These figures illustrate the β vs. x dependence of the tree function. The upper left shows $T_\beta(x;0)$, the upper right shows $T_\beta(x;0.3)$, the lower left shows $T_\beta(x;0.5)$, the lower right shows $T_\beta(x;0.7)$. In each case, x runs from 0 to 1 along the x axis, while β runs from 1 to 2 along the vertical axis. As before, green indicates where the tree function is unity, and black where it is zero. The tree functions shown in figure 18 correspond to horizontal slices in the first and last images. Note that many (possibly all??) of the green spikes in the upper-left image reach all the way down to the bottom, although they are mostly much thinner than a pixel and thus not rendered. The vague blue hazing near the spikes is an attempt at anti-aliasing, to highlight the sharpness.

Figure 21: Tree function Unified Visualization

This figure presents a unified visualization of figures 18, 19 and 20. That is, it depicts the $T_\beta(x; y)$ varying all three parameters. The parameter β runs from 1 at the bottom, to 2 at the top. The parameter x runs from 0 to 1, left to right. Because $T_\beta(x; y)$ is either zero or one, the color is used to represent the largest value of y for which $1 = T_\beta(x; y)$. The color coding corresponds to red for $y = 1$, green for $y = 0.5$, blue for $y = 0.25$ and black for $y = 0$. Thus, for example, figure 20 can be obtained directly from this, by setting a given color, "or darker", to black. The figure 19 represents a single fixed horizontal slice through this figure, with the height of the rectangles in figure 19 corresponding to the color in this figure.

or if

$$\left[\frac{i}{2^m}, \frac{i+1}{2^m}\right] \cap \left[\beta\left(\frac{j}{2^n} - \frac{1}{2}\right), \beta\left(\frac{j+1}{2^n} - \frac{1}{2}\right)\right] \neq \emptyset$$

In all other cases, the Haar wavelets completely fail to overlap, and thus the matrix elements are zero. In addition, only three pairs of wavelets overlap in a non-zero fashion. That is, for a fixed m, n and j, there are at most six different values of i for which the matrix elements are non-vanishing: the first three of these are the values for which

$$\frac{\beta j}{2^n} \in \left[\frac{i}{2^m}, \frac{i+1}{2^m}\right] \quad \text{or} \quad \frac{\beta\left(j+\frac{1}{2}\right)}{2^n} \in \left[\frac{i}{2^m}, \frac{i+1}{2^m}\right] \quad \text{or} \quad \frac{\beta(j+1)}{2^n} \in \left[\frac{i}{2^m}, \frac{i+1}{2^m}\right]$$

and likewise for three more. The observation is that the integral vanishes unless the first wavelet intersects an edge transition of the second wavelet.

The primary failure of this basis is that there is no obvious way to diagonalize the transfer operator in this basis. There is no obvious way of solving it, of finding it's eigenfunctions and eigenvalues, other than by brute-force numerical attack.

3.19 Julia Set
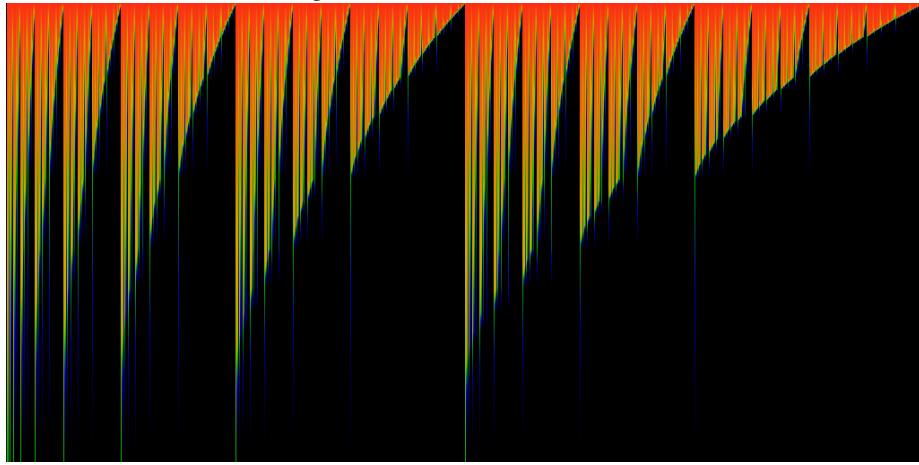
Consider the two iterators $a_0(y) = \min\left(\frac{\beta}{2}, \beta y\right)$ and $a_1(y) = \max\left(0, \beta y - \frac{\beta}{2}\right)$. Individually, they are the two arms of the beta shift. Here, they have been separated from each other, so that the full domain $0 \leq y \leq 1$ is allowed. Exploring all possible inter-iterations for these gives the Julia set for the transfer operator: it indicates where a point "came from", for the iterated transfer operator. There are several related ways to visualize this. One way is to fix y and then, given a bit-sequence (b_n) to compute

$$j((b_n)) = a_{b_o} \circ a_{b_1} \circ a_{b_2} \circ \cdots (y)$$

Figure 22 shows a visualization for finite bit-sequences: in essence, the very first few iterations. Although it is similar to figure 10, it is not the same.

For a related notion, consider the definition of "laps", from Jeffrey Lagerias *etal*.[25].

Figure 22: Julia Set visualization

Consider the binary tree of dyadic fractions: that is, the tree whose rows are 1/2, (1/4 3/4), (1/8 3/8 5/8 7/8), ... Consider a function J on this tree. For the head of the tree, set $J(1/2) = \beta$. For the next row, set $J(1/4) = a_0(J(1/2))$ and $J(3/4) = a_1(J(1/2))$. Iterate in this fashion so that $J((2k-1)/2^{n+1}) = a_0(J(k/2^n))$ and $J((2k+1)/2^{n+1}) = a_1(J(k/2^n))$ recursively. This produces a function J taking values on every dyadic fraction $k/2^n$.

In the above figure, β runs from 1 at the bottom to 2 at the top. A single horizontal slice through the image shows a color-coded version of J, with red coding values near 1, green coding values near 1/2 and blue, fading to black coding values of 1/4 and less. Note that there are many faint blue lines that extend quite far down, but not all the way down: these form a stair-step. The image is 1024 pixels wide: it shows the first ten rows of the binary tree. Although this image is similar to figure 10, it differs in many details.

4   Hessenberg basis

There is a set of Haar-like wavelets in which the transfer operator is of the form of a Hessenberg operator - that is, the operator becomes almost upper-diagonal, with only one diagonal, just below the main diagonal, that is non-zero. Explicitly, the transfer operator $\mathscr{L}_\beta$ has matrix entries $\left[\mathscr{L}_\beta\right]_{ij}$ such that $\left[\mathscr{L}_\beta\right]_{ij} = 0$ whenever $i > j+1$. A matrix having this form is called a Hessenberg matrix; such matrices have various interesting properties; among others, they generalize the Jacobi matrix. This chapter explicitly constructs an infinite-dimensional Hessenberg matrix, which may now be called a Hessenberg operator.

Hessenberg operators occur naturally in spectral measure theory; some of this will be reviewed in several later chapters. To get a flavor for what is to come: Given a Hessenberg operator, one may construct a system of orthogonal polynomials that provide a basis for square-integrable holomorphic functions on some domain of the complex plane. Such a space is called a Bergman space; in this sense it generalizes the Jacobi operator for real Borel measures. This basis of polynomials in turn allows the Hessenberg operator to be explicitly seen as a shift operator on that domain, with $[Hf](z) = zf(z)$ for H the Hessenberg operator and $f(z)$ a holomorphic function (specifically a Bergman function) on the Bergman domain. But all of this is for later chapters; its mentioned here only to whet the appetite.

TODO: More motivation is needed for this line of reasoning. After all, we have an explicit solution for the eigenfunctions, it is given in expression 21. So, why bother writing the transfer operator as a Hessenberg operator? What does this teach us?

4.1   Hessenberg wavelet basis

The transfer operator $\mathscr{L}_\beta$ can be fairly easily brought into Hessenberg matrix form. A sequence of of orthonormal functions is constructed in this section; when used as a basis, the transfer operator becomes almost upper-diagonal.

The trick to the construction is to define wavelets such that the transfer operator applied each wavelet causes the end-points of the wavelet to exactly line up with the end- or mid-points of previous wavelets, thus avoiding the nasty interval-overlap algebra required with the Haar basis. This is accomplished by carefully picking the midpoint of the next wavelet in the sequence to be located exactly at the discontinuity of the transfer operator applied to the previous wavelet.

The construction proceeds as follows. Let

$$\psi_0(x) = \begin{cases} \frac{1}{\sqrt{\beta/2}} & \text{for } 0 \leq x \leq \beta/2 \\ 0 & \text{for } \beta/2 < x \leq 1 \end{cases}$$

Consider $\mathscr{L}_\beta \psi_0$. It is the sum of two parts: two step-functions; one which is constant for $x \leq \beta/2$ and another that is constant for $\frac{x}{\beta} + \frac{1}{2} \leq \frac{\beta}{2}$. Solving explicitly for the location of the step, it is $x = \beta(\beta-1)/2$. For convenience, define $m_1 = \beta(\beta-1)/2$ and $m_0 = \beta/2$. These will anchor a series of midpoints, beginning with $m_{-1} = 0$. Using

the midpoint m_1, construct the wavelet

$$\psi_1(x) = \begin{cases} \frac{1}{m_1}\sqrt{\frac{m_1(m_0-m_1)}{m_0}} & \text{for } 0 \le x \le m_1 \\ \frac{-1}{m_0-m_1}\sqrt{\frac{m_1(m_0-m_1)}{m_0}} & \text{for } m_1 < x \le m_0 \\ 0 & \text{for } m_0 < x \le 1 \end{cases}$$

Note that this is normalized to unit length: $\int_0^1 |\psi_1(x)|^2 dx = 1$ and that it is explicitly orthogonal to the first: $\int_0^1 \psi_1(x)\psi_0(x)dx = 0$.

Consider $\mathscr{L}_\beta \psi_1$. As always, it is the sum of two parts. The midpoint of ψ_1 is at $m_1 = \beta(\beta-1)/2$ and this mid-point is mapped to one of two different places. If $m_1 < 1/2$ then it is mapped to $m_2 = \beta m_1$ else it maps to $m_2 = \beta(m_1 - 1/2)$. Thus, if $m_1 < 1/2$, define

$$\psi_2(x) = \begin{cases} 0 & \text{for } 0 \le x \le m_1 \\ \frac{1}{(m_2-m_1)}\sqrt{\frac{(m_2-m_1)(m_0-m_2)}{m_0-m_1}} & \text{for } m_1 \le x \le m_2 \\ \frac{-1}{(m_0-m_2)}\sqrt{\frac{(m_2-m_1)(m_0-m_2)}{m_0-m_1}} & \text{for } m_2 < x \le m_0 \\ 0 & \text{for } m_0 < x \le 1 \end{cases}$$

else define

$$\psi_2(x) = \begin{cases} \frac{1}{m_2}\sqrt{\frac{m_2(m_2-m_1)}{m_1}} & \text{for } 0 \le x \le m_2 \\ \frac{-1}{(m_1-m_2)}\sqrt{\frac{m_2(m_2-m_1)}{m_1}} & \text{for } m_2 \le x \le m_1 \\ 0 & \text{for } m_1 < x \le 1 \end{cases}$$

Because each end of the interval on which ψ_2 is non-zero lies entirely within one of the constant arms of ψ_1, one has, by construction, that $\int_0^1 \psi_2(x)\psi_1(x)dx = 0$ (and, of course, $\int_0^1 \psi_2(x)\psi_0(x)dx = 0$.)

The rest of the basis can be constructed iteratively, based on these examples. The midpoints are given by iterating 4 on $m_0 = \beta/2$, so that $m_p = T_\beta(m_{p-1}) = T_\beta^p(m_0)$ is the p'th iterate of $\beta/2$. Let m_l be largest midpoint smaller than m_p (and $l < p$); let m_u be the smallest midpoint larger than m_p (and $l < p$). Let $m_{-1} = 0$ initiate the sequence by providing the smallest-possible "midpoint"; $m_0 = \beta/2$ already provides the largest possible.

Then define

$$\psi_p(x) = \begin{cases} 0 & \text{for } 0 \le x \le m_l \\ \frac{C_p}{(m_p-m_l)} & \text{for } m_l \le x \le m_p \\ \frac{-C_p}{(m_u-m_p)} & \text{for } m_p < x \le m_u \\ 0 & \text{for } m_u < x \le 1 \end{cases} \tag{28}$$

By construction, this has the property that $\int_0^1 \psi_{p+1}(x)\psi_n(x)dx = 0$ for any $n < p+1$. The normalization constant is

$$C_p = \sqrt{\frac{(m_p - m_l)(m_u - m_p)}{m_u - m_l}}$$

which is determined by requiring that $\int_0^1 \left| \psi_p(x) \right|^2 dx = 1$.

4.2 Matrix Elements

The above-defined basis provides the Hessenberg representation for the transfer operator. Defining

$$\langle n | \mathscr{L}_\beta | m \rangle = \int_0^1 \psi_n(x) \left[\mathscr{L}_\beta \psi_m \right](x) \, dx \tag{29}$$

this has the expected Hessenberg form, in that

$$\langle n | \mathscr{L}_\beta | m \rangle = 0 \quad \text{for } n > m+1$$

This is just one diagonal short of being actually solvable. A visualization of the matrix elements is shown in figure 23.

4.3 Completeness

The Hessenberg basis construction gives a countable set of ψ_n that is an orthonormal basis on the unit interval: $\int_0^1 \psi_m(x) \psi_n(x) \, dx = \delta_{mn}$. Are they complete? Obviously the $\{\psi_n\}$ cannot be complete on the unit interval, as they all vanish for $\beta/2 < x$. Perhaps they are complete on the interval $[0, \beta/2]$, where they are already orthonormal: $\int_0^{\beta/2} \psi_m(x) \psi_n(x) \, dx = \delta_{mn}$.

A numerical exploration shows that the midpoints m_p are dense in the interval $(0, \beta/2)$, and so this suggests that the basis should be considered to be "sufficiently complete" on the interval $[0, \beta/2]$. The distribution of the m_p follow exactly the distribution of the invariant measure. Convergence is uniform to the same degree that the midpoints "eventually" fill in and become dense in some interval. Renyi[2] and Parry[3] do more: they show that the midpoint process is ergodic (Parry points out that it's weakly mixing), and provide a formal proof that the distribution is one and the same as the invariant measure.
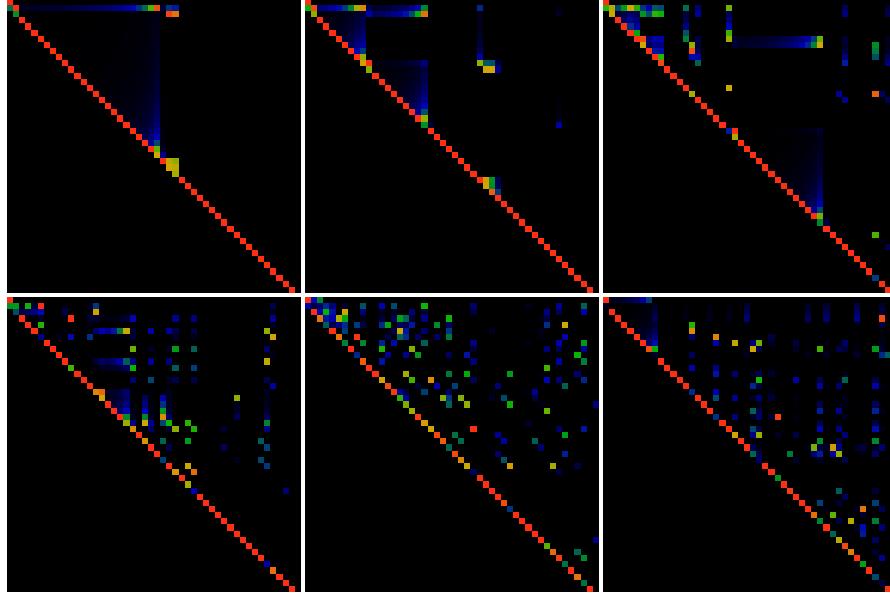
The above has some exceptions: there are some values of β for which the midpoint m_0 iterates $x = 1/2$, whereupon iteration stops (i.e. iterates to zero), or becomes cyclic (forming a periodic orbit). Which is which depends on how the point $1/2$ is treated by the map. These values of β are potential "trouble spots", and are explored in greater detail in the next chapter. They are dense in the interval $1 < \beta < 2$, but they form a countable set that can be taken to be of measure zero. Thus, most "most" values of β are not problematic. Excluding the trouble spots, the Hessenberg basis can be taken to be complete.

Clearly, the ψ_n span some subspace; do they span the Hilbert space $L_2[0, \beta/2]$ of square-integrable functions on the interval $[0, \beta/2]$? To what degree can one legitimately write

$$\delta(y-x) = \sum_{n=0}^{\infty} \psi_n(y) \psi_n(x)$$

as a resolution of the identity?

Figure 23: Hessenberg Operator Matrix Elements

Six illustrations of the absolute value of the matrix elements $\langle n | \mathscr{L}_\beta | m \rangle$ for the transfer operator $\mathscr{L}_\beta$ for (left to right, top to bottom) $\beta =$ 1.1, 1.2, 1.3, 1.6, 1.90, 1.998 and $0 \le n, m < 48$ in the Hessenberg basis. The red color represents values of 0.66 or larger, green represents values of 0.33 and blue and darker correspond to 0.16 or less. Almost all matrix elements are in fact precisely zero; black pixels in these images correspond to matrix elements that are zero. Note that the almost all of the diagonal matrix elements are exactly zero: that is $\langle n | \mathscr{L}_\beta | n \rangle = 0$ for most n. The bright-red pixels are just below the diagonal: for most n, one has that $\langle n+1 | \mathscr{L}_\beta | n \rangle \gtrsim 0.5$ with the occasional blueish pixel suggesting a smaller value. These two, taken together, suggests that the eigenvalue spectrum is rapidly decreasing. The first few images suggests a regular pattern that gets increasingly compressed and chaotic as β increases. More-or-less the sane structure prevails if one instead zooms out to look at the 600x600 submatrix; animating with fine-grained steps in β does not result in an interesting animated movie.

The question of completeness dogs some "obvious" assumptions one wants to make. For example, if the set of states is complete, and the resolution of the identity holds, then one expects that the transfer operator resolves to the iterated function:

$$\delta\left(y - (\beta x \mod 1)\right) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \psi_n(y) \left\langle n \left| \mathscr{L}_\beta \right| m \right\rangle \psi_m(x)$$

It is fun to verify that the world works as one expects it to work: the above can be verified to hold numerically, for sums limited to a finite cutoff.

4.4   Numerical Eigenvalues

Given the apparent sparsity visible in figure 23, one might think that the eigenvalue problem is fairly stable, numerically. It is not all that much. Numerical exploration suggests that the spectrum is on or near a circle lying in the complex plane[2], of radius $|\lambda| = 1/\beta$ (ignoring, that is, the leading eigenvalue of 1, which is easily found).

To be clear, this is a numerical exploration of the $N \times N$ principle submatrix of $\left\langle n \left| \mathscr{L}_\beta \right| m \right\rangle$. The eigenvalue problem being posed is to find a vector $\vec{v} = (v_k)_{k=0}^{N}$ that solves

$$\sum_{m=0}^{N} \left\langle n \left| \mathscr{L}_\beta \right| m \right\rangle v_m = \lambda v_n$$

for some constant λ (with the set of possible λ depending on N, of course).

There are various pitfalls in extrapolating from this to the $N \to \infty$ limit. For the next few paragraphs, consider only some notion of a "minimal" extension from finite N to the limit. That is, for each finite N, one has a finite set of eigenvalues and eigenvectors. In the limit, there may be accumulation points: points where the eigenvalues accumulate to a limit point, in a standard topological sense. What should that that topological space be? For finite N, all eigenvectors are explicitly summable, and thus can be taken to belong to any Banach space ℓ_p. One may as well take $p = 2$ the Hilbert space, and normalize the eigenvectors $\vec{v}$ so that $1 = \sum_{m=0}^{N} v_m^2$.

For finite N, it appears that "most" eigenvalues λ are "near" the circle $|\lambda| = 1/\beta$, and that they seem to be very uniformly distributed around this circle. The numerical results indicate that in the limit $N \to \infty$, that the scare-quotes "most" becomes "almost all" in the usual sense. Similarly, "near" appears to mean that for any given λ at finite N, one has that $|\lambda| - 1/\beta \sim \mathscr{O}(1/N)$. As to uniformity, it seems that the spacing between nearest neighbors is also $\mathscr{O}(1/N)$, and that there are no "premature" accumulation points: eigenvalues never get any closer than $\mathscr{O}(1/N)$, either.

Thus, the minimal closure, the minimal extrapolation to limit points strongly suggests that the limit points really do lie, uniformly distributed, on the circle $|\lambda| = 1/\beta$.

---

[2]This was confirmed with both GSL gsl_eigen_nonsymmv() and Lapack DHSEQR solvers, exploring the principle submatrix of various sizes, up to about 2000×2000 entries. Both systems agree to at least six decimal places, if not more. Both show sporadic eigenvalues off the circle, but these are not numerically stable; ergo, the only valid eigenvalues are those on the circle. The matrix entries were constructed using the midpoint algorithm, described in the last section. To verify that they are accurate, several techniques were used: numerical integration to confirm orthogonality, and the use of the GnuMP multi-precision library to push up accuracy.

Then, writing a given accumulation point as $\lambda = \beta^{-1} \exp 2\pi i \phi$, what the numerics do not reveal, or, at least, do not easily reveal, is whether the allowed values of ϕ are always rational, irrational or might have some other regular structure. The numerical exploration does suggest that the eigenvalues are dense on the circle. Certainly it is the case Hessenberg basis is countable, an so one would expect the eigenvalue spectrum obtained in this way to be at least countable, as well. Whether it is also uncountable seems unknowable in this naive sense.

This question is interesting because if only rational ϕ are allowed, then the decaying eigenfunctions belong to a cyclic group, and exhibit an exact form of Poincaré recurrence as they decay. If irrational ϕ are allowed, then the decaying eigenfunctions are at least ergodic.

For $\beta = 2$, the β-transform is the Bernoulli shift, the transfer operator is solvable, and the spectrum is exactly known. This has been explored by various authors[31]. I've written extensively about this spectrum and the eigenvalues in other texts[32, 33, 34]. To recap, it takes several forms, depending on the function space that one chooses to work in. If one restricts oneself to polynomial eigenfunctions, then the spectrum is real, non-negative (it has an extensive kernel) and has eigenvalues of 2^{-n} for all n. The eigenfunctions are the Bernoulli polynomials. Restricting to square-integrable eigenfunctions, the spectrum continuous, having eigenvalues on the unit disk in the complex plane. The continuous-spectrum eigenfunctions (for eigenvalues other than 2^{-n}) can be understood in several ways: if forced to be differentiable, then they are not bounded (they diverge) at the endpoints of the interval. If forced to be bounded, then they are fractal (non-smooth) over the entire interval. The unitary spectrum corresponds to differentiable-nowhere eigenfunctions (wait, or continuous-nowhere? I forget.)

A pair of plausible, numerically-extracted eigenfunctions are shown in image 24.

Presumably, the spectrum can be related to the lap-counting function, given by Lagarias[25].
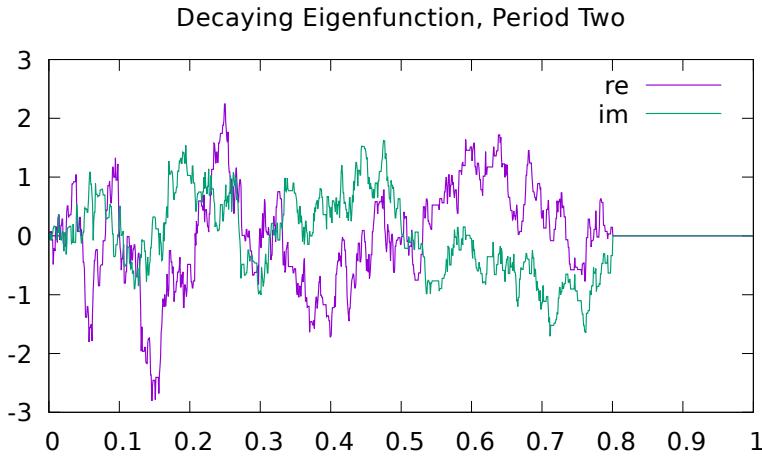
4.5   (Non-)Unitarity

The numerical results suggest a hypothesis that perhaps some fragment of $\mathscr{L}_\beta$ is unitary, as it is ordinarily the case that when eigenvalues appear on the unit circle, it is because an operator is unitary. That does not seem to be the case here. Specifically, define the Frobenius-Perron eigenvector ρ as the one satisfying $\mathscr{L}_\beta \rho = \rho$ and normalizing it to unit length, so that $\|\rho\| = 1$ in the Hilbert (mean-square) norm. Define the reduced operator $\mathscr{R}_\beta$ in terms of the matrix elements

$$\frac{1}{\beta} \langle n | \mathscr{R}_\beta | m \rangle = \langle n | \mathscr{L}_\beta | m \rangle - \langle \rho | n \rangle \langle \rho | m \rangle$$

That is, it is just the beta shift operator, with the Frobenius-Perron eigenvector removed, so that $\mathscr{R}_\beta \rho = 0$. Its rescaled, so that the remaining eigenvectors of $\mathscr{R}_\beta$ lie on the unit circle. Is this operator unitary in any way? That is, might either $\mathscr{R}_\beta \mathscr{R}_\beta^\dagger$ or $\mathscr{R}_\beta^\dagger \mathscr{R}_\beta$ be the identity? Here, the dagger $\dagger$ is just the transpose, as $\mathscr{R}_\beta$ is purely real. Numerical exploration clearly shows that $\mathscr{R}_\beta$ is neither unitary on the left nor on the right. Not a surprise, but does leave the understanding of $\mathscr{L}_\beta$ in a curious situation.

Figure 24: Decaying Eigenfunction, Period Two

Decaying Eigenfunction, Period Two

This shows a numerically-computed decaying eigenfunction of period two, for $\beta = 1.6$. It is period two, in that it corresponds to an eigenvalue of $\lambda = -1/\beta = -0.625$, so that after one iteration of $\mathscr{L}_\beta$, the sign flips. This can be confirmed, numerically: after one iteration, the sign really does flip, to within numerical errors. This was computed by numerically diagonalizing the 861×861 matrix given by the lowest terms of eqn 29, and then graphing the eigenvector closes to $\lambda = -0.625$ (The GnuMP library was used to provide the required level of precision in the calculations.)

Although this figure is drawn with curves labeled "real" and "imaginary", this is a bit fantastic, and is a numeric artifact. For any period-two eigenfunction, the real and imaginary parts would have no coupling, and would be independent of each other; either one could be set to zero and one would still have a valid eigenfunction. This differs from the case of period-three and higher, where the real and imaginary parts are expected to mix. (Nor are the two components orthogonal, as one might expect.) The eigenfunction is also fantastic in that only slightly different numerics result in a completely different eigenfunction being computed. Even the functions resulting from diagonalizing the 863×863 matrix differ fair amount from those arising from the 861×861 matrix; there's only a general resemblance. This is not entirely surprising: the magnitude of the basis coefficients decays very slowly; even at 861, that are still on the order of 10^{-3}, and thus contribute strongly.

Computed eigenfunctions for period-three are not dissimilar, nor are the ones for other values of β. They do seem to start having the general oscillatory character of $\sin(1/x)$ as $\beta \to 1$, but its not clear if this is a numeric artifact, or something real. The wildness of these functions contrast sharply with the seemingly tame $\lambda = 1$ eigenfunctions shown in figure 1. Compare to figure 14, which paves the way.

Perhaps it is not enough to subtract the invariant measure: The zeros of the formula 24 lying inside the disk must be subtracted as well. There seems to be a countable number of these; the subtraction won't be straight-forward.

4.6 Invariant Measure

Let v_n be the Ruelle-Frobenius-Perron eigenvector in the Hessenberg basis. That is, let v_n be the vector that solves

$$\sum_{m=0}^{\infty} \langle n | \mathscr{L}_\beta | m \rangle v_m = v_n \tag{30}$$

This is readily computed numerically, and it is straightforward to verify the numerics by confirming that

$$\rho(x) = \sum_{m=0}^{\infty} v_m \psi_m(x)$$

is the invariant measure of equations 18,19, with the $\psi_k(x)$ being the wavelets of eqn 28. This expansion seems to "make sense", as the discontinuities seen in the graph of $\rho(x)$ in figure 1 occur at exactly the midpoints m_p and the size of each discontinuity appears to get smaller as p gets larger. Given that the wavelet $\psi_p(x)$ has its central discontinuity at m_p and is bounded on left and right by midpoints of lower order, this expansion seems to be very natural. This is supported by the diagram 25, which depicts the values of v_n as a function of n for selected values of β. These values of v_n are real, positive, and quickly get small; there are no difficulties or issues of convergence.

Is there some simple expression for the values of v_n as a function of β? If so, it must be formed using some sort of fractal shift. Figure 26 illustrates v_1 through v_5.

The orbit of the midpoint is correlated with value of the coefficients, illustrated in figure 27. The midpoint polynomial for $m_p = T_\beta^p(\beta/2)$, given in eqn 40, is compared to $v_m \beta^m$. It can be seen to "line up". The two are somehow related; its not clear just how.
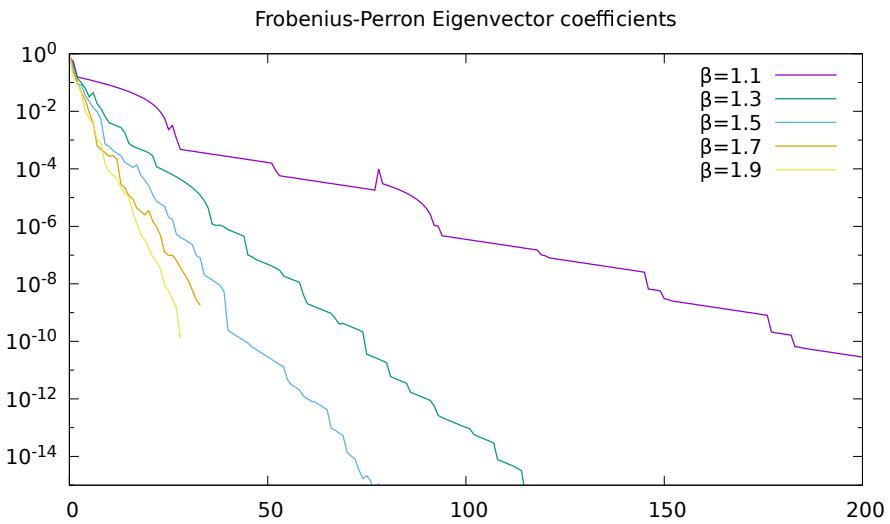
4.7 Generating Function

The truncated ordinary generating function associated with the eigenvector of eqn 30 is

$$G_N(z) = \sum_{m=0}^{N} v_m z^m$$
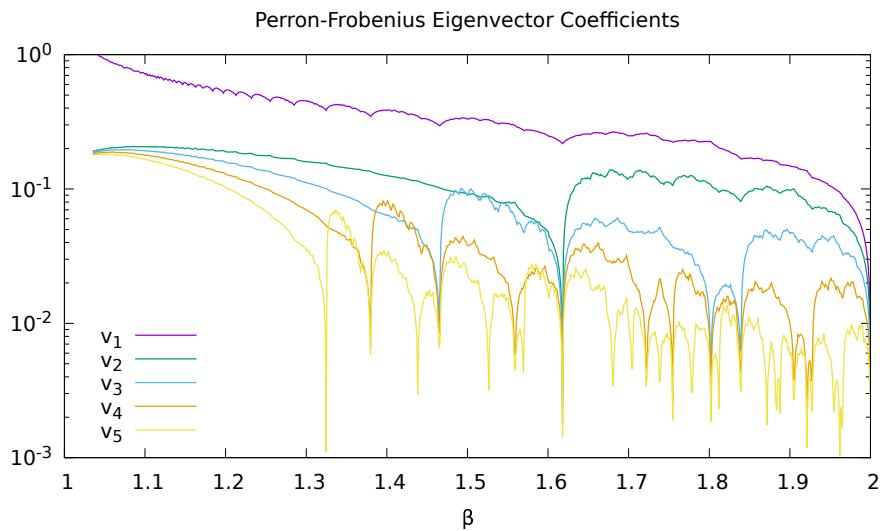
with the ordinary generating function being the limit $N \to \infty$. A numerical study of this function indicates that most of the N zeros of G_N are arranged approximately on a circle of radius β. The arrangement appear to be quite uniform, with more-or-less equidistant spacing of the zeros. As N increases, it seems that more of the zeros get closer to the circle, although the numerical instabilities associated with double-precision math make this difficult to control; an arbitrary-precision eigenvalue solver would be needed to confirm this behavior.

Figure 25: Frobenius-Perron Eigenvector Coefficients

Frobenius-Perron Eigenvector coefficients
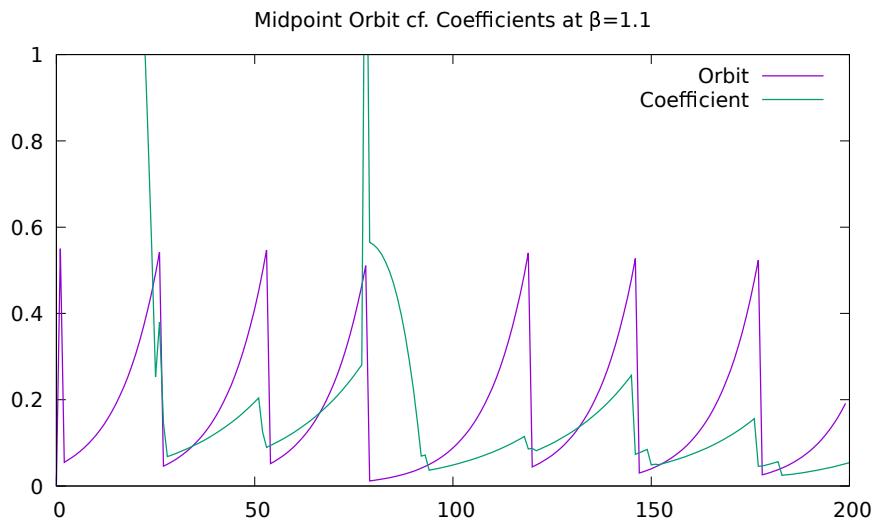
The coefficients v_n solving eqn 30 as a function of n, for various values of β. Note that the coefficients are all real and positive. These can be obtained in two different ways: either by numerically diagonalizing the matrix equation of 30 or by numerically integrating $\int_0^1 \rho(x) \psi_n(x) \, dx$. Either method gives the same results; diagonalization is far, far quicker. The slope appears to go as approximately $v_m \sim C\beta^{-m}$ with $C = 0.02$ roughly.

Figure 26: Perron-Frobenius Eigenvector Coefficients

This figure shows v_1 through v_5 as β is varied. The most prominent spike is located at $\beta = \varphi = 1.618\cdots$ the Golden Ratio. All spikes correspond to orbits that terminate in a fixed point after a finite number of iterations. The root cause and location of the spikes is shown in figure 28.

Figure 27: Orbits and Coefficients

Midpoint Orbit cf. Coefficients at β=1.1

This figure compares the midpoint orbit to the coefficients, providing evidence for the hypothesis stated in the text. The midpoint orbit is just $m_p = T_\beta^p (\beta/2)$. Because $\beta = 1.1$ in this figure, the discontinuities are infrequent and appear to be quasi-regular (they are ultimately fully chaotic), as the midpoint mostly just walks up to where it is knocked down again. The "coefficient" curve is a graph of $10 v_p \beta^p$ for p running along the horizontal axis. This is the same v_p as discussed in the text, and previously shown in figure 25. Here, its rescaled by its asymptotic behavior, and a constant of 10 to place it on the same vertical scale. The discontinuities clearly line up. The relationship is clearly non-trivial.

If this behavior persists, and it seems that it will, then the limit $N \to \infty$ cannot be taken, and the ordinary generating function, as an analytic function, can't exist, *per se*, as it would be uniformly zero inside the disk. Thus, the zeros already found by means of eqn 24 seem to come to the rescue: these are located inside the disk; perhaps these are masquerading as "numerical instabilities", and should be taken as actually existing, and not spurious.

In the next chapter, it will be seen that circles of zeros in the complex plane is a recurring theme. This suggests a hypothesis that somehow it might hold that

$$\sum_m^N v_m (\beta z)^m \sim z^{N+1} - \sum_k b_k z^k$$

as both sides have zeros arranged in circles of unit radius. The right hand side is defined and explored in detail in the next chapter. Superficially, this hypothesis is clearly false: coefficients on the left are all real and positive; coefficients on the right - the b_k, are bits, either zero or one. Yet both exhibit a circle of zeros.

XXX This section is awkward. Revise it or cut it.

4.8   Givens rotations

An open question: A Hessenberg matrix can be brought to solvable form by applying a sequence of Givens rotations. Is the sequence of angles that appear in these rotations meaningful in any way, or are they just some form of uninteresting junk?

5 Finite Orbits

The iteration of the midpoint $m_0 = \beta/2$, that is, the iterated series $m_n = T_\beta^n(\beta/2)$ is ergodic in the unit interval, for almost all values of β. However, for certain values of β, the midpoint iterates will hit the point $x = 1/2$ where the β-shift map has a discontinuity. Here, iteration stops: at the next step, this point is defined to iterate to zero, in eqn 4. Zero is a fixed point, and so there is nowhere further to go. This section explores these special values of β.

Aside from the definition in eqn 4, one can consider the modified map, where the less-than sign has been altered to a less-than-or-equals:

$$T_\beta^{\leq}(x) = \begin{cases} \beta x & \text{for } 0 \leq x \leq \frac{1}{2} \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} < x \leq 1 \end{cases}$$

In this map, the point $x = 1/2$ iterates to $\beta/2$, which is just the initial midpoint itself. In this case, the halted orbits become periodic orbits. There is a third possibility, to simply remove the points 0, 1 and 1/2 from the domain:

$$T_\beta^{<}(x) = \begin{cases} \beta x & \text{for } 0 < x < \frac{1}{2} \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} < x < 1 \end{cases}$$

In this case, if the midpoint iterates to 1/2, it can be taken to simply have wandered out of the domain of validity. The word "wander" is used here in the technical sense: The map $T_\beta(x)$ is dissipative in two different senses: first, in the obvious sense, the points $\beta/2 < x$ wander away after exactly one iteration; secondly, the mid-point iterates wander down to zero, never to return, unless they are made explicitly periodic with $T_\beta^{\leq}(x)$.

The β values at which the midpoint has a periodic or terminating orbit will be called "trouble spots", for lack of a better term. They can be imagined to be prototypes of a bifurcation point, "depending delicately on initial conditions": where two choices are possible, depending on infinitesimally small perturbations of m_0, or, alternately, of β.

All three variants can be considered together, so that the "true" beta shift is taken as the quotient space or identification space[35] of the three variants, in the strict topological sense of a quotient space. Thus, interestingly, for the beta shift, the periodic orbits and the fixed point both belong to the same equivalence class. This has some interesting implications when one compares the beta shift to other iterated maps, such as the logistic map, which have nontrivial stable regions. Topologically, it would seem that one can perform a kind of surgery, attaching stable regions exactly into those spots where, in the beta shift, one has an equivalence class. This solves (at least for me) the longstanding problem of exactly how to properly describe the topological conjugacy between different kinds of iterated maps.

5.1 The β-generalized Golden Ratio

Trouble spots occur whenever the p'th iterate $m_p = T_\beta^p(m_0)$ lands at the discontinuity, so that one may take either $m_p = 0$ or $m_p = m_0$. The iteration immediately before corresponds to $m_{p-1} = 1/2$. The length of the orbit is p.

The first finite orbit can be found when $\beta = \varphi = \left(1 + \sqrt{5}\right)/2$ the Golden Ratio. In this situation, one has that $m_0 = \varphi/2$ and $m_1 = 1/2$. The length of the orbit is $\nu = 2$. For $\nu = 3$, there are two such trouble spots, which occur when either $\beta^3 - \beta^2 - 1 = 0$ or when $\beta^3 - \beta^2 - \beta - 1 = 0$. These correspond to the values of $\beta = 1.465571231876768\cdots$ and $\beta = 1.839286755214161\cdots$.

Where else are such spots located? Consider, for example, $m_4 = T_\beta^4(m_0)$, and consider the movement of m_4 as β is swept through the range $1 < \beta < 2$. This is shown in figure 28. As made clear in the image, three new degenerate points appear. These are located at $\beta = 1.380327757\cdots$ and $\beta = 1.754877666\cdots$ and $\beta = 1.927561975\cdots$, which are the real roots of $\beta^4 - \beta^3 - 1 = 0$ and $\beta^4 - \beta^3 - \beta^2 - 1 = 0$ and $\beta^4 - \beta^3 - \beta^2 - \beta - 1 = 0$ respectively.

Following a suggestion by Dajani[4], the β numbers corresponding to the trouble spots may be called "generalized golden means". Unfortunately, the term "generalized golden mean" is in common use, and is applied to a variety of different systems. Not all are relevant to the present situation; one that is, is given by Hare *et al.*[36] who provide series expansions for the real roots of $\beta^p - \sum_{k=0}^{n-1} \beta^k = 0$; these are known as the n-bonacci constants (Fibonacci, tribonacci, tetranacci, *etc.*). Stakhov[37] considers $\beta^{p+1} - \beta^p - 1 = 0$ in general settings. Some, but not all of these numbers are known to be Pisot numbers or Salem numbers[14]. In what follows, these will be referred to as the "beta golden means", since all of the ones that appear here have explicit origins with the beta shift.

5.2 Counting Orbits

How many trouble spots are there? The table below shows the count M_ν of the number of "new" trouble spots, as a function of the orbit length ν.

ν	2	3	4	5	6	7	8	9	10	11	12
M_ν	1	2	3	6	9	18	30	56	99	186	335
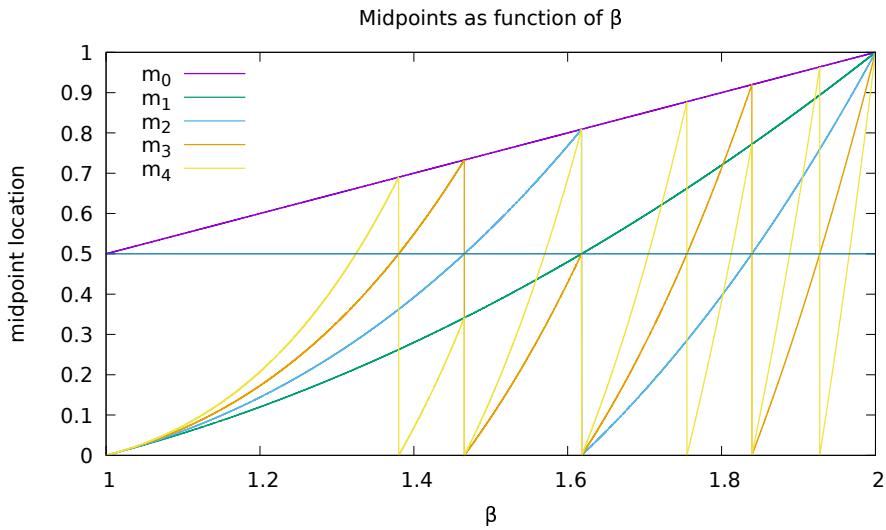
This appears to be Sloane's OEIS A001037 which has a number of known relationships to roots of unity, Lyndon words, and the number of orbits in the tent map. The values are given by Moreau's necklace-counting function. The trouble spots are the positive real roots of polynomials of the form

$$p_{\{b_k\}}(\beta) = \beta^\nu - \beta^{\nu-1} - b_1\beta^{\nu-2} - b_2\beta^{\nu-3} - \cdots - 1 = 0$$

with the $\{b_k\}$ being certain binary bit sequences. There is just one such (positive, real) root for each such polynomial. These polynomials are relatively prime, in the sense that a bit-sequence b_k is disallowed if it has the same root as some lower-order polynomial. For example, $\beta^4 - \beta^3 - \beta - 1$ is disallowed; it has the same root as $\beta^2 - \beta - 1$. Equivalently, the quadratic is a factor of the quartic; the quartic is not relatively prime with respect to the quadratic.

The reason for the appearance of the necklace-counting function is straightforward: it is counting the number of distinct orbits of a given length. An orbit of length ν is, by definition, a point x such that $x = T_\beta^\nu(x)$. Such an orbit generates a binary string, of

Figure 28: Location of Midpoints

This chart illustrates the location of the first five midpoints, $m_0, m_1, \cdots, m_4$ as a function of β. When $m_k = 0.5$, further iteration is ambiguous, as this is the location of the discontinuity in the shift map, and the next iteration leaves the midpoint bifurcated. These are the "trouble spots". The first trouble spot is visible for $m_1 = 1/2$, corresponding to $\beta = \varphi$ and a length of $v = 2$. The midpoint m_2 crosses $1/2$ (in the ascending direction) at $\beta = 1.465\cdots$ and $\beta = 1.839\cdots$, corresponding to orbits of length $v = 3$. It also crosses discontinuously downwards at $\beta = \varphi$. This crossing point has already been assigned to a shorter orbit. The midpoint m_3 has three new crossings. It also rises to touch $1/2$ at $\beta = \varphi$; but this has already been assigned to shorter orbits. The midpoint m_4 has six new crossings. Crossings will generally fall to the left and right of earlier crossings, and so are related in a bracketing relationship. The bracketing is not the full binary tree; it is pruned, as shorter orbit assignments can knock out longer one. For example, m_4 falls back down at the m_2 crossing at $\beta = 1.465\cdots$, and so, here, m_4 never even gets close to $1/2$; it won't bracket until later. This is a formalized version of figure 3, which shows midpoints of all orders.

length v corresponding to whether $T_\beta^j(x) < 1/2$ is true or not. A cyclic permutation of such a string still corresponds to the same orbit; a reversed permutation does not: thus, it is a necklace without reversal. The necklace-counting function gives the number of distinct, unique orbits of a given length that cannot be factored into shorter orbits. The beta transformation has the property that every possible orbit does occur; none are prohibited.

Yet, the bit-string defining the polynomial is not a necklace; it cannot be rotated. Each bit-string corresponds to a unique polynomial, having roots that differ from those of other polynomials. The polynomials also have a canonical order, fixed as the integer that generates the bit-string; they cannot be reordered. Ideas such as Lyndon words apply to the orbits, but not to the defining polynomials. The ordering of the polynomials is *not* the lexicographic ordering of the Lyndon words, and cannot be brought into this order.

The values of M_n are given explicitly by Moreau's necklace-counting function

$$M_n = \frac{1}{n} \sum_{d|n} 2^d \mu\left(\frac{n}{d}\right)$$

where the sum runs over all integers d that divide n and μ is the Möbius function. The generating function is

$$\frac{t}{\frac{1}{2} - t} = \sum_{n=1}^{\infty} n M_n \frac{t^n}{1 - t^n}$$

which has a radius of convergence of $|t| < 1/2$. For large n, the asymptotic behavior can be trivially deduced from the defining sum:

$$M_n = \frac{2^n}{n} - \mathscr{O}\left(\frac{2^{n/2}}{n}\right)$$

The above counting function is for necklaces with only two colors. In general, one can have necklaces with 3 or more colors; can that happen here? Yes, of course: if one considers the general β-transform for $2 < \beta$, then, in general, it can be taken as a "kneading transform" with $\lceil \beta \rceil$ branches or folds in it. The analogous trouble-spots again appear, and they can appear after an arbitrary finite-length orbit. Insofar as they correspond to periodic orbits, they are necessarily counted by the necklace-counting function. That is, one must consider all possible strings of $\lceil \beta \rceil$ letters, modulo a cyclic permutation: this is the very definition of a necklace (or "circular word"). The number of such necklaces is given by the necklace-counting function. Each such orbit is necessarily represented by a Lyndon word, which is a representative of the conjugacy class of the orbit.

The isomorphism of different systems described by necklace polynomials is a subject that gets some fair amount of attention. Golomb gives an isomorphism between the irreducible polynomials over $\mathbb{F}_p$, for p prime and necklaces built from Lyndon words.[38, 39] A number of other results exist, including [40, 41]. At any rate, a closer study of the beta-polynomials seems to be called for.

5.3 β-Golden Polynomials

The "trouble spots" occur whenever the k'th iterate $m_k = T_\beta^k(m_0)$ of the midpoint $m_0 = \beta/2$ lands on the starting midpoint $m_k = m_0$; alternately, when $m_{k-1} = 1/2$. Because of the piece-wise linear form of T_β, the k'th iterate will be a piece-wise collection of polynomials, each of order k, each of the form $p_{\{b_k\}}(\beta)$. These must be arranged such that $p_{\{b_k\}}(\beta) = 0$ at each discontinuity, as illustrated in figure 28. This constrains the polynomials that can appear; it constrains the possible coefficients $\{b_k\}$; not all bit-sequences appear. The sequences that do appear encode the orbit of the mid-point; see below.

The table below explicitly shows the polynomials for the first few orders. A polynomial is included in the table if it is an iterate of a previous polynomial, and if it's real root is bracketed by the roots of the earlier iterates. Adopting ordinal numbering, $p_n(\beta)$ must have the form

$$p_n(\beta) = \begin{cases} \beta\left(p_{n/2}(\beta) + 1\right) - 1 & \text{for } n \text{ even} \\ \beta p_{(n-1)/2}(\beta) - 1 & \text{for } n \text{ odd} \end{cases} \tag{31}$$

This recursion terminates at $p_0(\beta) = \beta - 1$.

The positive real root r_n satisfying $p_n(r_n) = 0$ is unique; the other $n-1$ roots are complex; they are arranged in a roughly evenly-spaced ring on the complex plane, not far from the unit circle, reminiscent of roots of unity. There is always a positive real root, which satisfies $1 \le r_n < 2$; the real roots and the polynomials are in one-to-one correspondence. The roots must be bracketed (to the left and right) by the roots of polynomials occurring earlier in the sequence; if the root is not bracketed, then the corresponding polynomial does not appear in the list.

The bracketing constraint can be represented by a recursive function $\theta_n(\rho)$ returning a boolean true/false value, as to whether a given polynomial is acceptable. It is

$$\theta_n(\rho) = \begin{cases} \Theta\left(r_{n/2} - \rho\right) \cdot \theta_{n/2}(\rho) & \text{for } n \text{ even} \\ \theta_{\lfloor n/2 \rfloor}(\rho) & \text{for } n \text{ odd} \end{cases} \tag{32}$$

The Heaviside $\Theta(x)$ used here is one for strictly positive $x > 0$, and is zero otherwise. This is important, as using $x \ge 0$ will not work. In numerical work, the test should be bounded away from zero. The recursion compares the candidate ρ to some root at each lower order. The recursion terminates at $\theta_0(\rho) = 1$. The index n corresponds to a valid polynomial, and thus a valid root, if and only if $\theta_n(r_n) = 1$. Effectively, this states that roots of higher-order polynomials must be less than a certain sequence of lower-order roots. This is visible in the location of the discontinuities in figure 28: new discontinuities at higher orders must occur to the left of earlier ones.

For example, the polynomial $\beta^3 - \beta - 1$ is excluded from the list simply because it is not an iterate of an earlier polynomial, even though it has the interesting real root $1.324717957244746\cdots$, the "silver constant". The numbering scheme does not even have a way of numbering this particular polynomial. Despite this, the silver constant does appear, but a bit later, as the root of $p_8 = \beta^5 - \beta^4 - 1$, which is an allowed polynomial.

117

The polynomial $p_5 = \beta^4 - \beta^3 - \beta - 1$ is excluded because it has $\varphi = 1.618\cdots$ as a root, which was previously observed by p_1. The polynomial $p_9 = \beta^5 - \beta^4 - \beta - 1$ is excluded because it's root, $r_9 = 1.497094048762796\cdots$ is greater than its predecessor r_2; the recursive algorithm does not compare it to r_4. Note that p_9 is relatively prime to the earlier polynomials, so irreducibility is not a sufficient criterion; the root must also be less.

The indexing has the property that, whenever $\theta_n(r_n) = 1$, the integer $2n+1$, expressed as a binary bitstring, encodes both the coefficients of the polynomial, and also the orbit of the midpoint. This can be taken as an alternate, non-recursive definition of θ_n: it is one if and only if the orbit of r_n encodes the bitsequence of $2n+1$.

The degree ν of the polynomial is identical to the length ν of the orbit; it is $\nu = \lceil \log_2(2n+1) \rceil$. The bits b_i of the bitstring $2n+1 = b_0 b_1 b_2 \cdots b_\nu$ correspond to the orbit as

$$b_i = \Theta\left(T_\beta^i\left(\frac{\beta}{2}\right) - \frac{1}{2} \right) = d_i\left(\frac{1}{2}\right) = k_i\left(\frac{\beta}{2}\right)$$

where the d_i are as given before, in eqn 23, and the k_i as in eqn 5. Note that $b_0 = 1$ always corresponds to $1/2 < \beta/2$, always. By convention, the last digit is always 1, also.

order ν	$p_n(\beta)$	n	binary	root r_n
0	1			
1	β		0	0
	$\beta - 1$	0	1	1
2	$\beta^2 - \beta - 1$	1	11	$\varphi = \frac{1+\sqrt{5}}{2} = 1.618\cdots$
3	$\beta^3 - \beta^2 - 1$	2	101	$1.465571231876768\cdots$
	$\beta^3 - \beta^2 - \beta - 1$	3	111	$1.839286755214161\cdots$
4	$\beta^4 - \beta^3 - 1$	4	1001	$1.380277569097613\cdots$
	$\beta^4 - \beta^3 - \beta^2 - 1$	6	1101	$1.754877666246692\cdots$
	$\beta^4 - \beta^3 - \beta^2 - \beta - 1$	7	1111	$1.927561975482925\cdots$
5	$\beta^5 - \beta^4 - 1$	8	10001	$1.324717957244746\cdots$
	$\beta^5 - \beta^4 - \beta^2 - 1$	10	10101	$1.570147312196054\cdots$
	$\beta^5 - \beta^4 - \beta^3 - 1$	12	11001	$1.704902776041646\cdots$
	$\beta^5 - \beta^4 - \beta^3 - \beta - 1$	13	11011	$1.812403619268042\cdots$
	$\beta^5 - \beta^4 - \beta^3 - \beta^2 - 1$	14	11101	$1.888518845484414\cdots$
	$\beta^5 - \beta^4 - \beta^3 - \beta^2 - \beta - 1$	15	11111	$1.965948236645485\cdots$

The next table lists the acceptable polynomial indexes for order 5, 6 and 7. Again, the coefficients appearing in the polynomial are encoded by the binary value of $2n+1$ in the sequence. This sequence has recently been published as OEIS A368747.

order ν	valid indexes
5	8,10,12,13,14,15
6	16,20,24,25,26,28,29,30,31
7	32,36,40,42,48,49,50,52,53,54,56,57,58,59,60,61,62,63

The properties of this sequence are briefly reviewed in the next section.

5.4 Properties of the theta sequence

Define the "validity set" of valid or acceptable indexes as

$$\Psi = \{n \in \mathbb{N} : \theta_n(r_n) = 1, p_n(r_n) = 0\} \tag{33}$$

This is just the list of indexes from the previous two tables; it is $\Psi=\{0,1, 2,3, 4,6,7, 8,10,12,13,14,15, 16,20,24,25,26,28,29,30,31, ...\}$. The θ_n is the acceptance function from eqn 32. By abuse of notation, write $\theta_n = \mathbb{1}_\Psi(n) = \theta(n)$ for the membership indicator function for this set. This function is visualized in figure 29.

The elements of the validity set Ψ are ordered; this is the validity sequence $\psi_m = \psi(m)$. It is convenient to start the sequence at $\psi_0 = 0$. This corresponds to $\beta = 2$ at the right; while, at the far left, for $\beta = 1$, write $\psi_{-1} = \infty$. The function ψ_m encodes the locations of one-bits in the bitmask θ, and so $\theta(\psi_m) = 1$.

The summatory function of the indicator is $S(k) = \sum_{n=1}^{k} \theta_n$. It counts the total number of one-bits below the location k. The validity sequence is the pullback of the summatory function. Each ψ_m is the smallest integer k for which $m = S(k)$ holds true, and so one has $m = S(\psi_m)$. The pullback can be expressed as $\psi(m) = \psi(S(\psi(m)))$.

A few additional properties may be noted:

- For all m, $\theta(2^m) = \theta(2^m - 1) = 1$.

- If $\theta(m) = 1$ then $\theta(2m) = 1$. By recursion, $\theta(2^n m) = 1$ for all n.

- If m is odd, and if $\theta(m) = 1$, then $\theta((m-1)/2) = 1$. This is reminiscent of the Collatz conjecture.

- If $\theta(m) = 0$ then $\theta(2^n(2m+1)) = 0$ for all n.

Each of these properties is visible in figure 29. The second bullet accounts for the stability of the comb-teeth, once they appear, while the last bullet accounts for the large spaces that open up, and stay open, never filling in.
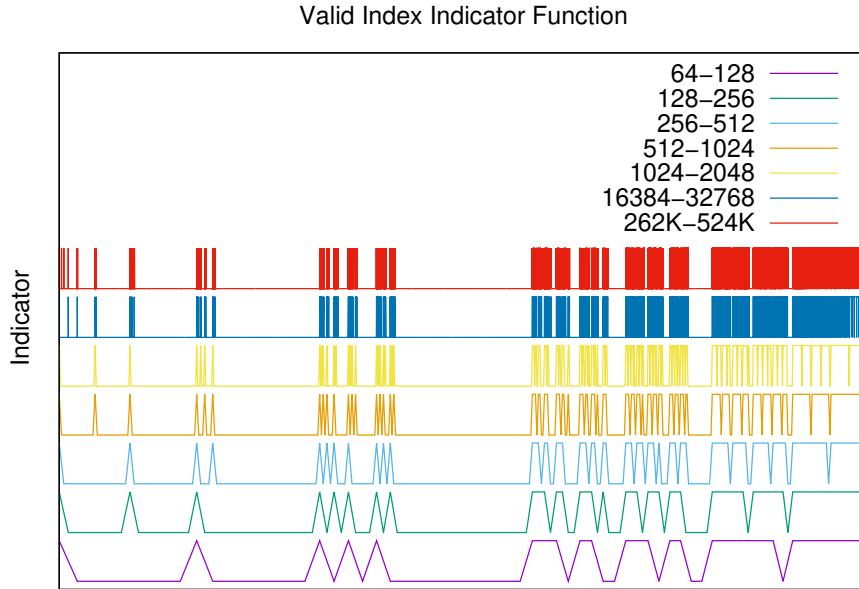
It is convenient to partition Ψ into ranks ν that correspond to the length of the orbits, or equivalently, the order of the defining polynomial. Examining the earlier tables, the rank of n is $\nu(n) = \lfloor \log_2 n \rfloor + 2$. The partition is then $\Psi = \bigcup_{\nu=1}^{\infty} \Psi_\nu$ with $\Psi_\nu = \{n \in \Psi : \nu(n) = \nu\}$. As before, it is convenient to extend the partition so that it can deal with the endpoints $\beta = 1$ and $\beta = 2$; this is a kind of (two-point) compactification of these and other various sequences and sets. The compactification here is to write $\nu(0) = 1$ and $\nu(-1) = \infty$, which allows components $\Psi_1 = \{0\}$ and $\Psi_\infty = \{-1\}$. The size of each component is $|\Psi_\nu| = M_\nu$ given by Moreau's necklace counting function M_ν.

The representation of θ as a real number is $\theta = \sum_{n=1}^{\infty} \theta_n 2^{-n} = 0.93258880035365 \cdots$. At this time, this does not appear in OEIS.

5.4.1 Leaders

An important subsequence consists of the leaders of the doubling sequences. These can be defined in several equivalent ways. One property of the bitmask is that if $\theta(m) = 1$, then $\theta(2^n m) = 1$ for all n. A leader λ is the smallest such m at the front of such a

Figure 29: Indicator function θ

Valid Index Indicator Function

This figure shows the indicator function θ_n, as defined in eqn 33, over the range of $2^6 \leq n \leq 2^{19}$. The function is approximately periodic as a function of the order $v = \lfloor \log_2 n \rfloor + 2$ of the corresponding polynomial; the order is just the length of a finite orbit. This figure just stacks seven such ranges, for orders v=6,7,8,9,10,14,18. One can view each row as a comb. The teeth of the comb are the members of Ψ_v. The number of teeth at each rank is given by Moreau's necklace-counting function M_v. The width of the teeth go as 2^{-v}, but the number of teeth goes as $M_v \sim 2^v/v$. In the limit $v \to \infty$, the teeth converge onto a set of measure zero. This can be understood as a representation of the set Ψ, taken with a real-valued index, rather than an integer index.

doubling sequence: it is either an odd number $\lambda = (2k+1)$ satisfying $\theta(\lambda) = 1$ or it is an even number of the form $\lambda = 2^h(2k+1), h > 0$ such that $\theta(2^h(2k+1)) = 1$ but $\theta(2^{h-1}(2k+1)) = 0$. The minimal power h defining the leader will be called the "height of the leader".

Equivalently, the leaders can be defined as a subsequence of ψ_m. By definition, one has that $\theta(\psi_m) = 1$. Thus, if ψ_m is odd, then ψ_m is a leader. If ψ_m is even, then it is a leader if and only if $\theta(\psi_m/2) = 0$.

Given a valid index ψ_m, it is useful to define a function Λ that provides the map

$$\Lambda(\psi_m) = 2^h(2\psi_m + 1) \tag{34}$$

This map is not monotonic: $\Lambda(1) = 3$ and $\Lambda(2) = 10$ but $\Lambda(3) = 7$. It is, however, one-to-one, as leaders are always a product of an odd number times some power of two.

Sorting the leaders into ascending order, the start of the sequence is 1, 3, 7, 10, 13, 15, 25, 29, 31, ... This sequence is not currently known to OEIS. It is handy to count from one, and to assign $\lambda_0 = 0$, so that $\lambda_1 = 1$ and $\lambda_2 = 3$ and so on.

The number of leaders of non-zero height ($h \neq 0$) at each order ν, starting at $\nu = 2$, is 0, 1, 1, 3, 3, 7, 10, 18, 28, 52, 83, 151, 256, 457, 798, 1439, 2549, 4627, 8328, 15190, 27659, 50800, 93270, 172330, 318717, 591996, ... This sequence is not currently known to OEIS.

The number of leaders of a given order ν can be defined as the size of the set $\Omega(\nu) = \{\lambda : \nu = \lfloor \log_2 \lambda \rfloor, \lambda = \Lambda(\psi_m)\}$. This can be understood as the necklace function, minus the elements of non-zero height, plus those promoted up from a lower grade. This grading is 1, 1, 3, 3, 9, 12, 26, 43, 87, 149, 295, 531, 1021, 1898, 3630, 6822, 13062, 24783, 47481, 90699, 174165, 334148, 643306, 1238619, 2390213, ... This sequence is not currently known to OEIS.

The number of leaders of a given order ν with a non-zero height is the size of the set $\{\lambda : \nu = \lfloor \log_2 \lambda \rfloor, \lambda = \Lambda(\psi_m), 0 < h\}$. This sequence is 0, 0, 1, 0, 3, 1, 6, 5, 16, 15, 43, 52, 116, 173, 348, 551, 1079, 1816, 3432, 6031, 11267, 20226, 37706, 68773, 128135, ... This sequence is not currently known to OEIS.
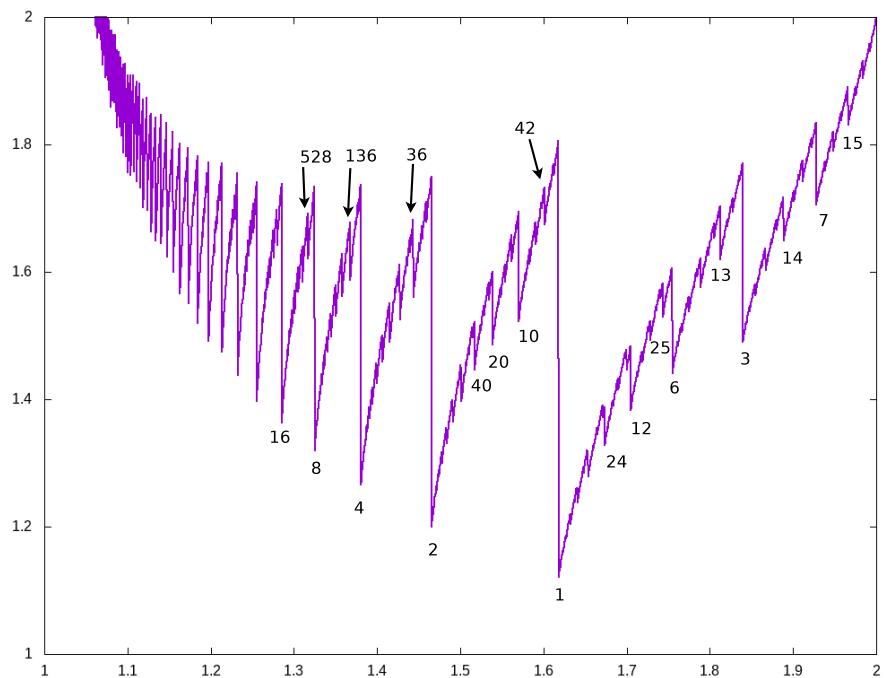
5.5   Location of β-Golden Roots

The location of the roots can be visualized by using the normalization of the Parry–Gelfond measure. The function in eqn 10 or more generally 21 can be integrated in a straightforward manner. One has

$$I(\beta; z) = \int_0^1 v_{\beta;z}(x)\, dx = \sum_{n=0}^{\infty} \frac{z^n}{\beta^n} \int_0^1 d_n(x)\, dx = \sum_{n=0}^{\infty} \frac{z^n}{\beta^n} T^n\left(\frac{\beta}{2}\right)$$

The result is a sawtooth, shown in figure 30. Each discontinuity corresponds to the real root of one of the polynomials. The first few are labeled by the integer labels from the previous table. The doubling sequences and their leaders are easy to identify.

121

Figure 30: Normalization Integral

This figure shows the integral $I(\beta) = \sum_{n=0}^{\infty} \beta^{-n} T^n \left(\frac{\beta}{2}\right)$ with $1 < \beta \leq 2$ running along the horizontal axis. Each discontinuity corresponds to the location of a real root of one of the β-Golden polynomials. Some of these are manually labeled by integers, corresponding to the polynomial labels from the previous polynomial table.

5.5.1 Bracketing intervals

That this figure is a self-similar fractal is presumably self-evident. Thus, for example, the graph to to right of $r_1 = \varphi = 1.618...$ repeats again between r_2 and r_1 and again between r_4 and r_2. Each such bracketed interval contains a unique largest discontinuity; it can be seen as being at the front of a doubling sequence. Each discontinuity is in one-to-one correspondence with a bracketing interval; the bracketing intervals are all self-similar to one-another.

A distinct notation for bracketed intervals is useful. Write $\ell \mapsto f \Leftarrow \rho$ for the discontinuity f bracketed on the left and right by ℓ and ρ. By "left" and "right", it is literally meant that the three roots are in order, with $r_\ell < r_f < r_\rho$ with the inequalities being strict. Not all ascending sequences of three roots form a valid bracket; valid brackets are obtained by recursive subdivision; this is given in the next section. But first, some examples.

Taking the liberty to write $r_0 = 2$, the interval to the right of r_1 is then $1 \mapsto 3 \Leftarrow 0$. The most prominent self-similar intervals are then $2 \mapsto 10 \Leftarrow 1$ and $4 \mapsto 36 \Leftarrow 2$ and $8 \mapsto 136 \Leftarrow 4$. In each of these examples, the front f was also a leader, with leadership as defined in the previous section. This is not always the case: the brackets $1 \mapsto 6 \Leftarrow 3$ and $1 \mapsto 12 \Leftarrow 6$ and $1 \mapsto 24 \Leftarrow 12$ are clearly visible; they are a part of an index-doubling sequence.

The extreme left side can be assigned the index of ∞ so that $r_\infty = 1$. Thus, the entire interval $1 \le \beta \le 2$ corresponds to the bracket $\infty \mapsto 1 \Leftarrow 0$.

5.5.2 The bracket tree

The brackets can be arranged into a binary tree, recursively defined. Any valid interval $\ell \mapsto f \Leftarrow \rho$ can be split into two: the left side and the right side. These two pieces are $\ell \mapsto 2f \Leftarrow f$ on the left, and $f \mapsto \Lambda(f) \Leftarrow \rho$ on the right, where Λ is the leader function given by eqn 34 in the previous section. Recursion starts with $\infty \mapsto 1 \Leftarrow 0$. A bracketing relationship is valid if and only if it appears in this recursive binary tree.

The left and right moves $L, \mathfrak{R}$ on the binary tree can then be written as

$$L : (\ell \mapsto f \Leftarrow \rho) \mapsto (\ell \mapsto 2f \Leftarrow f) \tag{35}$$
$$\mathfrak{R} : (\ell \mapsto f \Leftarrow \rho) \mapsto (f \mapsto \Lambda(f) \Leftarrow \rho)$$

The right-move is denoted with a fraktur $\mathfrak{R}$ instead of a roman R in order to distinguish between the action of the leader function, and the conventional dyadic map

$$L : m \mapsto 2m$$
$$R : m \mapsto 2m + 1$$

The issue is that not all dyadic R moves result in a valid index: having $\theta(m) = 1$ does not generally imply that $\theta(2m + 1)$ is one. However, the leader function does provide a successor that is always valid: $\theta(\Lambda(m)) = 1$ is guaranteed by construction.

Any sequence of $L, \mathfrak{R}$ moves is guaranteed to produce a valid interval; every location in the binary tree is mapped to a valid index by the bracket recursion moves. A general location on the bracket tree is

$$G = L^{a_1} \mathfrak{R}^{a_2} L^{a_3} \cdots \mathfrak{R}^{a_k}$$

and this has the property that if $\theta(m) = 1$ then $\theta(G(m)) = 1$. Equivalently, if $m \in \Psi$ then $G(m) \in \Psi$. This implies that the bracket recursion relations generate all of Ψ. Starting at $m = 1$, all left-right moves produce elements of Ψ; conversely, every element of Ψ can be expressed as some sequence of left-right moves applied to $m = 1$. It is convenient to take G as the function that produces only the "good" indexes. Thus

$$G : L^{a_1} R^{a_2} L^{a_3} \cdots R^{a_k} \mapsto L^{a_1} \mathfrak{R}^{a_2} L^{a_3} \cdots \mathfrak{R}^{a_k}$$

Labeling the nodes of the binary tree with the natural numbers $\mathbb{N}$ provides a bijection $G : \mathbb{N} \to \Psi$.

The bracket recursion relations split the interval $r_\ell < r_f < r_\rho$ into left and right pieces, as well. That is, each node of the binary tree is labeled by some (unique) r_f, with the root r_ℓ and r_ρ appearing as predecessors in the tree. Of course, the binary tree can also be mapped to the dyadic rationals; thus the bracketing recursion relations give a bijection between the dyadic rationals and the finite-length orbits of the β-map. The bracket tree and the dyadic tree are in one-to-one correspondence. The correspondence is shown in figure 31.

Perhaps the most interesting aspect to the figure is that it appears to be continuous. That is, the roots r_n appear to be dense in the interval $1 \le \beta \le 2$. Although the midpoints r_f in an interval $r_\ell < r_f < r_\rho$ are not evenly spaced, the midpoint always appears to be sufficiently far away from either endpoint so that the convergents are dense in the reals, which implies in turn that the map is continuous. It would appear that the map to infinite-length sequences can be taken, without pathologies.

5.5.3 The finite comb

The figure 29 shows each indicator rank Ψ_v as a comb. A better understanding is gained by mapping this to the infinite binary tree. Each rank v corresponds to a single horizontal row in the tree. The numbering that has been adopted is that $v = 2$ is at the root of the tree. The dyadic left and right moves on the comb are
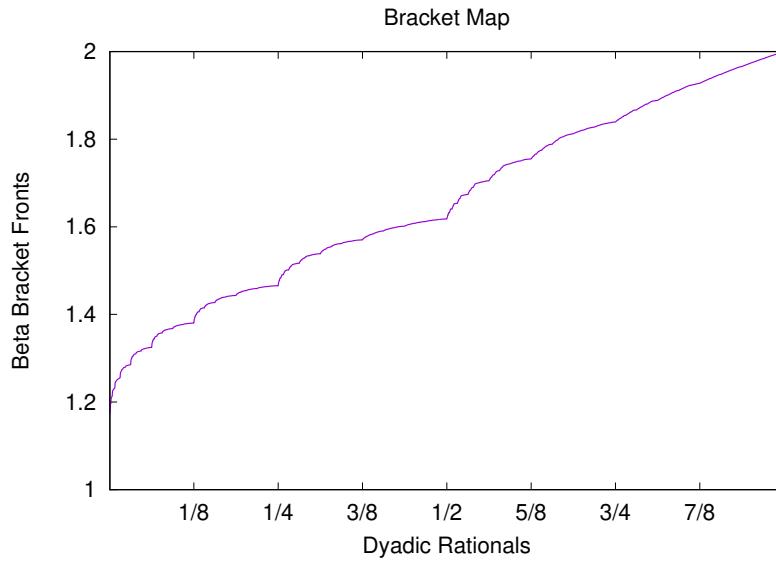
$$L : \theta_m \mapsto \theta_{2m}$$
$$R : \theta_m \mapsto \theta_{2m+1}$$

The indicator function marks each node in the tree with a zero or a one.

One property of this marking is that when a node is marked with a zero, all nodes in the right subtree are marked with a zero, as well: the presence of a zero removes the entire right-hand branch under that point. This follows from the property noted earlier, that if $\theta(m) = 0$, then $\theta(2^n(2m+1)) = 0$ for all n. Restating in terms of moves, if $\theta_m = 0$ then, taking $n = 0$, one has $R\theta_m = \theta_{2m+1} = 0$. Of course it, follows that $RR\theta_m = 0$. Taking $n = 1$, one has that $LR\theta_m = 0$, and so both left and right sub-branches are gone.

A converse marking is given by the property that, if $\theta(m) = 1$ then $\theta(2m) = 1$. On the tree, this means that left-branches of a node marked with a one are never trimmed: if $\theta_m = 1$, then $L\theta_m = 1$.

Figure 31: Beta Bracket Map

This figure shows the mapping from the dyadic rationals, interpreted as strings of left-right moves, and the corresponding $\beta = r_f$ root for the front (center) of the corresponding bracket. Thus, 1/2 maps to $r_1 = \varphi = 1.61803...$ and 1/4 maps to $r_2 = 1.46557...$ while 3/4 maps to $r_3 = 1.83929...$ It would appear that the curve is continuous. The curve does drop to $\beta = 1$ at the left, but it does so very slowly. The brackets $\ell \mapsto f \Leftarrow \rho$ remain quite wide, as the limit approaches $\beta = 1$. It does get there, though. An explicit derivation of this limit is given in a later section. The sharp takeoff at the left is reminiscent of the Minkowski question-mark mapping. The left generator for the bracket map is is quite close to left generator $x/(1+x)$ of the Minkowski fractal. The Minkowski fractal has $1/(2-x)$ as the right generator. Replacing this with $x + 1/2$ generates a curve resembling the bracket curve above, having Minkowski-like behavior to the left, an linear-like behavior to the right. The de Rham curve construction may be used to generate such curves, given arbitrary L,R maps. The L,R maps for the bracket curve are both fractal themselves, so finding affine generators would be a surprise.

Write $2^{<\omega}$ for the collection of all finite strings in the letters L, R and 2^ω as the set of all infinite strings. The indicator function is then a map $\theta : 2^{<\omega} \to \{0,1\}$ that indicates when a given finite, unbounded walk down the tree might have more branches, or definitely does not. This definitely-maybe marking defines a filter, and dually, an ideal. The filter/ideal can be taken on the integers, or equivalently, on the reals. The second possibility is what the figure 29 is hinting at.

The filter has the form that if $n \in \Psi_\nu$ then $2n \in \Psi_{\nu+1}$. The converse is not true: $10 \in \Psi_5$ but $5 \notin \Psi_4$. Written as a filter, this is $L\Psi_\nu \subset \Psi_{\nu+1}$ and the subset relation is strict. Written as and ideal, if $n \notin \Psi_\nu$ then $(2n+1) \notin \Psi_{\nu+1}$. This is more clearly stated with set complements at a given rank. Define the unit interval $I_\nu = \{n : 2^{\nu-2} \leq n < 2^{\nu-1}\}$ and the set complement $\overline{\Psi}_\nu = I_\nu \backslash \Psi_\nu$. The ideal is then $R\overline{\Psi}_\nu \subset \overline{\Psi}_{\nu+1}$ together with $\left(LR\overline{\Psi}_\nu \cup RR\overline{\Psi}_\nu\right) \subset \overline{\Psi}_{\nu+2}$.

5.5.4 The comb bijection

The good-index bijection $G : \mathbb{N} \to \Psi$ provides a mechanism to map the full binary tree into the trimmed tree, so that every node in the trimmed tree is in one-to-one correspondence with the full binary tree.

The leadership function provides the desired mapping. If $n \in \Psi_\nu$, then $\Lambda(n) \in \Psi_{\nu+1+h}$ where h was the height of the leader. Any valid index n gets "kicked upstairs" by the leadership function; one can write $\Lambda\Psi_\nu \subset \bigcup_{h=0}^\infty \Psi_{\nu+1+h}$. This just corresponds to the bracket move $\Re : \theta_m \to \theta_{\Lambda(m)}$ on the trimmed tree: given any location $\theta(m) = 1$ in the trimmed tree, the bracket right move returns the next valid right branch in the trimmed tree: $\Re(m) = \Lambda(m)$, since, by construction, $\theta(\Lambda(m)) = 1$ whenever $\theta(m) = 1$.

5.5.5 The infinite comb

The comb is mapped to the reals by defining open subsets bounded by the dyadic rationals. Let

$$I(m, \nu) = \{x \in \mathbb{R} : m < x2^{\nu-2} < (m+1); 0 \leq m < 2^{\nu-2}\} \tag{36}$$

so that $I(0,2) = \{x \in \mathbb{R} : 0 < x < 1\}$. The left and right moves are the obvious ones: $LI(m, \nu) = I(2m, \nu+1)$ and $RI(m, \nu) = I(2m+1, \nu+1)$. These are interpreted as the left and right halves of the (fat) Cantor set, on the reals; the fat Cantor set being taken as the reals with the dyadic rationals removed.

The filters on Ψ_ν become filters on $I(m, \nu)$ in the obvious way, by means of a commuting diagram. Write $I(\Psi_\nu) = \bigcup_{m \in \Psi_\nu} I\left(m - 2^{\nu-2}, \nu\right)$ as the union of intervals that cover the nonempty runs in the binary tree. This is effectively what is being graphed in the figure 29. The limit set is then

$$\overline{\theta} = \bigcap_{\nu=2}^\infty I(\Psi_\nu) \tag{37}$$

where the notation $\overline{\theta}$ is happily abused to denote the limit of the indicator function θ_n as a subset of the reals. This is possible precisely because θ_n can be mapped to the binary tree, which can then be partitioned as filters and ideals.

Since $2^{v-2} \in \Psi_v$ for all v, i.e. the leftmost branch is never trimmed, one easily concludes that $0 \in \overline{\theta}$. Since $(2^{v-1} - 1) \in \Psi_v$ for all v, the rightmost branch is never trimmed, and so one may conclude that $1 \in \overline{\theta}$.

It should be clear that the set $\overline{\theta}$ is isomorphic to the Cantor set. The comb bijection shows how to map points in $\overline{\theta}$ back into the untrimmed infinite tree; the full binary tree is isomorphic to the Cantor set. From this, one concludes that the set $\overline{\theta}$ is uncountable, and can be placed in bijection with the reals. This is explored in the next section. A measure can be assigned to $\overline{\theta}$. This is given in the section after next. A more formal examination of $\overline{\theta}$, tightening up some of the loose language above, will be given in a later section on eventually-periodic orbits.

5.5.6 Self-describing orbits

What is the meaning of the finite and infinite combs? The finite comb is a mapping of of the valid-index set Ψ to the dyadics. The infinite comb is the closure of the finite comb in the reals.

A polynomial index $n \in \Psi$ if and only if n encodes a self-describing finite orbit. That is, $n \in \Psi$ if and only if the real root r_n of $p_n(r_n) = 0$ iterates under the β-map such that the iterate reproduces the bit-sequence of n. That is, $n \in \Psi$ if and only if the bits b_i of the bitstring $2n + 1 = b_0 b_1 b_2 \cdots b_v$ are given by the (finite) orbit as

$$b_i = \Theta \left(T_\beta^i \left(\frac{r_n}{2} \right) - \frac{1}{2} \right) = k_i \left(\frac{r_n}{2} \right)$$

If n does not have this property, then $n \notin \Psi$. At each rank v, the elements of Ψ_v correspond to finite orbits of length v. In the limit of $v \to \infty$, one gets self-describing orbits of unbounded (infinite) length.

By construction, if $n \in \Psi_v$ then $Ln \in \Psi_{v+1}$ and $\Lambda n \in \Psi_{v+1+h}$. That is, if n is a self-describing orbit, then it appears as the prefix of longer self-describing orbits.

Polynomials of infinite order are holomorphic functions. Given an infinite bitsequence $\{b\} = b_0 b_1 \cdots$, define a holomorphic function

$$q^{\{b\}}(\zeta) = 1 - \sum_{j=0}^{\infty} b_j \zeta^{j+1}$$

If the bit-sequence is finite, in that all $b_j = 0$ when $j > k$, then this is related to the polynomials as

$$\zeta^{k+1} p_n \left(\frac{1}{\zeta} \right) = 1 - b_0 \zeta - b_1 \zeta^2 - \cdots - b_k \zeta^{k+1}$$

Given any arbitrary sequence $\{b\}$, the holomorphic function $q^{\{b\}}(\zeta)$ will have a single, unique real, positive root. To make contact with the polynomials, write this as the reciprocal, so that the root r satisfies $q^{\{b\}}(1/r) = 0$. This root will have some orbit, given by

$$a_i = \Theta \left(T_\beta^i \left(\frac{r}{2} \right) - \frac{1}{2} \right) = k_i \left(\frac{r}{2} \right)$$

127

Such an orbit is self-describing if and only if $\{a\} = \{b\}$.

The claim being pursued here is that the infinite comb contains all self-describing sequences, and conversely, every element of the comb corresponds to a self-describing bit-sequence. That is, if

$$x = \sum_{j=0}^{\infty} \frac{b_j}{2^j} \in \overline{\theta}$$

then

$$x = 2\left(1 - q^{\{b\}}\left(\frac{1}{2}\right)\right)$$

Every real number $x \in \overline{\theta}$ in the comb corresponds to such a self-describing orbit. A proof of these claims will be given in a later section, after the development of some formal definitions.

In the meanwhile, it can be noted that every rational number corresponds to a bitsequence that is ultimately periodic. After an initial unstable finite sequence, the bitstring settles down to a cyclic orbit. One task ahead is to examine the set $\mathbb{Q} \cap \overline{\theta}$: this is the set of self-describing eventually-periodic infinite-length orbits. It turns out these can be readily described as root of a finite polynomial. The holomorphic function $q^{\{b\}}(\zeta)$ to be factored into two finite polynomials, one describing the initial aperiodic segment, and a second describing the cyclic segment. Such orbits are examined in a later section.

5.6   Formal definitions

A sufficient number of distinct concepts have been introduced, that some basic house-keeping is in order. The definitions that follow are straight-forward and conventional. The goal is to provide a workable vocabulary for further discussion.

Let $\mathbb{B}$ denote the finite but unbounded binary tree, and $\overline{\mathbb{B}}$ it's closure as the infinite tree. The infinite tree is, of course, isomorphic to the Cantor space 2^{ω}; but this mechanism is not currently needed. A more careful definition of the finite tree is needed. Thus, let $\mathbb{B}$ the the graph of vertices and connecting edges $\mathbb{B} = \{v_i, e_{ij} : i \in \mathbb{N}, j \in \{2i, 2i+1\}\}$. Let $\eta : \mathbb{N} \to \mathbb{B}$ denote the canonical labeling of the binary tree by the positive integers, so that the root of the tree is given the label 1, the first row is 2,3 and the next row is 4,5,6,7. This is a bijection: every finite walk down the tree can be labeled with a positive integer. The walks are left and right moves, in the canonical sense: $L : \mathbb{N} \to \mathbb{N}$ with $L : m \mapsto 2m$ and likewise $R : \mathbb{N} \to \mathbb{N}$ with $R : m \mapsto 2m+1$. The pushforward of L, R provide the canonical walks on the tree, as a commuting diagram, so that $L \circ \eta = \eta \circ L$ and likewise $R \circ \eta = \eta \circ R$. It is useful to adjoin the pre-root elements $\{0, \infty\}$ so that $R : 0 \to 1$ and $L : \infty \to 1$.

Let $\mathbb{D}$ be the dyadic rationals, with the canonical bijection to the natural numbers $\delta : \mathbb{D} \to \mathbb{N}$ given by $\delta : (2n+1)/2^m \mapsto 2^{m-1} + n$. This labels the root of the tree with 1/2, and the first row under it as 1/4 and 3/4. The left and right moves are $L : (2n+1)/2^m \mapsto (4n+1)/2^{m+1}$ and $R : (2n+1)/2^m \mapsto (4n+3)/2^{m+1}$. This was set up so that $L \circ \delta = \delta \circ L$ and likewise $R \circ \delta = \delta \circ R$. The map $\delta^{-1} : \mathbb{N} \to \mathbb{D}$ is equally familiar: it is $n = b_0 b_1 \cdots b_k \mapsto \sum_{n=0}^{k} b_n 2^{-n-1}$ that interprets $n \in \mathbb{N}$ as a sequence of L, R moves in the binary tree, returning the fraction found at that location.

This allows the bracket tree and the good-index bijection to be specified more precisely. The validity set $\Psi \subset \mathbb{N}$ is defined in eqn 33 as the collection of natural number indexes corresponding to polynomials with self-describing roots. This set came with two functions $L, \mathfrak{R} : \Psi \to \Psi$ given by index-doubling $L : m \mapsto 2m$ and leadership (eqn34) $\mathfrak{R} : m \mapsto \Lambda(m)$. The good-index bijection $G : \mathbb{N} \to \Psi$ is then defined as the pullback $L \circ G = G \circ L$ and $\mathfrak{R} \circ G = G \circ R$.

The trimmed tree $\mathfrak{B} \subset \mathbb{B}$ is the image of Ψ under the mapping η, so that $\mathfrak{B} = \eta \Psi$. This consists of those nodes and edges in the finite binary tree that are labeled by integers from the validity set Ψ. The good-index function G places the trimmed tree and the finite tree in a bijection, so that $\mathfrak{B} = \eta G \eta^{-1} \mathbb{B}$.

The root map $r : \Psi \to [1,2]$ takes valid integer indexes, and maps them to the roots of the corresponding (finite-orbit) polynomials, so that $p_n(r_n) = 0$. The bracket map, depicted in figure 31, can now be written more precisely as the function $\rho = r \circ G \circ \delta^{-1} :$ $\mathbb{D} \to [1,2]$. It maps the dyadics to the β values that have finite orbits.

The notation for the finite comb θ is abused in several ways. In eqn 32 it is used to indicate whether a given polynomial root has a self-describing orbit. It is then defined as the indicator function for set membership $\theta = \mathbb{1}_\Psi : \Psi \to \{T, F\}$ and finally as the finite comb $\theta \subset \mathbb{D}$. Formally, the finite comb can now be written as $\theta = \delta^{-1} \Psi$.

The good-index bijection $G : \mathbb{N} \to \Psi$ can be commuted with the dyadic bijection to obtain the "good dyadic map" $\mathsf{G} = \delta \circ G \circ \delta^{-1} : \mathbb{D} \to \theta$. Given any dyadic fraction, this map returns another dyadic that lies within the finite comb. As all the other maps discussed so far, it is a bijection. It is depicted in figure 32.
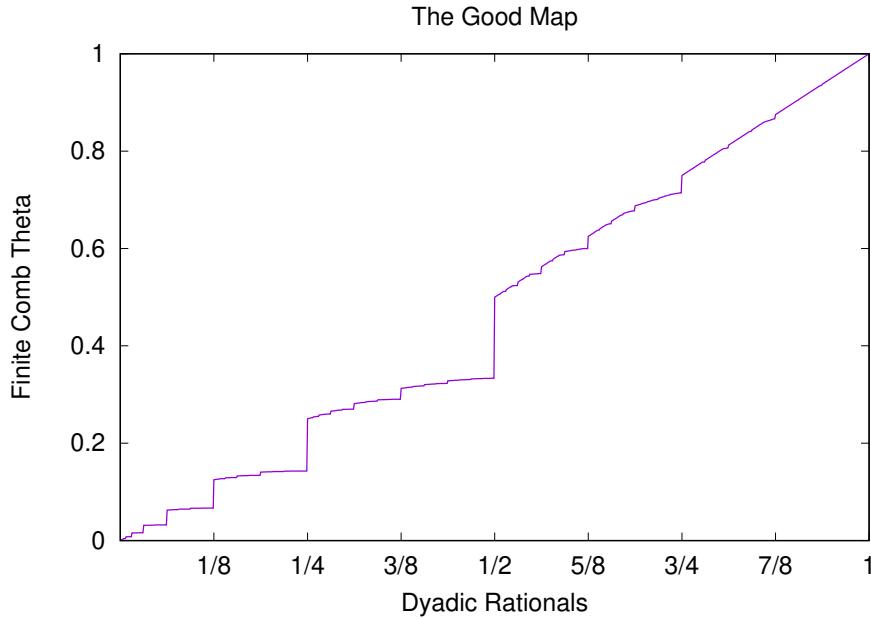
The infinite comb $\overline{\theta}$ was the closure of the finite comb, so that $\theta \subset \overline{\theta}$. The closure, defined in eqn 37, is constructed as the infinite intersection of open sets. Corresponding to this is the closure $\overline{\mathfrak{B}} \subset \overline{\mathbb{B}}$ of infinitely long paths in the trimmed tree. It is taken as the limit of the finite but unbounded-length paths in $\mathfrak{B}$. The dyadics can be closed in several ways; one way is to the rationals $\mathbb{Q}$, and then further to the reals $\mathbb{R}$. In the present case, we restrict attention to the unit interval $I = [0,1] \subset \mathbb{R}$ and to $\mathbb{Q}_I = \mathbb{Q} \cap [0,1]$. Proper diligence requires distinguishing $\overline{\mathfrak{B}}_\mathbb{Q}$ from $\overline{\mathfrak{B}}_\mathbb{R}$ and also $\overline{\theta}_\mathbb{Q}$ from $\overline{\theta}_\mathbb{R}$.

The rationals correspond to infinite length orbits that are ultimately periodic; these will be examined in a later section. A closure to the root map to the rationals appears naturally; it can be written as $\overline{r} : \overline{\theta} \to [1,2]$. Combined with the closure $\mathsf{G} : \mathbb{Q}_I \to \overline{\theta}$ this gives a closure of the bracket map $\overline{\rho} = \overline{r} \circ \mathsf{G} : \mathbb{Q}_I \to [1,2]$. Closing to the reals gives $\mathsf{G} : [0,1] \to \overline{\theta}$ and $\overline{\rho} = \overline{r} \circ \mathsf{G} : [0,1] \to [1,2]$. This is the primary achievement of this section: the definition of the bracket map as a continuous monotonic ascending bijection between the reals in the unit interval, and β values understood as self-describing orbits. Proving continuity will require some theoretical machinery and lemmas; these will be developed in a later section. However, the net result can already be seen in figure 31.

5.7 The Good Map and Measures

The formal definitions allow the description of the bracket map to be completed. This is a matter of reviewing the remaining congruences. The fundamental one, moving forward, is the "good map" $\mathsf{G} : \mathbb{D} \to \theta$ that is a correspondence between the dyadic rationals to the comb. The comb, in turn, can be understood as inducing a measure.

Figure 32: The Good Map

A graph of the "good map" $\mathsf{G} : \mathbb{D} \to \theta$. It is defined as the commutator $\mathsf{G} = \delta \circ G \circ \delta^{-1}$ of the valid-index map $G : \mathbb{N} \to \Psi$ by the canonical mapping $\delta : \mathbb{N} \to \mathbb{D}$ between the natural numbers and the dyadic rationals. The finite comb is just a mapping of the good indexes Ψ to the dyadics: $\theta = \delta^{-1}\Psi$.

The measure, combined with the good map, induces the "beta measure", which is just the bracket map. This closes the circle of commuting diagrams.

5.7.1 The good map
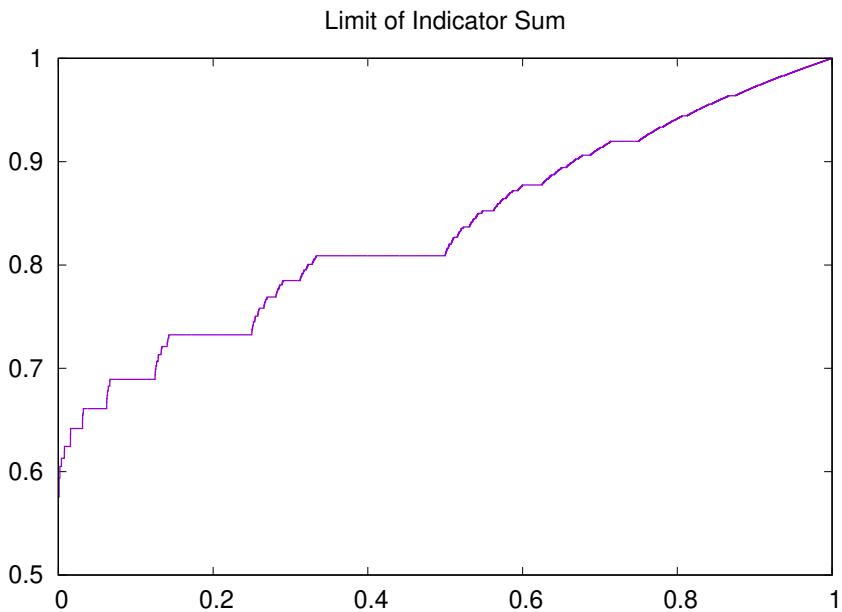
The validity map $G : \mathbb{N} \to \Psi$ can be commuted with the canonical mapping $\delta : \mathbb{N} \to \mathbb{D}$ between the natural numbers and the dyadic rationals. This defines a map $\mathsf{G} = \delta \circ G \circ \delta^{-1}$ that is a bijection between the dyadic rationals and the finite comb: $\mathsf{G} : \mathbb{D} \to \theta$. It is shown in figure 32.

5.7.2 The comb measure

The infinite comb was constructed as a closure or limit of the finite comb. An interesting trick is to interpret it as a measure, and so to integrate over it. This can be obtained as a limit over sums of ranks in the finite comb. The measure is depicted in figure 33.

The sum over the indicator function $S(k) = \sum_{n=1}^{k} \theta_n$ shows power-of-two periodicity, same as each rank in the finite comb. Each rank ν can be separated out as

Figure 33: Limit of Indicator Sum

Limit of Indicator Sum

This figure shows the limit of the indicator sum $A(x) = \lim_{v \to \infty} A_v(x)$ of eqn 38. More precisely, it shows $A_v(x)$ for $v = 20$. By this point, convergence is sufficient that any differences from the limit are not visible to the naked eye.

$$S_v(k) = S(k) - S\left(2^{v-2} - 1\right)$$

The intent is to isolate the range $2^{v-2} \leq k < 2^{v-1}$, and to split the sum $S(k)$ into a collection of ranks S_v.

At the end of the range, the sum S_v achieves Moreau's necklace-counting function: $S_v\left(2^{v-1} - 1\right) = M_v$. Dividing by M_v gives each rank the same vertical scale: zero to one. It is also useful to rescale the horizontal range, to run zero-to-one as well. This gives a normalized version

$$F_v(x) = \frac{S_v\left(\lfloor 2^{v-2}(1+x)\rfloor\right)}{M_v}$$

that runs from zero to one as x runs from zero to one. This function does not have an interesting limit as $v \to \infty$, as it slowly drops to zero over the entire unit interval. However, it does so at a fixed rate, and can be held constant with a radical. The sequence of functions

$$A_v(x) = \exp\left(M_v 2^{-v} \log F_v(x)\right) \tag{38}$$

converge rapidly and more-or-less uniformly to a limit $A(x) = \lim_{v\to\infty} A_v(x)$. It appears to be well-defined on the unit real interval $0 \leq x \leq 1$. This is the limit shown in figure 33. It is perhaps useful to keep in mind the asymptotic limit of Moreau's function, $M_v = 2^v/v - \mathcal{O}\left(2^{v/2}/v\right)$ and so the radical scales as $M_v 2^{-v} = 1/v - \mathcal{O}\left(2^{-v/2}/v\right)$.

The convergence to the limit appears to be uniform and rapid, except at $x = 0$, which proceeds slowly. This is easily demonstrated. The $x = 0$ limit is

$$A(0) = \lim_{v\to\infty} (F_v(0))^{M_v 2^{-v}} = \lim_{v\to\infty}\left(\frac{1}{M_v}\right)^{M_v 2^{-v}} = \lim_{v\to\infty}\left(\frac{v}{2^v}\right)^{1/v} = \frac{1}{2}\lim_{v\to\infty} v^{1/v} = \frac{1}{2}$$

The slow convergence is entirely due to the last limit, above.

The offsets to the sums and limit above were defined above, so as to avoid having to debate the meaning of $\lim_{v\to\infty} \sqrt[v]{0}$. Yet clearly, the intent is that $A(x)$ should provide a measure for the infinite comb. Should the infinite comb be thought of as having a large point-weight at $x = 0$? Perhaps not; thus, perhaps a more suitable measure is $\mu_\theta(x) = 2A(x) - 1$, which runs from zero to one over the unit interval.
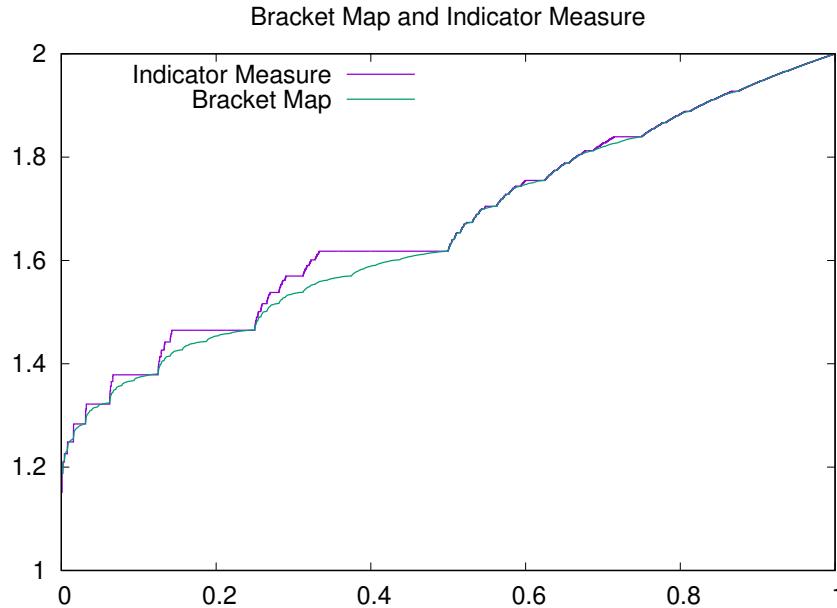
To summarize: The function $\mu_\theta(x)$ provides a measure for the infinite comb given in eqn 37. The non-flat sections are where the infinite-length self-describing sequences are accumulating.

5.7.3 The Beta measure

The significance of the comb measure is revealed by superimposing it's graph 33 on the bracket map 31, so that both appear side by side. This is shown in 34. The stair-treads line up with the blancmange dips in the bracket map. The good map can be used to eliminate the stair treads. The identity is

$$2A \circ \mathsf{G} = \rho = r \circ G \circ \delta^{-1} = r \circ \delta^{-1} \circ \mathsf{G}$$

Figure 34: The Beta Measure

This figure superimposes the two figures 31 and 33 into one. The indicator measure has been rescaled, so that the y-axis aligns with the interval $1 \leq \beta \leq 2$, as that is the nominal topic of discussion. The comb is also a map through β values, but taken sideways, as it were. This figure indicates visually what that correspondence is. When the stair-treads are removed with the "good map", the two curves are identical. This is surprising, as they have entirely different origins: the indicator measure is counting periodic orbits, while the bracket map is providing the locations.

Since $\mathsf{G} = \delta \circ G \circ \delta^{-1}$ is a bijection, it can be peeled off, to give

$$2A = r \circ \delta^{-1}$$

This is surprising. The map r is purely local: it takes integers n to the enumerated roots $p_n(r_n) = 0$. It is just specifying locations of roots. The map A is global: it is counting how many roots there are. It is the limit of a sum, a kind of peculiar integral, that captures information about all other β values, and how they behaved.

An alternative interpretation is that this provides a way of estimating the number of orbits of length v, satisfying $\beta < \alpha$ for some fixed $1 \leq \alpha \leq 2$. The total number of orbits of length v is given by Moreau's M_v. The counting function $S_v(k)$ returns the total number of orbits of length v that occurred at some β with $\beta \leq r_k$. The normalized version provides this same number slightly more elegantly: the total number of orbits of length v with $\beta < \rho(x)$ is given by $M_v F_v(x)$. The indirection with ρ is annoying; define $f_v(\alpha) = F_v(\rho^{-1}(\alpha))$, so that $f_v(\alpha)$ counts the fraction of all orbits of length

v occuring at some (any) $\beta \le \alpha$. This fraction is approximated as $f_v(\alpha) \approx (\alpha/2)^v$, which holds exactly in the limit $v \to \infty$.

This last again illustrates the local-global tie between these two: $f_v(\alpha)$ is a counting function, an integral of sorts, while α is just a number. It's quite rare to find such specific analytic results in this project. Very rare: this is the first.

5.7.4 The front sequence

The validity map or "good index" map $G : \mathbb{N} \to \Psi$ provides a one-to-one correspondence between the natural numbers and the valid indexes. This can be interpreted as a one-to-one correspondence between nodes in the binary tree, and the center points or "fronts" f of brackets $\ell \mapsto f \Leftarrow \rho$. The bracket recursion relations in eqn 35 can be restated as a commuting diagram

$$L : f_n \mapsto f_{2n} = 2f_n$$
$$R : f_n \mapsto f_{2n+1} = \mathfrak{R}(f_n) = \Lambda(f_n)$$

The sequence begins at $f_1 = 1$. This sequence contains the same integers as the validity sequence ψ_m, and so, for all n, $\theta(f_n) = 1$. Unlike ψ_m, it is not in sorted order. The start of the sequence is (1), (2), (3,4), (10,6,7,8), (36,20,42,12,13,14,15,16), (136,...,32), ... where the parenthesis is used to improve readability, by grouping a run of length 2^n.
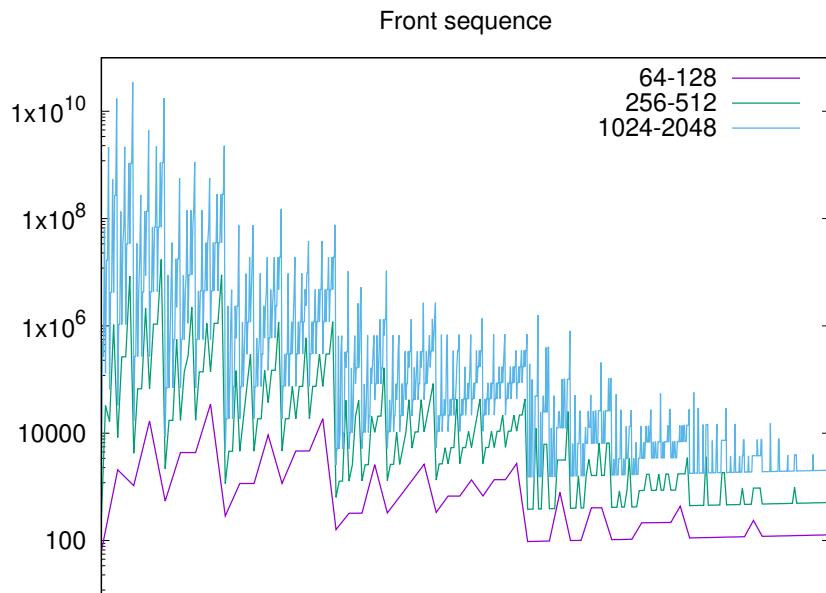
Note that $f_5 = 10$; this is an expression of the idea that $5 \notin \Psi$. Thus, it is bumped by the leadership function to the next index that is in Ψ. That is, $Rf_2 = f_5 = \Lambda(f_2) = 10$. Likewise, $f_9 = 36$; the height of the leadership function was four at this location. $f_{10} = 20$ because ten was "already taken"; but this is just a different way of thinking about $Lf_5 = f_{10} = 2f_5 = 20$.

As always, the series is quasi-self-similar across rank sets of $I_v = \{n : 2^{v-2} \le n < 2^{v-1}\}$. This is visualized in figure 35.

At each rank v, one has a corresponding set of β values for that rank. Denote these sets as $\beta_v = \{r(f_n) : n \in I_v\}$ where $r(m) = r_m$ is the positive real root of the polynomial p_m, as always: $p_m(r_m) = 0$. In the limit $v \to \infty$, these sets converge rapidly to the bracket map of figure 31. Not a surprise: this is, more or less, how that figure is generated.
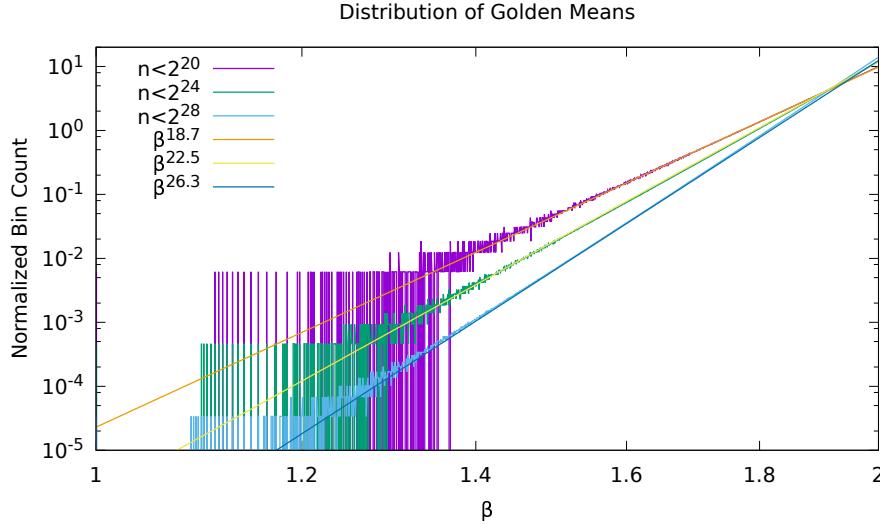
The convergence appears to be uniform! Uniform, in that a uniform bound is enough to give a worst-case bound. Convergence is slowest, just above a dyadic rational, and fastest, just below it, varying by huge orders of magnitude. By "convergence", it is meant how rapidly the brackets $\ell_n \mapsto f_n \Leftarrow \rho_n$ shrink. For $n \in I_v$, the two endpoints belong to I_{v-1}. The corresponding roots are $r(\ell_n) < r(f_n) < r(\rho_n)$ and so one is looking to see how the difference $|r(\rho_n) - r(\ell_n)|$ decreases as $v \to \infty$. The limit to convergence is at $\beta = 1$, where $n = 2^{v-2}$. Although $f_n \to 1$ as $v \to \infty$, it does so very slowly. At $v = 30$, one finds $r(f_n) \approx 1.08$ while at $v = 60$ one still has that $r(f_n) \approx 1.05$. This slowness is directly visible on the left side of figure 31, which visually has converged only to $r(f_n) \approx 1.2$ at $v = 10$. A bound can be obtained numerically; it is $r(f_n) < 1 + 0.5/\log v$, which holds at this worst-case location. This bound is pessimistic; its clear that better bounds exist.

Figure 35: Front Sequence

Front sequence

This figure shows the front sequence f_n over three rank ranges $v = 8, 10, 12$, for which $I_8 = \{64 \le n < 128\}$ and $I_{10} = \{256 \le n < 512\}$ and $I_{12} = \{1024 \le n < 2048\}$. The three ranges are rescaled, to run from left to right. Note the labels on the y-axis: the maximum value of f_n in a given range increases super-exponentially; super-factorially, even. For these three ranges, the maximum values occur at $f_{71} = 34952$, $f_{271} = 17318416$ and $f_{1055} = 34905131040$. Super-exponential means that for rank v, the max observed at that rank is approximately $\exp\left(v^{1.525}\right)$, which holds up to about $v = 16$, beyond which 64-bit numerics overflows. The rate of increase seem to be faster than factorial.

Figure 36: Distribution of Golden Means

The bin-counted distribution of roots of $p_n(\beta)$ for three different cutoffs, and the corresponding eyeballed fit. Bin-counting proceeds by dividing the range $1 < \beta < 2$ into 1303 equal–width bins. Proceeding methodically to find roots for all $n < 2^k$ for fixed k, each root is assigned to a bin. At the end of the counting process, the bin-counts are normalized by the width of the bin, and the total number of roots observed (*i.e.* by the Moreau counting function). For fixed k, the distribution appears to be approximately exponential (but not quite - there is a deviation from linearity in the graph above, just barely discernible by naked eye). Three different k's are shown, and three eyeballed fits. The general trend appears to be that, for fixed k, the distribution is approximately β^α with $\alpha \simeq k + 3 - \log_2 k \simeq \log_2 M_{k+3}$. Clearly, the $k \to \infty$ limit accumulates all the measure at $\beta = 2$.
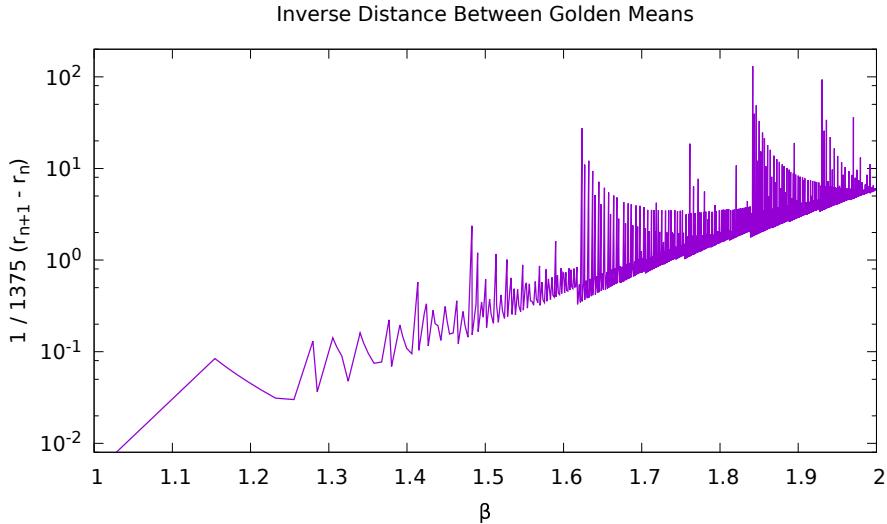
5.8 Distribution of β-Golden indices

It seems natural to ask how the indices n are distributed with respect to the real roots r_n of p_n. It appears that there is no such distribution, or rather, that it accumulates entirely at $\beta = 2$. Figure 36 shows the numerically computed (bin-counted) distribution of the zeros of $p_n(\beta)$ for $n < 2^k$ for three different values of k. This suggests that, in the limit of $k \to \infty$, almost all $p_n(\beta)$ have roots r_n that approach 2.

This is a relative statement, as it is comparing an implicit measure on the integers to the distribution of the roots. Thus, there are several (ambiguous) ways to phrase the result. One is to say that small integers (those $n < 2^k$ for any k) describe roots that accumulate at $\beta = 2$. The converse description is to say that for any fixed $\beta < 2$, the values of the indices n needed to label the roots r_n become unboundedly large.

The local distribution of roots can be sensed from the figure 37, which visualizes

136

Figure 37: Distance Between Means

This figure visualizes the inverse distance between golden means. A total of $1375 = \sum_{k=1}^{12} M_k$ roots were obtained, and then sorted into ascending order. Letting r_n represent the n'th root in this sequence, this shows the reciprocal distance $1/1375 \, (r_{n+1} - r_n)$. Increasing the number of roots essentially just rescales this graph, making it progressively more vertical. In essence, almost all of the roots accumulate near $\beta = 2$; roots become increasingly rare the smaller the β. In the limit, one might say that essentially all roots are at $\beta = 2$: although the roots are dense in the interval $1 < \beta < 2$, the counting measure indicates that they are accumulating at $\beta = 2$ only.

the distance between neighboring roots.

5.9 Complex Roots

What are the complex roots? Numerical work clearly indicates that they seem to be approximately cyclotomic in some sense or another. They seem to be more-or-less uniformly distributed in an approximate circle, always. The modulus of most of the complex roots appear to be less than one. This is violated for the complex roots of $p_{2^k}(\beta) = \beta^{k+2} - \beta^{k+1} - 1$, where some of the roots in the right-hand quadrant have a modulus larger than one. By contrast, the complex roots of $p_{2^k-1}(\beta) = \beta^{k+1} - \sum_{j=0}^{k} \beta^j$ seem to always have a modulus less than one. These two seem to be the extreme cases: in general, the polynomials appear to be "approximately cyclotomic". Its not clear how to make this statement more precise.

These numerical results can be argued heuristically: just divide the polynomial by

it's leading order. That is, a general polynomial of this form is

$$p_n(z) = z^{k+1} - \sum_{j=0}^{k} b_j z^{k-j}$$

with the convention that $b_0 = b_k = 1$, and the bit-sequence $2n + 1 = b_0 b_1 b_2 \cdots b_p$ corresponding to a terminating orbit. Dividing by z^{k+1} gives a series

$$1 - \frac{1}{z} - \frac{b_1}{z^2} - \frac{b_2}{z^3} - \cdots$$

Clearly, this can have a zero only when $|z| < 2$ as otherwise, the terms get too small too quickly.

5.10 β-Golden β-Fibonacci Sequences

It is well known that the golden ratio occurs as limit of the ratio of adjacent Fibonacci numbers:

$$\varphi = \lim_{m \to \infty} \frac{F_m}{F_{m-1}}$$

where $F_m = F_{m-1} + F_{m-2}$. There is a generalization of this, which also has received attention: the tribonacci, quadronacci, *etc.* sequences, whose limits are

$$\alpha_n = \lim_{m \to \infty} \frac{F_m^{(n)}}{F_{m-1}^{(n)}}$$

where $F_m^{(n)} = F_{m-1}^{(n)} + F_{m-2}^{(n)} + \cdots + F_{m-n}^{(n)}$.

Is it possible that the real roots of the polynomials $p_n(\beta)$ are also the roots of such sequences? But of course they are! Given a finite string of binary digits $\{b\} = b_0, b_1, \cdots, b_k$, write the beta-Fibonacci sequence as

$$F_m^{\{b\}} = b_0 F_{m-1}^{\{b\}} + b_1 F_{m-2}^{\{b\}} + \cdots b_k F_{m-k}^{\{b\}}$$

The name "beta-Fibonacci" is needed because the term "generalized Fibonacci sequence" is already in wide circulation for the special case of all bits being one. The ratio of successive terms is

$$\alpha^{\{b\}} = \lim_{m \to \infty} \frac{F_m^{\{b\}}}{F_{m-1}^{\{b\}}}$$

and is given as the (positive) real root of the polynomial

$$p_n(\beta) = \beta^{k+1} - b_0 \beta^k - b_1 \beta^{k-1} - \cdots - b_k = 0$$

These polynomials and their roots were already enumerated and tabulated in the previous section.

The beta-Fibonacci sequences do not appear by accident: these sequences have an ordinary generating function (OGF) given by the polynomial! That is,

$$\sum_{m=0}^{\infty} z^m F_m^{\{b\}} = \frac{z^k}{1 - b_0 z - b_1 z^2 - \cdots - b_k z^{k+1}} = \frac{1}{z p_n \left(\frac{1}{z}\right)}$$

The factor of z^k in the numerator serves only to initiate the sequence so that $F_0^{\{b\}} = \cdots = F_{k-1}^{\{b\}} = 0$ and $F_k^{\{b\}} = 1$.

These sequences are generic: they indicate how many different ways one can partition the integer m into elements of the set $\{b_0, 2b_1, 3b_2, \cdots, (k+1)b_k\}$. So, for example, the entry for $n = 12$ in the table below corresponds to OEIS A079971, which describes the number of ways an integer m can be partitioned into 1, 2 and 5 (or that $5m$ can be partitioned into nickels, dimes and quarters). This corresponds to the bit sequence $\{b\} = 11001$; that is, $\{b_0, 2b_1, 3b_2, \cdots, (k+1)b_k\} = \{1 \cdot 1, 2 \cdot 1, 3 \cdot 0, 4 \cdot 0, 5 \cdot 1\} = \{1, 2, 5\}$. From such partitions, it appears that one can build partitions of the positive integers that are disjoint, and whose union is the positive integers. This suggests a question: can these partitions be expressed as Beatty sequences?

The previous table is partly repeated below, this time annotated with the OEIS sequence references.

n	binary	root	root identity	sequence
0	1	1		
1	11	$\varphi = \frac{1+\sqrt{5}}{2} = 1.618\cdots$	golden ratio	Fibonacci A000045
2	101	$1.465571231876768\cdots$	OEIS A092526	Narayana A000930
3	111	$1.839286755214161\cdots$	tribonacci A058265	tribonacci A000073
4	1001	$1.380277569097613\cdots$	2nd Pisot A086106	A003269, A017898
6	1101	$1.754877666246692\cdots$	A109134	A060945
7	1111	$1.927561975482925\cdots$	tetranacci A086088	tetranacci A000078
8	10001	$1.324717957244746\cdots$	silver A060006	A003520, A017899
10	10101	$1.570147312196054\cdots$	Pisot A293506	A060961
12	11001	$1.704902776041646\cdots$		A079971
13	11011	$1.812403619268042\cdots$		A079976
14	11101	$1.888518845484414\cdots$		A079975
15	11111	$1.965948236645485\cdots$	pentanacci A103814	A001591
16	100001	$1.28519903324535\cdots$		A005708, A017900

All of these integer sequences and roots participate in a number of curious relations having a regular form; this is, of course, the whole point of listing them in the OEIS. This suggests a question: do the known relationships generalize to the beta-shift setting?

For example, the polynomials of the form

$$\beta^{k+1} - \beta - 1 = 0$$

are the Lamé polynomials, they arise as solutions to Lamé's equation, a kind of ellipsoidal harmonic differential equation. In the present notation, these correspond to polynomials $p_n(\beta) = 0$ for $n = 2^k$.

Another example is the Fibonacci-tribonacci-tetranacci sequence of "generalized golden means". These are the roots of the series for which all $b_k = 1$, that is, the roots of

$$\beta^{k+1} - \beta^k - \beta^{k-1} - \cdots - 1 = 0$$

In the present notation, these would be the polynomials $p_n(\beta) = 0$ for $n = 2^k - 1$. Such roots can be rapidly computed by a series provided by Hare, Prodinger and Shallit[36]:

$$\frac{1}{\alpha_k} = \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{\infty} \frac{1}{j} \binom{j(k+1)}{j-1} \frac{1}{2^{j(k+1)}}$$

This series is obtained by making good use of the Lagrange inversion formula. Here, α_k is the k'th generalized golden mean, i.e. the solution $p_{2^k-1}(\alpha_k) = 0$. Can the Hare series be extended to provide the roots r_n of $p_n(r_n) = 0$ for general n?

Another set of observations seem to invoke the theory of complex multiplication on elliptic curves, and pose additional questions. So:

The tribonacci root r_3 is given by

$$r_3 = \frac{1}{3}\left(1 + \sqrt[3]{19 + 3\sqrt{33}} + \sqrt[3]{19 - 3\sqrt{33}}\right) \simeq 1.839\cdots$$

The silver number (plastic number) r_8 is given by

$$r_8 = \frac{1}{6}\left(\sqrt[3]{108 + 12\sqrt{69}} + \sqrt[3]{108 - 12\sqrt{69}}\right) \simeq 1.324\cdots$$

The Narayana's cows number (supergolden ratio) r_2 is given by

$$r_2 = \frac{1}{6}\sqrt[3]{116 + 12\sqrt{93}} + \frac{2}{3\sqrt[3]{116 + 12\sqrt{93}}} + \frac{1}{3} \simeq 1.465\cdots$$

The root r_6 is related to the silver number r_8 as $r_8 = r_6(r_6 - 1)$ and is given by

$$r_6 = \frac{1}{6}\sqrt[3]{108 + 12\sqrt{69}} + \frac{2}{\left(\sqrt[3]{108 + 12\sqrt{69}}\right)^2} \simeq 1.754\cdots$$

Do the other roots have comparable expressions? To obtain them, is it sufficient to articulate the theory of "complex multiplication" on elliptic curves? Are these just disguised solutions to Lamé's ellipsoidal harmonic differential equation? The appearance of only the cube and square roots is certainly suggestive of an underlying process of points on elliptic curves.

Other factoids:

$$r_2 = \frac{1}{3} + \frac{2}{3}\cos\left(\frac{1}{3}\arccos\left(\frac{29}{2}\right)\right)$$

140

5.11 β-Fibonacci sequences as shifts

The nature of the β-Fibonacci sequences as shift sequences can be emphasized by noting that they arise from the iteration of the companion matrix for the polynomial $p_n(x)$. This is a $(k+1) \times (k+1)$ matrix in lower-Hessenberg form:

$$B = \left[\begin{array}{cccccc} b_0 & 1 & 0 & 0 & \cdots & 0 \\ b_1 & 0 & 1 & 0 & \cdots & 0 \\ b_2 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ b_{k-1} & 0 & 0 & 0 & \cdots & 1 \\ b_k & 0 & 0 & 0 & \cdots & 0 \end{array} \right] \tag{39}$$

Iteration produces a linear recursive sequence this is the β-Fibonacci sequence. The m'th element of the sequence is obtained from the m'th iterate B^m.

Define the exchange matrix as

$$J = \left[\begin{array}{ccccc} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \cdot\!\cdot\!\cdot & & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{array} \right]$$

This can be used to write the above in the more conventional companion-matrix form:

$$C = [JBJ]^{\mathrm{T}} = \left[\begin{array}{cccccc} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ b_k & b_{k-1} & b_{k-2} & b_{k-3} & \cdots & b_0 \end{array} \right]$$

Some explicit examples are in order. For the golden ratio, one has

$$B = \left[\begin{array}{cc} 1 & 1 \\ 1 & 0 \end{array} \right]$$

and the iterates are

$$B^2 = \left[\begin{array}{cc} 2 & 1 \\ 1 & 1 \end{array} \right], \; B^3 = \left[\begin{array}{cc} 3 & 2 \\ 2 & 1 \end{array} \right], \; B^4 = \left[\begin{array}{cc} 5 & 3 \\ 3 & 2 \end{array} \right], \; B^n = \left[\begin{array}{cc} F_n & F_{n-1} \\ F_{n-1} & F_{n-2} \end{array} \right]$$

with F_n being the n'th Fibonacci number, as usual. For the general case, after $m \geq k-1$

iterations, one gets the Hankel matrix

$$
B^m = \begin{bmatrix}
F_m^{\{b\}} & F_{m-1}^{\{b\}} & F_{m-2}^{\{b\}} & \cdots & F_{m-k+1}^{\{b\}} & F_{m-k}^{\{b\}} \\
F_{m-1}^{\{b\}} & F_{m-2}^{\{b\}} & F_{m-3}^{\{b\}} & \cdots & F_{m-k}^{\{b\}} & F_{m-k-1}^{\{b\}} \\
F_{m-2}^{\{b\}} & F_{m-3}^{\{b\}} & & \cdots & & \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
F_{m-k+1}^{\{b\}} & & & \cdots & & \\
F_{m-k}^{\{b\}} & & & \cdots & & F_{m-2k}^{\{b\}}
\end{bmatrix}
$$

so that the top row consists of the latest sequence values. When multiplied by the bits, this just generates the next iterate in the sequence. The upper-diagonal 1's just serve to shift columns over by one, with each iteration: that is why it's a shift!

The product $B^m J$ is a Toeplitz matrix.

The characteristic polynomial of this matrix is, of course, the polynomial p_n:

$$
\det[B - xI] = (-1)^k p_n(x)
$$

Thus, we can trivially conclude that the eigenvalues of B are given by the roots of $p_n(x)$. This matrix is in lower-Hessenberg form; this makes it obvious that it's a shift; a finite shift, in this case.

5.12 Equivalent labels for orbits

There are many equivalent ways of labeling the various expressions and properties under consideration. These are recapped here.

5.12.1 Orbits

For every given $1 < \beta < 2$ there is a unique orbit of midpoints $\{m_p\}$ given by $m_p = T_\beta(m_{p-1}) = T_\beta^p(m_0)$ and $m_0 = \beta/2$. The orbits are in one-to-one correspondence with β. The midpoints are the same as the Renyi–Parry sequence; namely $T_\beta^p(\beta/2) = (\beta/2)t_\beta^p(1)$, recalling here the notation of eqn 8 and 10. Some orbits are of finite length; the rest are either eventually periodic or are ergodic.

5.12.2 Orbit encoding

The midpoint generates a unique sequence of bits $\{b_0, b_{1,}, \cdots, b_k, \cdots\}$ given by the left-right moves of the mid-point, as it is iterated. That is, $b_k = \Theta(m_k - 1/2)$ so that b_k is one if the midpoint is greater than half, else b_k is zero. Each bit-sequence is in one-to-one correspondence with β. Finite orbits have finite-length sequences.

5.12.3 Monotonicity

The compressor function $w(\beta) = \sum_k b_k 2^{-k}$ is a monotonically increasing function of β, so that values of $w(\beta)$ are in one-to-one correspondence with β.

5.12.4 Polynomial numbering

If the orbit is finite, then there exists a polynomial $p_n(z) = z^{k+1} - b_0 z^k - b_1 z^{k-1} - \cdots - b_{k-1} z - 1$ with $k = 1 + \lfloor \log_2(2n+1) \rfloor$ being the length of the orbit. The positive real root r_n of $p_n(r_n) = 0$ is $\beta = r_n$. That is, the iteration of r_n will generate the finite-length bit-sequence $\{b\} = \{b_0, b_1,, \cdots, b_k\}$. The integer n is in one-to-one correspondence with the bit sequence, and with the value of β. The integer is explicitly given by $2n+1 = \sum_{j=0}^{k} 2^j b_j$.

If the orbit is not finite, there is a function $q^{\{b\}}(\zeta) = 1 - \sum_{j=0}^{\infty} b_j \zeta^{j+1}$ holomorphic on the unit disk, having one unique positive real zero $q^{\{b\}}(r_{\{b\}}) = 0$ where this $r_{\{b\}} = 1/\beta$ is the same β that generated the bit-sequence $\{b\}$. Iterating $r_{\{b\}}$ generates $\{b\}$. If $\{b\}$ is finite, then $q^{\{b\}}(\zeta) = \zeta^{k+1} p_n(1/\zeta)$, so these functions agree on finite-length sequences.

5.12.5 Brackets

If the orbit is finite, then there exists a unique bracketing relationship $\ell \mapsto n \Leftarrow \rho$ for which n is the polynomial index. The left and right bounds ℓ, ρ are strictly smaller indexes: $\ell < n$ and $\rho < n$, and even more strongly, $2\ell \leq n$ and $2\rho \leq n$ that have the property of bounding the positive real roots of the corresponding polynomials: $r_\ell < r_n < r_\rho$, with $p_\ell(r_\ell) = p_n(r_n) = p_\rho(r_\rho) = 0$.

5.12.6 Binary tree

Each bracket is in one-to-one correspondence with a node in the full, unbounded binary tree. Sub-brackets define left and right sub-intervals that are disjoint, and whose union makes up the whole interval. Every node in the full binary tree can be labeled with a unique sequence of left-right moves to get to that node. This places the brackets (and thus, the polynomials and the roots and the mid-point orbits) in unique, one-to-one correspondences with finite-length strings of L,R moves. Such strings are, in turn, in one-to-one correspondence with the dyadic rationals. The L,R strings are in one-to-one correspondence with the orbits $\{b\}$ but they are *not* numerically the same! This are distinct sequences! In particular, all possible L,R moves are allowed. Only a limited number of orbits $\{b\}$ are possible, as limited by necklace-counting considerations.

5.12.7 Baire sequences

If the orbit is finite, then there exists a unique integer sequence $[m_1, m_2, \cdots, m_k]$ such that the index is given by $\eta[m_1, m_2, \cdots, m_k]$. This is a bijection between all valid indexes and all possible finite-length sequences. Due to the bounding property of the brackets, limits of $k \to \infty$ can be taken, and these limits are unique. Thus, all β values, including those with non-finite orbits, can be placed in a one-to-one bijection with infinite-length sequences $[m_1, m_2, \cdots] \in \mathbb{N}^\omega$.

5.12.8 Beta-Fibonacci sequences

If the orbit is finite, then there exists a sequence of integers $F^{\{b\}}$, the beta-Fibonacci sequence, that is in one-to-one correspondence with the finite bit sequence $\{b\} = b_0, b_1, \cdots, b_k$, and with the value of β. There are also sequences for each infinite-length orbit $\{b\}$.

5.12.9 Shift matrix

If the orbit is finite, then the finite bit sequence $\{b\} = b_0, b_1, \cdots, b_k$ defines a lower-Hessenberg "golden shift" matrix B, as shown in eqn 39. The limit of $k \to \infty$ can be taken in a relatively straightforward manner, given below.

5.12.10 Summary

To summarize: any one of these: the integer n, the polynomial $p_n(x)$, the bracket location in the binary tree, a dyadic rational, a point in Baire space, the integer sequence $F_m^{\{b\}}$, the orbit of midpoints $m_p = T^p(\beta/2)$, the orbit encoding $\{b\}$, the shift matrix B, the value of the compressor function $w(\beta)$ and, of course, β itself can each be used as a stand-in for the others, as they are all in one-to-one correspondence. Specifying one determines the others; all uniquely map to one-another. The formulas that provide maps between each of these can all be given in closed form, except for the handful of recursively-defined formulas. The recursive formulas are all invertible; thus they are computable (decidable). They are all equivalent labels. Fashionably abusing notation, $n \equiv p_n(x) \equiv r_n \equiv \{b\} \equiv F_m^{\{b\}} \equiv m_p \equiv w(\beta) \equiv \beta \equiv B$.

An explicit expression relating the orbit encoding and the orbit can be read off directly from eqn 7. Plugging in,

$$m_p = T_\beta^{p+1}\left(\frac{\beta}{2}\right) = \frac{\beta}{2}\left[\beta^{p+1} - \sum_{j=0}^{p} b_j \beta^{p-j}\right] \tag{40}$$

for $p < k$ the length of the bit sequence, and $m_k = T_\beta^{k+1}(\beta/2) = \beta p_n(\beta)/2 = 0$ terminating, since β is the positive root of $p_n(x)$.

Four of the correspondences given above ask for finite orbits. Three of these can be extended to non-finite orbits in an unambiguous and uncontroversial way. The extensions are covered in the next two sections. The fourth is the numbering n of the finite orbits. These are countable; there is no way to extend the counting number n to the non-finite orbits. Indeed, there are too many: the non-finite orbits are uncountable.

5.13 Infinite-nacci integer sequences

The beta-Fibonacci integer sequence can be extended to arbitrary (*viz.* infinite) length bit sequences, as

$$F_m^{\{b\}} = \sum_{j=1}^{m} b_{j-1} F_{m-j}^{\{b\}}$$

starting with $F_0^{\{b\}} = 1$. The sum is always finite, but one cannot perform it without first knowing at least the first m bits of the (now infinite) bit-sequence $\{b\}$. The integer sequence still has the desirable property it had before:

$$\beta = \lim_{m \to \infty} \frac{F_m^{\{b\}}}{F_{m-1}^{\{b\}}}$$

Here, the β value is the one associated to $\{b\}$. So, as before, the real number β and the bit sequence $\{b\}$ label exactly the same orbit.

Remarkably, one can be sloppy in how one deals with finite orbits with this extension. One has two choices that are equivalent: One choice is to truncate, so that the bit-sequence ends with all-zeros, effectively rendering it of finite length. The alternative is to allow it to continue periodically, forever. Either form results in the same β-Fibonacci sequence!

As an example, consider $\beta = 1.6$, which is close to the golden ratio, but not quite. It has an infinite non-periodic (nonrecurring) bit-sequence $\{b\} = 10101001010010100000100\cdots$. The generated integer sequence is $F_m^{\{b\}} = 1,1,1,2,3,5,8,12,20,32,51,82,130,209,335,535,\cdots$ which undershoots the Fibonacci sequence (12 appears, where we expected 13, and 20 instead of 21, and so on). The ratio of the last two is $535/335 = 1.597\cdots$ and the previous is $335/209 = 1.603\cdots$ and the ratio of successive elements eventually converges to 1.6. By comparison, the Fibonacci sequence is generated by the bit-string 1010101010... of alternating ones and zeros.

The β-Fibonacci representation of the orbits has the remarkable property that one does not need some *a priori* mechanism to know if some orbit is finite or not. This dual representation of finite orbits is reminiscent of a property commonly seen in Cantor space 2^ω representations of the real number line, where the dyadic rationals (which are countable, of course) map to two distinct bit-sequences (one ending in all-ones, the other ending in all-zeros). A more general setting for this is given in symbolic dynamics, where the totally disconnected Bernoulli scheme N^ω can be used to represent elements of certain countable sets two different ways. For $N = 10$, one famously has that 1.000...=0.999... as an example. So likewise here, one can toggle between finite and infinite-periodic strings. So, given a finite string $\{b\} = b_0, b_1, \cdots, b_{k-1}, b_k$ which has, by definition, $b_k = 1$, create a new finite string that is twice as long: $\{b'\} = b_0, b_1, \cdots, b_{k-1}, 0, b_0, b_1, \cdots, b_k$ which necessarily has exactly the same beta-Fiboanacci sequence. That is, $F_m^{\{b'\}} = F_m^{\{b\}}$. Once can repeat this process *ad infinitum*, obtaining an infinite periodic string. The difference between these two is simply the difference between a less-than sign, and a less-than-or-equal sign used in the generation of the orbit, as noted at the very beginning of this chapter. We have proven: finite orbits are exactly the same as infinite periodic orbits, at least when represented by real numbers and by integer sequences. Conversely, the difference between using $<$ and $\leq$ during iteration is immaterial for describing convergents.

5.14 Infinite β-Polynomials

An infinite polynomial is, of course, an analytic function. The goal here is to extend the finite framework. The definition of the polynomials above requires a finite bit

sequence. This can be extended to an asymptotic series, by writing first

$$p_n(z) = z^{k+1} \left(1 - b_0 z^{-1} - b_1 z^{-2} - \cdots - b_k z^{-k-1} \right)$$

Set $\zeta = 1/z$ to get

$$\zeta^{k+1} p_n \left(\frac{1}{\zeta} \right) = 1 - b_0 \zeta - b_1 \zeta^2 - \cdots - b_k \zeta^{k+1}$$

which extends to the holomorphic function

$$q^{\{b\}}(\zeta) = 1 - \sum_{j=0}^{\infty} b_j \zeta^{j+1}$$

This is manifestly holomorphic on the unit disk, as each coefficient is either zero or one. It has a positive real zero, of course: $q^{\{b\}}(1/\beta) = 0$. Comparing to eqn 24, we see that this is exactly the same function, or rather, it's negative. Indeed, following the definition, $b_n = d_n(1/2)$ and so $E(\beta; \zeta) = -q^{\{b\}}(\zeta)$.

This at last provides a foot in the door for correctly describing the eigenvalues of the β-transfer operator: they are in one-to-one correspondence with the zeros of $q^{\{b\}}(\zeta)$. As before, though, this only exposes a discrete spectrum in the region $1/\beta < |\lambda| \leq 1$; if there is any spectrum outside this region, the methods here cannot access it.

5.15 Tschirnhaus transformation

What happens is the Tschirnhaus transformation is applied to the polynomials? Where do they go to?

5.16 β-Hessenberg operator

Extending the golden shift matrix B of eqn 39 to an infinite-dimensional operator is a bit trickier. Of course, one could just declare the matrix elements of the operator to be this-and-such, but these matrix elements are with respect to what basis? Is the operator even bounded? The answer to the second question is obviously "no".

The characteristic equation of B is $\det(B - \lambda I) = (-1)^k p_n(\lambda) = 0$; the Frobenius-Perron eigenvalue $\beta > 1$ is too large, although the $k - 1 = \lfloor \log_2 n \rfloor$ other roots are conveniently arranged near the unit circle, more-or-less equidistant from one another. The solution is to rescale B by $1/\beta$. The Frobenius-Perron eigenvalue is now one, and the remaining eigenvalues distributed near or on a circle of radius $1/\beta$. We may as well take the transpose as well, so that $\mathscr{H}_\beta = B^T/\beta$ is in upper-Hessenberg form. Rescaled in this way, it now seems safe to declare, by fiat, that the operator $\mathscr{H}_\beta$ is the correct extension of the matrix B to infinite dimensions. Just to be explicit: given the bit-sequence $\{b\}$, the operator $\mathscr{H}_\beta$ has the matrix elements

$$\langle 0 | \mathscr{H}_\beta | j \rangle = \frac{b_j}{\beta}$$

$$\langle j+1 | \mathscr{H}_\beta | j \rangle = \frac{1}{\beta}$$

with all other entries being zero. This is clearly in upper-Hessenberg form, with the subdiagonal providing the shift.

Comparing to the upper-Hessenberg form of $\mathscr{L}_\beta$ of eqn 29, and the numerical results on it's eigenvalues, it seems clear that $\mathscr{H}_\beta$ and $\mathscr{L}_\beta$ must surely be similar. That is, there must be an operator S such that

$$\mathscr{L}_\beta = S^{-1}\mathscr{H}_\beta S$$

The invariant measure $\mathscr{L}_\beta \rho = \rho$ is mapped to $\sigma = S\rho$, where $\mathscr{H}_\beta \sigma = \sigma$ is the FP-eigenvector. It is easy to write down σ explicitly: $\sigma = \left(1, \beta^{-1}, \beta^{-2}, \cdots\right)$, that is, $\sigma_j = \beta^{-j}$. This is obviously so: the subdiagonal entries of $\mathscr{B}_\beta$ act as a shift on σ and the top row is just

$$1 = \sum_{j=0}^{\infty} \left\langle 0 \left| \mathscr{H}_\beta \right| j \right\rangle \sigma_j = \sum_{j=0}^{\infty} b_j \beta^{-j-1} = 1 - q^{\{b\}}\left(\frac{1}{\beta}\right) = 1$$

Although $\mathscr{H}_\beta$ is no more solvable than $\mathscr{L}_\beta$ in the wavelet basis is, it is certainly much, much easier to work with. It also reaffirms the Ansatz 22 for eigenfunctions. To be explicit: if v is a vector satisfying $\mathscr{H}_\beta v = \lambda v$, with vector elements v_j, then the function

$$v(x) = \sum_{j=0}^{\infty} d_j(x) v_j$$

is an eigenfunction of the transfer operator: that is, $\left[\mathscr{L}_\beta v\right](x) = \lambda v(x)$, or, explicitly:

$$\frac{1}{\beta}\left[v\left(\frac{x}{\beta}\right) + v\left(\frac{x}{\beta} + \frac{1}{2}\right)\right] \Theta\left(\frac{\beta}{2} - x\right) = \lambda v(x) \tag{41}$$

which is just eqn 17. So, for $\lambda = 1$, this is just $v = \sigma$ which is just eqn 22 for $z = 1$, the invariant measure, as always. But it also says more: the *only* solutions to $\mathscr{H}_\beta v = \lambda v$ are necessarily of the form $v = \left(1, (\lambda\beta)^{-1}, (\lambda\beta)^{-2}, \cdots\right)$, because the subdiagonal forces this shift. To satisfy the the top row of $\mathscr{H}_\beta$, one must have that

$$\lambda = \sum_{j=0}^{\infty} \left\langle 0 \left| \mathscr{H}_\beta \right| j \right\rangle v_j = \frac{1}{\beta} \sum_{j=0}^{\infty} \frac{b_j}{(\lambda\beta)^j} = \lambda\left(1 - q^{\{b\}}\left(\frac{1}{\lambda\beta}\right)\right) = \lambda$$

and so the eigenvalue λ is exactly the eigenvalue that solves the β-series $q^{\{b\}}(1/\lambda\beta) = 0$. This effectively concludes a proof: the solutions to this series are the only eigenvalues of the β-transfer operator; there are no others.

To recap: finite orbits have an associated shift matrix B; this extends naturally to a shift operator $\mathscr{H}_\beta$ for non-finite orbits. The shift operator has a sufficiently simple form that it's eigenvectors can be explicitly written down in closed form; they are necessarily coherent states[3]. The top row of the shift operator defines a holomorphic function $q^{\{b\}}$ whose zeros correspond to eigenstates of the shift operator. The holomorphic function

---

[3]The term "coherent states" comes from quantum optics: a state that can be written as an analytic series over states/vectors expressed in a different basis.

is determined by the binary digit sequence $\{b\}$. The binary digit sequence is obtained from the iterated midpoint, as $b_j = d_j(1/2)$ where $d_j(x) = 1$ if $x < T^n(\beta/2)$. This is enough to prove eqn 41 holds for the special value $x = 0$ (for *any* eigenvalue λ). It was previously proven that the vanishing of $q^{\{b\}}$ is independent of x, *i.e.* that eqn 41 holds for any x.

5.17 Hamburger moment problem

Closely related is the Hamburger moment problem. Given the sequence $F_m^{\{b\}}$ of generalized Fibonacci numbers, one can write the corresponding Hankel matrix, and then ask what corresponding measure corresponds to that sequence.

Perhaps the Hausdorff moment problem is more appropriate in this context... it asks for the measure on the unit interval, instead of the whole real-number line.

5.18 Eigenfunctions from finite orbits

To recap: eigenstates of the transfer operator correspond with the zeros of $q^{\{b\}}(\zeta)$, or, more precisely, the zeros for which $|\zeta| \le 1$. The reason for this limitation is that the eigenstates are explicitly given by

$$v(x) = \sum_{m=0}^{\infty} d_m(x) \zeta^m$$

for $\zeta = 1/\beta\lambda$; this is absolutely convergent only for $|\zeta| < 1$. One might hope to analytically continue this to the entire complex plane, but the continuation depends on the digit sequence $d_m(x)$. One might expect that an analytic continuation is impossible, as the $d_m(x)$ are ergodic, and thus throws up some kind of essential singularity at $|\zeta| = 1$ that cannot be continued past. We are lacking in tools and language to discuss this situation. Perhaps some insight can be gleaned by examining the periodic orbits...

5.18.1 Case n=1

Consider first $\beta = \varphi = 1.6180\cdots$ the golden ratio. The corresponding finite beta-polynomial is $q^{\{11\}}(\zeta) = 1 - \zeta - \zeta^2$; the infinite series is

$$q^{\{1010101\cdots\}}(\zeta) = 1 - \zeta - \zeta^3 - \zeta^5 - \cdots = \left(1 - \zeta - \zeta^2\right) / \left(1 - \zeta^2\right)$$

which has a positive real zero at $\zeta = 1/\varphi$ and poles at $\zeta = \pm 1$. The zero corresponds to the FP eigenvalue of one. The invariant measure is

$$v(x) = \sum_{m=0}^{\infty} \frac{d_m(x)}{\varphi^m} = \begin{cases} \varphi & \text{for } 0 \le x < \frac{1}{2} \\ 1 & \text{for } \frac{1}{2} \le x < \varphi \\ 0 & \text{for } \varphi \le x \end{cases}$$

There is a negative real zero at $\zeta = -\varphi$, but the eigenfunction summation is not convergent here.

5.18.2 Case n=2

The $n = 2$ case has the finite bitstring $\{b\} = 101$ and the periodic bitstring $\{b\} = 1001001\cdots$. The corresponding finite beta-polynomial is $q^{\{101\}}(\zeta) = 1 - \zeta - \zeta^3$; the infinite series is

$$q^{\{1001\cdots\}}(\zeta) = 1 - \zeta - \zeta^4 - \zeta^7 - \cdots = \left(1 - \zeta - \zeta^3\right) / \left(1 - \zeta^3\right)$$

which has a positive real zero at $\zeta = 1/\beta = 0.6823\cdots$ and three poles on the unit circle. The FP eigenvalue provides $\beta = 1.4655\cdots$. The invariant measure is

$$v(x) = \sum_{m=0}^{\infty} \frac{d_m(x)}{\beta^m} = \begin{cases} \frac{\beta}{\beta-1} & \text{for } 0 \le x < T\left(\frac{\beta}{2}\right) \\ \frac{1}{\beta-1} & \text{for } T\left(\frac{\beta}{2}\right) \le x < \frac{1}{2} \\ \frac{1/\beta}{\beta-1} & \text{for } \frac{1}{2} \le x < \beta \\ 0 & \text{for } \beta \le x \end{cases}$$

There are many equivalent ways to write the invariant measure; the above just selected some representatives from the coset of equivalent expressions. For example, the third entry could be written as $\beta = 1/\beta\,(\beta - 1)$.

5.18.3 Case n=3

The $n = 3$ case has the finite bitstring $\{b\} = 111$ and the periodic bitstring $\{b\} = 1101101\cdots$. The corresponding finite beta-polynomial is $q^{\{111\}}(\zeta) = 1 - \zeta - \zeta^2 - \zeta^3$; the infinite series is

$$q^{\{110110\cdots\}}(\zeta) = 1 - \zeta - \zeta^2 - \zeta^4 - \cdots = \left(1 - \zeta - \zeta^2 - \zeta^3\right) / \left(1 - \zeta^3\right)$$

which has a positive real zero at $\zeta = 1/\beta = 0.5436\cdots$ and three poles on the unit circle. The FP eigenvalue gives $\beta = 1.8392\cdots$. The invariant measure is
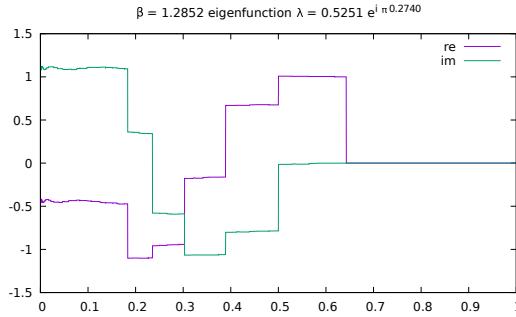
$$v(x) = \begin{cases} \frac{\beta}{\beta-1} & \text{for } 0 \le x < \frac{1}{2} \\ \beta & \text{for } \frac{1}{2} \le x < T\left(\frac{\beta}{2}\right) \\ \frac{1}{\beta-1} & \text{for } T\left(\frac{\beta}{2}\right) \le x < \beta \\ 0 & \text{for } \beta \le x \end{cases}$$

5.18.4 Case n=4,6,7

The pattern gets repetitive. There is no case $n = 5$, as this is not one of the allowed orbits. The bitstrings are those previously listed in tables; they are $\{b\} = 1001$ $\{b\} = 1101$ and $\{b\} = 1111$. The infinite series is $q^{\{b\cdots\}}(\zeta) = q^{\{b\}}(\zeta) / \left(1 - \zeta^4\right)$. The zeros are as previously listed. The $n = 4$ plateaus are at $\frac{1}{\beta-1}\left[\ \beta,\ \ 1,\ \ \frac{1}{\beta},\ \ \frac{1}{\beta^2}\ \right]$. The $n = 6$ plateaus are at $\left[\ \frac{\beta}{\beta-1},\ \ \frac{\beta^2}{(\beta^2-1)(\beta-1)},\ \ \beta,\ \ \frac{\beta}{(\beta^2-1)(\beta-1)}\ \right]$. The $n = 7$ plateaus are at $\left[\ \frac{\beta}{\beta-1},\ \ \beta,\ \ \frac{\beta+1}{\beta(\beta-1)},\ \ \frac{1}{\beta-1}\ \right]$. Again, the values at the plateaus can be written in many different ways, given the finite polynomial.

149

5.18.5 Case n=16

The $n = 16$ polynomial is the first one to have complex zeros inside the unit disk. The finite bitstring is $\{b\} = 100001$ and so the polynomial is $q^{\{100001\}}(\zeta) = 1 - \zeta - \zeta^6$. The positive real root is $\zeta = 0.7780895986786\cdots$ and so $\beta = 1/\zeta = 1.28519903324535\cdots$. The complex zeros are located at $\zeta = 0.965709509\cdots\exp\pm i\pi0.2740452363\cdots$ which corresponds to the eigenvalues $\lambda = 0.525107\cdots\pm i0.611100\cdots = 0.805718\cdots\exp\pm i\pi0.274045\cdots$. The corresponding eigenfunction is shown immediately below.

$\beta = 1.2852$ eigenfunction $\lambda = 0.5251\ e^{i\ \pi\ 0.2740}$

The order of $q^{\{b\}}$ is six, and this has six almost-plateaus; they are not quite flat, although they are close to it, presumably because ζ is close to one.

5.18.6 The general case

Generalizing from the above, one finds the following:

- For a period-k orbit, the infinite series is $q^{\{b\cdots\}}(\zeta) = q^{\{b\}}(\zeta)/(1 - \zeta^k)$.

- The first label n for which $q^{\{b\}}(\zeta)$ has a complex zero within the disk is $n = 16$. As a general rule, it seems that complex zeros inside the disk only appear for $\beta < \varphi$ (I believe; have not carefully checked. This seems reasonable, as later chapters show that the region of $\beta < \varphi$ behaves very differently from larger values.)

- The invariant measure has k plateaus. The plateau boundaries are given by $T^m\left(\frac{1}{2}\right)$ for $m = \{0, \cdots, k-1\}$ (so that $T^0\left(\frac{1}{2}\right) = \frac{1}{2}$ and $T^1\left(\frac{1}{2}\right) = \frac{\beta}{2}$, and so on).

- The leftmost plateau (of the invariant measure) is at $\beta/(\beta - 1) = \sum_{n=0}^{\infty} 1/\beta^n$.

- The other plateaus appear to be at simple rational functions of β, but a precise expression is elusive.

To solve the last issue, perhaps one can find tools in Galois theory. Let $\mathbb{R}[\zeta]$ be the ring of polynomials in ζ and consider the quotient ring $L = \mathbb{R}[\zeta]/q^{\{b\}}(\zeta)$. This L is a field extension of $\mathbb{R}$ and so one expects a Galois group $\text{Gal}(L/\mathbb{R})$. The plateaus of the invariant measure are presumably associated with the group elements. This seems like a promising direction to go in: perhaps this is just enough to explain the length

of an orbit, the sequence of points in the orbit, the reason that some polynomials are forbidden (they don't generate prime ideals), the appearance of Moreau's necklace-counting function, *etc*. This remains an unfinished exercise.

5.19 Factorization

The polynomials factorize. Let r_n denote the real positive root of $p_n(x)$ – that is, $p_n(r_n) = 0$. Then one has the factorizations (dropping the subscript on r for readability)

$$p_1(x) = x^2 - x - 1 = (x-r)(x+r-1) = (x-r)(x+p_0(r))$$

where $p_0(x) = x - 1$. Likewise, there are two order-3 polynomials. They factor as

$$p_2(x) = x^3 - x - 1 = (x-r)\left(x^2 + xp_0(r) + rp_0(r)\right)$$

while

$$p_3(x) = x^3 - x^2 - x - 1 = (x-r)\left(x^2 + xp_0(r) + p_1(r)\right)$$

Continuing in this way, there are three order-4 polynomials. They factor as

$$p_7(x) = x^4 - x^3 - x^2 - x - 1 = (x-r)\left(x^3 + x^2 p_0(r) + xp_1(r) + p_3(r)\right)$$

and

$$p_6(x) = x^4 - x^3 - x^2 - 1 = (x-r)\left(x^3 + x^2 p_0(r) + xp_1(r) + rp_1(r)\right)$$

and (noting that there is no p_5 that occurs in the series)

$$p_4(x) = x^4 - x^3 - 1 = (x-r)\left(x^3 + x^2 p_0(r) + xrp_0(r) + r^2 p_0(r)\right)$$

There's clearly a progression, but perhaps a bit difficult to grasp. It can be more clearly seen by writing $p_n = q_{2n+1}$ and then writing out $2n+1$ in binary. So, once again, from the top:

$$p_1(x) = q_{11}(x) = (x-r)(x+q_1)$$

where $q_1 = q_1(r)$ which adopts the shorthand that the q polynomials on the right-hand side always have r as an argument, which can be dropped for clarity. Note also that $q_0(r) = r$ was already previously observed, in an earlier section. That is, using the dropped-r convention, $q_0 = r$. Next

$$p_2(x) = q_{101}(x) = (x-r)\left(x^2 + xq_1 + q_{01}\right)$$

where, by definition, $q_{01}(x) \equiv rq_1(x)$. Next,

$$p_3(x) = q_{111}(x) = (x-r)\left(x^2 + xq_1 + q_{11}\right)$$

is the second factorization of order 3. For order 4, one has

$$p_4(x) = q_{1001}(x) = (x-r)\left(x^3 + x^2 q_1 + xq_{01} + q_{001}\right)$$

where, this time, $q_{001}(x) = xq_{01}(x) = x^2 q_1(x)$. Continuing,

$$p_6(x) = q_{1101}(x) = (x-r)\left(x^3 + x^2 q_1 + xq_{11} + q_{011}\right)$$

where, by definition, $q_{011}(x) \equiv xq_{11}(x)$. Finally,

$$p_7(x) = q_{1111}(x) = (x-r)\left(x^3 + x^2 q_1 + xq_{11} + q_{111}\right)$$

It is worth doing one more, just to clinch that the reversal of the bit sequence is indeed correct. For this purpose, $p_{12} = q_{11001}$ should serve well. One has

$$p_{12}(x) = q_{11001}(x) = (x-r)\left(x^4 + x^3 p_0(r) + x^2 p_1(r) + xrp_1(r) + r^2 p_1(r)\right)$$
$$= (x-r)\left(x^4 + x^3 q_1 + x^2 q_{11} + xq_{011} + q_{0011}\right)$$

The general pattern should now be clear. Given one of the admissible bit sequences $b_0 b_1 b_2 \cdots b_{k-1} b_k$ and recalling that $b_k = 1$ always, (and that $b_0 = 1$ always) one has

$$p_n(z) = q_{b_0 b_1 b_2 \cdots b_{k-1} b_k}(z) = z^{k+1} - b_0 z^k - b_1 z^{k-1} - \cdots - b_{k-1} z - 1$$

which has the factorization, with bits reversed:

$$q_{b_0 b_1 b_2 \cdots b_{k-1} b_k}(z) = (z-r)\left(z^k + z^{k-1} q_{b_0} + z^{k-2} q_{b_1 b_0} + z^{k-3} q_{b_2 b_1 b_0} + \cdots + q_{b_{k-1} b_{k-2} \cdots b_1 b_0}\right)$$

where, as already noted, each q is a polynomial in the root r. Although, notationally, the root r was taken as the real root, the above factorization works for any root.

The trick can be repeated. Although at first it might seem daunting, the pattern is uniform: every power of z occurred in the above. Let $s \neq r$ be some other root. Then

$$q_{b_0 b_1 b_2 \cdots b_{k-1} b_k}(z) = (z-r)(z-s)\left(z^{k-1} + (s + q_{b_0}) z^{k-2} + (s^2 + sq_{b_0} + q_{b_0 b_1}) z^{k-3} + \cdots\right)$$

The coefficient of the z^{k-4} term is $s^3 + s^2 q_{b_0} + sq_{b_0 b_1} + q_{b_0 b_1 b_2}$ and so on down the line. From this point on, this becomes just an ordinary factorization of polynomials... well, but so was the first step, as well. What made the first step interesting was that, because the coefficients at that first step were explicitly either zero or one, the corresponding reversal of the bit sequence became manifest.

One may as well bring this detour to a close. There's nothing particularly magic in the above factorization, other than the combinatorial rearrangement of the polynomial labels. A generic polynomial factorization looks like the below, for comparison. If

$$p(x) = x^{n+1} + c_0 x^n + c_1 x^{n-1} + \cdots c_n$$

and if r is a root of $p(x)$ *viz* $p(r) = 0$ then

$$p(x) = (x-r)\left(x^n + (r + c_0) x^{n-1} + (r^2 + c_0 r + c_1) x^{n-2} + \cdots\right)$$
$$= (x-r)\left(x^n + a_0 x^{n-1} + a_1 x^{n-2} + \cdots\right)$$

with

$$a_k = r^{k+1} + \sum_{j=0}^{k} c_j r^{k-j}$$

152

So the above works for any polynomial. The present case was special since the c_j were all either zero or minus one. The general form of this factorization is the one associated with the Vandermonde matrix.

TODO: Perhaps writing down the above as matrix equations will shed some insight? Maybe the matrix form will reveal some structure that is hidden, when working algebraically?

There are some notable values occurring in the factorization. These are shown in the table below:

v	n	bin	root r	q polynomial	OEIS	root of
2	1	11	$\varphi = 1.618\cdots$	$q_1 = 0.618\cdots$		
3	2	101	$1.465571\cdot$	$q_1 = 0.46557123187676\cdots$	A088559	$q^3 + 2q^2 + q - 1$
				$q_{01} = 0.68232780382801\cdots$	A263719	$q^3 + q - 1$
	3	111	$1.839286\cdot$	$q_1 = 0.83928675521416\cdots$		$q^3 - 2q^2 - 2$
				$q_{11} = 0.54368901269207\cdots$	A192918	$q^3 + q^2 + q - 1$
4	4	1001	$1.380277\cdot$	$q_1 = r - 1$		
				$q_{01} = 0.52488859865640\cdots$	A072223	$q^4 - 2q^2 - q + 1$
				$q_{001} = 0.72449195900051\cdots$		
	6	1101	$1.7548776\cdot$	$q_1 = r - 1$	A075778	$q^3 + q^2 - 1$
				$q_{11} = 0.32471795724474\cdots$	silver - 1	
				$q_{011} = 0.56984029099805\cdots$		
	7	1111	$1.9275619\cdot$	$q_1 = r - 1$		
				$q_{11} = 0.78793319384471\cdots$		
				$q_{111} = 0.51879006367588\cdots$		

As may be seen, some of these constants are already notable for various reasons. Many are also the real roots of yet other polynomials, of a not entirely obvious form. (Well, the q_1 polynomials will always be obvious expansions in binomial coefficients). The suggestion here is that these are all in turn part of some filigreed partially-ordered set of intertwining polynomials. Exactly how to express that intertwining in any sort of elegant or insightful way is not obvious.

6 Decaying Eigenfuncs

This section examines the general set of solutions to $\mathscr{L}_\beta \psi = \lambda \psi$ for $|\lambda| < 1$. Two general classes of solutions are presented. One consists of a discrete spectrum of (piecewise) polynomial eigenfunctions, occurring at the β values associated with the finite orbits. The other class consists of (piece-wise) fractal eigenfunctions, which can be constructed for any complex λ with $|\lambda| < 1$. That is, the continuous spectrum.

The discrete spectrum is closely analogous to the discrete spectrum of the polynomial eigenfunctions of the Bernoulli operator. Recall that the Bernoulli operator is the beta transform, for $\beta = 2$. It has a discrete spectrum with $\lambda = 2^{-n}$ for all nonnegative integers n; the corresponding eigenfunctions are the Bernoulli polynomials. The Bernoulli polynomials are orthonormal to one-another. In the $n \to \infty$ limit, they converge to either $\sin 2\pi x$ or $\cos 2\pi x$, depending on whether n is even or odd. Both of these trig functions, and any function odd about 1/2 lie in the kernel of the operator.

The Bernoulli operator also has a continuous spectrum that has a countable number of distinct eigenfunctions for each λ. [32] As these all share the same eigenvalue, one is free to choose the basis as desired. There are two generic possibilities. In one choice, the basis functions are square integrable, but are not differentiable at the dyadic rationals. These are built up as coherent states of underlying periodic waves. They resemble the blancmange curve; the periodic spectrum appears here, in the same way that the blancmange curve becomes a parabola for a specific parameter value.

A different choice of basis provides eigenfunctions that are smooth over the open interval $0 < x < 1$ but diverge at $x = 0, 1$ and are thus not square-integrable, due to these divergent endpoints. Equivalently, they are not smooth at $x = 0, 1$. The discrete spectrum emerges as those special values of λ for which the endpoints are finite, thus both smooth and square-integrable. Explicit constructions of both sets of basis functions, and the transformation between them, is given in [32].

This section presents explicit analogous solutions for the beta operator, for the β values corresponding to finite orbits. That is, both kinds of eigenfunctions are presented: piecewise-polynomial eigenfunctions, and fractal (coherent state) eigenfunctions. The development below is long and verbose, but never rises above the complexity level of basic algebra. There does not seem to be much need for abstract conceptualization here; it is mostly a matter of turning the crank and verifying identities.

6.0.1 Stability and numerics

On a tangent to the main topic, an interesting behavior was observed in numerical work. All of the eigenfunctions presented below can be implemented on a discrete lattice, and verified that they behave as expected. The primary utility of doing this is to double-check algebra and create the occasional pretty picture. However, such simulations lead to their own interesting behaviors.

Of primary note is that the eigenfunctions are remarkably unstable. One can set one down on a lattice, say, of 10K or 100K points, and iterate the transfer operator on it, point by point, for dozens or hundreds of iterations. Any given eigenfunction rarely survives a dozen iterations. Of course, finer lattices allow greater stability, but

that a 20K-point lattice is only good for a mixing time of only 10 or 15 steps indicates a strong coupling.

This is the usual case. In a few exceptional cases, stability has been observed: poor solutions converge on a good one, and stay there indefinitely. This suggests the lattice dynamics of the beta map have a life of their own.

In addition to the rare case of a stable attractor, a number of cases show a form of Poincare recurrence on extremely short timescales (dozens of iterations). It's not clear what the ultimate cause and explanation is, or what analogy to use. The recurrences resemble bet frequencies, Moire patterns, aliasing and sampling effects. The presence of a discrete lattice necessarily infects the system with p-adic and cyclotomic behaviors. Perhaps there is some resemblance to the Fermi–Pasta–Ulam–Tsingou problem, as the transfer operator does couple across the lattice, in an ergodic, possibly mixing kind of way.

6.1 Definitions

The conventions used in this section follow those used elsewhere in this text. Define $m_0 = \beta/2$ and $m_1 = T_\beta(m_0) = \beta(\beta - 1)/2$. Then the transfer operator can be written as

$$[\mathcal{L}_\beta f](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right)\Theta(m_0 - y) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta(m_1 - y)\right]$$

The task is to exhibit eigenfunctions with $\lambda < 1$; that is, to obtain solutions to

$$\lambda\beta f(y) = f\left(\frac{y}{\beta}\right)\Theta(m_0 - y) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta(m_1 - y)$$

The presentation works through a sequence of examples, generalizing at each step.

6.2 Examples

Case by case. This proceeds with the conventional polynomial and orbit numbering, so that n is a "valid index" number, and v is the length of the orbit.

6.2.1 Case n=1 (order $v = 2$)

Consider, for example, $\beta = \varphi \approx 1.618034$ the golden ratio. Then

$$\rho(y) = \begin{cases} \varphi y - \frac{1}{2} & \text{for } 0 \le y < \frac{1}{2} \\ y - \frac{1}{2} & \text{for } \frac{1}{2} \le y < \frac{\varphi}{2} \\ 0 & \text{for } \frac{\varphi}{2} \le y \le 1 \end{cases}$$

satisfies $\mathcal{L}_\varphi\rho = \rho/\varphi$ and so its a decaying eigenfunction. It has the L_1 norm $\int_0^1 |\rho(y)| dy = 1/4\varphi$. Note that $B_1(y) = y - 1/2$ is the first Bernoulli polynomial.

Orthogonality   The question of orthogonality is interesting. Well, interesting only because it is prone to thought-crime, and so an effort to educate seems needed. Note that $\int_0^1 \rho(y)\,dy = 0$; it this were not so, then ρ would not be orthogonal to the invariant measure μ. Here, the invariant measure is (un-normalized)

$$\mu(y) = \begin{cases} \varphi & \text{for } 0 \le y < \frac{1}{2} \\ 1 & \text{for } \frac{1}{2} \le y < \frac{\varphi}{2} \\ 0 & \text{for } \frac{\varphi}{2} \le y \le 1 \end{cases}$$

Now the subtle point arises. The integral $\int_0^1 \rho(y)\mu(y)\,dy \ne 0$, but this is not the correct way to compute orthogonality! The problem is the Jacobian is missing. The correct integral to determine the inner product $\langle f, g \rangle$ between f and g with respect to the invariant measure ν is

$$\langle f, g \rangle = \int_0^1 f(y)g(y)\left| \nu'\left(\nu^{-1}(y)\right)\right|^{-1} dy = \int_0^1 f(\nu(x))g(\nu(x))\,dx$$

In the present case, $\nu'(x) = \mu(x)$ ... Note the prime. Alas, this is a victim of the usual convention for naming things. Plugging through the rest of the way with $f = \rho$ and $g = \mu$ has g and the Jacobian cancel, leaving $\int_0^1 \rho(y)\,dy = 0$ as the correct orthogonality result. This works for the present, but requires the reinsertion of the Jacobian, if the product was of some other two functions.

6.2.2   Case n=2 (order $\nu = 3$)

Let r_2 be the positive real root of $p_2(\beta) = \beta^3 - \beta^2 - 1$. Some useful identities and factoids

- $r_2 \approx 1.465571231876768$

- $m_0 = r_2/2$ is the midpoint and its in the theta of the first term in the xfer

- $r_2(r_2 - 1)/2 = 1/2r_2 = T_{r_2}(r_2/2) = m_1$ appears in the second theta in the xfer.

- $m_1 < 1/2$.

Then

$$\rho_2(y) = \begin{cases} \beta^2 y - \frac{1}{2} & \text{for } 0 \le y \le m_1 \\ \beta y - \frac{1}{2} & \text{for } m_1 < y \le \frac{1}{2} \\ y - \frac{1}{2} & \text{for } \frac{1}{2} < y \le m_0 \\ 0 & \text{for } m_0 < y \le 1 \end{cases}$$

solves $\mathscr{L}_\beta \rho = \rho/\beta$ for $\beta = r_2$. Here's a sketchy proof.

- Write four intervals $I_0 = [m_0, 1]$ and $I_1 = \left[\frac{1}{2}, m_0\right]$ and $I_2 = \left[m_1, \frac{1}{2}\right]$ and $I_3 = [0, m_1]$

- Assume $f_1(y) = y - \frac{1}{2}$ on I_1 and some unknown functions $f_2(y)$ on I_2 and $f_3(y)$ on I_3.

- Then $\mathscr{L}\rho|_{I_1} = \frac{1}{\beta}\rho\left(\frac{y}{\beta}\right)\Big|_{I_1}$ because the second term is vanishes on the interval I_1. By definition, $\mathscr{L}\rho = \frac{1}{\beta}\rho(y)$ so $\rho\left(\frac{y}{\beta}\right)\Big|_{I_1} = \rho(y)|_{I_1}$. However, $I_1/\beta = I_2$ and so conclude that $f_2(y/\beta) = f_1(y)$ or $f_2(y) = f_1(\beta y)$.

- The next part proceeds recursively, as before, this time with $\mathscr{L}\rho|_{I_2} = \frac{1}{\beta}\rho\left(\frac{y}{\beta}\right)\Big|_{I_2}$ because the second term vanishes on I_2. Thus, much as before, $\rho\left(\frac{y}{\beta}\right)\Big|_{I_2} = \rho(y)|_{I_2}$. The interval arithmetic is that $I_2/\beta = \left[\frac{\beta-1}{2}, \frac{1}{2\beta}\right] \subset I_3$ and so conclude that $f_3(y/\beta) = f_2(y)$ or $f_3(y) = f_2(\beta y)$. All three functions have been determined.

- Verify that the last segment works as desired. This time, all terms participate, so $\mathscr{L}\rho|_{I_3} = \frac{1}{\beta}\rho\left(\frac{y}{\beta}\right) + \rho\left(\frac{y}{\beta} + \frac{1}{2}\right)\Big|_{I_3}$. The interval arithmetic is $I_3/\beta \subset I_3$ while $I_3/\beta + 1/2 = I_1$ and so one expects to find that $f_3(y) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$. Plugging through, this is satisfied when $\beta^3 - \beta^2 - 1 = 0$. But this is the defining eqn for r_2, so everything worked out.

That concludes the proof.

The L_1 norm is $\int_0^1 |\rho_2| = 1/4r_2^2$. Note that $\rho_2(m_0) = \rho_2(1/2) = \rho_2(m_1) = 1/2r_2^2$, so its a saw tooth, with the teeth aligned.

6.2.3 Case n=3 (order $v = 3$)

Let r_3 be the positive real root of $p_3(\beta) = \beta^3 - \beta^2 - \beta - 1$. Some useful identities and factoids

- $r_3 \approx 1.839286755214161$ and $m_0 \approx 0.9196$ and $m_1 \approx 0.7718$ and $1/r_3 \approx 0.5436$.

Write four intervals $I_0 = [m_0, 1]$ and $I_1 = [m_1, m_0]$ and $I_2 = \left[\frac{1}{2}, m_1\right]$ and $I_3 = \left[0, \frac{1}{2}\right]$. This works a bit different than the $n = 2$ case.

$$
\rho_3(y) = \begin{cases} 0 & \text{on } I_0 \\ y - \left(m_1 - \frac{1}{4\beta}\right) & \text{on } I_1 \\ \left(\frac{\beta+1}{\beta}\right)y - m_1 & \text{on } I_2 \\ \beta y - \left(m_1 - \frac{1}{4\beta}\right) & \text{on } I_3 \end{cases}
$$

solves $\mathscr{L}_\beta \rho = \rho/\beta$ for $\beta = r_3$. Written in this final form, it is not at all "obvious". Let's look at the proof mechanism for insight.

- Starting with the midpoint (endpoint) $m_0 = \beta/2$, iterate to obtain $T_\beta(m_0) = m_1$ and $T_\beta(m_1) = 1/2$. This last iteration is the terminus. Thus, we have a collection of intervals $0 < \frac{1}{2} < m_1 < m_0 < 1$.

- Assign numberings to these intervals, from right to left, so that $I_0 = [m_0, 1]$ and $I_1 = [m_1, m_0]$ and $I_2 = \left[\frac{1}{2}, m_1\right]$ and $I_3 = \left[0, \frac{1}{2}\right]$ (just as before).

157

- As a notational convenience, define restrictions on each interval, so that $\rho(y)|_{I_k} = f_k(y)$. Note that $f_0 = 0$.

- Treat $\mathscr{L}\rho|_{I_k} = \frac{1}{\beta}\rho(y)|_{I_k}$ as a recurrence relation, so as to relate $\rho(y)|_{I_k}$ to other intervals. This is $\rho(y)|_{I_k} = \rho\left(\frac{y}{\beta}\right)\Big|_{I_k}\Theta(m_0 - y) + \rho\left(\frac{y}{\beta} + \frac{1}{2}\right)\Big|_{I_k}\Theta(m_1 - y)$.

- The recurrence relation above requires two interval mappings. Define them as the "left map" $L : I_k \mapsto \frac{I_k}{\beta}$ and the "right map" $R : I_k \mapsto \frac{I_k}{\beta} + \frac{1}{2}$.

- Lemma: $LI_k \subseteq I_\ell$ for some $\ell > k$; likewise for $RI_k \subseteq I_r$ for some $k > r$. The specific values of ℓ, r are TBD, as well as whether the subset relation is strict, or an equality.

- Starting recursion with $\rho(y)|_{I_0} = 0$ gives no usable conclusions.

- The next interval is I_1. The left map is $L : I_1 \mapsto \frac{I_1}{\beta} = \left[\frac{\beta-1}{2}, \frac{1}{2}\right] \subset I_3$; the right map is $R : I_1 \mapsto \frac{I_1}{\beta} + \frac{1}{2} = \left[\frac{\beta}{2}, 1\right] = I_0$.

- By recursion and restriction, deduce that $f_1(y) = f_3\left(\frac{y}{\beta}\right) + f_0\left(\frac{y}{\beta} + \frac{1}{2}\right)$. Since $f_0 = 0$, the second term vanishes, and f_3 is determined by f_1 (and vice-versa).

- Repeat the process above for I_2 to get $L : I_2 \mapsto \frac{I_2}{\beta} = \left[\frac{1}{2\beta}, \frac{\beta-1}{2}\right] \subset I_3$ and $R : I_2 \mapsto \frac{I_2}{\beta} + \frac{1}{2} = \left[\frac{\beta+1}{2\beta}, \frac{\beta}{2}\right] = I_1$. The bound follows from $(\beta + 1)/2\beta = \beta(\beta - 1)/2 = m_1$ which follows from the defining polynomial $p_3(\beta) = 0$.

- By recursion and restriction, $f_2(y) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$. This gives f_2 in terms of f_1. Thus, all three restrictions f_1, f_2 and f_3 are determined by one-another.

- Repeating again for I_3 gives $L : I_3 \mapsto \frac{I_3}{\beta} = \left[0, \frac{1}{2\beta}\right] \subset I_3$ and $R : I_3 \mapsto \frac{I_3}{\beta} + \frac{1}{2} = \left[\frac{1}{2}, \frac{\beta+1}{2\beta}\right] = I_2$.

- By recursion and restriction, $f_3(y) = f_3\left(\frac{y}{\beta}\right) + f_2\left(\frac{y}{\beta} + \frac{1}{2}\right)$. This gives an additional constraint between the functions.

- Make the Ansatz that $f_1(y) = y - c$ for some constant c, to be determined.

- Plugging through gives f_2 and f_3. The final constraint fixes the value of c. The result, reassembled, gives ρ_3 as reported above.

The above is a rather long process, but is readily mechanized, and appears to describe an algorithm that will work for any finite-order polynomial index. There are only a few "tricky bits". These are:

- The need to sort the midpoint iterates $T_\beta^j(m_0)$ into ascending order. This is needed, in order to obtain a list of intervals I_k.

- The determination of $LI_k \subseteq I_j$ and likewise for RI_k. This seems to require some sort of brute-force elbow-grease; it there is a pattern, it is not yet clear.

The general algo is taken up below. But first one more special case.

6.2.4 Case n=4 (order $\nu = 4$)

Perhaps having the midpoint less than 1/2 makes things better? Let r_4 be the positive root of $\beta^4 - \beta^3 - 1$.

- $r_4 \approx 1.380277569097613$ with $m_0 \approx 0.6901$, $m_1 \approx 0.2624$, $m_2 \approx 0.3622$ and $1/r_4 \approx 0.7245$.

This works as easily as the $n = 2$ case. As before, write five intervals $I_0 = [m_0, 1]$ and $I_1 = \left[\frac{1}{2}, m_0\right]$ and $I_2 = \left[m_2, \frac{1}{2}\right]$ and $I_3 = [m_1, m_2]$ and $I_4 = [0, m_1]$. Then

$$
\rho_4(y) = \begin{cases} 0 & \text{on } I_0 \\ y - \frac{1}{2} & \text{on } I_1 \\ \beta y - \frac{1}{2} & \text{on } I_2 \\ \beta^2 y - \frac{1}{2} & \text{on } I_3 \\ \beta^3 y - \frac{1}{2} & \text{on } I_4 \end{cases}
$$

solves $\mathscr{L}_\beta \rho = \rho/\beta$ for $\beta = r_4$.

Unlike the $n = 3$ case, the regular pattern of the $n = 2$ case is reprised. One can conclude that higher values of n are not necessarily messy, but can exhibit recurring patterns seen at lower orders.

6.3 Quadratic eigenfunctions

The above examples all used piece-wise linear segments. One can also pursue parabolas. The general case is handled in the next section. A few specific cases are written out below.

6.3.1 Case n=1 (order $\nu = 2$)

For the $n = 1$ case, with $\beta = \varphi$, the function

$$
\rho(y) = \begin{cases} \varphi y^2 - y + \frac{1}{8} & \text{for } 0 \leq y < \frac{1}{2} \\ y^2 - y + \frac{\varphi}{8} & \text{for } \frac{1}{2} \leq y < \frac{\varphi}{2} \\ 0 & \text{for } \frac{\varphi}{2} \leq y \leq 1 \end{cases}
$$

satisfies $\mathscr{L}_\varphi \rho = \lambda \rho$ for $\lambda = \varphi^{-2}$, and so its a decaying eigenfunction. This generalizes to polynomial segments of any order. For a polynomial of order k, the eigenvalue will be $\lambda = \varphi^{-k}$. The general case is presented the next subsection.

The roots are

$$0 = \varphi y^2 - y + \frac{1}{8} \text{ solved by } y = \frac{\varphi - 1}{2} \left[1 \pm \sqrt{\frac{\varphi - 1}{2\varphi}} \right]$$

$$0 = y^2 - y + \frac{\varphi}{8} \text{ solved by } y = \frac{1}{2} \left[1 \pm \sqrt{\frac{\varphi - 1}{2\varphi}} \right]$$

with

$$\frac{\varphi - 1}{2} \approx 0.30901699$$

$$\sqrt{1 - \frac{\varphi}{2}} \approx 0.43701602$$

6.3.2 Case n=2 (order $\nu = 3$)

For the $n = 2$ case, it is computed explicitly below. Copying the result, it is

$$\rho(y) = \begin{cases} \frac{\beta-1}{\beta}y^2 - \frac{\beta-1}{\beta}y + \frac{\beta(\beta-1)}{4(\beta+1)} & \text{for } 0 \le y < m_1 = \frac{\beta(\beta-1)}{2} \\ (\beta-1)y^2 - \frac{\beta-1}{\beta}y + \frac{\beta-1}{4(\beta+1)} & \text{for } m_1 \le y < m_2 = \frac{1}{2} \\ y^2 - \frac{\beta-1}{\beta}y + \frac{\beta-1}{4\beta(\beta+1)} & \text{for } \frac{1}{2} \le y < m_0 = \frac{\beta}{2} \\ 0 & \text{for } m_0 \le y \le 1 \end{cases}$$

6.4 General case

The general case was mostly described above, in the $n = 3$ subsection. Let's abstract it in full for general finite orbits of order ν.

- Starting with the midpoint $m_0 = \beta/2$, obtain it's first iterate $m_1 = \beta(\beta - 1)/2$, and then the rest of them: $m_j = T_\beta^j(m_0)$. The iteration is presumed to stop with $m_\nu = 1/2$ for an orbit of order ν. This is exactly as before, for finite orbits.

- Sort the midpoints into ascending order, so that $0 < ... < m_0$.

- Encode the order with a sequence L,R if $m_{j+1} < m_j$ or not ... (why? Not needed.)

- Define the intervals I_k from above, ordered so that $I_\nu < I_{\nu-1} < \cdots < I_1 < I_0 = [m_0, 1]$.

- The recursive interval maps are the "left map" $L : I_k \mapsto \frac{I_k}{\beta}$ and the "right map" $R : I_k \mapsto \frac{I_k}{\beta} + \frac{1}{2}$.

- Lemma: $LI_k \subseteq I_\ell$ for some $\ell > k$; likewise for $RI_k \subseteq I_r$ for some $k > r$. The specific values of ℓ, r are TBD, as well as whether the subset relation is strict, or an equality. It is always the case that $LI_\nu \subset I_\nu$.

- As a notational convenience, define the restriction $\rho(y)|_{I_k} = f_k(y)$. There will be ν of these, not counting $f_0 = 0$.

- Treat $\mathscr{L}\rho|_{I_k} = \lambda\,\rho(y)|_{I_k}$ as a recurrence relation, so as to relate $\rho(y)|_{I_k}$ to other intervals. Note the (re-)appearance of λ as the explicit eigenvalue; it was implicitly taken as $\lambda = 1/\beta$ in the earlier examples. The recurrence relation, with this generalized eigenvalue, is then $\lambda\beta\,\rho(y)|_{I_k} = \rho\left(\frac{y}{\beta}\right)\Big|_{I_k}\Theta(m_0 - y) + \rho\left(\frac{y}{\beta} + \frac{1}{2}\right)\Big|_{I_k}\Theta(m_1 - y)$.

- Employ the recurrence relation to obtain ν relations between the interval maps f_k. Of these, $\nu - 1$ interrelate the f_k to one-another, while the last serves to fix any constants.

- This recurrence relation is explicitly $\lambda\beta f_k(y) = f_{Lk}\left(\frac{y}{\beta}\right)\Theta(m_0 - y) + f_{Rk}\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta(m_1 - y)$, where $Lk = j$ when $I_j \subseteq LI_k$ and likewise for Rk.

- Make the Ansatz that $f_k(y) = \sum_j a_{kj}y^j$ for some unknown constants a_{kj} to be determined. Plugging through, and collecting terms by powers of y gives equations for the a_{kj}. The leading term gives an expression for λ.

- Lemma: if $f_1(y)$ is of order k then the eigenvalue will be $\lambda = \beta^{-k}$.

There are three unresolved issues with the above: (1) the need to sort midpoints, (2) the precise form of the interval inclusion relations, and (3) the daunting algebra for fixing the a_{kj} and λ. The tables in the next section try to gain insight by writing out explicit cases. Unfortunately, they shine no light on the situation. Ignore them, and skip them and move to the section after.

6.5   Interval Inclusion Tables

The tables below try to gain insight by writing out explicit cases, but they shine no light on the situation. Ignore them, and hop to the next section.

Perhaps some insight can be gained into the midpoint cycles problem by examining the special cases again? These are given in the table below.

v	n	m_j	$2n+1$	cycle
2	1	1,0	11	-
3	2	1,2,0	101	-
	3	2,1,0	111	(12)
4	4	1,2,3,0	1001	-
	6	2,3,1,0	1101	(123)
	7	3,2,1,0	1111	(13)
5	8	1,2,3,4,0	10001	-
	10	3,1,4,2,0	10101	(1342)
	12	2,3,4,1,0	11001	(1234)
	13	2,4,1,3,0	11011	(1243)
	14	3,4,2,1,0	11101	(1324)
	15	4,3,2,1,0	11111	(14)(23)

Legend:

- The column labeled m_j gives the index order for j such that $0 < \cdots < m_j < \cdots < m_0$.

- The column $2n+1$ shows the midpoint bit-sequence, as usual.

- The cycle column shows the permutation of the midpoints in cycle notation.

Hmm. Nothing compelling in that table.

Perhaps some insight can be gained into the interval inclusion problem by examining the special cases? These are given in the table below.

v	n	m_j	LI_5	LI_4	LI_3	LI_2	LI_1	RI_5	RI_4	RI_3	RI_2	RI_1
2	1	1,0				$\subset I_2$	$\subset I_2$				$= I_1$	$= I_0$
3	2	1,2,0			$\subset I_3$	$\subset I_3$	$= I_2$			$= I_1$	$\subset I_0$	$\subset I_0$
	3	2,1,0			$\subset I_3$	$\subset I_3$	$\subset I_3$			$= I_2$	$= I_1$	$= I_0$
4	4	1,2,3,0		$\subset I_4$	$\subset I_4$	$= I_3$	$= I_2$		$= I_1$	$\subset I_0$	$\subset I_0$	$\subset I_0$
	6	2,3,1,0		$\subset I_4$	$\subset I_4$	$\subset I_3$	$\subset I_3$		$= I_2$	$\subset I_1$	$\subset I_1$	$= I_0$
	7	3,2,1,0		$\subset I_4$	$\subset I_4$	$\subset I_4$	$\subset I_4$		$= I_3$	$= I_2$	$= I_1$	$= I_0$
5	8	1,2,3,4,0	$\subset I_5$	$\subset I_5$	$= I_4$	$= I_3$	$= I_2$	$= I_1$	$\subset I_0$	$\subset I_0$	$\subset I_0$	$\subset I_0$
	10	3,1,4,2,0	$\subset I_5$	$\subset I_5$	$\subset I_5$	$= I_4$	$= I_3$	$= I_2$	$= I_1$	$\subset I_0$	$\subset I_0$	$\subset I_0$
	12	2,3,4,1,0	$\subset I_5$	$\subset I_5$	$= I_4$	$\subset I_3$	$\subset I_3$	$= I_2$	$\subset I_1$	$\subset I_1$	$\subset I_1$	$= I_0$
	13	2,4,1,3,0	$\subset I_5$	$\subset I_5$	$\subset I_5$	$\subset I_5$	$= I_4$	$= I_3$	$= I_2$	$= I_1$	$\subset I_0$	$\subset I_0$
	14	3,4,2,1,0	$\subset I_5$	$\subset I_5$	$\subset I_4$	$\subset I_4$	$\subset I_4$	$= I_3$	$\subset I_2$	$\subset I_2$	$= I_1$	$= I_0$
	15	4,3,2,1,0	$\subset I_5$	$\subset I_5$	$\subset I_5$	$\subset I_5$	$\subset I_5$	$= I_4$	$= I_3$	$= I_2$	$= I_1$	$= I_0$

The columns labeled LI_k and RI_k indicate what happens when L, R are applied to I_k. The result is marked as equality or strict inclusion. Although there is a pattern to the above table, it is difficult to express in words. More progress can be made by encoding the binary relationships as matrices. See next section.

6.5.1 Interval inclusion incidence matrices

Attempt to gain insight into the interval inclusions by writing matrices and graphs. The matrices are square matrices B, C, having matrix entries $B_{kj} = \Theta(I_j \subseteq LI_k)$ where

$$\Theta(I_j \subseteq LI_k) = \begin{cases} 1 & \text{if } I_j \subseteq LI_k \\ 0 & \text{if } I_j \nsubseteq LI_k \end{cases}$$

is the inclusion membership function. Likewise $C_{kj} = \Theta(I_j \subseteq RI_k)$. In tabular form:

ν	n	m_j	B	C	seq	ind	P	D
2	1	1,0	·1 ·1	·· 1·	2 1	L R	·1 1·	·1 11
3	2	1,2,0	·1· ··1 ··1	··· ··· 1··	2 3 1	L L R	·1· ··1 1··	·1· ··1 1·1
	3	2,1,0	··1 ··1 ··1	··· 1·· ·1·	3 1 2	L R R	··1 1·· ·1·	··1 1·1 ·11
4	4	1,2,3,0	·1·· ··1· ···1 ···1	···· ···· ···· 1···	2 3 4 1	L L L R	·1·· ··1· ···1 1···	·1·· ··1· ···1 1··1
	6	2,3,1,0	·1·· ··1· ···1 ···1	···· 1··· 1··· ·1··	3 1 4 2	L R L R	··1· 1··· ···1 ·1··	··1· 1·1· ···1 ·1·1
	7	3,2,1,0	···1 ···1 ···1 ···1	···· 1··· ·1·· ··1·	4 1 2 3	L R R R	···1 1··· ·1·· ··1·	···1 1··1 ·1·1 ··11
5	8	1,2,3,4,0	·1··· ··1·· ···1· ····1 ····1	····· ····· ····· ····· 1····	2 3 4 5 1	L L L L R	·1··· ··1·· ···1· ····1 1····	·1··· ··1·· ···1· ····1 1···1
	10	3,1,4,2,0	··1·· ···1· ····1 ····1 ····1	····· ····· ····· 1···· ·1···	3 4 5 1 2	L L L R R	··1·· ···1· ····1 1···· ·1···	··1·· ···1· ····1 1···1 ·1··1
	12	2,3,4,1,0	··1·· ·1··· ···1· ····1 ····1	····· 1···· 1···· 1···· ·1···	3 1 4 5 2	L R L L R	··1·· 1···· ···1· ····1 ·1···	··1·· 1·1·· ···1· ····1 ·1··1
	13	2,4,1,3,0	···1· ····1 ····1 ····1 ····1	····· ····· 1···· ·1··· ··1··	4 5 1 2 3	L L R R R	···1· ····1 1···· ·1··· ··1··	···1· ····1 1···1 ·1··1 ··1·1
	14	3,4,2,1,0	···1· ···1· ···1· ····1 ····1	····· 1···· ·1··· ·1··· ··1··	4 1 2 5 3	L R R L R	···1· 1···· ·1··· ····1 ··1··	···1· 1··1· ·1·1· ····1 ··1·1
	15	4,3,2,1,0	····1 ····1 ····1 ····1 ····1	····· 1···· ·1··· ··1·· ···1·	5 1 2 3 4	L R R R R	····1 1···· ·1··· ··1·· ···1·	····1 1···1 ·1··1 ··1·1 ···11

The structure of the B and C matrices have no strong, easy-to-describe pattern that jumps out. It is all more subtle. First, some general comments. These all follow from the nature of the interval inclusions, and are fairly easily derived.

- The sum $D = B + C$ never has any entries that have 2 in it. This follows from the L, R maps.

- The first column of B is always zero. This follows because L always shifts down; the topmost interval can never be a subinterval of some other interval.

- The first column of C always has at least one entry that is one. This follows because R always shifts up, and some interval must always be above the one or more lower intervals.

- There are one or more non-zero entries in the last column of B.

- The last column of C is always all-zeros.

- Every row in B has exactly one non-zero entry. This follows, because every interval is a subinterval of one and only one other interval.

- The top row of C is always all-zeros.

- It is always the case that $B(v, v) = 1$ and $B(v-1, v) = 1$.

A more opaque but important pattern is the shift structure across the two. For every row of B and C, the following holds: If $B(k, j) = 1$, then one and only one of the following possibilities hold:

- $B(k+1, j+1) = 1$ and $C(k+1, 1) = 0$

- $B(k+1, j+1) = 1$ and there does not exist any i such that $C(k, i) = 1$ and $C(k+1, i+1) = 1$

- $B(k+1, j+1) = 0$ and $C(k+1, 1) = 1$

- $B(k+1, j+1) = 0$ and there exists i such that $C(k, i) = 1$ and $C(k+1, i+1) = 1$

- $j = v$ and $C(k+1, 1) = 1$

- $j = v$ and there exists i such that $C(k, i) = 1$ and $C(k+1, i+1) = 1$

Since for every row k, there is always some column j for for which $B(k, j) = 1$, the above holds for all rows. It is saying that there is always a shift, either in the B matrix, or in the C matrix, and never both at the same time, for every row.

The six bullet points above are expressed as logic statements written in English. To be useful for calculations, they need to be converted into formulas. This will be needed later, in the section on the continuous spectrum. The structure above can be captured by two distinct functions. One is a sequencing function: given interval k (row k), return the next shift $j + 1$; the other function is the sequence indicator function: given interval k, was the shift obtained on the left or right? The table above shows these two functions, labeled as "seq" and "ind".

The "seq" function is a permutation; the corresponding permutation matrix is shown in the P column. The "sporadic" matrix $S = B + C - P$ is what is left over, after removing P.

The indicator function "ind" shows how the permutation matrix is assembled from B and C. One has that

$$P(k,j) = \Theta(\text{ind}(k) = L)B(k,j) + \Theta(\text{ind}(k) = R)C(k,j)$$
$$P_{kj} = \delta_{\text{ind}(k),L}B_{kj} + \delta_{\text{ind}(k),R}C_{kj}.$$

where $\Theta(x) = 1$ if x evaluates to true, else $\Theta(x) = 0$. The δ_{ab} is the Kronecker delta.

The sequence function "seq" is the permutation, expressed as a cycle:

$$P_{kj} = \delta_{\text{seq}(k),j} = \Theta(\text{seq}(k) = j)$$
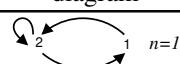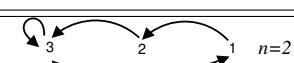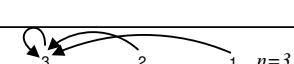
It is currently unclear how to obtain any of these matrices and functions, other than to brute-force iterate the midpoint, sort the midpoints into order, compute the intervals, and compute the overlaps. This is an algorithmic process that does not provide much insight.
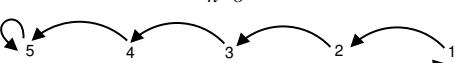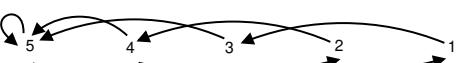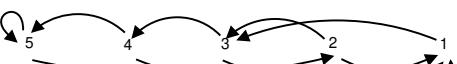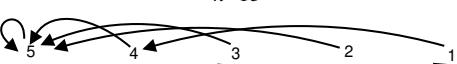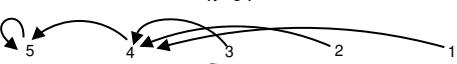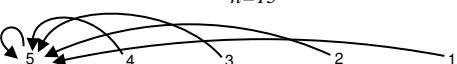
The permutation matrices are orthogonal, in that $P^{-1} = P^T$. There are some additional sporadic identities. For $P = P_n$ for index n, note that $P_2^T = P_3$ and $P_7^T = P_4$ and $P_{15}^T = P_8$ and $P_{13}^T = P_{10}$. Unclear how these interchanges might be meaningful; there's no corresponding relationship between P_{12} and P_{14}.

The product $P^T D$ is lower-triangular, and thus explicitly solvable. $P^T D$ is NOT upper-triangular.

The D matrix encodes the generating polynomial. That is, $\det[D - \lambda I] = p_n(\lambda)$ so that $\det[D - \beta I] = 0$.

Each matrix D can be taken to be the adjacency matrix for a graph. These are shown below; they show how intervals mix into one-another.

v	n	D	diagram
2	1	$\begin{bmatrix} \cdot\ 1 \\ 1\ 1 \end{bmatrix}$	
3	2	$\begin{bmatrix} \cdot\ 1\ \cdot \\ \cdot\ \cdot\ 1 \\ 1\ \cdot\ 1 \end{bmatrix}$	
	3	$\begin{bmatrix} \cdot\ \cdot\ 1 \\ 1\ \cdot\ 1 \\ \cdot\ 1\ 1 \end{bmatrix}$	
4	4	$\begin{bmatrix} \cdot\ 1\ \cdot\ \cdot \\ \cdot\ \cdot\ 1\ \cdot \\ \cdot\ \cdot\ \cdot\ 1 \\ 1\ \cdot\ \cdot\ 1 \end{bmatrix}$	

ν	n	D	diagram
	6	$\begin{bmatrix} \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & 1 \end{bmatrix}$	*n=6*
	7	$\begin{bmatrix} \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & 1 & 1 \end{bmatrix}$	*n=7*
5	8	$\begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & 1 \end{bmatrix}$	*n=8*
	10	$\begin{bmatrix} \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & 1 \end{bmatrix}$	*n=10*
	12	$\begin{bmatrix} \cdot & \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & 1 \end{bmatrix}$	*n=12*
	13	$\begin{bmatrix} \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot & 1 \end{bmatrix}$	*n=13*
	14	$\begin{bmatrix} \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot & 1 \end{bmatrix}$	*n=14*
	15	$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & 1 \end{bmatrix}$	*n=15*

Left-moving arrows indicate $LI_k \mapsto I_j = I_k/\beta$ while right-moving arrows indicate $RI_k \mapsto I_j = I_k/\beta + 1/2$. If only one arrow lands on an interval, then the map is an

167

equality (one-to-one and onto); otherwise it is an inclusion. The arrow-heads are drawn in proper interval sequence; the L and R maps never alter the overall sequence. Due to the self-inclusion $LI_v \mapsto I_v = I_v/\beta$ for the left-most interval, for all orders v, this map is always an inclusion (recursively splitting the interval into ever-smaller pieces). These in turn are mapped back to the right, eventually creating a set of enpoint mappings that is dense in the unit interval. Note also that there is always a map $LI_{v-1} \mapsto I_v$ for all orders v.

Some open questions:

- Why these graphs, and not others?

- Are these somehow the prime graphs, in that all other (measure-preserving) graphs can be constructed from these? For example, all Markovian two-point graphs can be taken as linear combinations of the $n = 0$ and $n = 1$ graphs. In this example, $M_{ij} = a\delta_{i1}\delta_{j1} + b\delta_{i2}\delta_{j2} + c\left(\delta_{i1}\delta_{j1} + \delta_{i2}\delta_{j1} + \delta_{i1}\delta_{j2}\right)$ represents all possible 2×2 graphs, where as much material flows into a vertex, as flows out of it. The a and b terms are just the $n = 0$ graph, while the c term is the $n = 1$ graph. (A Markov normalization additionally requires that $a + c = b + c = 1$.)

Another table; this one shows that DP^T is lower-triangular, but $P^T D$ is not.

v	n	D	S	P	DP^T	$P^T D$
2	1	$\begin{bmatrix} \cdot & 1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot \\ \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & 1 \\ 1 & \cdot \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ \cdot & 1 \end{bmatrix}$
3	2	$\begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \\ 1 & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & 1 \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \end{bmatrix}$
3	3	$\begin{bmatrix} \cdot & \cdot & 1 \\ 1 & \cdot & 1 \\ \cdot & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 \\ \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot \\ 1 & 1 & \cdot \\ 1 & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & 1 \\ \cdot & 1 & 1 \\ \cdot & \cdot & 1 \end{bmatrix}$
4	4	$\begin{bmatrix} \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}$
4	6	$\begin{bmatrix} \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot \\ \cdot & 1 & 1 & \cdot \\ \cdot & \cdot & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & 1 \end{bmatrix}$
4	7	$\begin{bmatrix} \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & 1 \end{bmatrix}$

v	n	D	S	P	DP^T	$P^T D$
5	8	$\begin{bmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$
	10	$\begin{bmatrix}\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&1&\cdot\\\cdot&\cdot&1&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$
	12	$\begin{bmatrix}\cdot&\cdot&1&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot\\\cdot&1&1&\cdot&\cdot\\\cdot&1&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&1&\cdot&\cdot\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&1\end{bmatrix}$
	13	$\begin{bmatrix}\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&1&1&\cdot&\cdot\\\cdot&1&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$
	14	$\begin{bmatrix}\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&1&\cdot\\\cdot&\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&1&\cdot\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&1\end{bmatrix}$
	15	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&1\\\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}$

As before, there is no particularly obvious way to obtain any of these matrices, except by brute-force iteration of the midpoints.

Another table. This one labels intervals according to the iteration. In this labeling, call it J, the interval J_k is the interval to the left of the midpoint iterate m_k. Define Q as the permutation that brings these "iteration-order" intervals into "sorted order" intervals. Nothing particularly interesting shows up here, and it seems that $Q^T DQ$ is just messy. The progression 4,6,7 seems to have a zipper pattern, but that zipper breaks at 10, which looks a mess.

v	n	m_j	D	P	Q	$Q^T DQ$
2	1	1,0	$\begin{bmatrix}\cdot&1\\1&1\end{bmatrix}$	$\begin{bmatrix}\cdot&1\\1&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot\\\cdot&1\end{bmatrix}$	$\begin{bmatrix}\cdot&1\\1&1\end{bmatrix}$

v	n	m_j	D	P	Q	$Q^T D Q$
3	2	1,2,0	$\begin{pmatrix}\cdot&1&\cdot\\\cdot&\cdot&1\\1&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&1&\cdot\\\cdot&\cdot&1\\1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot\\\cdot&\cdot&1\\\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&1\\1&1&\cdot\\\cdot&1&\cdot\end{pmatrix}$
	3	2,1,0	$\begin{pmatrix}\cdot&\cdot&1\\1&\cdot&1\\\cdot&1&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&1\\1&\cdot&\cdot\\\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot\\\cdot&1&\cdot\\\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&1\\1&\cdot&1\\\cdot&1&1\end{pmatrix}$
4	4	1,2,3,0	$\begin{pmatrix}\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&1&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\end{pmatrix}$
	6	2,3,1,0	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&1&\cdot\\1&\cdot&\cdot&1\\\cdot&1&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1\\\cdot&1&1&\cdot\\1&\cdot&1&\cdot\end{pmatrix}$
	7	3,2,1,0	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1\\\cdot&1&\cdot&1\\\cdot&\cdot&1&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1\\\cdot&1&\cdot&1\\\cdot&\cdot&1&1\end{pmatrix}$
5	8	1,2,3,4,0	$\begin{pmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{pmatrix}$
	10	3,1,4,2,0	$\begin{pmatrix}\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&1&\cdot&\cdot\end{pmatrix}$
	12	2,3,4,1,0	$\begin{pmatrix}\cdot&1&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&1&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&1&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\end{pmatrix}$
	13	2,4,1,3,0	$\begin{pmatrix}\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&1&1\\1&\cdot&1&\cdot&\cdot\end{pmatrix}$
	14	3,4,2,1,0	$\begin{pmatrix}\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&1&\cdot&\cdot\\\cdot&1&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\\cdot&\cdot&\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&1&\cdot\\\cdot&\cdot&1&1&\cdot\\\cdot&1&\cdot&1&\cdot\end{pmatrix}$
	15	4,3,2,1,0	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{pmatrix}$	$\begin{pmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\end{pmatrix}$	$\begin{pmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&1\end{pmatrix}$

170

As before, there is no particularly obvious way to obtain any of these matrices, except by brute-force iteration of the midpoints.

6.5.2   Change of basis

The matrix D encodes the the eigenvalues, in that $\det(D - \lambda I) = p_n(\lambda)$ and, as always, the largest real eigenvalue is $\lambda = \beta$. The eigenvectors are interpreted as distributions, in the basis where $e_k \mapsto \Theta(x \in I_k)$: that is, a function that is one on the interval I_k but zero elsewhere. By contrast, the Gelfond–Parry until invariant measure was stated in terms of $\Theta(x \le m_j)$. Explicitly, it is $\mu(x) = F^{-1} \sum_{k=0}^{\infty} \beta^{-j} \Theta(x \le m_j)$. We now wish to explore a change of basis from $e_k \mapsto \Theta(x \in I_k)$ to $u_j \mapsto \Theta(x \le m_j)$. This is not hard, it's just a triangular matrix. The painful part is to provide an adequate notation for all of this.

As always, the midpoints are $m_k = T^k(m_0)$ with $m_0 = \beta/2$. They are sorted so that $0 < \cdots < m_j < \cdots < m_0 < 1$. These define intervals, numbered right to left, with $I_0 = [m_0, 1]$ and $I_1 = [m_\pi, m_0]$ and $I_2 = [m_{\pi^2}, m_\pi]$ and so $I_k = [m_{\pi^k}, m_{\pi^{k-1}}]$ where π is the permutation that places the midpoints in sorted order. It is the same as the permutation matrix P given earlier, except for an off-by-one in the numbering. For order v there are a total of $v + 1$ intervals, with the left-most being $I_v = [0, m_{\pi^{v-1}}]$ since $m_{\pi^v} = 0$ by definition. That is, the midpoints in sorted order are

$$0 = m_{\pi^v} < m_{\pi^{v-1}} < m_{\pi^{v-2}} < \cdots < m_j = m_{\pi^i} < \cdots < m_\pi < m_0 < 1$$

With this notation, the union of intervals is

$$[0, m_j] = \bigcup_{k:m_k \le m_j} [m_{\pi k}, m_k]$$

where πk is the midpoint immediately to the left, in sorted order, not to be confused with π^k which is the permutation iterated k times. The notation is killing me. For a fixed j, the midpoint m_j appears at location i in the sort order: there is some integer i such that it is given by the permutation $\pi^i = j$. Thus, the right-most interval containing m_j is $I_{i+1} = [m_{\pi j}, m_j] = [m_{\pi^{i+1}}, m_{\pi^i}]$. The union runs over a total of $v - i$ intervals to the left of m_j:

$$[0, m_j] = \bigcup_{k=i+1; j=\pi^i}^{v} I_k$$

Fiddling with equivalent notation,

$$\Theta(x \le m_j) = \sum_{k=i+1; j=\pi^i}^{v} \Theta(x \in I_k)$$

which is the change of basis

$$u_j = \sum_{k=i+1; j=\pi^i}^{v} e_k$$

A concrete example follows.

Change of basis, explicit example So as not to lose track of what we are doing here, consider the case of $v = 2, n = 1$ which has $\beta = \varphi$ the golden ratio and $0 < m_1 = \varphi(\varphi - 1)/2 < m_0 = \varphi/2 < 1$ with intervals $I_2 = [0, m_1]$ and $I_1 = [m_1, m_0]$. In the union basis, the intervals are $U_1 = [0, m_1] = I_2$ and $U_0 = [0, m_0] = I_2 \cup I_1$ and the corresponding vectors are $u_1 = e_2$ and $u_0 = e_1 + e_2$. In the interval basis, we solve for $D\psi = \lambda\psi$ with

$$D = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

This has the solution

$$\psi = \begin{bmatrix} 1 \\ \varphi \end{bmatrix} = e_1 + \varphi e_2 = \Theta(x \in I_1) + \varphi\Theta(x \in I_2) = \begin{cases} \varphi & \text{for } 0 \leq x < m_1 \\ 1 & \text{for } m_1 \leq x < m_0 \end{cases}$$

which is just the usual Gelfond–Parry invariant measure, up to a normalization constant. In the union basis, this is

$$\psi = u_0 + \frac{1}{\varphi}u_1 = \Theta(x \leq m_0) + \frac{1}{\varphi}\Theta(x < I_1) = \begin{cases} \varphi & \text{for } 0 \leq x < m_1 \\ 1 & \text{for } m_1 \leq x < m_0 \end{cases}$$

since, of course, $1 + 1/\varphi = \varphi$ for the golden mean. These are equivalent descriptions, with the union form having the familiar shape

$$\psi = \sum_{j=0}^{v-1} \frac{1}{\beta^j}\Theta(x \leq m_j) = \sum_{j=0}^{v-1} \frac{1}{\beta^j}u_j$$

The change of basis can be written as $e_k = M_{kj}u_j$ with the matrix M being

$$M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad M^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

In the union basis, we have

$$H = M^{-1}DM = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

satisfying $H\psi = \varphi\psi$ with $\psi = u_0 + u_1/\varphi$ exactly as before.

Summary Table Armed with appropriate notation and a worked example, a table of explicit values for the remaining low-order cases provides some insight. The inverse M^{-1} has an interesing form: it appears to be a copy of the permutation Q from before, minus a column-shifted copy of Q. This is perhaps not a surprise.

ν	n	m_j	D	M	M^{-1}	$H = M^{-1}DM$	$K = M^{-1}CM$
2	1	1,0	$\begin{bmatrix}\cdot&1\\1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot\\1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot\\-1&1\end{bmatrix}$	$\begin{bmatrix}1&1\\1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot\\1&\cdot\end{bmatrix}$
3	2	1,2,0	$\begin{bmatrix}\cdot&1&\cdot\\\cdot&\cdot&1\\1&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot\\1&\cdot&1\\1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot\\\cdot&-1&1\\-1&1&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&1\\1&\cdot&\cdot\\\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot\\1&\cdot&\cdot\\\cdot&\cdot&\cdot\end{bmatrix}$
	3	2,1,0	$\begin{bmatrix}\cdot&\cdot&1\\1&\cdot&1\\\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot\\1&1&\cdot\\1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot\\-1&1&\cdot\\\cdot&-1&1\end{bmatrix}$	$\begin{bmatrix}1&1&1\\1&\cdot&\cdot\\\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot\\1&\cdot&\cdot\\\cdot&1&\cdot\end{bmatrix}$
4	4	1,2,3,0	$\begin{bmatrix}\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot\\1&\cdot&\cdot&1\\1&\cdot&1&1\\1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot\\\cdot&\cdot&-1&1\\\cdot&-1&1&\cdot\\-1&1&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\end{bmatrix}$
	6	2,3,1,0	$\begin{bmatrix}\cdot&1&\cdot&\cdot\\1&\cdot&1&\cdot\\1&\cdot&\cdot&1\\\cdot&1&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot\\1&1&\cdot&1\\1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot\\-1&1&\cdot&\cdot\\\cdot&\cdot&-1&1\\\cdot&-1&1&\cdot\end{bmatrix}$	$\begin{bmatrix}1&1&\cdot&1\\1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot\end{bmatrix}$
	7	3,2,1,0	$\begin{bmatrix}\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1\\\cdot&1&\cdot&1\\\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot\\1&1&1&\cdot\\1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot\\-1&1&\cdot&\cdot\\\cdot&-1&1&\cdot\\\cdot&\cdot&-1&1\end{bmatrix}$	$\begin{bmatrix}1&1&1&1\\1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot\\\cdot&\cdot&1&\cdot\end{bmatrix}$
5	8	1,2,3,4,0	$\begin{bmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1&1\\1&\cdot&1&1&1\\1&1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&-1&1\\\cdot&\cdot&-1&1&\cdot\\\cdot&-1&1&\cdot&\cdot\\-1&1&\cdot&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\end{bmatrix}$
	10	3,1,4,2,0	$\begin{bmatrix}\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&\cdot&1&\cdot&1\\1&1&1&\cdot&1\\1&1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&-1&1&\cdot\\-1&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&-1&1\\\cdot&-1&1&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&\cdot&1&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\end{bmatrix}$
	12	2,3,4,1,0	$\begin{bmatrix}\cdot&1&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot\\1&1&\cdot&1&\cdot\\1&1&\cdot&1&1\\1&1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\-1&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&-1&1\\\cdot&\cdot&-1&1&\cdot\\\cdot&-1&1&\cdot&\cdot\end{bmatrix}$	$\begin{bmatrix}1&1&\cdot&1&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\end{bmatrix}$
	13	2,4,1,3,0	$\begin{bmatrix}\cdot&\cdot&\cdot&1&\cdot\\\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot\\1&1&\cdot&1&\cdot\\1&1&\cdot&1&1\\1&1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\\cdot&-1&1&\cdot&\cdot\\\cdot&\cdot&\cdot&-1&1\\-1&1&\cdot&\cdot&\cdot\\\cdot&\cdot&-1&1&\cdot\end{bmatrix}$	$\begin{bmatrix}1&1&\cdot&1&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$
	14	3,4,2,1,0	$\begin{bmatrix}\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&1&\cdot\\\cdot&1&\cdot&1&\cdot\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot\\1&1&1&\cdot&\cdot\\1&1&1&\cdot&1\\1&1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\-1&1&\cdot&\cdot&\cdot\\\cdot&-1&1&\cdot&\cdot\\\cdot&\cdot&\cdot&-1&1\\\cdot&\cdot&-1&1&\cdot\end{bmatrix}$	$\begin{bmatrix}1&1&1&\cdot&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\end{bmatrix}$
	15	4,3,2,1,0	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&\cdot&1\\\cdot&1&\cdot&\cdot&1\\\cdot&\cdot&1&\cdot&1\\\cdot&\cdot&\cdot&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot\\1&1&1&\cdot&\cdot\\1&1&1&1&\cdot\\1&1&1&1&1\end{bmatrix}$	$\begin{bmatrix}1&\cdot&\cdot&\cdot&\cdot\\-1&1&\cdot&\cdot&\cdot\\\cdot&-1&1&\cdot&\cdot\\\cdot&\cdot&-1&1&\cdot\\\cdot&\cdot&\cdot&-1&1\end{bmatrix}$	$\begin{bmatrix}1&1&1&1&1\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$	$\begin{bmatrix}\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&\cdot&\cdot&\cdot\\\cdot&1&\cdot&\cdot&\cdot\\\cdot&\cdot&1&\cdot&\cdot\\\cdot&\cdot&\cdot&1&\cdot\end{bmatrix}$

173

Visible in the above is the general form, which is exactly what were expecting. It is

$$H = \begin{bmatrix} b_0 & b_1 & \cdots & b_{v-1} & b_v \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

This is exactly $H = \beta \mathcal{H}_\beta = B^T$ from before, in the β-Hessenberg section. This is transpose of eqn 39. Eigenvectors are exacly like before: $H\psi = \lambda\psi$ is solved by λ being any root to $p_n(\lambda) = 0$ and the eigenvector is $\psi = \sum_{j=0}^{v-1} \lambda^{-j} u_j$. For $\lambda = \beta$ this is identical to the Gelfond–Parry solution. So this was a very round-about check of an expected result.

The column $K = M^{-1}CM$ seems to almost have a predictable pattern to it. If it did, then the rest of the table could be filled in from that pattern. But that's the core issue: all of these structures seem to have some almost regular, predictable pattern, but exactly what that is remains elusive. Every variant seems to require explicit midpoint iteration; a simpler algorithmic or recursive description remains unmanifested.

6.6 Recursive Rework

Clearly, writing out the intervals, sorting them and labeling them blocks the path to generalizing to the non-finite orbits. But perhaps the sorting is not needed, and instead can be inferred directly from the recursive moves. The key ingredients were

- This recurrence relation $\lambda\beta f_k(y) = f_{Lk}\left(\frac{y}{\beta}\right)\Theta(m_0 - y) + f_{Rk}\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta(m_1 - y)$, where $Lk = j$ when $I_j \subseteq LI_k$ and likewise for Rk.

- Since the f_k are defined as restrictions, the extra Θ terms are superfluous: the interval restrictions already enforce the relations. Thus, one can safely write $\lambda\beta f_k(y) = f_{Lk}\left(\frac{y}{\beta}\right) + f_{Rk}\left(\frac{y}{\beta} + \frac{1}{2}\right)$.

- The Ansatz that $f_k(y) = \sum_j a_{kj} y^j$ for some unknown constants a_{kj} to be determined. Plugging through, and collecting terms by powers of y gives equations for the a_{kj}. The leading term gives an expression for λ.

Starting with only f_1 it is enough to work with the empty label 1 and then move to $L1$ and $R1$ and thence $L^21, LR1, RL1, R^21$ for the next set of intervals.

Based on this observation, it seems that a better numbering scheme is to use the visitation labels (below).

6.6.1 Case n=2 (order $v = 3$) quadratic

Apply the above algo for the $n = 2$ case, look at the steps explicitly, and see how it goes. This is very verbose.

- Polynomial: $p_2(\beta) = \beta^3 - \beta^2 - 1$

- Root: $r_2 \approx 1.465571231876768\cdots$

- Midpoints:

$$m_0 = \beta/2 \approx 0.732785615938384$$
$$m_1 = \beta\left(\beta - 1\right)/2 \approx 0.3411639019140096$$
$$m_2 = \beta^2\left(\beta - 1\right)/2 = 1/2$$

Thus $0 < m_1 < m_2 < m_0 < 1$.

- Intervals $I_3 = [0, m_1] < I_2 = [m_1, m_2] < I_1 = [m_2, m_0] < I_0 = [m_0, 1]$.

- The recursive interval maps are the "left map" $L : I_k \mapsto \frac{I_k}{\beta}$ and the "right map" $R : I_k \mapsto \frac{I_k}{\beta} + \frac{1}{2}$.

- Interval inclusions:

v	n	m_j	LI_3	LI_2	LI_1	LI_0	RI_3	RI_2	RI_1
3	2	1,2,0	$\subset I_3$	$\subset I_3$	$= I_2$	$\subset I_1$	$= I_1$	$\subset I_0$	$\subset I_0$

- The segments are $f_k\left(y\right) = \sum_j a_{kj} y^j$ so that

$$f_0\left(y\right) = 0$$
$$f_1\left(y\right) = a_{10} + a_{11}y + a_{12}y^2$$
$$f_2\left(y\right) = a_{20} + a_{21}y + a_{22}y^2$$
$$f_3\left(y\right) = a_{30} + a_{31}y + a_{32}y^2$$

- The recurrence relation is $\lambda \beta f_k\left(y\right) = f_{Lk}\left(\frac{y}{\beta}\right) + f_{Rk}\left(\frac{y}{\beta} + \frac{1}{2}\right)$, where $Lk = j$ when $I_j \subseteq LI_k$ and likewise for Rk. Based on the inclusion table, this is

$$\lambda \beta f_1\left(y\right) = f_2\left(\frac{y}{\beta}\right)$$
$$\lambda \beta f_2\left(y\right) = f_3\left(\frac{y}{\beta}\right)$$
$$\lambda \beta f_3\left(y\right) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

Note this is a linear equation.

- Plugging through,

$$\lambda \beta \left(a_{10} + a_{11}y + a_{12}y^2\right) = a_{20} + \frac{a_{21}}{\beta}y + \frac{a_{22}}{\beta^2}y^2$$

$$\lambda \beta \left(a_{20} + a_{21}y + a_{22}y^2\right) = a_{30} + \frac{a_{31}}{\beta}y + \frac{a_{32}}{\beta^2}y^2$$

$$\lambda \beta \left(a_{30} + a_{31}y + a_{32}y^2\right) = a_{30} + \frac{a_{31}}{\beta}y + \frac{a_{32}}{\beta^2}y^2 + a_{10} + a_{11}\left(\frac{y}{\beta} + \frac{1}{2}\right) + a_{12}\left(\frac{y}{\beta} + \frac{1}{2}\right)^2$$

$$= a_{30} + \frac{a_{31}}{\beta}y + \frac{a_{32}}{\beta^2}y^2 + a_{10} + \frac{a_{11}}{2} + \frac{a_{11}}{\beta}y + \frac{a_{12}}{4} + \frac{a_{12}}{\beta}y + \frac{a_{12}}{\beta^2}y^2$$

Collecting terms,

$$\lambda \beta a_{10} = a_{20} \qquad \lambda \beta a_{11} = \frac{a_{21}}{\beta} \qquad \lambda \beta a_{12} = \frac{a_{22}}{\beta^2}$$

$$\lambda \beta a_{20} = a_{30} \qquad \lambda \beta a_{21} = \frac{a_{31}}{\beta} \qquad \lambda \beta a_{22} = \frac{a_{32}}{\beta^2}$$

$$\lambda \beta a_{30} = a_{30} + a_{10} + \frac{a_{11}}{2} + \frac{a_{12}}{4}$$

$$\lambda \beta a_{31} = \frac{1}{\beta}\left(a_{31} + a_{11} + a_{12}\right)$$

$$\lambda \beta a_{32} = \frac{1}{\beta^2}\left(a_{32} + a_{12}\right)$$

Reducing, last one first:

$$\lambda \beta a_{32} = \frac{1}{\beta^2}\left(a_{32} + a_{12}\right)$$

$$= \frac{1}{\beta^2}\left(a_{32} + \frac{a_{22}}{\lambda \beta^3}\right)$$

$$= \frac{1}{\beta^2}\left(a_{32} + \frac{a_{32}}{\lambda^2 \beta^6}\right)$$

$$\lambda \beta^3 = 1 + \frac{1}{\lambda^2 \beta^6}$$

$$0 = \lambda^3 \beta^9 - \lambda^2 \beta^6 - 1$$

$$\lambda = \frac{1}{\beta^2}$$

Plugging in for λ for readability:

$$a_{10} = \beta a_{20} \qquad a_{11} = a_{21} \qquad a_{12} = \frac{a_{22}}{\beta}$$

$$a_{20} = \beta a_{30} \qquad a_{21} = a_{31} \qquad a_{22} = \frac{a_{32}}{\beta}$$

$$a_{30} = \beta \left(a_{30} + a_{10} + \frac{a_{11}}{2} + \frac{a_{12}}{4} \right)$$

$$0 = a_{11} + a_{12}$$

$$a_{32} = \frac{1}{\beta} \left(a_{32} + a_{12} \right)$$

- The above were regular and algorithmic, and can be gotten by uniform operations. The operations below are ad hoc, solving the assorted relations in some arbitrary order. They are sloppy Gaussian elimination, sloppy because not explicitly manifested. But for now, its quick and easy. So lets do it. Working back,

$$a_{12} = a_{32} \frac{\beta - 1}{\beta}$$

$$a_{22} = a_{32} \left(\beta - 1 \right)$$

$$a_{11} = a_{21} = a_{31} = -a_{12}$$

and finally, the constant term:

$$a_{30} = \beta \left(a_{30} + a_{10} + \frac{a_{11}}{2} + \frac{a_{12}}{4} \right)$$

$$= \beta \left(a_{30} + \beta^2 a_{30} - a_{32} \frac{\beta - 1}{4\beta} \right)$$

$$0 = a_{30} \left(\beta^3 + \beta - 1 \right) - a_{32} \frac{\beta - 1}{4}$$

$$a_{30} = a_{32} \frac{\beta - 1}{4 \left(\beta^3 + \beta - 1 \right)}$$

$$= a_{32} \frac{\beta - 1}{4\beta \left(\beta + 1 \right)}$$

and

$$a_{20} = a_{32} \frac{\beta - 1}{4 \left(\beta + 1 \right)}$$

$$a_{10} = a_{32} \frac{\beta \left(\beta - 1 \right)}{4 \left(\beta + 1 \right)}$$

Yuck, what a mess.

- The second from last step can be written as a matrix equation. It is:

$$
\lambda
\begin{bmatrix}
a_{10} \\ a_{20} \\ a_{30} \\ a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32}
\end{bmatrix}
= \frac{1}{\beta}
\left[
\begin{array}{ccc|ccc|ccc}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & \frac{1}{2} & 0 & 0 & \frac{1}{4} & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & \frac{1}{\beta} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{\beta} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\beta} & 0 & \frac{1}{\beta} & \frac{1}{\beta} & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\beta^2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\beta^2} \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\beta^2} & 0 & \frac{1}{\beta^2}
\end{array}
\right]
\begin{bmatrix}
a_{10} \\ a_{20} \\ a_{30} \\ a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32}
\end{bmatrix}
$$

  and so indeed, the lower-right corner is solved first. The off-diagonal blocks are coupling between the different orders.

- Anyway, the above 9×9 matrix is the representation of the FP operator $\mathscr{L}_\beta$, in the monomial basis, for this specific value of β.

- The shift-like structure arises from the interval inclusion table. Alas, the interval inclusion table is cryptic.

Lets summarize the results. Setting $a_{32} = 1$ as an overall scale factor, the following is the eigenfunction, obtained above:

$$
\rho(y) =
\begin{cases}
\frac{\beta-1}{\beta} y^2 - \frac{\beta-1}{\beta} y + \frac{\beta(\beta-1)}{4(\beta+1)} & \text{for } 0 \le y < m_1 = \frac{\beta(\beta-1)}{2} \\
(\beta-1) y^2 - \frac{\beta-1}{\beta} y + \frac{\beta-1}{4(\beta+1)} & \text{for } m_1 \le y < m_2 = \frac{1}{2} \\
y^2 - \frac{\beta-1}{\beta} y + \frac{\beta-1}{4\beta(\beta+1)} & \text{for } \frac{1}{2} \le y < m_0 = \frac{\beta}{2} \\
0 & \text{for } m_0 \le y \le 1
\end{cases}
$$

So there we have it.

6.6.2 Case n=2 (order $v = 3$) arbitrary power

The above computation can be repeated to arbitrary power. The interval inclusion table gave the relations for a 3×3 block of interval relations. These were

$$
\lambda \beta f_1(y) = f_2\left(\frac{y}{\beta}\right)
$$

$$
\lambda \beta f_2(y) = f_3\left(\frac{y}{\beta}\right)
$$

$$
\lambda \beta f_3(y) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)
$$

In matrix form, this block has the shape

$$\lambda \begin{bmatrix} f_1(y) \\ f_2(y) \\ f_3(y) \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_1\left(\frac{y}{\beta}\right) \\ f_2\left(\frac{y}{\beta}\right) \\ f_3\left(\frac{y}{\beta}\right) \end{bmatrix} + \frac{1}{\beta} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_1\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ f_2\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ f_3\left(\frac{y}{\beta}+\frac{1}{2}\right) \end{bmatrix}$$

with the arguments to the f_k suitably twizzled.

The next step is to expand $f_k(y) = \sum_{j=0}^{\infty} a_{kj} y^j$ in place, and collect terms. The first two rows are easy; they are shifts:

$$\lambda \beta \left(\sum_j a_{1j} y^j \right) = \sum_j \frac{a_{2j}}{\beta^j} y^j$$

$$\lambda \beta \left(\sum_j a_{2j} y^j \right) = \sum_j \frac{a_{3j}}{\beta^j} y^j$$

The last row is more complex; it requires expansion using binomial coefficients:

$$\lambda \beta \left(\sum_j a_{3j} y^j \right) = \sum_j \frac{a_{3j}}{\beta^j} y^j + \sum_j a_{1j} \left(\frac{1}{2} + \frac{y}{\beta} \right)^j$$

$$= \sum_j \frac{a_{3j}}{\beta^j} y^j + \sum_j a_{1j} \sum_{m=0}^{j} \binom{j}{m} y^m \frac{1}{\beta^m} \cdot \frac{1}{2^{j-m}}$$

$$= \sum_{m=0}^{\infty} y^m \frac{1}{\beta^m} \left(a_{3m} + 2^m \sum_{j=m}^{\infty} a_{1j} \frac{1}{2^j} \binom{j}{m} \right)$$

The algebraic form above is opaque; writing it out in blocks is revealing:

$$\lambda \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ \vdots \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} D & \frac{1}{2}C & \frac{1}{4}C & C_{03} & \cdots \\ 0 & \frac{1}{\beta}D & \frac{1}{\beta}C & C_{13} & \cdots \\ 0 & 0 & \frac{1}{\beta^2}D & C_{23} & \\ 0 & 0 & 0 & \frac{1}{\beta^3}D & \ddots \\ \vdots & \vdots & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ \vdots \end{bmatrix}$$

with

$$A_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad D = B + C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

The off-diagonal blocks are scalar multiples $C_{jm} = c_{jm} C$ with

$$c_{jm} = \left(\frac{2}{\beta} \right)^m \frac{1}{2^j} \binom{j}{m}$$

179

This gives the final desired result: the operator $\mathscr{L}_\beta$ for this particular fixed β is that upper-diagonal block operator. It can be truncated at any fixed j, giving a $3j \times 3j$ matrix, with eigenvalue of β^{-j} and an eigenfunction consisting of three piece-wise polynomial parts. Setting $j = 0$ returns the Gelfond–Parry invariant measure, for this fixed β. Setting $j = 2$ returns the result of the previous section. Truncating at fixed j also allows for a hunt of complex eigenvalues.

6.6.3 Case n=3 (order $v = 3$)

It is not yet clear how the interval inclusions are to be interpreted. Thus, a few more special cases are worth looking at. This is a shortened repetition of the $n = 2$ case.

- Polynomial: $p_3(\beta) = \beta^3 - \beta^2 - \beta - 1$

- Root: $r_3 \approx 1.839286755214161\cdots$

- Midpoints:

$$m_0 = \beta/2 \approx 0.91964337760708$$
$$m_1 = \beta(\beta - 1)/2 \approx 0.771844506346038$$
$$m_2 = \beta(\beta^2 - \beta - 1)/2 = 1/2$$

Thus $0 < m_2 < m_1 < m_0 < 1$.

- Intervals $I_3 = [0, m_2] < I_2 = [m_2, m_1] < I_1 = [m_1, m_0] < I_0 = [m_0, 1]$.

- The recursive interval maps are the "left map" $L : I_k \mapsto \frac{I_k}{\beta}$ and the "right map" $R : I_k \mapsto \frac{I_k}{\beta} + \frac{1}{2}$.

- Interval inclusions:

v	n	m_j	LI_3	LI_2	LI_1	RI_3	RI_2	RI_1
3	3	2,1,0	$\subset I_3$	$\subset I_3$	$\subset I_3$	$= I_2$	$= I_1$	$= I_0$

- The recurrence relation is $\lambda\beta f_k(y) = f_{Lk}\left(\frac{y}{\beta}\right) + f_{Rk}\left(\frac{y}{\beta} + \frac{1}{2}\right)$, where $Lk = j$ when $I_j \subseteq LI_k$ and likewise for Rk. Based on the inclusion table, this is

$$\lambda\beta f_1(y) = f_3\left(\frac{y}{\beta}\right)$$
$$\lambda\beta f_2(y) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$
$$\lambda\beta f_3(y) = f_3\left(\frac{y}{\beta}\right) + f_2\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

- This has the twizzled form

$$\lambda \begin{bmatrix} f_1(y) \\ f_2(y) \\ f_3(y) \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_1\left(\frac{y}{\beta}\right) \\ f_2\left(\frac{y}{\beta}\right) \\ f_3\left(\frac{y}{\beta}\right) \end{bmatrix} + \frac{1}{\beta} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} f_1\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ f_2\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ f_3\left(\frac{y}{\beta}+\frac{1}{2}\right) \end{bmatrix}$$

- The blocks are

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad D = B + C = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

- Expanding $f_k(y) = \sum_{j=0}^{\infty} a_{kj} y^j$ and gathering together monomials yields the same block form as before:

$$\lambda \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \vdots \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} D & C_{01} & C_{02} & \cdots \\ 0 & \frac{1}{\beta}D & C_{12} & \cdots \\ 0 & 0 & \frac{1}{\beta^2}D & \\ \vdots & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \vdots \end{bmatrix}$$

with $C_{jm} = c_{jm}C$ as before, with

$$c_{jm} = \left(\frac{2}{\beta}\right)^m \frac{1}{2^j} \binom{j}{m}$$

as before, and of course

$$A_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{bmatrix}$$

Comparing to the previous section, the generic form is now visible.

6.7 General case, redux

The above provides sufficient notation to discern the general case. From the top:

- The goal is to solve $\mathscr{L}_\beta f = \lambda f$ to obtain the eigenvalue λ and eigenfunction f. This will be done for β values that have finite orbits. These values are given by the largest real root of the corresponding polynomial.

- Fix n, selecting a polynomial $p_n(\beta)$ of order v. Solve $p_n(\beta) = 0$ to obtain a root $r_n = \beta$.

- Iterate the midpoint $m_0 = \beta/2$ to get $m_1 = T_\beta(m_0) = \beta(\beta-1)/2$, ..., $m_k = T_\beta^k(m_0)$, ..., $m_v = 1/2$.

- Use the midpoints to split the interval $[0,1]$ into ν distinct sub-intervals. Provide a label k for each interval, so that the intervals I_k can be referred to by their labels. It is generally easiest to sort these into ascending order, but this is not mandatory.

- Define the recursive interval maps $L : I_k \mapsto \frac{I_k}{\beta}$ and $R : I_k \mapsto \frac{I_k}{\beta} + \frac{1}{2}$.

- Starting with $I_0 = [m_0, 1]$, define an interval inclusion graph.

- The interval inclusion graph determines incidence matrices. These are square matrices B, C, having matrix entries $B_{kj} = \Theta(I_j \subseteq LI_k)$ where

$$\Theta(I_j \subseteq LI_k) = \begin{cases} 1 & \text{if } I_j \subseteq LI_k \\ 0 & \text{if } I_j \nsubseteq LI_k \end{cases}$$

  is the inclusion membership function. Likewise $C_{kj} = \Theta(I_j \subseteq RI_k)$.

- For each interval k, define the component $f_k(y) = f(y)|_{I_k}$ as the restriction of f to the interval I_k.

- On each interval, the pushforward $\mathcal{L}_\beta$ takes the form of a recurrence relation $\lambda \beta f_k(y) = f_{Lk}\left(\frac{y}{\beta}\right) + f_{Rk}\left(\frac{y}{\beta} + \frac{1}{2}\right)$, where $Lk = j$ when $I_j \subseteq LI_k$ and likewise for Rk.

- Explicitly inserting the membership function Θ into the recurrence relation gives the linear equation

$$\lambda f_k(y) = \frac{1}{\beta} \sum_{m=1}^{\nu} \left[\Theta(LI_k \subseteq I_m) f_m\left(\frac{y}{\beta}\right) + \Theta(RI_k \subseteq I_m) f_m\left(\frac{y}{\beta} + \frac{1}{2}\right) \right]$$

- This is more visually transparent when written in block form a

$$\lambda \begin{bmatrix} f_1(y) \\ \vdots \\ f_\nu(y) \end{bmatrix} = \frac{1}{\beta} B \begin{bmatrix} f_1\left(\frac{y}{\beta}\right) \\ \vdots \\ f_\nu\left(\frac{y}{\beta}\right) \end{bmatrix} + \frac{1}{\beta} C \begin{bmatrix} f_1\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ \vdots \\ f_\nu\left(\frac{y}{\beta} + \frac{1}{2}\right) \end{bmatrix}$$

  with the square matrices B, C as just defined.

- For each interval k, write $f_k(y) = \sum_j a_{kj} y^j$ for unknown constants a_{kj} to be solved for. Thus, the eigenfunction will be piece-wise polynomial. The above forms a linear system that is solvable.

- Insert the polynomial pieces, and gather terms into monomials y^j. This requires the use of binomial coefficients to write

$$\left(\frac{y}{\beta} + \frac{1}{2}\right)^j = \sum_{m=0}^{j} \binom{j}{m} y^m \frac{1}{\beta^m} \cdot \frac{1}{2^{j-m}}$$

so that

$$f_k\left(\frac{y}{\beta}+\frac{1}{2}\right) = \sum_j a_{kj}\left(\frac{1}{2}+\frac{y}{\beta}\right)^j$$

$$= \sum_{m=0}^{\infty} y^m \left(\frac{2}{\beta}\right)^m \sum_{j=m}^{\infty} a_{kj}\frac{1}{2^j}\binom{j}{m}$$

- Grouping into monomials results in the block-triangular form

$$\lambda\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \vdots \end{bmatrix} = \frac{1}{\beta}\begin{bmatrix} D & C_{01} & C_{02} & \cdots \\ 0 & \frac{1}{\beta}D & C_{12} & \cdots \\ 0 & 0 & \frac{1}{\beta^2}D & \\ \vdots & \vdots & & \ddots \end{bmatrix}\begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \vdots \end{bmatrix}$$

with $D = B+C$ and with $C_{jm} = c_{jm}C$ with

$$c_{jm} = \left(\frac{2}{\beta}\right)^m \frac{1}{2^j}\binom{j}{m}$$

and block vectors

$$A_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{vj} \end{bmatrix}$$

- In other words,

$$\mathscr{L}_\beta = \frac{1}{\beta}\begin{bmatrix} D & C_{01} & C_{02} & \cdots \\ 0 & \frac{1}{\beta}D & C_{12} & \cdots \\ 0 & 0 & \frac{1}{\beta^2}D & \\ \vdots & \vdots & & \ddots \end{bmatrix}$$

is a matrix representation of $\mathscr{L}_\beta$ in a monomial-basis form. It is upper-triangular. It is solvable. It can be truncated at any finite j to give a finite $vj \times vj$ matrix.

- For any fixed v, the eigenfunction consists of v pieces, each piece being a polynomial of the same degree j.

- For such a truncated, finite-size matrix, the spectrum of $\mathscr{L}_\beta$ is determined entirely by the spectrum of the block D, as solvability implies that only the bottom-right block matters. Thus, one concludes that the spectrum of $\mathscr{L}_\beta$ is just the spectrum of D times β^{-j}.

- The spectral radius of D is one. In general, it will have complex eigenvalues; these are bounded below by $1/\beta$. Much as in the $j = 0$ case, these accumulate onto the circle of radius $1/\beta$ as v gets large.

183

- Given this block-triangular structure, the consequence is that the largest real eigenvalue of $\mathscr{L}_\beta$ will be β^{-j} and there will in general be additional complex eigenvalues λ satisfying $\beta^{-j-1} < |\lambda| < \beta^{-j}$. These are determined entirely by the spectrum of D.

This is a complete description for such piece-wise polynomial solutions. The solution decouples the behavior with respect to the monomials, which is captured in the D,C block matrix, from the behavior with respect to the intervals, which is captured in the individual blocks B,C. This allows limiting behavior can be explored in two distinct ways. One is the replacement of the monomial basis by some other basis, possibly not polynomial. Another is the $\nu \to \infty$ limit, where the number of plateaus becomes unbounded.

The factorization into block-triangular form indicates that the spectrum of $\mathscr{L}_\beta$ is determined entirely by the spectrum of D, and thus it is this spectrum that should be the subject of focus; the rest is just decorative trimmings. The only irregular or "difficult" part to the above is the calculation of the interval inclusions, needed to obtain the left and right incidence matrices B and C.

That's it. That's the complete solution for the general case.

6.8   General case, distillation

Lets distill some more. The ingredients seem to be:

- A collection of intervals I_k generated by midpoint iteration.

- A pair of interval maps $L : I_k \mapsto \frac{I_k}{\beta}$ and $R : I_k \mapsto \frac{I_k}{\beta} + \frac{1}{2}$.

- A pair of inclusion relations $B = \Theta\left(I_j \subseteq LI_k\right)$ and $C = \Theta\left(I_j \subseteq RI_k\right)$

- A spectrum determined entirely by $D = B + C$.

In a certain sense, the pushforward is determined entirely by D and nothing else matters. The monomial basis is some decorative tail-wagging by the monomials, demonstrating their prowess in introducing the binomial coefficient into combinatorial problems. But it illuminates nothing fundamental about the pushforward.

Notable properties:

- As square matrices, both B and C are singular.

- The matrix B has exactly one non-zero entry in each row.

- The matrices B and C never share a common non-zero entry. Equivalently, $D = B + C$ has entries that are only zero or one.

- The matrix D is unimodular, i.e. has determinant ± 1, and thus a spectral radius of 1.

- The spectrum of D is bounded below by $1/\beta$.

Questions:

- What other bases, besides the monomial basis, are there, that could be interesting to look at?

The coherent state is in the section after next.

6.9 Bounds estimates

The $j \to \infty$ limit is interesting. It presents the opportunity to describe piece-wise analytic eigenfunctions, although, based in the spectrum, these would all presumably live in the kernel.

This last statement should be qualified. The limiting behavior of of $\mathcal{L}$ is given by the diagonal sequence $\beta^{-j} D$ and since $\mathcal{L}$ is block-diagonal and solvable, it would seem that the eigenvalues are given by the spectrum of D times β^{-j}. If the spectrum of D is bounded; indeed, if the spectral radius is one, then the $j \to \infty$ limit implies that the high-j solutions approach the kernel. If, however, it could be arranged that the spectrum of D becomes unbounded, then the conclusion about the kernel no longer holds.

The off-diagonal blocks have coefficients

$$c_{jm} = \left(\frac{2}{\beta}\right)^m \frac{1}{2^j} \binom{j}{m}$$

The first few rows are

$$c_{0j} = \frac{1}{2^j}$$

$$c_{1j} = \frac{j}{\beta 2^{j-1}}$$

$$c_{2j} = \frac{j(j-1)}{\beta^2 2^{j+1}}$$

and subsequent rows introduce ever-greater falling factorials. In the $j \to \infty$ limit, these behave tamely.

For fixed m or even $m = o\left(\sqrt{j}\right)$, the binomial coefficient is estimated as[42]

$$\binom{j}{m} = \frac{j^m}{m!} \left(1 + o\left(\frac{m^2}{j}\right)\right)$$

and so

$$c_{jm} = \left(\frac{2}{\beta}\right)^m \frac{1}{2^j} \binom{j}{m} = \frac{1}{m!} \left(\frac{2}{\beta}\right)^m \frac{j^m}{2^j} \left(1 + o\left(\frac{m^2}{j}\right)\right)$$

and so the $j \to \infty$ limit is dominated by the $j^m 2^{-j}$ term, which grows for a bit before being stomped out. Similar behavior is seen near the diagonal, where it is handy to write $m = j - n$ for small enough n, so that $n = o\left(\sqrt{j}\right)$. Then

$$c_{j,j-n} = \left(\frac{2}{\beta}\right)^{j-n} \frac{1}{2^j} \binom{j}{n} = \frac{1}{n!} \left(\frac{\beta}{2}\right)^n \frac{j^n}{\beta^j} \left(1 + o\left(\frac{n^2}{j}\right)\right)$$

and so this is suppressed by a $j^n \beta^{-j}$ term.

The worst-case behavior is at the "middle coefficient", where $j = 2m$ and the binomial coefficients are maximal. The asymptotics are

$$\binom{2m}{m} = \frac{2^{2m}}{\sqrt{\pi m}}(1 + o(1))$$

Thus, along the half-diagonal, the matrix terms blow up:

$$c_{2m,m} = \left(\frac{2}{\beta}\right)^m \frac{1}{\sqrt{\pi m}}(1 + o(1))$$

since $1 < 2/\beta$. This really is the worst-case location. Close to this half-diagonal, taking $j \to \infty$ and writing $m = pj$ for $0 \ll p \ll 1$, the binomial coefficient converges to the normal distribution, so that

$$\log_2 \binom{j}{pj} = jH(p)(1 + o(1))$$

with $-H(p) = p\log_2 p + (1-p)\log_2(1-p)$ the binary entropy function. This gives

$$c_{j,pj} = \left(\frac{2}{\beta}\right)^{pj} \frac{2^{H(p)}}{\sqrt{\pi pj}}(1 + o(1))$$

as the estimate for the blowup at the near-half-diagonal.

6.10 Coherent States

Coherent functions can be used to construct the continuous spectrum solving $\mathscr{L}_\beta \psi = \lambda \psi$ for any complex $|\lambda| < 1$. These are exhibited in this section.

Given any function g, write the sawtooth $s(x) = g(x \mod \beta/2)$ so as to be explicitly periodic. Then define the coherent sawtooth as

$$\psi_{w,\alpha}(x) = \sum_{n=0}^{\infty} w^n s\left(\alpha T_\beta^n(x)\right)$$

By convention, $g \in \ker \mathscr{L}_\beta$ so as to exclude linear combinations of existing, known eigenfunctions constructed previously. However, this restriction does not seem to be explicitly needed. The α is a scaling constant; the intent is to use it as an aid to stitch pieces together. Note that for $\alpha = 1$, the modulo in the sawtooth is not needed. For $\alpha = \beta$, together with the mod, it is as if one more shift is introduced. There is also another variant to explore: $s(x) = g(x \mod \gamma)$ for some other value of γ.

We will need expressions for $\psi_{w,\alpha}(y/\beta)$ and for $\psi_{w,\alpha}(y/\beta + 1/2)$, so get them now. The shift T_β was defined in eqn 4 as

$$T_\beta(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases}$$

186

thus,

$$T_\beta\left(\frac{y}{\beta}\right) = y\Theta(m_0 - y) + (y - m_0)\Theta(y - m_0)$$

The second term is ignorable, as it will always be the case that $y < m_0$ and so this simplifies to

$$T_\beta\left(\frac{y}{\beta}\right) = y$$

Similarly

$$T_\beta\left(\frac{y}{\beta} + \frac{1}{2}\right) = y$$

Plugging through gives

$$\psi\left(\frac{y}{\beta}\right) = s\left(\frac{\alpha}{\beta}y\right) + w\sum_{n=0}^{\infty} w^n s\left(\alpha T_\beta^n(y)\right)$$

$$= s\left(\frac{\alpha}{\beta}y\right) + w\psi(y)$$

The above holds for all $0 \le y \le \beta/2$. Similarly

$$\psi\left(\frac{y}{\beta} + \frac{1}{2}\right) = s\left(\frac{\alpha}{\beta}y + \frac{\alpha}{2}\right) + w\psi(y)$$

The above holds for all $0 \le y \le 1/2$.

This is the point of coherent states: they scale when they shift.

6.10.1 Case n=1 (order $v = 2$)

For the $n = 1$ case, $p_1(\beta) = \beta^2 - \beta - 1$ has a root $\beta = r_1 = \varphi$. The midpoints are $m_0 = \beta/2$ and $m_1 = 1/2$ so $0 < m_1 < m_0 < 1$ and $I_2 = [0, 1/2]$ and $I_1 = [1/2, m_0]$. The interval inclusion map is given by

$$\Theta(I_j \subseteq LI_k) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \Theta(I_j \subseteq RI_k) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

The interval relationships that need to be solved are

$$\lambda\beta f_1(y) = f_2\left(\frac{y}{\beta}\right)$$

$$\lambda\beta f_2(y) = f_2\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

Assume that $f_k(y) = a_k \psi_{w,\alpha_k}(y)$ for unknown scale factors a_k and α_k. Simplify the notation, and write $\psi_k = \psi_{w,\alpha_k}$. Then the goal is to solve

$$\lambda \beta a_1 \psi_1(y) = a_2 \psi_2\left(\frac{y}{\beta}\right)$$

$$\lambda \beta a_2 \psi_2(y) = a_2 \psi_2\left(\frac{y}{\beta}\right) + a_1 \psi_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

Use the shift to pull out a sawtooth factor:

$$\lambda \beta a_1 \psi_1(y) = a_2 s\left(\frac{\alpha_2}{\beta} y\right) + w a_2 \psi_2(y)$$

$$\lambda \beta a_2 \psi_2(y) = a_2 s\left(\frac{\alpha_2}{\beta} y\right) + a_1 s\left(\frac{\alpha_1}{\beta} y + \frac{\alpha_1}{2}\right) + w a_2 \psi_2(y) + w a_1 \psi_1(y)$$

then plug in for ψ_1 (Gaussian elimination step):

$$\lambda \beta a_2 \psi_2(y) = a_2 s\left(\frac{\alpha_2}{\beta} y\right) + a_1 s\left(\frac{\alpha_1}{\beta} y + \frac{\alpha_1}{2}\right) + w a_2 \psi_2(y) + \frac{w}{\lambda \beta} a_2 s\left(\frac{\alpha_2}{\beta} y\right) + \frac{w^2}{\lambda \beta} a_2 \psi_2(y)$$

Collecting terms,

$$0 = a_2 s\left(\frac{\alpha_2}{\beta} y\right) + a_1 s\left(\frac{\alpha_1}{\beta} y + \frac{\alpha_1}{2}\right) + \frac{w}{\lambda \beta} a_2 s\left(\frac{\alpha_2}{\beta} y\right)$$

$$\lambda \beta = w + \frac{w^2}{\lambda \beta}$$

The second eqn has the solution $\lambda = w$, which is great, its exactly what is seen for the Bernoulli shift; its what we expected.

The first eqn reduces to

$$0 = a_2 \beta s\left(\frac{\alpha_2}{\beta} y\right) + a_1 s\left(\frac{\alpha_1}{\beta} y + \frac{\alpha_1}{2}\right)$$

The task is to find parameter combinations to solve this. This is not so easy; some variants fail to give valid overall solutions.

Failing variant This subsection documents a failed solution. The failure mode is subtle; we record it here as a lesson.

Recall the sawtooth was defined as $s(x) = g(x \mod \beta/2)$. The modulo will kill the additive term, if $(\alpha_1/2) \mod \beta/2 = 0$ which is achieved by $\alpha_1 = k\beta$ for any integer k. This leads the conclusion that $\alpha_2 = \alpha_1 = k\beta$ and that $a_1/a_2 = -\beta$, and this seems like it should be a solution, since the algebra worked out.

This is incorrect. Numeric checks reveal that, for $\alpha_2 = \alpha_1$

$$\lambda \beta a_1 \psi_1(y) \neq a_2 \psi_2\left(\frac{y}{\beta}\right)$$

It won't work. Let's look more closely. After reduction, one gets

$$\lambda \beta a_1 \psi_1(y) = a_2 s \left(\frac{\alpha_2}{\beta} y \right) + w a_2 \psi_2(y)$$

The question is: are there any possible values of α_1 and α_2 that could allow this to hold, as an equality, never mind the other assorted scaling factors? Well, try it. We would need to solve

$$\psi_1(y) - c \psi_2(y) = \sum_{n=0}^{\infty} w^n \left[s \left(\alpha_1 T_\beta^n(x) \right) - cs \left(\alpha_2 T_\beta^n(x) \right) \right]$$
$$= ds(ey)$$

for some constants $\alpha_1, \alpha_2, c, d, e$. This has the trivial solution $c = 1, \alpha_1 = \alpha_2, d = 0$ which implies $a_1 = a_2 = 0$ and so that won't do.

This is solvable by setting $\alpha_1 = 1$ and $\alpha_2 = \beta$, which allows a factor of s to be pulled out of the sum: $s \left(\beta T_\beta^n(x) \right) = s \left(T_\beta^{n+1}(x) \right)$ and then it does all go through. Rather than trying to rescue the above, a fresh start is made, below.

Working variant   Armed with the failure mode above, we look for solutions to the equation

$$\lambda \beta a_1 \psi_1(y) = a_2 \psi_2 \left(\frac{y}{\beta} \right)$$

with $\alpha_2 \neq \alpha_1$. One way this can be satisfied is by setting $\alpha_1 = 1$ and $\alpha_2 = \beta$. This has the effect of introducing one more shift into ψ_2, so that $\psi_2(y) = \psi_1 \left(T_\beta(y) \right)$. This allows the terms in the coherent series to be lined up. First, lets double-check this above claim.

$$\psi_2(y) = \sum_{n=0}^{\infty} w^n g \left(\left(\alpha_2 T_\beta^n(y) \right) \mod \frac{\beta}{2} \right)$$
$$= \sum_{n=0}^{\infty} w^n g \left(T_\beta \left(T_\beta^n(y) \right) \right)$$
$$= \sum_{n=0}^{\infty} w^n g \left(T_\beta^{n+1}(y) \right)$$
$$= \sum_{n=0}^{\infty} w^n g \left(\left(\alpha_1 T_\beta^{n+1}(y) \right) \mod \frac{\beta}{2} \right)$$
$$= \psi_1 \left(T_\beta(y) \right)$$

So that works out, as claimed. Lets proceed. Place this back into normal form:

$$\psi_1 \left(T_\beta(y) \right) = \sum_{n=0}^{\infty} w^n s \left(T_\beta^{n+1}(y) \right)$$
$$= \frac{1}{w} \left[-s(y) + \psi_1(y) \right]$$

189

The next equation to solve is

$$\lambda \beta a_2 \psi_2(y) = a_2 \psi_2\left(\frac{y}{\beta}\right) + a_1 \psi_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

Plugging through the various forms gives

$$\lambda \beta a_2 \psi_1(T_\beta(y)) = \lambda \beta a_1 \psi_1(y) + a_1 \psi_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

$$\frac{\lambda \beta a_2}{w}(-s(y) + \psi_1(y)) = \lambda \beta a_1 \psi_1(y) + a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) + w a_1 \psi_1(y)$$

Collecting terms requires

$$\frac{\lambda \beta a_2}{w} = \lambda \beta a_1 + w a_1$$

The above can be satisfied by setting $\lambda = w$ and $a_2 = w\beta a_1$. This leaves the sawtooth term. It is

$$-\frac{\lambda \beta a_2}{w} s(y) = a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

$$-w\beta^2 s(y) = s\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

The definition was $s(x) = g(x \mod \beta/2)$ and since $y \le \beta/2$ is guaranteed *a priori* we can drop the modulo term and hunt for a $g(x)$ that satisfies this. Such $g(x)$ can be found. When they are found, the resulting eigenfunction will have the form

$$f(y) = \begin{cases} f_2(y) = a_2 \psi_{w,\beta}(y) & \text{for } 0 \le y \le \frac{1}{2} \\ f_1(y) = a_1 \psi_{w,1}(y) & \text{for } \frac{1}{2} < y \le \frac{\beta}{2} \\ f_0(y) = 0 & \text{for } \frac{\beta}{2} < y \le 1 \end{cases}$$

with ψ as defined earlier, and with $\lambda = w$ and $a_2 = w\beta a_1$ in all cases.

Polynomial generators   Lets try the first few cases by hand. The first is $g(x) = 1$. This forces $w = -1/\beta^2$.

Continuing. In the linear domain, assume $g(x) = x + b$ so that

$$0 = w\beta^2(y+b) + \frac{y}{\beta} + \frac{1}{2} + b$$

The linear term forces $w = -1/\beta^3$ and $b = -1/2\left(1 + w\beta^2\right)$. For $\beta = \varphi$ this reduces to $b = -(1+\varphi)/2$

The quadratic form $g(x) = x^2 + bx + c$ gives

$$0 = w\beta^2 \left(y^2 + by + c\right) + \left(\frac{y}{\beta} + \frac{1}{2}\right)^2 + b\left(\frac{y}{\beta} + \frac{1}{2}\right) + c$$

$$0 = w\beta^2 + \frac{1}{\beta^2}$$

$$0 = b\left(w\beta^2 + \frac{1}{\beta}\right) + \frac{1}{\beta}$$

$$0 = c\left(w\beta^2 + 1\right) + \frac{1}{4} + \frac{b}{2}$$

For $\beta = \varphi$, these reduce to

$$w = -\frac{1}{\varphi^4} = 5 - 3\varphi$$

$$b = -(\varphi + 1)$$

$$c = \frac{3\varphi + 2}{4}$$

It appears that this pattern can be followed, thus obtaining an eigenfunction with eigenvalue $\lambda = w = -1/\beta^{k+2}$ for polynomial generators $g(x)$ of degree k.

Arbitrary generators   The general case requires solutions to

$$-w\beta^2 g(y) = g\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

that hold for all $0 \leq y \leq \beta/2$. But this is easy. The map takes $[0, 1/2]$ to $[1/2, \beta/2]$, which are non-overlapping. Thus, it is sufficient to set

$$g(y) = \begin{cases} h(y) & \text{for } 0 \leq y \leq \frac{1}{2} \\ -w\beta^2 h\left(\beta\left(y - \frac{1}{2}\right)\right) & \text{for } \frac{1}{2} < y \leq \frac{\beta}{2} \end{cases}$$

for any h whatsoever, and any complex w with $|w| < 1$. This will give a valid eigenfunction with eigenvalue $\lambda = w$. So, for $\beta = \varphi$ the golden mean, this provides the general solution.

6.10.2   Case n=2 (order $\nu = 3$)

As before, now that we know which direction to move in. The interval inclusion table gave the relations for a 3×3 block of interval relations. These were

$$\lambda \beta f_1(y) = f_2\left(\frac{y}{\beta}\right)$$

$$\lambda \beta f_2(y) = f_3\left(\frac{y}{\beta}\right)$$

$$\lambda \beta f_3(y) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

191

The above are not solvable with the old assumption that $f_k(y) = a_k \psi_{w,\alpha_k}(y)$ for unknown scale factors a_k and α_k. To repeat the earlier solution process would require $\alpha_1 = 1, \alpha_2 = \beta$ which is OK, but then $\alpha_3 = \beta$ leads to the "failed derivation" (because $\alpha_2 = \alpha_3$ gives the same failure mode) while the "natural" choice $\alpha_3 = \beta^2$ gives a term with a buried β that cannot be commuted out. The new trick, below, is to insert T instead of α.

Try this: $f_1(y) = \psi_{w,\alpha}(y) \equiv \psi_1(y)$ and then $\psi_2(y) = \psi_1(T_\beta(y))$ and $\psi_3(y) = \psi_2(T_\beta(y))$. The goal is to solve

$$\lambda \beta a_1 \psi_1(y) = a_2 \psi_2\left(\frac{y}{\beta}\right)$$

$$\lambda \beta a_2 \psi_2(y) = a_3 \psi_3\left(\frac{y}{\beta}\right)$$

$$\lambda \beta a_3 \psi_3(y) = a_3 \psi_3\left(\frac{y}{\beta}\right) + a_1 \psi_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

Using $T_\beta(y/\beta) = y$ causes dramatic simplification in the first two rows:

$$\lambda \beta a_1 = a_2$$

$$\lambda \beta a_2 = a_3$$

$$\lambda \beta a_3 \psi_3(y) = a_3 \psi_2(y) + a_1 \psi\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

For the last row, use

$$\psi_2(y) = \psi(T_\beta(y)) = \frac{1}{w}[-s(y) + \psi(y)]$$

$$\psi_3(y) = \psi\left(T_\beta^2(y)\right) = \frac{1}{w}\left[-s(T_\beta(y)) + \frac{1}{w}[-s(y) + \psi(y)]\right]$$

$$\psi\left(\frac{y}{\beta} + \frac{1}{2}\right) = s\left(\frac{y}{\beta} + \frac{1}{2}\right) + w\psi(y)$$

Plugging through gives

$$\lambda \beta a_3 \left[\frac{1}{w^2}\psi(y) - \frac{1}{w}s(T_\beta(y)) - \frac{1}{w^2}s(y)\right] = \frac{a_3}{w}[\psi(y) - s(y)] + a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) + a_1 w\psi(y)$$

Collecting the coefficients of ψ gives

$$\frac{\lambda \beta}{w^2} a_3 = \frac{1}{w}a_3 + a_1 w$$

$$\frac{\lambda \beta}{w^3} a_3 = \frac{1}{w^2}a_3 + a_1$$

Trying to slot in the identity $\beta^3 = \beta^2 + 1$ into the above gives a solution $a_3/w^2 = \beta^2 a_1$ and $\lambda = w$. The sawtooth term is

$$\lambda \beta a_3 \left[\frac{1}{w} s\left(T_\beta \left(y \right) \right) + \frac{1}{w^2} s\left(y \right) \right] = \frac{a_3}{w} s\left(y \right) - a_1 s \left(\frac{y}{\beta} + \frac{1}{2} \right)$$

$$\lambda \beta^3 \left[ws\left(T_\beta \left(y \right) \right) + s\left(y \right) \right] = \beta^2 ws\left(y \right) - s \left(\frac{y}{\beta} + \frac{1}{2} \right)$$

The right-most term has the implicit assumption that $y \leq 1/2$ and in this range, $T_\beta \left(y \right) = \beta y$ and so

$$w^2 \beta^3 s\left(\beta y \right) + ws\left(y \right) + s \left(\frac{y}{\beta} + \frac{1}{2} \right) = 0$$

where $\beta^2 \left(\beta - 1 \right) = 1$ is used to simplify.

To summarize, the midpoints iterated out to $0 < m_1 < m_2 < m_0 < 1$. The eigenfunc is then

$$f\left(y \right) = \begin{cases} f_3\left(y \right) = w^2 \beta^2 \psi_w \left(T_\beta^2 \left(y \right) \right) & \text{for } 0 \leq y \leq m_1 \\ f_2\left(y \right) = w\beta \, \psi_w \left(T_\beta \left(y \right) \right) & \text{for } m_1 < y \leq \frac{1}{2} \\ f_1\left(y \right) = \psi_w \left(y \right) & \text{for } \frac{1}{2} < y \leq m_0 \\ f_0\left(y \right) = 0 & \text{for } m_0 < y \leq 1 \end{cases}$$

This has not yet been double-checked numerically.

Polynomial generators   Look for solutions with $s\left(x \right) = g\left(x \mod \beta/2 \right)$.
For $g\left(x \right) = 1$ this gives

$$w^2 \beta^3 + w + 1 = 0$$

Yikes! This has no real solutions; it does have a pair of complex solutions $w = \left(-1 \pm \sqrt{1 - 4\beta^3} \right) / 2\beta^3$.

For the general analytic case, $g\left(x \right) = x^k + \sum_{j=0}^{k-1} c_j x^j$ and the leading term requires

$$w^2 \beta^{k+3} + \beta^2 w \left(\beta - 1 \right) + \beta^{-k} = 0$$
$$w^2 \beta^{2k+3} + w\beta^{k+2} \left(\beta - 1 \right) + 1 = 0$$
$$w^2 \beta^{2k+3} + w\beta^k + 1 = 0$$

which is still quadratic in w and so again has a pair of complex solutions. This has not yet been double-checked numerically.

Fractal generators   Analogous to the $n = 1$ case, we look for solutions to

$$w^2 \beta^3 s\left(\beta y \right) + ws\left(y \right) + s \left(\frac{y}{\beta} + \frac{1}{2} \right) = 0$$

Since $s\left(x \right) = g\left(x \mod \beta/2 \right)$, the search can be limited to $0 \leq y \leq \beta/2$. Just as before, the third term maps $[0, 1/2]$ to $[1/2, \beta/2]$ and thus can be used to cancel out any

behavior. Thus, it is sufficient to set

$$g(y) = \begin{cases} h(y) & \text{for } 0 \le y \le \frac{1}{2} \\ -wh\left(\beta\left(y - \frac{1}{2}\right)\right) - w^2\beta^3 h\left(\beta^2\left(y - \frac{1}{2}\right)\right) & \text{for } \frac{1}{2} < y \le \frac{\beta}{2} \end{cases}$$

for any h whatsoever, and any complex w with $|w| < 1$. This will give a valid eigenfunction with eigenvalue $\lambda = w$. Again, this provides the general solution.

6.10.3 Case n=3 (order $\nu = 3$)

This case exhibits several novel features, so lets look at it directly. Write the twizzled form as

$$\beta\lambda \begin{bmatrix} f_1(y) \\ f_2(y) \\ f_3(y) \end{bmatrix} = B \begin{bmatrix} f_1\left(\frac{y}{\beta}\right) \\ f_2\left(\frac{y}{\beta}\right) \\ f_3\left(\frac{y}{\beta}\right) \end{bmatrix} + C \begin{bmatrix} f_1\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ f_2\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ f_3\left(\frac{y}{\beta} + \frac{1}{2}\right) \end{bmatrix}$$

with

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad D = B + C = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{for } n = 3$$

In non-matrix form, the relations are

$$\lambda\beta f_1(y) = f_3\left(\frac{y}{\beta}\right)$$

$$\lambda\beta f_2(y) = f_3\left(\frac{y}{\beta}\right) + f_1\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

$$\lambda\beta f_3(y) = f_3\left(\frac{y}{\beta}\right) + f_2\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

Write $\psi_k(y) = \psi\left(T_\beta^{k-1}(y)\right)$. The direct assignment of f_k to ψ_k doesn't provide a solution of the form seen in the previous $n = 1$ and $n = 2$ cases. (Try it! You'll get a form that is satisfied only if everything is zero.) But it seems that a permutation should work. Assume that there exists a permutation matrix R such that the association of f_k to ψ_k is passes through that permutation. Assume the form

$$\begin{bmatrix} f_1(y) \\ f_2(y) \\ f_3(y) \end{bmatrix} = R \begin{bmatrix} a_1\psi_1(y) \\ a_2\psi_2(y) \\ a_3\psi_3(y) \end{bmatrix}$$

where R^T is the transpose, $RR^T = I$ and so

$$\beta\lambda \begin{bmatrix} a_1\psi_1(y) \\ a_2\psi_2(y) \\ a_3\psi_3(y) \end{bmatrix} = R^T BR \begin{bmatrix} a_1\psi_1\left(\frac{y}{\beta}\right) \\ a_2\psi_2\left(\frac{y}{\beta}\right) \\ a_3\psi_3\left(\frac{y}{\beta}\right) \end{bmatrix} + R^T CR \begin{bmatrix} a_1\psi_1\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ a_2\psi_2\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ a_3\psi_3\left(\frac{y}{\beta} + \frac{1}{2}\right) \end{bmatrix}$$

Then apply the shift, for $k > 0$,

$$\psi\left(T_\beta^k\left(\frac{y}{\beta}\right)\right) = \psi\left(T_\beta^k\left(\frac{y}{\beta} + \frac{1}{2}\right)\right) = \psi\left(T_\beta^{k-1}(y)\right)$$

while for $k = 0$

$$\psi\left(\frac{y}{\beta}\right) = s\left(\frac{y}{\beta}\right) + w\psi(y)$$

$$\psi\left(\frac{y}{\beta} + \frac{1}{2}\right) = s\left(\frac{y}{\beta} + \frac{1}{2}\right) + w\psi(y)$$

to get

$$\beta\lambda \begin{bmatrix} a_1\psi_1(y) \\ a_2\psi_2(y) \\ a_3\psi_3(y) \end{bmatrix} = R^T BR \begin{bmatrix} a_1\psi_1\left(\frac{y}{\beta}\right) \\ a_2\psi_1(y) \\ a_3\psi_2(y) \end{bmatrix} + R^T CR \begin{bmatrix} a_1\psi_1\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ a_2\psi_1(y) \\ a_3\psi_2(y) \end{bmatrix}$$

What is left can be reduced by plugging in the general form:

$$\psi\left(T_\beta^k(y)\right) = \sum_{n=0}^{\infty} w^n s\left(T_\beta^{n+k}(x)\right)$$

$$= \frac{1}{w^k}\left[\psi(y) - \sum_{n=0}^{k-1} w^n s\left(T_\beta^n(x)\right)\right]$$

The result requires collecting similar terms. The ψ terms give

$$\beta\lambda \begin{bmatrix} a_1 \\ a_2 w^{-1} \\ a_3 w^{-2} \end{bmatrix} = wR^T DR \begin{bmatrix} a_1 \\ a_2 w^{-1} \\ a_3 w^{-2} \end{bmatrix}$$

where $D = B + C$. This is an eigenvalue equation, and since $p_n(\beta) = \det[D - \beta I] = 0$ conclude that $\lambda = w$. Lets take a moment to solve this. It has the form

$$0 = [D - \beta I]R \begin{bmatrix} a_1 \\ a_2 w^{-1} \\ a_3 w^{-2} \end{bmatrix} \qquad \text{or} \qquad 0 = \begin{bmatrix} -\beta & 0 & 1 \\ 1 & -\beta & 1 \\ 0 & 1 & 1-\beta \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

or

$$e_3 = \beta e_1 \qquad \text{and} \qquad e_2 = \beta(\beta - 1)e_1$$

The generating terms (the s terms) give

$$-\beta \begin{bmatrix} 0 \\ a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = -R^T DR \begin{bmatrix} 0 \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix} + R^T BR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ 0 \end{bmatrix} + R^T CR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ 0 \\ 0 \end{bmatrix}$$

195

where $S_1(y) = s(y) + ws(T_\beta(y))$.

The desired form requires a permutation matrix R that keeps the $s\left(\frac{y}{\beta} + \frac{1}{2}\right)$ term alive, so that it can be used, just as in the previous $n = 1$ and $n = 2$ cases, to get an equation of the form

$$s\left(\frac{y}{\beta} + \frac{1}{2}\right) = \sum_k c_k s\left(\beta^k y\right)$$

Since $s(x) = g(x \mod \beta/2)$, the search can be limited to $0 \le y \le \beta/2$. As before, the left-hand side maps $[0, 1/2]$ to $[1/2, \beta/2]$ and thus can be used to cancel out any behavior on the right-hand side. The general form of the generator is then

$$g(y) = \begin{cases} h(y) & \text{for } 0 \le y \le \frac{1}{2} \\ -\sum_k c_k h\left(\beta^{k+1}\left(y - \frac{1}{2}\right)\right) & \text{for } \frac{1}{2} < y \le \frac{\beta}{2} \end{cases}$$

for any h whatsoever, and any complex w with $|w| < 1$, to obtain the general coherent state eigenfunction. This provides a complete solution, if an appropriate rotation matrix R can be found.

Failed guess As we are lazy, we can get to guessing. One obvious guess is that $R = P^T$ with P the shift permutation matrix discovered earlier. It turns out this won't work. This subsection documents this failed attempt. Perhaps reasons for the failure become clear, and thus the path to the fix. So, for $n = 3$, the permutation is gotten from the earlier table, as

$$P = \begin{bmatrix} \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \end{bmatrix} \quad \text{and} \quad P^T = \begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \end{bmatrix}$$

Then

$$PBP^T = \begin{bmatrix} \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

and

$$PCP^T = \begin{bmatrix} \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & 1 \\ 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

The initial shift form does not work out:

$$-\beta \begin{bmatrix} 0 \\ a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = -\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ 0 \\ 0 \end{bmatrix}$$

196

Very disappointing that the $s\left(\frac{y}{\beta}+\frac{1}{2}\right)$ term is killed; this was vital for the construction of the coherent states. The loss of this term is fatal for the general theory, and so it seems the guess that $R = P^T$ is incorrect. Lets look at the train-wreck. Multiplying out gives

$$\beta \begin{bmatrix} 0 \\ a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = \begin{bmatrix} a_3 w^{-1} s(y) \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} a_1 s\left(\frac{y}{\beta}\right) \\ a_1 s\left(\frac{y}{\beta}\right) \\ a_1 s\left(\frac{y}{\beta}\right) \end{bmatrix}$$

This looks dubious; perhaps there is a way to solve this, but certainly the simple form discovered earlier doesn't work.

Required form The goal is to find a rotation matrix R such that

$$\beta R \begin{bmatrix} 0 \\ a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = DR \begin{bmatrix} 0 \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix} - BR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ 0 \end{bmatrix} - CR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ 0 \\ 0 \end{bmatrix}$$

falls into the form for which one row has the structure

$$\sum_k c_k s\left(\beta^k y\right) + s\left(\frac{y}{\beta}+\frac{1}{2}\right) = 0$$

and the other rows do not prevent a solution (typically, by forcing $a_1 = 0$). Perhaps the collapsed form is easier to think about?

$$\beta R \begin{bmatrix} 0 \\ a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = BR \begin{bmatrix} -a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix} + CR \begin{bmatrix} -a_1 s\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix}$$

Try a brute-force approach.

$$CR = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix}$$

$$BR = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} r_{31} & r_{32} & r_{33} \\ r_{31} & r_{32} & r_{33} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

Curious. CR shifts down by one row, discarding it. But BR takes that discarded row, and reproduces it. The top row is

$$\beta r_{12} a_2 s(y) + \beta r_{13} a_3 w^{-1} S_1(y) = -r_{31} a_1 s\left(\frac{y}{\beta}\right) + r_{33} a_3 w^{-1} s(y)$$

But $S_1(y) = s(y) + ws(T_\beta(y))$ and when $y < 1/2$ the shift reduces to $s(T_\beta(y)) = s(\beta y)$
Plugging through gives the mess

$$\beta w r_{12} a_2 s(y) + \beta r_{13} a_3 (s(y) + ws(\beta y)) = -r_{31} a_1 ws\left(\frac{y}{\beta}\right) + r_{33} a_3 s(y)$$

$$(\beta w r_{12} a_2 + \beta r_{13} a_3 - r_{33} a_3) s(y) = -\beta r_{13} a_3 ws(\beta y) - r_{31} a_1 ws\left(\frac{y}{\beta}\right)$$

This is not solvable for general $s(y)$ unless $r_{13} = r_{31} = 0$ and $0 = \beta w r_{12} a_2 - r_{33} a_3$.
That leaves

$$R = \begin{bmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & r_{23} \\ 0 & r_{32} & r_{33} \end{bmatrix}$$

Next, verifying that $R^T R = I$ imposes additional constraints. Explicitly verifying,

$$R^T R = I = \begin{bmatrix} r_{11} & r_{21} & 0 \\ r_{12} & r_{22} & r_{32} \\ 0 & r_{23} & r_{33} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & r_{23} \\ 0 & r_{32} & r_{33} \end{bmatrix}$$

$$= \begin{bmatrix} r_{11}^2 + r_{21}^2 & r_{11}r_{12} + r_{21}r_{22} & r_{23}r_{21} \\ r_{11}r_{12} + r_{21}r_{22} & r_{12}^2 + r_{22}^2 + r_{32}^2 & r_{22}r_{23} + r_{32}r_{33} \\ r_{21}r_{23} & r_{22}r_{23} + r_{32}r_{33} & r_{23}^2 + r_{33}^2 \end{bmatrix}$$

Satisfying this requires either $r_{21} = 0$ or $r_{23} = 0$. But the former choice forces $r_{12} = 0$ which breaks the earlier relation. So take $r_{23} = 0$ to get

$$R = \begin{bmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

That leaves behind

$$BR = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad CR = \begin{bmatrix} 0 & 0 & 0 \\ r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \end{bmatrix}$$

Finally,

$$\beta \begin{bmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \end{bmatrix} \begin{bmatrix} -a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ 0 \\ a_3 w^{-1} s(y) \end{bmatrix}$$

The top row was already taken care of. Avoid the busywork by killing it, and then killing everything that multiplies to zero. This leaves

$$\beta \begin{bmatrix} r_{22} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} a_3 w^{-1} s(y) \end{bmatrix} - \begin{bmatrix} r_{11} \\ r_{21} \end{bmatrix} \begin{bmatrix} a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) \end{bmatrix}$$

198

Multiplying out,

$$\beta \begin{bmatrix} r_{22}a_2 s(y) \\ a_3 w^{-1} S_1(y) \end{bmatrix} = \begin{bmatrix} a_3 w^{-1} s(y) \\ a_3 w^{-1} s(y) \end{bmatrix} - \begin{bmatrix} r_{11}a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) \\ r_{21}a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right) \end{bmatrix}$$

It would appear that there is a common solution only if the upper and lower row are multiples of one-another. This requires

$$r_{21}\left(w\beta r_{22}a_2 s(y) - a_3 s(y)\right) = r_{11}a_3\left(\beta S_1(y) - s(y)\right)$$

The earlier constraint gave $a_3 = \beta w r_{12} a_2$, so

$$r_{21}\left(r_{22} - r_{12}\right) s(y) = r_{11}r_{12}\left(\beta S_1(y) - s(y)\right)$$

The sticky wicket is that $S_1(y) = s(y) + ws\left(T_\beta(y)\right)$ which forces

$$r_{21}\left(r_{22} - r_{12}\right) s(y) = r_{11}r_{12}\left(s(y)(\beta - 1) + ws\left(T_\beta(y)\right)\right)$$

or

$$\left[r_{21}\left(r_{22} - r_{12}\right) - r_{11}r_{12}(\beta - 1)\right] s(y) = r_{11}r_{12}ws\left(T_\beta(y)\right)$$

which is unsolvable for general s unless $r_{12} = r_{21} = 0$ and $r_{11} = r_{22} = 1$. However, this snowballs in all the wrong ways, forcing $a_3 = 0$. The eigenvalue equation then implies that $a_1 = a_2 = 0$ so we conclude that this is unsolvable, unless we accept a self-similar s that solves

$$Ks(y) = s(\beta y)$$

for some constant K. This has no general form, but setting $s(y) = y^p$ forces $K = \beta^p$ and so

$$r_{11}r_{12}w\beta^p = r_{21}\left(r_{22} - r_{12}\right) - r_{11}r_{12}(\beta - 1)$$
$$r_{11}r_{12}\left(w\beta^p + \beta - 1\right) = r_{21}\left(r_{22} - r_{12}\right)$$

Is this solvable? Well, $r_{11} = r_{22} = \cos\theta$ and $r_{12} = -r_{21} = \sin\theta$ so

$$w\beta^p + \beta = \frac{r_{12}}{r_{11}} = \tan\theta$$

which is solvable for any p and w. Fixing w, this gives a one-parameter set of solutions, with p being any complex number. We'll have to take what we can get. A one-parameter family is still plenty enough from which to construct a (complete?) orthonormal basis for any fixed w.

Final form So lets work our way back home, now. The general form was

$$\beta r_{22}a_2 s(y) = a_3 w^{-1} s(y) - r_{11}a_1 s\left(\frac{y}{\beta} + \frac{1}{2}\right)$$

199

and $a_3 = \beta w r_{12} a_2$ so

$$\beta a_2 (r_{12} - r_{22}) s(y) = r_{11} a_1 s \left(\frac{y}{\beta} + \frac{1}{2} \right)$$

This is solvable, as the non-overlapping domains allow

$$g(y) = \begin{cases} h(y) & \text{for } 0 \le y \le \frac{1}{2} \\ \frac{\beta a_2 (r_{12} - r_{22})}{r_{11} a_1} h\left(\beta \left(y - \frac{1}{2} \right) \right) & \text{for } \frac{1}{2} < y \le \frac{\beta}{2} \end{cases}$$

The constraint equation forced $h(y) = y^p$ for any complex p and so the above becomes

$$g(y) = \begin{cases} y^p & \text{for } 0 \le y \le \frac{1}{2} \\ \frac{a_2}{a_1} \beta (w\beta^p + \beta - 1) h\left(\beta \left(y - \frac{1}{2} \right) \right) & \text{for } \frac{1}{2} < y \le \frac{\beta}{2} \end{cases}$$

The ratio a_2/a_1 was fixed by the eigenvalue equation

$$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w^2 \\ wa_2/a_1 \\ a_3/a_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \beta(\beta - 1) \\ \beta \end{bmatrix}$$

Taking the upper row,

$$\begin{aligned} \frac{a_2}{a_1} &= \frac{1 - w^2 \cos\theta}{w \sin\theta} \\ &= \frac{1}{w \tan\theta} \left(\frac{1}{\cos\theta} - w^2 \right) \\ &= \frac{1}{w \tan\theta} \left(\sqrt{1 + \tan^2\theta} - w^2 \right) \\ &= \frac{1}{w(w\beta^p + \beta)} \left(\sqrt{1 + (w\beta^p + \beta)^2} - w^2 \right) \end{aligned}$$

Blech–yuck. No simplification. Oh well.

In conclusion, there is a one-parameter family of solutions to $\mathcal{L}_\beta f = wf$, for any complex $|w| < 1$. These all have the form

$$f(y) = \begin{cases} f_3(y) & \text{for } 0 \le y \le \frac{1}{2} \\ f_2(y) & \text{for } \frac{1}{2} < y \le m_1 \\ f_1(y) & \text{for } m_1 < y \le m_0 \end{cases}$$

where

$$\begin{bmatrix} f_1(y) \\ f_2(y) \\ f_3(y) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \psi_1(y) \\ a_2 \psi_2(y) \\ a_3 \psi_3(y) \end{bmatrix}$$

and $\psi_k(y) = \psi\left(T_\beta^{k-1}(y) \right)$ and

$$\psi(y) = \sum_{n=0}^{\infty} w^n s\left(T_\beta^n(y) \right)$$

200

and $s(x) = g(x \mod \beta/2)$ with $g(x)$ as given above. This is a one-parameter family of solutions, valid for any complex p. By imposing orthonormality constraints, this should be enough to obtain a countable basis.

To conclude: a solution is found, but the road to obtain it was twisty and narrow.

6.10.4   General case

The general case, for any order ν, can be sketched as below. The twizzled form is

$$
\beta\lambda \begin{bmatrix} f_1(y) \\ \vdots \\ f_\nu(y) \end{bmatrix} = B \begin{bmatrix} f_1\left(\frac{y}{\beta}\right) \\ \vdots \\ f_\nu\left(\frac{y}{\beta}\right) \end{bmatrix} + C \begin{bmatrix} f_1\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ \vdots \\ f_\nu\left(\frac{y}{\beta}+\frac{1}{2}\right) \end{bmatrix}
$$

Following the structure of the $n = 3$ case, assume that there is a rotation matrix R such that

$$
\begin{bmatrix} f_1(y) \\ \vdots \\ f_\nu(y) \end{bmatrix} = R \begin{bmatrix} \psi_1(y) \\ \vdots \\ \psi_\nu(y) \end{bmatrix}
$$

In index notation, this is $f_k(y) = R_{kj}a_j\psi_j(y)$ with $\psi_k(y) = \psi\left(T_\beta^{k-1}(y)\right)$. Several identities will be needed. One is

$$
\psi\left(T_\beta^k(y)\right) = \sum_{n=0}^{\infty} w^n s\left(T_\beta^{n+k}(y)\right)
$$

$$
= \frac{1}{w^k}\left[\psi(y) - \sum_{n=0}^{k-1} w^n s\left(T_\beta^n(y)\right)\right]
$$

Also, for $k > 0$,

$$
\psi\left(T_\beta^k\left(\frac{y}{\beta}\right)\right) = \psi\left(T_\beta^k\left(\frac{y}{\beta}+\frac{1}{2}\right)\right) = \psi\left(T_\beta^{k-1}(y)\right)
$$

while for $k = 0$

$$
\psi\left(\frac{y}{\beta}\right) = s\left(\frac{y}{\beta}\right) + w\psi(y)
$$

$$
\psi\left(\frac{y}{\beta}+\frac{1}{2}\right) = s\left(\frac{y}{\beta}+\frac{1}{2}\right) + w\psi(y)
$$

Plugging through, this gives

$$
\beta\lambda R \begin{bmatrix} a_1\psi_1(y) \\ \vdots \\ a_\nu\psi_\nu(y) \end{bmatrix} = BR \begin{bmatrix} a_1\psi_1\left(\frac{y}{\beta}\right) \\ a_2\psi_1(y) \\ \vdots \\ a_\nu\psi_{\nu-1}(y) \end{bmatrix} + CR \begin{bmatrix} a_1\psi_1\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ a_2\psi_1(y) \\ \vdots \\ a_\nu\psi_{\nu-1}(y) \end{bmatrix}
$$

Setting $D = B + C$ (as before) and continuing,

$$\beta\lambda R \begin{bmatrix} a_1\psi_1(y) \\ \vdots \\ a_\nu\psi_\nu(y) \end{bmatrix} = DR \begin{bmatrix} a_1 w\psi_1(y) \\ a_2\psi_1(y) \\ \vdots \\ a_\nu\psi_{\nu-1}(y) \end{bmatrix} + BR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ \vdots \\ 0 \end{bmatrix} + CR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Shorten the notation for the recurrence relation:

$$\psi_{k+1}(y) = \psi\left(T_\beta^k(y)\right) = \frac{1}{w^k}\left[\psi(y) - \sum_{n=0}^{k-1} w^n s\left(T_\beta^n(x)\right)\right]$$

$$= \frac{1}{w^k}\left[\psi(y) - S_{k-1}(y)\right]$$

and then plug through and collect distinct terms. The ψ terms give

$$\beta\lambda R \begin{bmatrix} a_1 \\ a_2 w^{-1} \\ \vdots \\ a_\nu w^{1-\nu} \end{bmatrix} = wDR \begin{bmatrix} a_1 \\ a_2 w^{-1} \\ \vdots \\ a_\nu w^{1-\nu} \end{bmatrix}$$

which has the form of an eigenvalue equation. The solution is given by

$$\det\left[wD - \beta\lambda I\right] = 0$$

and since $\det[D - \beta I] = 0$ conclude that $\lambda = w$. Note that this is independent of the matrix R.

The generating terms, those having only s in them, give

$$\beta R \begin{bmatrix} 0 \\ a_2 S_0(y) \\ \vdots \\ a_\nu w^{2-\nu} S_{\nu-2}(y) \end{bmatrix} = DR \begin{bmatrix} 0 \\ 0 \\ a_3 w^{-1} S_0(y) \\ \vdots \\ a_\nu w^{2-\nu} S_{\nu-3}(y) \end{bmatrix} - BR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}\right) \\ 0 \\ \vdots \\ 0 \end{bmatrix} - CR \begin{bmatrix} a_1 s\left(\frac{y}{\beta}+\frac{1}{2}\right) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The goal is to find a rotation matrix that allows the above to be solved. The rotation matrix has $\nu(\nu-1)/2$ free parameters, whereas the above provides only ν constraints. This would appear to allow for a fair amount of freedom.

An additional constraint is that the desired form requires a rotation R that keeps the $s\left(\frac{y}{\beta}+\frac{1}{2}\right)$ term alive, so that it can be used, just as in the previous $n = 1$ and $n = 2$ cases, to get an equation of the form

$$s\left(\frac{y}{\beta}+\frac{1}{2}\right) = \sum_k c_k s\left(\beta^k y\right)$$

for some constants c_k. Since $s(x) = g(x \mod \beta/2)$, the search is limited to $0 \le y \le \beta/2$. The left-hand side maps $[0, 1/2]$ to $[1/2, \beta/2]$ and thus can be used to cancel out any behavior at all, including the right-hand side. Thus, it is sufficient to set

$$g(y) = \begin{cases} h(y) & \text{for } 0 \le y \le \frac{1}{2} \\ -\sum_k c_k h\left(\beta^{k+1}\left(y - \frac{1}{2}\right)\right) & \text{for } \frac{1}{2} < y \le \frac{\beta}{2} \end{cases}$$

This works for any h whatsoever, unless h is constrained in some way, as it was for the $n = 3$ solution. It is not clear if the general form will be similarly constrained or not. Still, a one-parameter family of solutions were obtained (for fixed w), which is plenty enough to develop a countable orthonormal basis (for fixed w).

Presuming such a rotation matrix R can be found, this gives a family of solutions to $\mathcal{L}_\beta f = wf$ for any complex w with $|w| < 1$.

Commentary This wraps up the general case. The $n = 1$ and $n = 2$ cases were relatively straight-forward; the $n = 3$ case caused much grief, but was eventually solvable. It seems clear that all three cases will provide a complete set of orthonormal solutions to $\mathcal{L}_\beta f = wf$ for any complex w with $|w| < 1$. Explicit formulas for this set remain out of reach with the given toolset. Clearly, the $n = 1$ and $n = 2$ cases could be brute-forced, if desired.

It is not hard to hypothesize that the required form of R always exists. The initial hypothesis that R was the identity, and that R was a permutation matrix, were both broken by the $n = 3$ case. In general, R could be taken to be unitary, instead of orthogonal, but nothing in the $n = 3$ case even whispered that this extra freedom was needed. Complex conjugates never appear.

The coherent states provide a general solution. Presumably, this aligns with the piece-wise polynomial solutions, given earlier. This requires a review of what happens at $w = \beta^{-k}$.

The coherent state construction is performed for the countable set of finite orbits, enumerated by $p_n(\beta) = 0$. These give a collection of β that are dense in the interval $1 \le \beta \le 2$. Can the construction be extended to all β? Doing so would require some way of obtaining the interval inclusion map that the B and C matrices encode. This inclusion map is implicit in the sofic shift, but none of the above constructions work at an abstraction level that was appropriate for exposing it. There's some Borel algebra here, but the above calculations were not sufficient to make it explicit.

There are unexplored tools. For example, most real numbers x have an uncountable number of distinct β-expansions for most values of β. This is the opposite of the so-called "univoke" expansions. Nowhere in the above constructions was this fact made use of, or touched upon. Yet it is implicitly there: the coherent state is iterating with $T_\beta^k(x)$ but it was never questioned or debated that this may result in different iterates, and thus different coherent states. This is bad: it was implicitly assumed that the coherent state is unique and well-defined, whereas perhaps each coherent state should be taken to be an uncountable family of such states. Whether they all come together in the end, or remain distinct, is entirely unclear.

Clearly, we've pried the lid off of a jar, and peered inside. It seems to hold a bottomless collection of tasks, each engendering new and ever-more-complex hypothesis.

Of course, this is the way of mathematics, and one must be careful how one spends one's time on Earth. Still, it is all a bit unnerving.

6.11   Algorithms

The meta problem is one of term rewriting: the recursive equations are all solvable, but their general form is opaque and unenlightened. We want an algorithmic description of the solution to the recurrence relations that is somehow transparent. There seem to be multiple interacting algorithms:

- The use of multiplication and modulo to compute midpoints. When midpoints are expressed as base-two bit-sequences, the algorithm for multiplication is quite complex, but since we're indoctrinated into thinking that multiplication is simple, this is overlooked.

- After the midpoints are computed, they are placed into sorted order. This requires a sort algorithm.

- Interval inclusion relations are required. These can be algorithmically determined. The algorithm itself is not complicated, but the generated inclusion sequences are opaque and cryptic.

- The interval inclusion relations induce a linear matrix form relating the f_k on the different intervals. This matrix is solvable.

- Hypothesizing the piece-wise polynomial terms f_k on each interval I_k, the transfer operator $\mathscr{L}$ takes the form of recurrence relations that can be expanded and collected up by powers. The expansion deploys the binomial coefficients inside the polynomials. The expansion via binomial coefficient is purely algorithmic.

- Collecting up the powers once again gives a solvable matrix. Since it is solvable, it can be solved by Gaussian elimination (diagonalization). Do, purely algorithmically.

- The final matrix is just the matrix representation of $\mathscr{L}_\beta$ in the monomial basis, for this fixed value of β.

How do things start, and where do we want to go?

- There are two starting points: the integer n defining the "valid" polynomial p_n, or some arbitrary β from some external source.

- If the start is a polynomial, then a root-finding algo must run, so as to generate a bit-sequence for β to some indeterminate number of bits.

- The number of needed bits is only enough to perform the mid-point sort. If the β sources is "internal", as the root of a polynomial, then only a handful of bits are needed. The number of bits are determined by the index n of the polynomial p_n. For external sources of β, there might be a countable infinity of midpoints, and so an unbounded number of binary digits may be needed.

- Explicit sorting can be avoided by using the visitation tree, which implicitly encodes a left-to-right order.

- For finite orbits, the visitation tree is finite; it splits the unit interval into finitely many pieces.

- The interval inclusion map is "symbolic" or "relational"; can be treated with an indicator function showing inclusion. Since all inclusions follow from iteration of the initial interval, it can be expressed as a map (injection) from binary tree to binary tree. This is related to the visitation tree, below.

- Polynomials assigned to intervals can be treated with symbolic methods. This includes expansion via binomial coefficient.

- The "only reason" for expansion via binomials is to collect terms, and obtain a matrix representation for the shift in the monomial basis. Some other basis could be chosen; the monomial basis is the simplest, in terms of combinatorics.

What, exactly, is the source of unease? Lets review the issues:

- The precise form for the matrix representation of $\mathscr{L}_\beta$ in the monomial basis is tangled and opaque; the closed form is available, but requires algorithms to present it in a classical analytic form, where all matrix entries are rational functions of β.

- More precisely, we do not have a purely symbolic recursive relation for the matrix entries for matrices B,C. We just have a prescription of "compute midpoints and sort them and compute inclusions" which is very far from a recursive algo. We want that recursive algo.

- Even if a nice computer GUI was developed to show each of the individual entries of the matrix as rational functions of β, the overall symmetry and structure would remain hidden and opaque.

- For any fixed β, what interesting things about $\mathscr{L}_\beta$ might one be able to say? We can explicitly give its dimension (it is a product of the orbit length times the degree of the piece-wise polynomial.) We can point out how it is shift-like. We can point out that it's spectrum is discrete, and is explicitly $\lambda = \beta^{-n}$ for degree n piece-wise pieces.

- What else would we like to be able to say, if explicit expressions were available?

- The cascade of algorithmic operations above, how should we think about them? We can hand-wave and say "these are the algorithms of the beta transform", but that is vague and overly general. We can talk about homeomorphic algorithms, and ask about their homotopic structure. For example, are there other algorithms that are crypto-equivalent, that produce the beta transform, but are not obtainable by homeomorphic deformation of the above "obvious" form?

7 Visitation Tree

The above commentary about stability and instability of functions under iteration suggests that a slightly more comprehensive viewpoint should be adopted. There are several distinct problems. First, how does the iterated transfer operator $\mathscr{L}^n$ act on functions on the real interval, and secondly, how does it act on real intervals approximated by finite collections of points? And, for such finite collections, are the points evenly spaced? Or should we generate them by midpoint iteration? Or in some other way? Lets muddle about.

This is retracing some old steps; it starts as a stream-of-consciousness meander through a concept salad, and concludes with a specific mapping of midpoints to a tree, the visitation tree. It appears that the visitation tree is a potentially useful tool.

7.1 Space of functions

Define the space of functions on the unit interval $\mathscr{F} = \{\rho : [0,1] \to \mathbb{R}\}$. This is too large and vague a set; in practice, we'll want to work with piece-wise smooth functions (chopped up polynomials) or maybe piece-wise real-analytic, or perhaps square integrable, or perhaps continuous, etc. Lots of choices.

With this setting, the transfer operator is a function $\mathscr{L} : \mathscr{F} \to \mathscr{F}$, and we are iterating it: so working with $\mathscr{L}^n$. It has an obvious fixed point, the invariant measure. There is a basin of attraction: some region of $\mathscr{F}$ on which $\mathscr{L}$ is contracting in some way, so that the region $\mathscr{L}\mathscr{F}$ is somehow "strictly smaller", so that the Banach fixed-point theorem can be applied.

Questions arise:

- What is the shape of this basin? The answer to this seems straightforward: it is the set of functions for which $\int_0^1 \rho(x)\,dx = c \neq 0$. Furthermore, $\lim_{n\to\infty} \mathscr{L}^n \rho = c\nu$ with ν the Gelfond–Parry measure.

- How do we characterize "strictly smaller"? Certainly, any of the Banach norms provide a metric, but only L^1 seems to be compatible with the above definition of the basin. Or ... wait... are we being hasty? Why not general L^p? Hmmm.

- Is the basin regular at the fixed point? That is, is it a stable fixed point? Are there functions $\rho \in \mathscr{F}$ for which $\mathscr{L}^n \rho$ diverges? Since we seem to have picked L^1 as the proper norm, the question becomes, are there $\rho \in \mathscr{F}$ such that $\|\mathscr{L}\rho\|^1 > \|\rho\|^1$. We recognize this last question as being "is $\mathscr{L}$ a bounded operator"? I'm pretty sure the answer is yes, it is bounded, but oddly, before now, never got around to asking this, or exploring the answer. Huh. Blind spots.

- Claim: $\|\mathscr{L}\|^p = 1$ for all p!? For now, a working assumption. What's more, there is only one invariant measure ν and so $\|\mathscr{L} - I\nu\|^p < 1$ so that $\mathscr{L}$ is contracting in all other directions, except ν. There are no unitary eigenvalues, is that right?

The above does seem to settle one issue: the largest reasonable set of functions $\mathscr{F}$ are those that are Borel-meaureable. It seems that being able to talk about $\int_0^1 \rho(x)\,dx$ is a prerequisite; there's not a lot left if we can't talk about that.

The characterization of the basin of attraction above also tells us exactly what the wandering set is: it is the set of functions ρ for which $\int_0^1 \rho(x)\,dx = 0$. To talk about these, it seems we do need an L^p norm for some p. Right?

There is only a single invariant measure, and so define $\mathscr{F}^- = \mathscr{F} \backslash \nu$ so that for any $\rho \in \mathscr{F}$ there is a $\rho^- = \rho - c\nu$ where $\int_0^1 \rho(x)\,dx = c$ as before. Well, we seem to be walking down a conventional path of LRU-style decomposition but applied to an operator. Harrumpf.

7.2   Motivation for Visitation Tree

Instead of working with $\mathscr{F} = \{\rho : [0,1] \to \mathbb{R}\}$ lets work with $\mathscr{F} = \{\sigma : \mathscr{B} \to \mathbb{R}\}$ where $\mathscr{B}$ is the Borel set on the unit interval. So what we've really got is the inverse map $T^{-1} : \mathscr{B} \to \mathscr{B}$ composed with the standard measure $\sigma(A) = \int_A \rho$ for any $A \in \mathscr{B}$. That is, $\sigma = \int \rho$ and we should have been writing ρ' all along. Oh well. Anyway, this gives $\mathscr{L}\sigma = \sigma \circ T^{-1}$.

The Borel set $\mathscr{B}$ can be obtained willy-nilly, but the whole point is that it is easiest to generate it by mid-point iteration. That is, we want to generate a tree of midpoints, and map that into the standard Cantor tree. How does that work? See below.

The idea is that the midpoints, arranged on a tree, make it isomorphic to the Cantor tree. It provides a natural sigma-algebra, in the reference frame of the mid-point iterates. That is, the map is converts the canonical cylinder sets on $\Delta = 2^\omega$ to cylinder sets that are bounded by mid-point iterates. It provides a measure. It allows working with $\mathscr{F} = \{\rho : \Delta \to \mathbb{R}\}$ which is easier than $\mathscr{F} = \{\sigma : \mathscr{B} \to \mathbb{R}\}$.
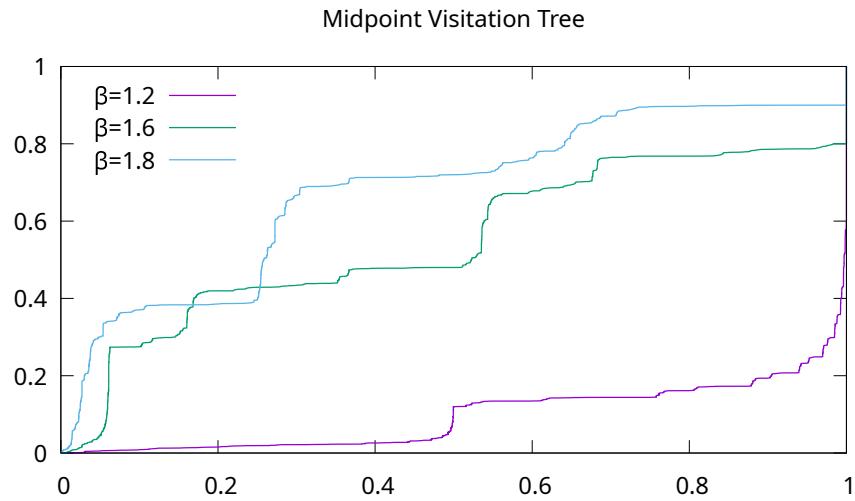
7.3   Visitation Tree

Start with $m_0 = \beta/2$ and iterate once to get $m_1 = \beta(\beta - 1)/2$ which we place at the top of the tree. So m_1 splits the interval $J_1 = [0, m_0]$ into two, a left-half and a right half. Call these two intervals $LJ_1 = J_2 = [0, m_1]$ and $RJ_1 = J_3 = [m_1, m_0]$. Then (assuming a non-finite orbit), each subsequent m_p lands inside of some earlier interval, and splits it in two. Where does it land? That is, $m_p \in J_k$ for some "earlier" k. The intended interval numbering scheme is one that preserves total order on the tree. For each interval $J_k = [a, b]$, there exists the smallest possible p such that $m_p \in J_k$. The sub-intervals are recursively defined as $LJ_k = J_{2k} = [a, m_p]$ and $RJ_k = J_{2k+1} = [m_p, b]$.

The "visitation function" is $v : \mathbb{N} \to \mathbb{N}$ that maps each iterate to the interval it subdivides, so that $m_p \in J_{v(p)}$. For non-finite orbits, this function is a bijection: every interval gets subdivided, sooner or later. For the finite orbits, it collapses: subdivision stops. For finite orbits of length v, the visitation function is defined only up to v.
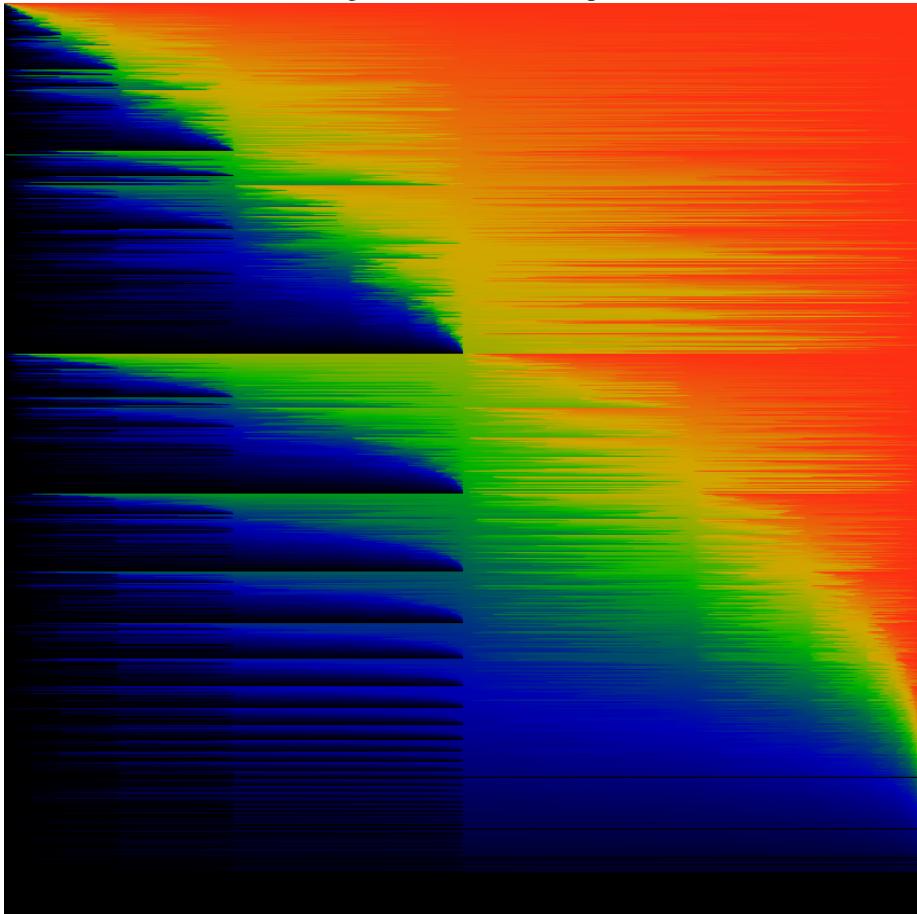
If the orbit iteration is infinite, then this subdivision process places all midpoints somewhere into a binary tree. We'll call this the "visitation tree", to indicate where the midpoint visits. Since it is a tree, we can graph it w.r.t. the dyadic tree. Three such maps are shown in figure 38. These can be understood as three horizontal slices through figure 39.

Figure 38: Visitation Tree

This figure shows the Borel-set-generating visitation map v, or more precisely, the map $m \circ v^{-1} \circ \delta^{-1} : \mathbb{D} \to \mathbb{R}$, for three different values of β. The visitation function is a map $v : \mathbb{N} \to \mathbb{N}$ that records which interval a midpoint lands in, so that $m_p \in J_{v(p)}$. Conversely, fixing the interval to be J_k, one has $m_{v^{-1}(k)} \in J_k$. The intervals are numbered canonically, so that they can be placed in one-to-one correspondence with the canonical dyadic tree. Thus, given a dyadic fraction $(2i+1)/2^n \in \mathbb{D}$, it is mapped to $\delta^{-1} : (2i+1)/2^n \mapsto 2^{n-1} + i = k$. This is mapped to the iteration number $p = v^{-1}(k)$ of the midpoint for that interval. Finally, $m : p \mapsto m_p = T^p(\beta/2)$ as the actual numerical value of the p'th iterate. By construction, this map is necessarily monotonically increasing. The goal of this map is to convert intervals defined by dyadic fractions (along he x-axis) to intervals defined by midpoints (y-axis), thus providing "natural" Borel sets to work with.

Figure 39: Visitation Map

This figure shows the same visitation map as figure 38, but arranged so that $1 \leq \beta \leq 2$ bottom to top, so that a horizontal slice is a constant-β slice. The color coding is the same as used elsewhere: black is approx zero, green is approx 0.5 and red is approx 1.0. Note the hard takeoff for smaller values of beta: the blue-black region at the bottom lies between $\beta = 1$ and about $\beta = 1.2$, after which the march of redness begins diagonally. This is reminiscent of the bracket map, shown in figure 31.

7.4 Finite Visits

The infinite tree above is very nice, but we also need a strong consistent story for the finite orbits. Fir each orbit index n wen know exactly what it's length v should be, and where the indexes go, before they stop. Lets do a table.

v	n	v_j	Paths	$2n+1$
2	1	0, 1	-	11
3	2	0, 1, 3	R	101
	3	0, 1, 2	L	111
4	4	0, 1, 3, 7	R, RR	1001
	6	0, 1, 2, 5	L, LR	1101
	7	0, 1, 2, 4	L, LL	1111
5	8	0, 1, 3, 7, 15	R, RR, RRR	10001
	10	0, 1, 3, 2, 6	R, L, RL	10101
	12	0, 1, 2, 5, 11	L, LR, LRR	11001
	13	0, 1, 2, 3, 5	L, R, LR	11011
	14	0, 1, 2, 4, 9	L, LL, LLR	11101
	15	0, 1, 2, 4, 8	L, LL, LLL	11111

Legend:

- The v_j column shows the entire sequence of midpoint visits. So, v_0 always corresponds to $m_0 = \beta/2$ and v_1 is always $m_1 = \beta(\beta-1)/2$, with m_1 always being the root of the tree. Then the canonical number continues: so v_2 is L and v_3 is R.

- The "paths" column shows exactly the same thing, but encoded as L,R paths. It's got a vaguely suggestive relationship to the $2n+1$ column but the "obvious" relation breaks at n=13. The $2n+1$ column records bits as to whether or not $m_p < 1/2$, whereas the leading symbol of paths records whether or not $m_p < m_1$. So there's similarity, but not really, in the end.

- Sorting is easy; the L,R notation tells us exactly when to swap.

- If the path begins with an R, then $m_1 = \beta(\beta-1)/2 < m_p$ and therefore $\beta/2 < m_p/\beta + 1/2$ and so it is out-of-bounds, and cannot contribute to $\mathscr{L}$.

8 Fractal Eigenfunctions

The Bernoulli operator has the Blancmange curve as an eigenfunction; this is part of a class of fractal eigenfunctions that form the continuous spectrum of the transfer operator. Linear combinations of these can be resumed to define a continuous spectrum of smooth c^∞ eigenfunctions, smooth everywhere except at the endpoints, where they diverge. This section reviews that construction, and applies it to the β-function. The resulting coherent states are not eigenstates; the obstruction is the usual one.

8.1 Bernoulli transform

The Bernoulli transform is the β-transform, at $\beta = 2$. The transfer operator acts as

$$[\mathscr{L}f](x) = \frac{1}{2}\left[f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right)\right]$$

The shift is $\tau(x) = 2x \mod 1$. A family fractal eigenfunctions are the coherent functions of any master wavelet $s \in \ker\mathscr{L}$. Examples include $s(x) = \sin 2\pi x$ and sawtooth function

$$s(x) = \begin{cases} x - \frac{1}{4} & \text{for } x < \frac{1}{2} \\ \frac{3}{4} - x & \text{for } x \geq \frac{1}{2} \\ s(x \mod 1) & \text{for } x \notin [0,1] \end{cases}$$

It is easy to verify that $\mathscr{L}s = 0$ for either of these two examples. Any function s that is odd about 1/2 and periodic with period one will be in the kernel $s \in \ker\mathscr{L}$ and so can be used to build the eigenfunctions.

The family is the coherent-wave sum

$$\psi_{w,l}(x) = \sum_{n=0}^{\infty} w^n s\left((2l+1)\,\tau^n(x)\right)$$

where $0 < w < 1$ is any real number (or, more generally, any complex number within the unit disk) and $l \in \mathbb{N}$ is any non-negative integer. For s the sawtooth, and $l = 0$ this is just the Blancmange curve.

It is not hard to verify that $\mathscr{L}\psi_{w,l} = w\psi_{w,l}$. The proof is straight-forward: $\mathscr{L}(f \circ \tau) = f$ since the transform and the transfer operator are adjoint to one-another. Then, since $\mathscr{L}s = 0$, the coherent-wave sum is just itself, again.

8.2 Beta Transform

The previous construction will build coherent states for the β-transform, but these are not eigenfunctions. The obstruction is that $\mathscr{L}(f \circ \tau) \neq f$ and so the proof fails to go through. This is reviewed below.

The transfer operator is given in eqn 17, copied below.

$$[\mathscr{L}_\beta f](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\right]\Theta\left(\frac{\beta}{2} - y\right)$$

For the case where $f(y)$ vanishes when $y > \beta/2$, as would be hold for an eigenfunction, then it is convenient to use the identity $f(y) = f(y)\Theta(m_0 - y)$ with $m_0 = \beta/2$ and $m_1 = \beta(\beta - 1)/2$ to write

$$\left[\mathscr{L}_\beta f\right](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right)\Theta(m_0 - y) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta(m_1 - y)\right]$$

The shift is $\tau = T_\beta$ as defined in eqn 4, copied below

$$T_\beta(x) = \begin{cases} \beta x & \text{for } 0 \le x < \frac{1}{2} \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} \le x \le 1 \end{cases}$$

The expression for $\mathscr{L}(f \circ \tau)$, for any $f(y) = f(y)\Theta(m_0 - y)$ is obtained by plugging through:

$$\left[\mathscr{L}_\beta(f \circ T_\beta)\right](y) = \frac{1}{\beta}\left[f\left(\frac{y}{\beta}\right)\Theta\left(\frac{\beta}{2} - \frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right)\Theta\left(\frac{\beta}{2} - \left(\frac{y}{\beta} + \frac{1}{2}\right)\right)\right]$$

$$= \frac{1}{\beta}f(y)\left[1 + \Theta(m_1 - y)\right]$$

The step function means the transfer function is unable to undo the action of the shift. For $\beta = 2$, the iterate is $m_1 = 1$ and the Bernoulli result $\mathscr{L}_2(f \circ T_2) = f$ is recovered.

One can still construct the coherent states, as before. They just won't be eigenstates. Given any $v \in \ker\mathscr{L}_\beta$, write the sawtooth $s(x) = v(x \mod 1)$ so as to be explicitly periodic. Then define

$$\psi_{w,l}(x) = \Theta(m_0 - x)\sum_{n=0}^{\infty} w^n s\left((2l+1)T_\beta^n(x)\right) \tag{42}$$

Applying the transfer operator to this gives

$$\left[\mathscr{L}_\beta \psi_{w,l}\right](y) = \left[\mathscr{L}_\beta s \circ (2l+1)\right](y) + \sum_{n=1}^{\infty} w^n\left[\mathscr{L}_\beta s \circ (2l+1)T_\beta^{n-1}T_\beta\right](y)$$

$$= \left[\mathscr{L}_\beta v\right](y) + \frac{w}{\beta}\left[1 + \Theta(m_1 - y)\right]\sum_{n=0}^{\infty} w^n s\left((2l+1)T_\beta^n(y)\right)$$

$$= s\left(\frac{(2l+1)y}{\beta}\right)\Theta(y - m_1) + \frac{w}{\beta}\psi_{w,l}(y)\left[1 + \Theta(m_1 - y)\right]$$

For $l = 0$, the first term is vanishing. For $\beta = 2$, one has $m_1 = 1$ and the Bernoulli form is recovered.

8.3 The Kernel

This section provides a brief description of the kernel.

A function $v \in \ker \mathscr{L}_\beta$ if and only if

$$\mathscr{L}_\beta v = \frac{1}{\beta}\left[v\left(\frac{x}{\beta}\right) + v\left(\frac{x}{\beta} + \frac{1}{2}\right)\right] = 0$$

which immediately implies $v(x) + v(x + 1/2) = 0$. In addition, if v is zero on $[\beta/2, 1]$ then it must also vanish on the interval $[(\beta - 1)/2, 1/2]$.

Equivalently, v is in the kernel if it has the form

$$v(x) = \begin{cases} f(x) & \text{for } 0 \le x < \frac{\beta-1}{2} \\ 0 & \text{for } \frac{\beta-1}{2} < x < \frac{1}{2} \\ -f\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} \le x < \frac{\beta}{2} \\ 0 & \text{for } \frac{\beta}{2} < x \le 1 \end{cases}$$

for any function f.

In the subsequent examples, v generated by

$$f(x) = x - \frac{\beta - 1}{4} \tag{43}$$

will be used. This is a sawtooth, (or cross-section of a volcano) and the corresponding $\psi_{w,l}(x)$ resembles the Blancmange curve.
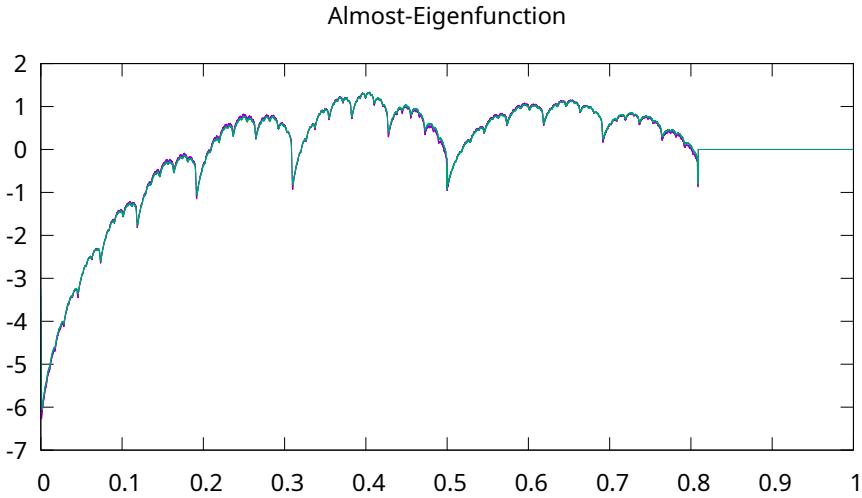
8.4   Anything else?

Define the coherent state $\psi_{w,l}(x)$ as in eqn 42. Then there's a vaguely interesting identity

$$\left[\mathscr{L}_\beta \psi_{w,l} \Theta\right](y) = \frac{1}{\beta} \psi_{w,l}\left(\frac{y}{\beta}\right) \Theta(m_0 - y)$$

where the Θ on the left-hand-side is $\Theta(m_1 - x)$ so that everything above 1/2 is knocked out. Note that the right-hand side has no explicit w terms. It feels as if it could be used to build something clever, but ... No, its just half of the grand total.

Figure 40: Almost-Resonance

Almost-Eigenfunction

Example of an almost-resonance. This is obtained numerically, by constructing a coherent state $\psi_{w,l}(x)$, as defined in eqn 42, for $l = 0, \beta = \varphi, w = 0.7$ and $v(x)$ generated from eqn 43. The result is an "almost eigenfunction"; its OK, but not a very good. However, $\mathscr{L}^9 \psi$ and $\mathscr{L}^{10} \psi$ generate the figure above. The eigenvalue is $\lambda \approx 0.728$, good to three decimal places. This is specifically for $\mathscr{L}^{10} \psi = \lambda \mathscr{L}^9 \psi$.

9 Almost-resonances

Numerical exploration of the coherent states constructed in the previous section reveals a number of almost-eigenfunctions. One is shown in figure 40. Selecting some $\varepsilon > 0$, these are (normalized) functions f for which there is some λ

$$\int_0^1 |\lambda f(x) - [\mathscr{L}f](x)|^p \, dx < \varepsilon$$

(after normalizing $\int_0^1 |f(x)|^p \, dx = 1$ so that the ε makes sense.) If there is one of these for any positive $\varepsilon > 0$ then, by compactness/completeness the whole thing will converge to an eigenfunction. So this question is just as much about almost-resonances: are there small-but-finite ε that prevent almost-resonances from becoming true resonances? Where are they? What are they like?

9.1 Iteration Confusion

Iterating on $\mathscr{L}$ causes exposes cyclic behavior, sometimes seeming periodic, sometimes chaotic. This can be confusing if one is not prepared for it. The underlying cause is that, in general, $\mathscr{L}$ has complex-valued eigenvalues, approximately cyclic.

This results in cyclically-recurring almost-eigenvectors. The symptom and explanation presented below.

This is explored numerically. Some initial vector f is chosen, such that $\int_0^1 |f(x)| \, dx = 1$ so that f is normalized, but also with $\int_0^1 f(x) \, dx = 0$, so that f is orthogonal to the invariant measure (see note on orthogonality below). Setting $f_0 = f$, $\mathscr{L}$ is iterated, so that $f_{n+1} = \mathscr{L} f_n$. After each step, $a_n = \int_0^1 |f_n(x)| \, dx$ is computed. The sequence a_{n+1}/a_n appears to be sometimes chaotic, sometimes ergodic, sometimes periodic. What is this motion?

The confusion comes from assuming that the eigenvalue spectrum of $\mathscr{L}$ is purely real. If $\mathscr{L}$ had a purely real eigenvalue spectrum (as is the case for the Bernoulli shift), then one would expect that repeated iteration would cause eigenvectors associated with smaller eigenvalues to vanish geometrically (by powers) – to iterate away. They don't.

9.1.1 Cyclic Blocks

Here's a model for explaining what's going on. The operator is filled with quasi-cyclic blocks, for example

$$
\begin{bmatrix}
0 & 0 & a \\
b & 0 & 0 \\
0 & c & 0
\end{bmatrix}
$$

Iterating on above will cause pretty much any vector to bounce around in some quasi-cyclic, quasi-ergodic fashion. The above has one real, two complex eigenvalues. The characteristic polynomial is $\lambda^3 - abc = 0$, and so one real eigenvalue at $\lambda = \sqrt[3]{abc}$ and two more at complex $\lambda = -\sqrt[3]{abc}\left(1 \pm i\sqrt{3}\right)/2$. If we get lucky and find the eigenvector for the real eigenvalue, all is well. Anything else will bounce in a cycle.

10   Generating Functions

The β-polynomials can be defined via the mask bits θ_n or via the orbit summatory function ψ_n. Corresponding to each is a unique positive real root $\beta_n = r_n$. This section takes a very short look at the generating functions for these sequences.

For any sequence c_k, the ordinary generating function is

$$\mathrm{OG}\left(c_k;z\right) = \sum_{k=1}^{\infty} c_k z^k$$

The exponential generating function is

$$\mathrm{EG}\left(c_k;z\right) = \sum_{k=1}^{\infty} c_k \frac{z^k}{n!}$$

Numerical results include $\mathrm{OG}\left(\theta_k;\frac{1}{2}\right) = 1.93258880035365\cdots$ This is not in OEIS.
Next, $\mathrm{OG}\left(r_k;\frac{1}{2}\right) = 3.0832181425255\cdots$. This also is not in OEIS.
Next, $\mathrm{OG}\left(\psi_k;\frac{1}{2}\right) = 4.14955396300387\cdots$. This also is not in OEIS.

Exploring these visually, on the complex plane, only the EGF is visually interesting. For both θ_n and r_n, the EGF shows some randomly, uniformly distributed zeros, showing no particular structure. Ordinary double precision allows calculations to about $|z| < 800$ before numeric overflow kills things.

The EGF for ψ_n is much more interesting. Here, the zeros are concentrated into rings, and uniformly distributed about any given ring. A zero-free ray extends along the positive real axis. These rings have a radius of approximately $|z| \approx 35, 55, 100, 190, 350$ with these figures good to about one-and-a-half digits. The rings are wide, and have structure. This is shown in figure 41. The rings are presumably due to ψ_n being concentrated at powers of two, which would be reflected in ring radii of 32, 64, 128, 256... which is approximately what is seen.
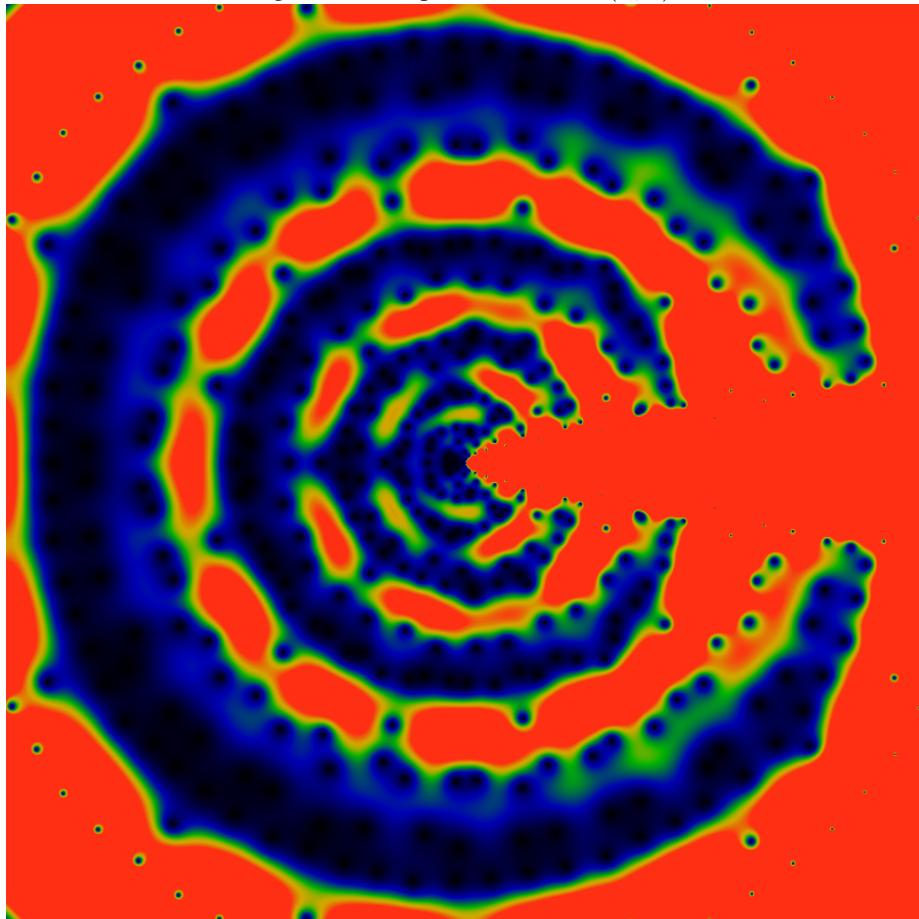
Inspired by the Lambert series, which arranges poles corresponding to the roots of unity along the edge of the unit disk, one may do the same for the golden polynomials. Thus, define the series

$$\lambda\left(z\right) = \sum_{k=1}^{\infty} \theta_n \frac{z^n}{p_n\left(z\right)}$$

This results in a disk that is sufficiently crazy and unusual, that it is worth showing. It arranges poles and zeros into nearby pairs, shown in figure 42.
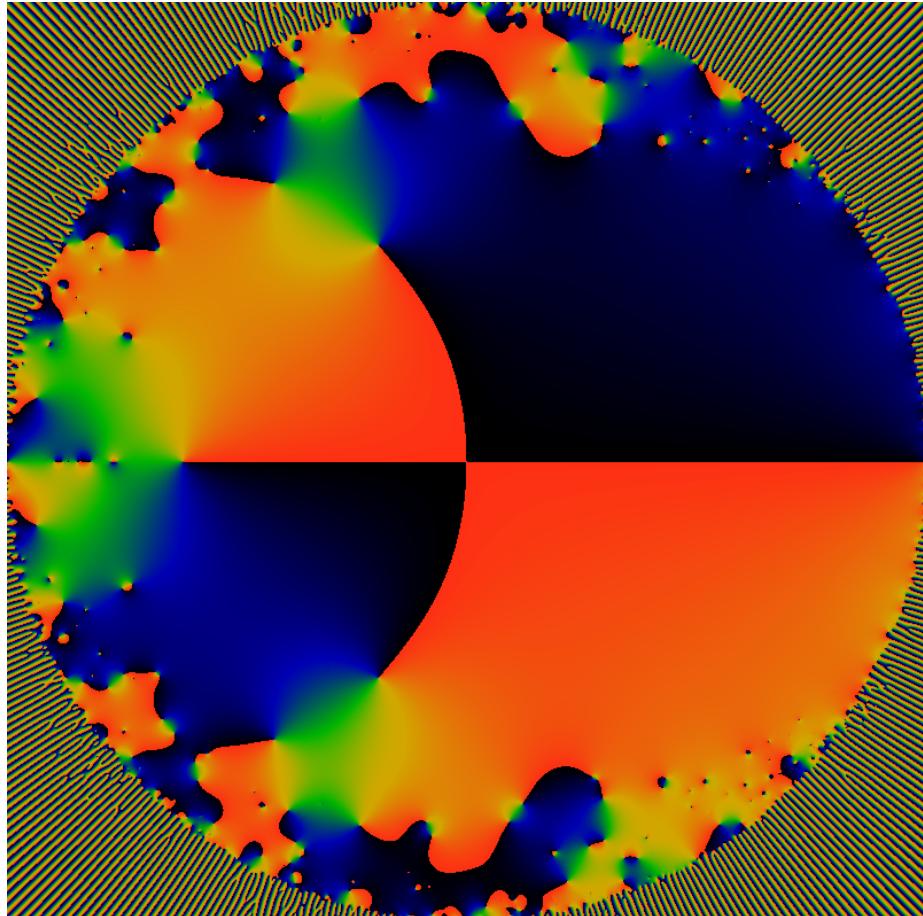
The conclusion to be drawn here is that, although there are lots of generating functions, and possible relations, none are particularly promising.

Figure 41: Complex Zeros of EG $(a_k; z)$

This figure shows the zeros of EG $(a_k; z)$ in the complex plane. The width of the figure is about 800, so that the outermost circle is at about $|z| \approx 350$. The color coding is such that the zeros appear in black, while red denotes a large magnitude. The asymptotic behavior is very approximately $|\text{EG} (a_k; z)| \sim e^{|z|} / |z|$, presumably governed by the asymptotic behavior of Moreau's necklace-counting function. The ring structure presumably reflects the fact that a_n is concentrated just below powers of two.

Figure 42: Lambert-inspired disk

The above illustrates the phase plot of the Lambert-inspired series $\lambda(z)$. The colors encode $\arg \lambda(z)$ with black corresponding to $-\pi$, green to zero, and red to $+\pi$. Note both right-handed and left-handed gradients. The left-handed ones are zeros, the right-handed ones are poles. It is quite unusual to see such structure.

11 Beta Odometer

It is clear that the figure 30 has some self-similarity properties. This section attempts to describe that self-similarity, by identifying sequences of discontinuities, and the regions that they bracket. This allows a labeling system to be developed for each discontinuity, and a bracketing relationship to be given for each self-similar interval.

As always, discontinuities correspond to β values that have finite orbits, and that each finite orbit is associated with a polynomial $p_n(x)$ for which $\beta = r_n$ is the unique positive real root $p_n(r_n) = 0$. Associated to each polynomial is a mask function $\theta_n = \theta_n(r_n)$ given in eqn 32. Discontinuities are associated to indexes n only when $1 = \theta_n$; the mask function acts both to identify the finite orbits and also the endpoints of a self-similar intervals.

The labeling will be for the form of a sequence or string $s = [m_1, m_2, m_3, \cdots]$ of non-negative integers. An example is shown in in figure 43, which can be compared to figure 30. The location β of the discontinuity will be written as $\beta([\cdots]) = \beta[\cdots]$, sometimes dropping parenthesis if the meaning is clear from context. Associated to each sequence is an integer index $n = \eta([\cdots]) = \eta[\cdots]$. By definition, indexes get labels if and only if $1 = \theta_n$; the mask function is the final arbiter for the labels. By construction, the identity $p_{\eta(s)}(\beta(s)) = 0$ must hold for all finite strings s. The task of this section is to expose and describe the bijection between the strings s and the corresponding indexes $n = \eta(s)$.
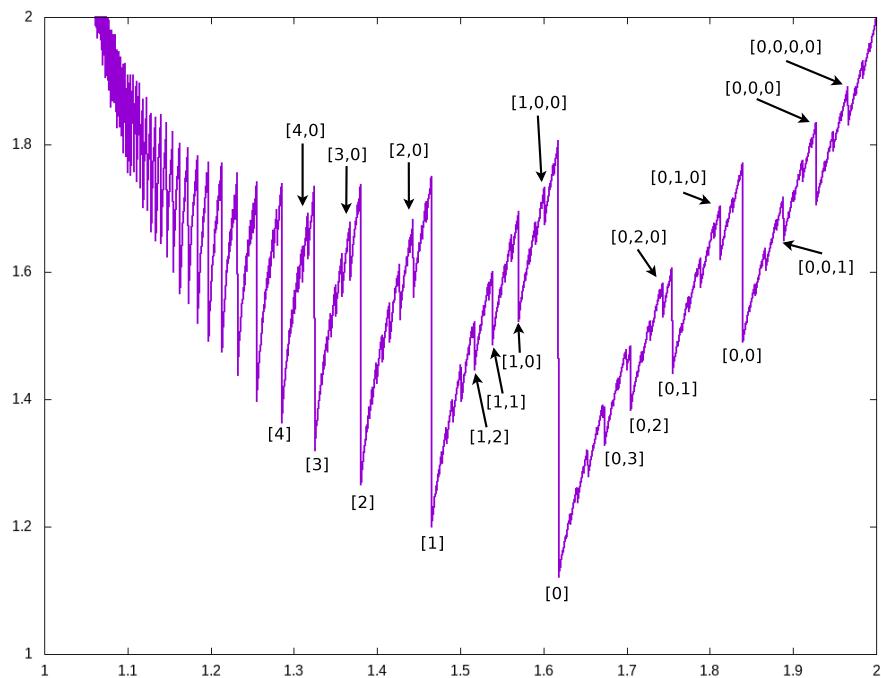
The development can be extended to infinite-length strings, as well. The roots r_n are dense in the interval $1 \leq \beta \leq 2$. The mapping of self-similar intervals provided by the bracketing relationship guarantees that sequences of intervals converge to a point. The labeling system guarantees that all labels within an interval share a common prefix with the endpoints. Thus, as the prefix gets longer, the bracketed interval gets smaller, converging to a unique limit. This is reminiscent of the way in which finite-length continued fractions converge to a unique limit. However, unlike continued fractions, each finite-length β label can be unambiguously associated to one unique infinite-length string. As a result, the mapping is a bijection between infinite-length strings $s \in \mathbb{N}^\omega$ in Baire space, and all of the reals in the interval $1 < \beta \leq 2$. This is a true bijection, so one-to-one and onto. This is very unlike the situation with continued fractions, which cannot represent the rationals with infinite-length strings. Continued-fraction mappings have "holes" at the rationals; the beta mapping has no holes.

11.1 The bracketing relation

The primary task is to exhibit the function $\eta : \mathbb{N}^{<\omega} \to \mathbb{N}$ that maps finite-length strings to indexes. This is given through a recursion relation that depends on the index mask θ_n given previously, in eqn 32. The mask plays a central role; it encodes information about the polynomials and the mid-point orbits; it also encodes the bracketing and self-similarity properties of the system. All this implies that, in a certain sense, θ is chaotic and ergodic; it captures the complexities of the β-transform in full. In particular, this implies that there is no closed-form expression for θ; it can only be defined only recursively, even in principle.

Comparing figure 43 to 30, the labeling begins with the main sequence of peaks,

Figure 43: Discontinuity Labels

This figure shows the integral $I(\beta) = \sum_{n=0}^{\infty} \beta^{-n} T^n \left(\frac{\beta}{2} \right)$ with $1 < \beta \le 2$ running along the horizontal axis. Each discontinuity corresponds to the location of a real root of one of the β-Golden polynomials. Prominent discontinuities are labeled with a finite sequence of non-negative integers following a regular pattern.

shown as 2^k in figure 30 and as $[k]$ in 43. At the center of the diagram is $[0]$, which corresponds to $\beta = \varphi$, root of the polynomial $p_1(\beta) = \beta^2 - \beta - 1 = 0$. Reading off the indexes for the main sequence, these are labeled as $n = 2^k = \eta[k]$. The labeling is defined such that this holds as an identity. The label $[-1]$ will be taken to correspond to $\beta = 2$ in upper right of the figure. Using a negative number for this point already breaks the idea that the labels will consist of non-negative integers. Later on, it will be argued that $[-1]$ can be understood as an infinite string of zeros; but, for now, it can be treated as a funny special value.

There seems to be a clear visual correspondence between the interval $\varphi = \beta[0] < \beta[-1] = 2$ and the other intervals $\beta[k] < \beta[k-1]$. But also it would seem that each of these should correspond to the entire interval $1 = \beta[\infty] < \beta[-1] = 2$. With this in mind, the largest discontinuity to the right of $[0]$ should correspond to $[0]$ itself: call this discontinuity "the leader" or "the front" f. Write the bracketing relationship as $\ell \mapsto f \Leftarrow \rho$ with ℓ and ρ with the left and right endpoints of the interval bracketing f.

With this, several identities emerge: first, that $\beta(\ell) < \beta(f) < \beta(\rho)$. It also appears that $\eta(\ell) < \eta(f) > \eta(\rho)$ will hold, so that the bracketed leader will always have an index greater than indexes on either side. With this notation, one has the following intervals

Bracket	$n_\ell \mapsto n_f \Leftarrow n_r$	$\beta_\ell < \beta_f < \beta_r$
$[\infty] \mapsto [0] \Leftarrow [-1]$	$\infty \mapsto 1 \Leftarrow 0$	$1 < \varphi = 1.61803... < 2$
$[\infty] \mapsto [1] \Leftarrow [0]$	$\infty \mapsto 2 \Leftarrow 1$	$1 < 1.46557... < \varphi$
$[\infty] \mapsto [2] \Leftarrow [1]$	$\infty \mapsto 4 \Leftarrow 2$	$1 < 1.38028...$
$[\infty] \mapsto [3] \Leftarrow [2]$	$\infty \mapsto 8 \Leftarrow 4$	$1 < 1.32472...$

The right-most major interval is $[0] \mapsto [0,0] \Leftarrow [-1]$ so that the labeling rule is to extend the string by appending a zero. The appended zero denotes the leader for that bracket. On the right, a sequence of labels appear, as shown below:

Bracket	$n_\ell \mapsto n_f \Leftarrow n_r$	$\beta_\ell < \beta_f$
$[0] \mapsto [0,0] \Leftarrow [-1]$	$1 \mapsto 3 \Leftarrow 0$	$\varphi < 1.83929...$
$[0] \mapsto [0,1] \Leftarrow [0,0]$	$1 \mapsto 6 \Leftarrow 3$	$\varphi < 1.75488...$
$[0] \mapsto [0,2] \Leftarrow [0,1]$	$1 \mapsto 12 \Leftarrow 6$	$\varphi < 1.7049...$
$[0] \mapsto [0,3] \Leftarrow [0,2]$	$1 \mapsto 24 \Leftarrow 12$	$\varphi < 1.67365...$

From this table, it can be inferred that the recursive sequence index numbering is given by $\eta[0,m] = 3 \cdot 2^m$. This corresponds to the mask identity $1 = \theta(3 \cdot 2^m)$ for all m.

This is sufficient to give a flavor of how bracketing works, and to expose an example of index-doubling within a bracketed interval. The general form of the bracketing relationship is

$$[m_1, m_2, \cdots, m_k] \mapsto [m_1, m_2, \cdots, m_{k+1}] \Leftarrow [m_1, m_2, \cdots, m_k - 1] \qquad (44)$$

The case where $m_k = 0$, resulting in a trailing -1 in the sequence, is handled by means of a "parlor trick". Sequences with a trailing -1 shorten in length, so that $[\cdots, m, -1] =$

221

$[\cdots, m-1]$. The final comma disappears, and the trailing digit is decremented by one. This may seem like a peculiar rule, but is really just a shorthand to simplify the notation for the intervals.

This perhaps most easily understood by plotting actual β values for interval endpoints. One bracketing relationship is $[1,0,0] \Mapsto [1,0,0,0] \Longleftarrow [0]$ which corresponds to $42 \Mapsto 170 \Longleftarrow 1$. The parlor trick just allows writing the right-hand side as $[1,0,-1] = [1,-1] = [0]$. That this is the correct right-hand side is revealed by plotting

$$\beta[1,0,0] = 1.60135... < \beta[1,0,0,0] = 1.61193... < \beta[0] = 1.61803...$$

and observing that $[0]$ is necessarily the right-hand limit for the interval, as there is no other appropriate discontinuity for $[1,0,0,0]$ to be bounded by.

The formula for η is recursive. The trailing digit of a sequence always encodes an index doubling. This allows a general sequence to be reduced to one with a trailing zero.

$$\eta[m_1, m_2, \cdots, m_k, n] = 2^n \eta[m_1, m_2, \cdots, m_k, 0]$$

The above is written with an explicit η, which will be dropped from subsequent expressions. The context should be clear enough: each bracketed form is in one-to-one correspondence with it's integer index.

Truncating the trailing zero is more difficult; this is where the index mask θ comes to the forefront. The recursion rule to truncate the trailing zero is

$$[m_1, m_2, \cdots, m_k, 0] = 2^p (2[m_1, m_2, \cdots, m_k] + 1)$$

where p is the smallest integer for which $\theta(2^p(2[m_1, m_2, \cdots, m_k] + 1)) = 1$. Finding this value of p requires a recursive evaluation of the expression for θ. Although $n = [m_1, m_2, \cdots, m_k]$ is necessarily a valid index, by construction, it is not generally the case that $2n+1$ is valid. The simplest example is $4 = \eta[2]$, for which one has $2 \cdot 4 + 1 = 9$ but $\theta(9) = 0$. A glance at the earlier table of valid sequences shows that $\theta(18) = 0$ but $\theta(36) = 1$ and so one has that $[2,0] = 2^2(2[2]+1) = 2^2(2^3+1) = 36$. As luck would have it, this rule generalizes, so that $[k,0] = 2^k(2[k]+1)$; however, such generalizations are not always possible. Recursion is controlled by θ.

It is not hard to find many examples for which $p = 0$. There are many other examples for which $p = m_1$. In yet other cases, one discovers that $p = \sum_i m_i$ seems to hold. However, none of these are general rules; there is no closed-form expression for p as a function of $[m_1, m_2, \cdots, m_k]$. This may seem frustrating, but in the end, should not be surprising. There is no closed-form expression for the mask function θ that is not recursive. The mask θ encodes the mid-point orbits; the mid-point orbits encode polynomials; the real root of the polynomials equals the β value that is generating that orbit. The orbits are in general chaotic, the progress of the mask θ is chaotic as well. However, thanks to the recursive definition of the polynomials given in eqn 31 and illustrated in figure 28, an appropriately bracketed root will eventually appear, at some finite integer p. In general, but not always, p will be small. The root will always lie in between the two endpoints of the bracketed interval, given by eqn 44.

222

11.2 Bijection to Baire Space

The bracketing relationship allows a simple proof that infinite sequences (the non-finite betas), have a unique convergent. Let $\beta = \beta\,[m_1, m_2, \cdots, m_k]$ be the root of the polynomial determined by $[m_1, m_2, \cdots, m_k]$. This root can be directly computed as the convergent of a β-Fibonacci sequence, as will be demonstrated in an upcoming section. The relevant point here is that, as numerical values, one has that

$$\beta\,[m_1, m_2, \cdots, m_{k-1}] < \beta\,[m_1, m_2, \cdots, m_k] < \beta\,[m_1, m_2, \cdots, m_k - 1]$$

The inequalities are strict. Thus, any given finite-length sequence is bounded above and below by two convergents. This allows infinite-length sequences $[m_1, m_2, \cdots]$ to be defined without ambiguity: they are always bounded above and below, by increasingly tight bounds, and so have have a unique convergent.

The parlor trick turns out to be just the image of $\beta = 2$ under self-similarity transformations. The far-right bound is

$$[0, 0, 0, \cdots] = [-1]$$

The shifted version provides

$$[m_1, m_2, \cdots, m_k, 0, 0, \cdots] = [m_1, m_2, \cdots, m_k, -1]$$

which can be interpreted as the two points on either end of a discontinuity. These identities are meant to be interpreted as the values of limiting sequences. For example, for the second identity, on has that, for a 1 followed by n zeros, that

$$[1, 0, \cdots, 0] \to \beta^{2n+3} - \beta^{2n+2} - \beta^{2n} - \cdots - 1$$
$$= \frac{\left(\beta^2 - \beta - 1\right)\beta^{2n+3} + 1}{\beta^2 - 1}$$

and so in the limit of large $n \to \infty$, the root of this polynomial approaches φ.

The extension of finite sequences into infinite ones provides a bijection (one-to-one and onto) of all real values $1 < \beta \leq 2$ into Baire space $\mathbb{N}^\omega$, where $\mathbb{N} = \mathbb{N}_0$ are the non-negative integers.

11.3 Index distribution
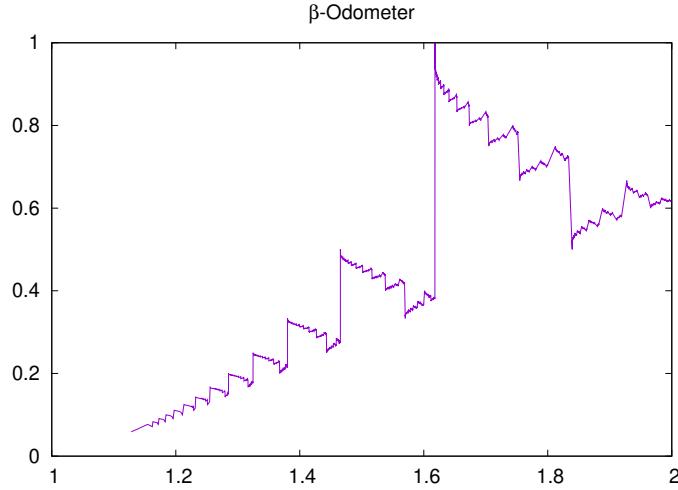
Not all index values appear in the sequence. For any given order, the number expected is given by Moreau's necklace-counting function. But what is the actual distribution?

Attempted graph, its ugly, spiky, unilluminating.

11.4 β-Odometer

The continued fraction expansion can be seen to be an odometer that is of characteristic zero. This is unusual, as odometers are conventionally described in characteristic p. A quick review of the concept of an odometer is in order.

Figure 44: Beta odometer

This visualizes the β-odometer in terms of continued fractions. Any given sequence $[m_1, m_2, \cdots, m_k]$ has an associated β value; this is graphed on the horizontal axis. It also can be interpreted as a continued fraction, this is shown on the vertical axis. It is clearly self-similar. Some of the vertical joins are drawn not quite vertical; this is due to insufficient convergence during graph generation.

An odometer of characteristic p is defined as the set of all infinite-length integer sequences $(a_0, a_1, \cdots)$ with each a_k drawn from the cyclic group $\mathbb{Z}/p\mathbb{Z}$. This set is endowed with a transition function T given by the map

$$T(a_0, a_1, \cdots) \mapsto \begin{cases} (a_0 + 1, a_1, \cdots) & \text{if } a_0 \neq p - 1 \\ (0, \cdots, 0, a_k + 1, a_{k+1}, \cdots) & \text{if } (a_0, a_1, \cdots) = (p-1, \cdots, p-1, a_k, a_{k+1}, \cdots) \end{cases}$$

That is, it increments a_0 by one, and if that rolls over, then a carry bit is propagated to the next term, and so on. This is just p-adic addition, treated as a dynamical system. It can also be interpreted as a map of the unit interval, to itself, when the sequence $(a_0, a_1, \cdots)$ is interpreted as $x = \sum_n a_n p^{-n-1}$. As a map of the unit interval, it permutes a countable sequence of intervals of decreasing size.

The β-odometer is defined as the bijection that arises by interpreting a sequence $[m_1, m_2, \cdots, m_k]$ either as specifying a β value or a continued fraction. The β value is of course, the root of the corresponding polynomial; this root can be computed directly, as a convergent of a β-Fibonacci series, as shown in a later section. The corresponding continued fraction is

$$c = c([m_1, m_2, \cdots]) = \cfrac{1}{m_1 + 1 + \cfrac{1}{m_2 + 1 + \cfrac{1}{\cdots}}}$$

224

where this differs from the conventional definition, being off-by-one, because the m_k can be zero. It is visualized in figure 44.

That the mapping is a bijection is not hard to demonstrate. Certainly, for finite-length sequences, continued fractions are rationals; each such sequence determines a unique rational, and thus the finite β orbits are mapped to the rationals in this way (thus, the map is an injection). Most orbits of the β-transform are not finite, however; these map to infinitely long sequences of bits in the bit expansion. The bracketing relation 44 guarantees that an infinitely sequence is bounded above and below by pairs of finite sequences. The bound applies not only to the sequences, but also to the corresponding β values. The bounds are increasingly tight as the sequence lengthens; by standard epsilon-delta arguments, the convergent is unique for infinite-length series. That is, each distinct non-finite value of β corresponds to a single, unique infinite sequence $[m_1, m_2, \cdots]$. It is not hard to see that all possible such sequences occur. Likewise, every irrational number corresponds to a unique infinite-length continued fraction; again, every possible sequence occurs. Thus, the odometer provides a bijection between irrationals and non-finite orbits. The only difficulty is presented by the rationals: any given rational number has several inequivalent continued fraction expansions, and is thus not strictly invertable. This difficulty is easily swept away by adopting an expansion convention: *e.g.* by taking the shortest possible continued fraction representation for any given rational. In this way, one sees that the odometer is a bijection.

The use of the word "odometer" is perhaps a bit cryptic. The rationale for this can be clarified by looking at how it "rolls over". The prototypical roll-over occurs at the juncture of $\beta = \varphi = 1.618\cdots$. One observes the sequence

$$\beta\,[0,1] = 1.75488\cdots$$
$$\beta\,[0,2] = 1.70490\cdots$$
$$\vdots$$
$$\beta\,[0,16] = 1.61816\cdots$$
$$\beta\,[0,N] > \varphi \text{ for } N \to \infty$$
$$\beta\,[0] = \beta\,[1,-1] = \beta\,[1,0,0,\cdots] = \varphi$$
$$\beta\,[1,0,0,\cdots,1,\cdots] < \varphi$$

That is, the sequence $[0,N]$ for $N \to \infty$ rolls over to $[1,0,0,\cdots]$. It is this peculiar behavior that merits the name "odometer".

Unlike the p-adic case, there is no way to increment the odometer by one unit. The best one can do is to write out the expansion for β, decrement by some arbitrarily small ε and notice that the expansion for $\beta - \varepsilon$ has rolled over at some location. The increment is, formally speaking, an infinitesimal.

This can be compared to how continued fractions roll over. Using our off-by-one notation, write

$$c\,([0,0,N]) = \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{N+1}}} = 0.5 + \varepsilon$$

As the limit $N \to \infty$ is taken, this rolls over to

$$c\left([1,M]\right) = \cfrac{1}{2 + \cfrac{1}{M+1}} = 0.5 - \varepsilon$$

which rolls over to which is, in a sense, backwards: the sequence $[1,N]$ for $N \to \infty$ does not roll over to $[2,0,0,\cdots]$ as one might expect some kind of plausible odometer to do, but instead to $[0,0,N]$.

The p-adic odometers are interval maps; they rearrange countable sequences of subintervals of the unit interval. Here, likewise, the beta odometer rearranges intervals, at least in the sense that rearrangements are disjoint. So, the interval $(\varphi,2)$ is mapped to $\left(\frac{1}{2},1\right)$ albeit in a fractal manner, and likewise each subinterval bounded by the descending series of roots to $\beta^n - \beta^{n-1} - 1 = 0$ is mapped to the interval $\left(\frac{1}{n}, \frac{1}{n-1}\right)$. By self-symmetry, the sub-intervals will rearrange as well.

11.5   Odometer symmetry and the Vandermonde matrix

This section is 98% junk. The bracketing binary tree gives the appropriate symmetry for the betas. This section tries to do the same thing, except using the sequence labels instead. There are two issues below: first, some off-by-one errors in the initial write-up, and some general confusion, and issue two: what the heck are we doing here, that isn't in the binary tree given earlier? The question mark and the interval encoding are similar, but don't really line up; see irred-tree.c and irred-tree.gplot for details. So this section is mostly junk.

The self-symmetry of continued fractions is described by the dyadic monoid.[33] General elements of the monoid have the form $\gamma = g^{a_1} \circ r \circ g^{a_2} \circ \cdots \circ g^{a_N}$ which act on an individual sequences $[x]$ as

$$\gamma : [x] \mapsto [a_1, a_2 - 1, a_3 - 1, \cdots, a_{N-1} - 1, a_N - 2, x]$$

All of the $a_k > 0$ are taken as positive integers, except possibly the first.

When the sequence $[x]$ is understood as a continued fraction, i.e. when applying $c : [x] \mapsto c\left([x]\right)$, then the two generators g, r can be understood as Möbius transformations acting on the unit real interval: $g_M(x) = x/(x+1)$ and the reflection $r_M(x) = 1 - x$. That is, there is a commuting diagram $c \circ g = g_M \circ c$ and likewise $c \circ r = r_M \circ c$. That is, one can go back and forth between the unit interval, and the sequence notation by using the continued-fraction mapping.

Of course, the beta odometer offers another interpretation of sequences as real numbers. Presumably, there are functions $g_B : \mathbb{R} \to \mathbb{R}$ and $r_B : \mathbb{R} \to \mathbb{R}$ that commute with the beta map: that is, $\beta \circ g = g_B \circ \beta$ and likewise $\beta \circ r = r_B \circ \beta$.

It is not clear if there is an analytic form for g_B. Writing

$$g_B(x) = \frac{h(x-1)}{1 + h(x-1)}$$

and experimentally hunting for h reveals an approximation that is good to one percent:

$$h(x) \approx x\left(1 - x/\varphi^2\right)$$

suggesting that an analytic series might be possible. Pursing this requires the inversion of a Vandermonde matrix applied to a shift sequence; this appears to be possibly ill-conditioned.

XXX Confusion abounds! Need more care and a possible redo!

These elements can be put in one-to-one correspondence with the infinite binary tree, i.e. the dyadic monoid. Of course, γ can also be written as a sequence itself, so the dyadic self-symmetry is just an action of finite-length sequences on arbitrary-length sequences, and they are more-or-less nothing more than the act of pre-pending a prefix string. All the confusion and etc. is with avoiding making off-by-one errors during the prefixing operation, and with making explicit identification with the matching Möbius transformations, etc. A non-even number of r's are also a pain in the neck. XXX FIXME. (and fix in the beta paper, too. This is silly-muddled.)

11.5.1 Some continued fraction identities

A short aside to refresh our acquaintance with some continued fraction identities. Rationals have two inequivalent finite continued-fraction expansions:

$$c\left([m_1,m_2,\cdots,m_k,0]\right) = c\left([m_1,m_2,\cdots,m_k+1]\right)$$

If one allows ∞ in an expansion, then one gets a sequence truncation:

$$c\left([m_1,m_2,\cdots,m_k,\infty,m_{k+2},\cdots]\right) = c\left([m_1,m_2,\cdots,m_k]\right)$$

The rollover discussed above corresponds to

$$c\left([m_1,m_2,\cdots,m_k,0,\infty]\right) = c\left([m_1,m_2,\cdots,m_k+1]\right)$$

Defining extended Baire space as $(\mathbb{N}\cup\infty)^\omega$, then the continued fraction mapping will hit all of the rationals. The mapping is not one-to-one, since the first occurrence of ∞ terminates the show, and all strings following this collapse onto the same rational.

Allowing -1 into the mix provides some oddball identities, including a no-op:

$$c\left([-1,m_1,m_2,\cdots]\right) = c\left([m_1,m_2,\cdots]\right)$$

and erasure of the last digit:

$$c\left([m_1,m_2,\cdots,m_k,-1]\right) = c\left([m_1,m_2,\cdots,m_{k-1}]\right)$$

In the middle of the string, one gets a drop:

$$c\left([m_1,m_2,\cdots,m_k,-1,m_{k+2},\cdots]\right) = c\left([m_1,m_2,\cdots,m_k+m_{k+2},\cdots]\right)$$

11.5.2 Vandermonde

An explicit expression for g and r for the odometer is not obvious. Numeric work indicates that $g(x) \approx x\left(1-x/\varphi^2\right)$. This works because the roots almost line up.

Let $f_n(\beta) = \beta^{n+1} - \beta^n - 1$ and let φ_n be the (one and only) real root of $f_n(\varphi_n) = 0$. Finally, let $\alpha_n = \varphi_n - 1$. The remapping requires that $g(\alpha_n) = \alpha_{n+1}$ must hold for $n \geq 0$

as an exact identity. The above approximate g approximately obeys this identity. That is, g acts as a (right-)shift on the root sequence.

Certainly, an analytic series can be obtained; its not clear if this series is adequate for describing the self-symmetry. Let's proceed anyway.

Write

$$g(x) = x + \sum_{n=2}^{\infty} g_n x^n$$

This can be solved subject to the above constraints by inverting

$$\begin{bmatrix} 1 & 1 & 1 & \cdots \\ \alpha_1 & \alpha_1^2 & \alpha_1^3 & \cdots \\ \alpha_2 & \alpha_2^2 & \alpha_2^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ g_2 \\ g_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \end{bmatrix}$$

This matrix is (more or less) a Vandermonde matrix. It can be brought into a more explicit form by writing $g_1 = 1$ and $g_0 = 0$ and $\alpha_0 = 1$. Then one has

$$\begin{bmatrix} 1 & \alpha_0 & \alpha_0^2 & \alpha_0^3 & \cdots \\ 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 & \cdots \\ 1 & \alpha_2 & \alpha_2^2 & \alpha_2^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \end{bmatrix}$$

which is explicitly an (infinite-dimensional) Vandermonde matrix. Any given finite Vandermonde matrix has a non-vanishing determinant, and is thus invertable. Explicit expressions for the inverse can be readily found on the internet. As a general rule, Vandermonde matrices are known to be highly ill-conditioned, and difficult to invert numerically.[43]. In principle, we can avoid such difficulties, as we have precise definitions for all quantities involved; thus, simpler notation, such as that provided by Rawashdeh[44] overpowers other concerns. Still, it is far too complicated.

The shift nature of this can be more directly illustrated. Let

$$A = \begin{bmatrix} 1 & \alpha_0 & \alpha_0^2 & \alpha_0^3 & \cdots \\ 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 & \cdots \\ 1 & \alpha_2 & \alpha_2^2 & \alpha_2^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

and

$$T = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

and

$$\alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \end{bmatrix}$$

228

Then the above eqn takes the form

$$Ag = T\alpha = TA \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

Presuming that A is invertable, then

$$g = A^{-1}TA \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

So the desired solution picks out one column from a shift that uses a Vandermonde matrix as the similarity transform.

The rows of A are the iterates of the monomials in a quotient ring. So, first of all, the polynomials $f_n(\beta)$ are irreducible (over $\mathbb{Q}$; they have one and only one real root; that root is not a rational number). The quotient ring $\mathbb{Q}[\beta]/f_n$ is thus a field. It's also a vector space, spanned by the monomials $1, \beta, \beta^2, \cdots, \beta^n$ as basis vectors, where modulo f_n means that $\beta^{n+1} = \beta^n + 1$. Of course, the root φ_n is an algebraic element of $\mathbb{Q}[\beta]/f_n$ and so L_n is a simple extension of $\mathbb{Q}$ generated by φ_n. The "first ring isomorphism theorem" states that $\mathbb{Q}[\beta]/f_n$ and L_n are isomorphic.

(Note, BTW, that we've broken our convention for n used in the rest of the paper. Everywhere where we have n in this section, it should actually be 2^n to be consistent with earlier notation.)

What is the orbit of φ_n^m for $m \in \mathbb{N}$? Numerically, since $\varphi_n > 1$, taking its powers will clearly blow up. The corresponding β-Fibonacci sequence is $F_{m+1}^{(n)} = F_m^{(n)} + F_{m-n}^{(n)}$ and so one can immediately write $\varphi_n^m = F_m^{(n)} \varphi_n^n + F_{m-1}^{(n)} \varphi_n^{n-1} + \cdots + F_{m-n}^{(n)} \varphi_n^0$ as the orbit of φ_n^m in L_n.

In the present case, we have $\alpha_1 = \varphi_1 - 1 = \varphi - 1$ and so $\alpha_1^2 = 1 - \alpha_1 = 2 - \varphi$, and onward: $\alpha_1^3 = 2\alpha_1 - 1$ and $\alpha_1^4 = 2 - 3\alpha_1$ and $\alpha_1^5 = 5\alpha_1 - 3$ and $\alpha_1^6 = 5 - 8\alpha_1$ and so clearly the Fibonacci sequence is showing up. For the next row, we have that

$\varphi_2^3 - \varphi_2^2 - 1 = 0$, and so the second row has

$$\alpha_2 = \varphi_2 - 1$$
$$\alpha_2^2 = \varphi_2^2 - 2\varphi_2 + 1$$
$$\alpha_2^3 = -2\varphi_2^2 + 3\varphi_2$$
$$\alpha_2^4 = 3\varphi_2^2 - 3\varphi_2 - 2$$
$$\alpha_2^5 = -3\varphi_2^2 + \varphi_2 + 5$$
$$\alpha_2^6 = \varphi_2^2 + 4\varphi_2 - 8$$
$$\alpha_2^7 = 4\varphi_2^2 - 12\varphi_2 + 9$$
$$\alpha_2^8 = -12\varphi_2^2 + 21\varphi_2 - 5$$
$$\alpha_2^8 = 21\varphi_2^2 - 26\varphi_2 - 7$$
$$\alpha_2^9 = -26\varphi_2^2 + 19\varphi_2 + 28$$
$$\alpha_2^{10} = 19\varphi_2^2 + 9\varphi_2 - 54$$
$$\alpha_2^{11} = 9\varphi_2^2 - 63\varphi_2 + 73$$

I am not enlightened by staring at that. Brute force does not appear to work on this.

It's still an interesting problem: each row of the Vandermonde matrix is associated with the polynomial $\beta^{n+2} - \beta^{n+1} - 1 = 0$ and so each row will be some chaotic, iterated polynomial – that is, the β-Fibonacci sequence that delivers φ_n, except that since $\alpha_n = \varphi_n - 1$, each row is multiplied by the Pascal matrix (the matrix of binomial coefficients). That is,

$$\begin{bmatrix} 1 & \varphi_0 & \varphi_0^2 & \varphi_0^3 & \cdots \\ 1 & \varphi_1 & \varphi_1^2 & \varphi_1^3 & \cdots \\ 1 & \varphi_2 & \varphi_2^2 & \varphi_2^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1 & \cdots \\ 0 & 1 & -2 & 3 & \cdots \\ 0 & 0 & 1 & -3 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & & & \vdots & \ddots \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \end{bmatrix}$$

However, this is ill-conditioned. For the matrix A, we had $1 > \alpha_n^m$ for all m, n and even $\alpha_n > \alpha_{n+m}$ for all m, n so A seems well-behaved .... I guess that means A^{-1} is poorly behaved. Here, its worse; one has $1 < \varphi_n^m$ for all m, n so everything blows up.

What else can we do? Well, there's this: write $f_n(\beta) = \beta^{n+1} - \beta^n - 1$ and note that these are easily generated recursively: $f_{n+1}(\beta) = \beta f_n(\beta) + \beta - 1$. The ordinary generating function is easy to find:

$$\sum_{n=0}^{\infty} y^n f_n(\beta) = \frac{y + \beta - 2}{(1 - y)(1 - y\beta)}$$

12 Eventually-periodic orbits

The previous section characterized the set of β values which have midpoint orbits of finite length. Another interesting class is the set of eventually-periodic orbits: orbits of infinite length, settling down to a stable, periodic cycle after an initial bout of chaotic motion. These β values occur as the roots of a slightly different set of self-describing polynomials, as a sum of two parts: one for the initial chaotic motion, and a second polynomial for the cyclic motion. These can be (monotonically) paired with the rational numbers, with infinite paths through the binary tree, and with locations on the comb function. They naturally extend the bracket mapping from the dyadic rationals to all rational numbers.

12.1 Bitsequence polynomials

Consider the set of all eventually-periodic bit-sequences. These consist of a leading chaotic prefix of length L followed by a periodic orbit of length N. Such sequences can be placed in one-to-one correspondence with the rationals, in the conventional fashion. Select such a sequence $\{b_k\}$. The cyclic condition has that $b_k = b_{k+N}$ for all $k \geq L$. Split this into two parts: a finite length-L digit sequence $d_k = b_k$ for $k < L$ and a finite length-N cyclic bit sequence $c_k = b_{k+L}$ for $k < N$. Associated to this is a rational number $x = \sum_{k=0}^{\infty} b_k 2^{-k-1}$. To get it's value, write

$$
\begin{aligned}
S(\beta) &= \sum_{k=0}^{\infty} b_k \beta^{-k} \\
&= \sum_{k=0}^{L-1} b_k \beta^{-k} + \sum_{k=L}^{L+N-1} b_k \beta^{-k} + \sum_{k=L+N}^{L+2N-1} b_k \beta^{-k} + \cdots \\
&= \sum_{k=0}^{L-1} b_k \beta^{-k} + \frac{\beta^N}{\beta^N - 1} \sum_{k=L}^{L+N-1} b_k \beta^{-k} \\
&= \frac{1}{\beta^{L-1}} \sum_{k=0}^{L-1} d_k \beta^{L-k-1} + \frac{1}{\beta^{L-1}(\beta^N - 1)} \sum_{k=0}^{N-1} c_k \beta^{N-k-1}
\end{aligned}
$$

Plugging in $\beta = 2$ gives ratios of whole numbers: a rational that corresponds to this bit-sequence.

Any given $1 \leq \beta \leq 2$ generates a mid-point orbit bit-sequence. Sadly, we've introduced too many different but equivalent notations for this. Starting with the mid-point $x = m_0 = \beta/2$, the characteristic bit-sequence is

$$
b_n = \Theta\left(m_n - \frac{1}{2}\right) = \Theta\left(T_\beta^n\left(\frac{\beta}{2}\right) - \frac{1}{2}\right) = d_n\left(\frac{1}{2}\right) = k_n\left(\frac{\beta}{2}\right) = \varepsilon_n\left(\frac{1}{\beta}\right)
$$

Such orbits have the self-describing property, that $\beta = S(\beta)$. This follows from eqn 6, as well as other identities. For the eventually-periodic orbits, $S(\beta)$ is a ratio of polynomials. It can clearly be placed in one-to-one correspondence with the rationals.

It is not hard to show that it has one unique real root $1 < \beta \leq 2$, and so these are in one-to-one correspondence as well.

This can be related to earlier notation. Writing $\beta - S(\beta) = 0$ and then multiplying through by the denominator,

$$\left(\beta^N - 1\right)\left(\beta^L - \sum_{k=0}^{L-1} d_k \beta^{L-1-k}\right) - \sum_{k=0}^{N-1} c_k \beta^{N-1-k} = 0$$

This touches the earlier notation for the polynomials describing finite orbits:

$$p^{\{d_0 \cdots d_{v-1}\}}(\beta) = \beta^v - \sum_{k=0}^{v-1} d_k \beta^{v-1-k}$$

and so the eventually-periodic orbits are the roots of

$$\left(\beta^N - 1\right) p^{\{d_0 \cdots d_{L-1}\}}(\beta) + p^{\{c_0 \cdots c_{N-1}\}}(\beta) - \beta^N = 0$$

Finite-length orbits have $N = 1$ and $c_0 = 0$, so that the cyclic term vanishes, and the earlier form is recovered, after ignoring an extra factor of $\beta - 1$.

Finite-length orbits can be converted to infinite-length periodic orbits simply by setting the last bit to zero, and then repeating the bit sequence cyclically. Recall, finite orbits always have the last bit equal to one, so this operation is always unambiguous. Restating this explicitly: if $b_0 b_1 \cdots b_{v-1}$ is a finite-length orbit, then set $L = 0$ and $N = v$ and $c_k = b_k$ for $k < v - 1$ and finally $c_{v-1} = 0$. Plugging through just yields the usual polynomial for finite orbits.

Collecting terms of the same order, write

$$0 = \beta^{N+L} - \sum_{k=0}^{N+L-1} a_k \beta^{N+L-1-k}$$

The coefficients are then

$$a_k = d_k + c_{k-L} + \delta_{k-N-1}^0 - d_{k-N}$$

with δ the Kronecker delta. The only coefficients that can appear are in the set $a_k \in \{-1, 0, 1, 2\}$, and at most one coefficient can be 2; it is specifically a_{N-1}, which can only ever be 1 or 2. This can be best understood visually, lining up columns. For $L > N$, get a_k by summing the columns:

$$
\begin{array}{ccccccccc}
 & d_0 & \cdots & & \cdots & d_{L-1} & c_0 & \cdots & c_{N-1} \\
\hline
+ & 0 & \cdots & 0 & 1 & \text{-}d_0 & & \cdots & \text{-}d_{L-1} \\
\hline
 & a_0 & & \cdots & & a_N & & \cdots & a_{N+L-1}
\end{array}
$$

The top row shows the prefix string and the first run of the cyclic string. The second row shows minus the prefix string, shifted all the way to the right, establishing alignment at the right side. Just before it is a lone 1, and then padded on the left with zeros. The bottom row is just the sum of the two rows above it. Clearly, the coefficient a_{N-1} can only ever be 1 or 2. It is exactly the same for $L < N$, with the bottom row being shorter:

$$
\begin{array}{cccccccccc}
 & d_0 & \cdots & d_{L-1} & c_0 & & \cdots & & c_{N-1} \\
+ & 0 & & \cdots & & 0 & 1 & -d_0 & \cdots & -d_{L-1} \\
\hline
 & a_0 & & \cdots & & a_N & & \cdots & & a_{N+L-1}
\end{array}
$$

12.2 Examples

Not all possible bit-sequences $\{b_k\}$ are allowed. They must be self-describing, so that the root of $\beta - S(\beta) = 0$, when iterated, generates $\{b_k\}$. For orientation, some examples are shown below.

$\frac{p}{q}$	$\frac{p+q}{2q}$	m,c	$b_0 b_1 \cdots$	L	N	polynomial	β
0	1/2	1,0	$1^{\bar{}}$	1	1	$\beta - 1 = 0$	1
1	1	0,1	$\bar{1}$	0	1	$\beta - 2 = 0$	2
1/2	3/4	3,0	$11^{\bar{}}$	2	1	$\beta^2 - \beta - 1 = 0$	$\varphi = 1.61803\cdots$
1/3	2/3	0,2	$\overline{10}$	0	2	above	above
2/3	5/6	1,2	$1\overline{10}$	1	2	$\beta^3 - \beta^2 - 2\beta + 1 = 0$	$1.80193773... = 2\cos\frac{\pi}{7}$
5/6	11/12	3,2	$11\overline{10}$	2	2	$\beta^4 - \beta^3 - 2\beta^2 + 1 = 0$	$1.905166167754018...$
11/12	23/24	7,2	$111\overline{10}$	3	2	$\beta^5 - \beta^4 - 2\beta^3 + 1 = 0$	$1.954683120048208...$
						$\cdots$	
1/4	5/8	5,0	$101^{\bar{}}$	3	1	$\beta^3 - \beta^2 - 1$	$1.465571231876768...$
1/7	4/7	0,4	$\overline{100}$	0	3	$\beta^3 - \beta^2 - 1$	above
4/7	11/14	1,4	$1\overline{100}$	1	3	$\beta^4 - \beta^3 - \beta^2 - \beta + 1 = 0$	1.722083805739041
11/14	25/28	3,4	$11\overline{100}$	2	3	$\beta^5 - \beta^4 - \beta^3 - 2\beta^2 + \beta + 1 = 0$	1.87134931301491
25/28	53/56	7,4	$111\overline{100}$	3	3	$\beta^6 - \beta^5 - \beta^4 - 2\beta^3 + \beta + 1 = 0$	1.939924487935805
2/7	9/14	1,2	$1\overline{010}$	1	3	$\beta^4 - \beta^3 - 2\beta + 1 = 0$	1.55897987798175
9/14	23/28	3,2	$11\overline{010}$	2	3	$\beta^5 - \beta^4 - \beta^3 - \beta^2 + 1 = 0$	1.77847961614338
4/5	9/10	1,12	$1\overline{1100}$	1	4	$\beta^5 - \beta^4 - \beta^3 - \beta^2 - \beta + 1 = 0$	1.88320350591352

The rationals in the first column are from the set $\frac{0}{1} \le \frac{p}{q} \le \frac{1}{1}$. The β bit-sequences always have $b_0 = 1$ and thus, the second column shows $S(2) = (p+q)/2q \ge 1/2$. The third column m,c shows the prefix and the cyclic part as integers. The fourth column shows the actual bitsequence. An overline is drawn over the repeating digits. If there is one repeating digit, and it is zero, it is written as $\bar{}$ so as to keep things a bit more readable. The prefix must always start with a 1, and so the integer m is unique. The length of the prefix is in the L column. The cycle might have leading zeros, and so specifying c is not enough; a cycle length is required, given in the N column. The corresponding polynomial and it's root are given in the last two columns.

Notes:

- The sequence $1\overline{10}$ is described in OEIS A160389.

- The sequence $10\overline{10}$ is obviously not allowed, as it is the same as $\overline{10}$.

- The sequence $100\overline{10} = 100\overline{01}$ is not allowed. In the unreduced form, replacing $\overline{10}$ by $11\overline{0}$ gives $100\overline{10} \mapsto 10011 = 19$, which wasn't allowed as a finite orbit.

- The sequence $101\overline{10}$ is not allowed; it reduces to a finite form $101\overline{10} \mapsto 10111 = 23$ that is a disallowed finite orbit.

- These last two observations are a coincidence, and do not hold in the general case. There are good periodic orbits that have disallowed finite versions, and vice-versa.

- The sequence $10\overline{100}$ is not allowed, because it is reducible: $10\overline{100} = 1\overline{010}$.

- 4/7 gives $1\overline{100}$ seems to be OEIS A289917 which is $\left(1 + \sqrt{13} + \sqrt{2\sqrt{13} - 2}\right)/4$.

- 4/5 gives $1\overline{1100}$ seems to be OEIS A289915 which is $\left(1 + \sqrt{2} + \sqrt{2\sqrt{2} - 1}\right)/2$.
  The polynomial factors as $\beta^5 - \beta^4 - \beta^3 - \beta^2 - \beta + 1 = (\beta + 1)\left(\beta^4 - 2\beta^3 + \beta^2 - 2\beta + 1\right)$.

- Rationals with prefix $m = 0$ correspond to the finite orbits.

- It seems the prefix $m = 2^{L-1}$ never occurs, when $L > 1$.

- The prefix is often but not always odd. Counterexamples include $7/12 = 110\overline{01}$ and $13/20 = 110\overline{1001}$.

- The first time that $m = 5$ occurs is for $9/28 = 101\overline{010}$.

Not all rationals give valid bit-sequences. The set of valid rationals are exactly the ones that are in the set $\mathbb{Q} \cap \overline{\theta}$ where $\overline{\theta}$ is the infinite comb, given in eqn 37, and depicted visually in figure 29. The comb is extremely fractal, and working directly with $\mathbb{Q} \cap \overline{\theta}$ would be a chore. Fortunately, the bracket map saves the day: it is a map of the entire unit interval onto $\overline{\theta}$. The bracket map, shown in figure 31, maps rationals to rationals; it provides a map $\mathbb{Q} \cap [0, 1] \to \mathbb{Q} \cap \overline{\theta}$. Thus, it provides a better, though more indirect way, of describing the ultimately-periodic orbits. Some examples are reviewed in the section after next.

12.3 Rational to periodic binary

As a practical matter, it is computationally useful to convert a given fraction p/q into the binary prefix m, it's length L, and the cyclic part c, and its length N. This is not

hard, but also not entirely easy, and so is presented here. Write

$$\frac{p+q}{2q} = \frac{m}{2^L} + \frac{1}{2^L}\left(\frac{c}{2^N} + \frac{c}{2^{2N}} + \frac{c}{2^{3N}} + \cdots\right)$$

$$\frac{1}{2^L}\left(m + \frac{c}{2^N}\left(1 + \frac{1}{2^N} + \frac{1}{2^{2N}} + \cdots\right)\right)$$

$$\frac{1}{2^L}\left(m + \frac{c}{2^N - 1}\right)$$

$$\frac{1}{2^L} \cdot \frac{m\left(2^N - 1\right) + c}{2^N - 1}$$

In the last line, both numerator and denominator are integers. This allows the following algorithm:

1. Find $\gcd(p+q, 2q)$ and so reduce to lowest terms $a/b = (p+q)/2q$.

2. Factor $b = 2^L b'$ to obtain $L \geq 0$.

3. Solve $(2^N - 1) \mod b' = 0$ for the smallest positive $N \geq 1$.

4. Define $r = (2^N - 1) a/b'$.

5. Solve $(r - c) \mod (2^N - 1) = 0$ for the smallest c such that $0 \leq c < 2^N$.

6. Define $m = (r - c)/(2^N - 1)$.

This provides all four integers m, L, c, N that define the cyclic expansion.

12.4   The Bracket Map

The bracketing relationship provided a mapping to the dyadic rationals, via moves on the binary tree. What do ultimately-periodic moves correspond to, on that tree?

The left and right moves L, R on the valid-index binary tree were $L : (\ell \mapsto f \Leftarrow \rho) \mapsto (\ell \mapsto 2f \Leftarrow f)$ and $R : (\ell \mapsto f \Leftarrow \rho) \mapsto (f \mapsto \Lambda(f) \Leftarrow \rho)$. (The symbol R is used in place of $\mathfrak{R}$ in this section, for ease of typography. They are isomorphic.) Consider starting at $\infty \mapsto 1 \Leftarrow 0$ and applying a sequence of alternating left and right moves. This is shown in the table below:

position	move	bracket	β
1/2	-	$\infty \mapsto 1 \Leftarrow 0$	1.61803398...
1/4	L	$\infty \mapsto 2 \Leftarrow 1$	1.46557123...
3/8	LR	$2 \mapsto 10 \Leftarrow 1$	1.57014731...
5/16	LRL	$2 \mapsto 20 \Leftarrow 10$	1.53849659...
11/32	$LRLR$	$20 \mapsto 82 \Leftarrow 10$	1.56175206...
21/64	$LRLRL$	$20 \mapsto 164 \Leftarrow 82$	1.55392112...
43/128	$LRLRLR$	$164 \mapsto 658 \Leftarrow 82$	1.55970265...
85/256	$LRLRLRL$	$164 \mapsto 1316 \Leftarrow 658$	1.55767530...
171/512	$LRLRLRLR$	$1316 \mapsto 5266 \Leftarrow 658$	1.55917021...
1/3	$\overline{LR}$	—	1.55897978...

The column labeled "position" is the location in the dyadic tree. By convention, 1/2 is at the top, and 1/4 lies to the left, and 3/4 lies to the right. The column labeled "move" consists of the sequence of left-right moves to get to a given tree position. For n moves, encoded as a binary integer m, the corresponding position in the dyadic tree is $(2m+1)/2^n$. The bracket is likewise a position in the valid-index binary tree. The string of moves to get to that location are written in reverse applicative order, so that the first letter in the string is the first move. If the moves are treated as functions to be composed and applied, then the string needs to be reversed to get the applicative order. The beta is the beta value at the center of that bracket. Based on the beta value, we conclude that $\mathsf{G}: 1/3 \mapsto 2/7$.

A second example helps cement the idea.

position	move	bracket	β
1/2	-	$\infty \Mapsto 1 \Lleftarrow 0$	1.61803398...
1/4	L	$\infty \Mapsto 2 \Lleftarrow 1$	1.46557123...
1/8	LL	$\infty \Mapsto 4 \Lleftarrow 2$	1.38027756...
3/16	LLR	$4 \Mapsto 36 \Lleftarrow 2$	1.44326879...
5/32	$LLRL$	$4 \Mapsto 72 \Lleftarrow 36$	1.42705896...
11/64	$LLRLR$	$72 \Mapsto 580 \Lleftarrow 36$	1.43949911...
21/128	$LLRLRL$	$72 \Mapsto 1160 \Lleftarrow 580$	1.43591015...
43/256	$LLRLRLR$	$1160 \Mapsto 9284 \Lleftarrow 580$	1.43866733...
85/512	$LLRLRLRL$	$1160 \Mapsto 18568 \Lleftarrow 9284$	1.43784133...
1/6	$L\overline{LR}$	—	1.43841656...

Based on the convergent, we conclude that $\mathsf{G}: 1/6 \mapsto 2/15$, deduced below. In this case, $1/6 \notin \overline{\theta}$ and so the raw fraction 1/6 does not generate a self-describing orbit. But 2/15 is self-describing. Of course, that is the entire intent of the bracket map. Originally formulated as the good-index map $G: \mathbb{N} \to \Psi$ which maps natural numbers to brackets, the extension G is the "good rational" map $\mathsf{G}: \mathbb{Q} \to \mathbb{Q} \cap \overline{\theta}$, which maps rationals to those that are self-describing. This is the same map as visualized in figure 31.

12.5 Bracketing Examples

A curated collection of examples of rationals passed through the bracketing function is given below. It serves mostly to give a sense of the patterns that develop, as well as counterexamples that defy simple patterns.

$\frac{a}{b}$	Ψ	moves	$\frac{p}{q} = \mathsf{G}\frac{a}{b}$	id	orbit	β
0/1	-1	ε	0/1	y	$\overline{1}$	1
$1/2^n - \varepsilon$		$\approx RL \cdots L\overline{R}$	$1/\left(2^{n+1}-1\right) - \delta$	-	$\approx 1\overline{0 \cdots 01}$	
$1/2^n$	2^{n-1}	$RL \cdots L\overline{R}$	$1/2^n$	y	$1\overline{0 \cdots 01}$	

236

$\frac{a}{b}$	Ψ	moves	$\frac{p}{q} = \mathsf{G}\frac{a}{b}$	id	orbit	β
$1/8-\varepsilon$		$\approx RLLL\overline{R}$	$1/15-\delta$	-	$\approx 1\overline{0001}$	<1.380277
$1/8$	4	RLL	$1/8$	y	1001	1.38027756
$1/7$		$\overline{RLLR}$	$4/31$	-	$10\overline{0100}$	1.42109608
$1/6$		$R\overline{LLR}$	$2/15$	-	$1\overline{0010}$	1.43841656
$1/5$		$R\overline{LLRR}$	$12/85$	-	$100\overline{100100}$	1.45394278
$1/4-\varepsilon$		$\approx RLL\overline{R}$	$1/7-\delta$	-	$\approx 1\overline{001}$	<1.4655712
$1/4$	2	RL	$1/4$	y	$10\overline{1}$	1.46557123
$2/7$		$\overline{RLRL}$	$4/15$	-	$10\overline{100}$	1.52626195
$1/3$		$R\overline{LR}$	$2/7$	-	$1\overline{010}$	1.55897987
$2/5$		$R\overline{LRRL}$	$20/63$	-	$10\overline{10100}$	1.58925391
$3/7$		$R\overline{LRR}$	$10/31$	-	$10\overline{1010}$	1.60022189
$1/2-\varepsilon$		$\approx RL\overline{R}$	$1/3-\delta$	-	$\approx 1\overline{01}$	$< 1.61803\cdots$
$1/2$	1	R	$1/2$	y	$11\overline{}$	$\varphi = 1.61803\cdots$
$6/11$			$8900/16383$	-		1.69971346
$5/9$		$\overline{RRLLLRR}$	$40/73$	-		1.70348856
$4/7$		$R\overline{RLL}$	$4/7$	y	$1\overline{100}$	1.72208380
$3/5$		$R\overline{RLLR}$	$25/42$	-	$11\overline{001100}$	1.74720863
$9/14$		$RR\overline{LRL}$	$9/14$	y	$11\overline{010}$	1.77847961
$2/3$		$R\overline{RL}$	$2/3$	y	$1\overline{10}$	1.80193773
$7/10$		$RR\overline{LRRL}$	$43/62$	-		1.82000973
$5/7$		$R\overline{RLR}$	$7/10$	-	$110\overline{110}$	1.82651577
$3/4$	3	RR	$3/4$	y	$111\overline{}$	1.83928675
$7/9$		$R\overline{RRLLLR}$	$227/292$	-		1.86625406
$11/14$		$RR\overline{RLL}$	$11/14$	y	$11\overline{100}$	1.87134931
$4/5$		$R\overline{RRLL}$	$4/5$	y	$11\overline{100}$	1.88320350
$5/6$		$RR\overline{RL}$	$5/6$	y	$11\overline{10}$	1.90516616
$6/7$		$R\overline{RRL}$	$6/7$	y	$1\overline{110}$	1.92128960
$8/9$		$R\overline{RRRLLL}$	$8/9$	y	$1\overline{111000}$	1.93762945
$25/28$		$RRR\overline{RLL}$	$25/28$	y	$111\overline{100}$	1.93992448
$9/10$		$R\overline{RRRLL}$	$9/10$	y	$11\overline{1100}$	1.94470383
$10/11$			$10/11$	y		1.94988340
$11/12$		$RRR\overline{RL}$	$11/12$	y	$111\overline{10}$	1.95468312
$1/1$	0	$\overline{R}$	$1/1$	y	$\overline{1}$	2

Table legend.

- The first column shows selected interesting rationals. Every possible rational is allowed here, and will generate an ultimately-periodic orbit.

- The second column shows the corresponding index Ψ; only seven are shown, for the finite orbits that correspond to the dyadic rationals; all others have an infinite limit.

- The third column shows the bracket moves generated by that rational. These are obtained as the binary expansion of the fraction a/b, or, more properly, the expansion for $(a+b)/2b$. Thus, the first move is always R, to arrive from the far left to the center of the tree. The expansions for 6/11 and 10/11 were left blank; these are pointlessly long strings that would have cluttered the table.

- The fourth column shows the result of the "good rational" map $\mathsf{G} : \mathbb{Q} \to \mathbb{Q} \cap \overline{\theta}$ that takes $\mathsf{G} : \frac{a}{b} \mapsto \frac{p}{q}$. Thus, by definition, $p/q \in \overline{\theta}$. The fractions are those obtained by extending the "good index" map $G : \mathbb{N} \to \Psi$ to it's limit points. A graph of the fourth column, relative to the first, is given in figure 32. Although that figure was prepared for the finite orbits, it is exactly the same for the periodic orbits.

- Some of the discontinuities visible in 32 are shown here, in the rows with $\delta, \varepsilon \ll 1$ in them. For example, the top of the tree is 1/2, which has a finite orbit $1\overline{0}$. It can also be written as an infinite cyclic orbit $1\overline{01} = 1/3$ and these two distinct binary strings are mapped to distinct fractions; thus, the discontinuities. Such a discontinuity appears at every finite orbit.

- Notable rows are 6/11, which has a horribly large denominator; and 7/9 is the runner-up. The large denominator is due to leader heights that are greater than one, that are encountered during the expansion. Compare rows 7/9 to 8/9. Both have long move strings, but $\mathsf{G}(7/9) = 7/9$. Similarly, 6/11 and 10/11 have intolerably long move strings, but $\mathsf{G}(10/11) = 10/11$.

- The fifth column has a 'y' in it, whenever $\mathsf{G}(a/b) = a/b$; when the first and fifth column are equal. Note that most 'y''s occur at the end of the table, rather than at the beginning.

- The sixth column shows the bitsequence $b_0 b_1 \cdots$ for the binary expansion of p/q. Some rows are blank; these rows had horribly long expansions that provide no intuition.

- The seventh and rightmost column shows the corresponding β value for the orbit in the previous column.

Perhaps most notable is that the good map G maps rationals to rationals. Reassuringly, it extends from the dyadic rationals, to all rationals. Formal lemmas and proofs follow in a later section.

13 Formal Work

Theorems, lemmas, proofs. A number of observations, claims and assumptions have been made. They seem "obviously true" from numerical work, but are lacking a formal proof. These are sketched below. As sketches, this chapter is still quite informal; the difference is that it uses a more sophisticated vocabulary to say things in a more abstract manner.

- Theorem: The finite orbits are dense in $1 \leq \beta \leq 2$. Proof: Provided by the bracketing relation, which states that between any two endpoints, there's another orbit strictly in the middle. We even have more: the estimate $v^{1/v}$ from the limit diagram, showing how they don't want to accumulate near beta=1.

- Lemma: The bracketed roots are always in strict ascending order, with $r_\ell < r_f < r_\rho$.

- Lemma: The valid-index map G is a bijection. Proof: Leadership function is monotonically increasing.

- Theorem: The good map G always maps rationals to rationals. More precisely, those rationals are always inside of $\overline{\theta}$ so that $G : \mathbb{Q}_I \to \overline{\theta} \cap \mathbb{Q}$ is a bijection, except at the dyadics, where we have a countable freedom of finite orbits and one infinite-periodic orbit to choose from.

- Theorem: $G : [0,1] \to \overline{\theta}$ is a bijection for all reals $[0,1] \subset \mathbb{R}$. That is, the closure to reals works as expected. This follows because the roots form a countable dense subset, the reals are separable, and the function is continuous.

- Proper diligence requires distinguishing $\overline{\mathfrak{B}}_{\mathbb{Q}}$ from $\overline{\mathfrak{B}}_{\mathbb{R}}$ and also $\overline{\theta}_{\mathbb{Q}}$ from $\overline{\theta}_{\mathbb{R}}$.

The remainder of this section consists of fragments of proofs of the above theorems and lemmas. If these were stitched together properly, whole proofs should emerge. As it is, its more an outline or sketch.

13.1 The Rational Comb

The dyadic comb, of figure 29, shows where valid finite orbits can occur. It's limit to infinite rank gives the infinite comb θ of eqn 37. The only valid rationals, that is, those giving self-describing orbits, are those that belong to $\mathbb{Q} \cap \theta$.

This can be understood as a limiting procedure. Every rational can be approximated by a dyadic rational; each set is dense in the other. The dyadic rationals are finite walks down the binary tree; the non-dyadic rationals are infinite walks. The infinite comb corresponds to the trimmed tree, described earlier. The trimming always maintains branches of unbounded length. These have a limit; in the limit, some of these will correspond to rationals. These are exactly the rationals that correspond to the ultimately-periodic orbits. The correspondence goes in both directions: given an infinite path down the tree, it is sufficient to truncate it to be of finite length. By construction, the truncated path corresponds to a valid dyadic rational. To recap: going

in one direction, there is a sequence of finite orbits that converge, in the limit, onto an eventually-periodic orbit. Conversely, given such an orbit, every truncated version thereof is valid.

13.2   Closures

The ultimately-periodic orbits live inside the closure $\overline{\mathfrak{B}} \subset \overline{\mathbb{B}}$. Write $\mathfrak{C} \subset \overline{\mathbb{B}}$ for the cyclic (ultimately periodic) orbits; these are infinite-length strings describing moves down the infinite binary tree. As already noted, $\mathfrak{C} \cap \mathbb{B} \subset \mathfrak{B}$; that is, every infinite-length cyclic orbit, when truncated to finite length, is a valid finite orbit. The closure $\overline{\mathfrak{B}}$ consists of all self-describing bit-strings obtained as solutions to $\beta = S(\beta)$ where $S(\beta) = \sum_{k=0}^{\infty} b_k \beta^{-k}$. The cyclic orbits are just a special case: $\mathfrak{C} \subset \overline{\mathfrak{B}}$. There are presumably many more, uncountably many chaotic orbits in $\overline{\mathfrak{B}}$.

Just a little bit more machinery is needed. Let 2^ω be Cantor space, the space of infinitely-long binary strings. Let $\chi : 2^\omega \to [0,1]$ be the canonical mapping of Cantor space to the unit interval: $\chi : (b_0 b_1 \cdots) \mapsto \sum_k b_k 2^{-k-1}$. Note that this mapping manages to miss all the dyadic rationals, as these correspond to finite strings, of which there aren't any in the Cantor space. Let $2^{<\omega} = 2^*$ be the set of finite-length binary strings, and allow the symbol χ to perform double-duty, by writing $\chi : 2^* \to \mathbb{D}$ to map finite strings to dyadics. These binary strings can also be interpreted as strings in the symbols L, R, so that they are moves on the tree. Write $\iota : 2^* \to \mathbb{B}$ that maps the empty string to the root of the tree, and likewise $\iota : 2^\omega \to \overline{\mathbb{B}}$ that maps infinite strings to the leaves of the infinite tree.

The comb was a subset of the unit interval: $\overline{\theta} \subset [0,1]$. The claim is that $\overline{\theta} = \chi \iota \overline{\mathfrak{B}}$. The mapping is onto, but not one-to-one: The finite orbits all have corresponding cyclic orbits, all of which are distinct elements in $\overline{\mathfrak{B}}$ but map to the (finite-orbit) polynomial, and thus the same root.

The good-index function provided a bijection $\mathfrak{B} = \eta G \eta^{-1} \mathbb{B}$ between the trimmed and untrimmed finite, unbounded trees. The claim is that this can be extended to a bijection $\overline{\mathfrak{B}} = \overline{G}\,\overline{\mathbb{B}}$. This requires a short detour into topology. The standard (weak) topology on Cantor space is the product topology. The base for the topology may be taken as the set of all finite strings, followed by an infinite number of don't-care markers. These are the open sets; the full topology is the finite intersection and infinite union of the open sets in this base.

Every set in the base of the topology is represented by a finite string. The function η^{-1} maps this to an integer. The function G is defined on all integers; it returns an integer, which is mapped by η to a finite string and thus an open set. To conclude, $\eta G \eta^{-1}$ maps open sets to open sets, and it is defined on *every* open set in the topology. Thus, it is safe to write $\overline{G} = \eta G \eta^{-1}$, as it is defined everywhere; all of $\overline{\mathbb{B}}$ is in it's domain. The range of $\overline{G}$ can be taken to be the definition of the closure $\overline{\mathfrak{B}}$. This has several benefits: it avoids having to make funny arguments that pass through the reals via the comb, and it also makes clear that $\overline{G}$ is a bijection. Note $\overline{G}$ is not continuous, in the context that $\overline{\mathfrak{B}} \subset \overline{\mathbb{B}}$, since $\overline{G}^{-1}$ is defined almost nowhere on $\overline{\mathbb{B}}$.

The points r_n are dense in the interval $[1,2]$. Within each rank ν they are monotonically increasing. The bracketing relation guarantees that, for all r_n in the subtree, these

are all contained within the endpoints of the bracket, and that they are always totally ordered, by the natural tree-ordering. The notion of open sets is compatible. Given a basic open set in $\overline{\mathbb{B}}$, the function $\rho = r \circ G \circ \eta^{-1}$ maps it to an open set $U \subset [1,2]$. Claim that the inverse map also maps open sets to open sets, and thus ρ is continuous. This is established as follows. Since the r_n are dense in the interval $[1,2]$, any open set $U \subset [1,2]$ can be written as a countable union of brackets. The function ρ^{-1} is defined on all brackets, and by bracketing, it is defined on all $U \subset [1,2]$ and by bracketing $\rho^{-1}U$ is an open set in the product topology on $\overline{\mathbb{B}}$.

The goal is to extend this reasoning to $\rho : \mathbb{D} \to [1,2]$, given by $\rho = \rho \circ \eta \circ \delta^{-1} = r \circ G \circ \delta^{-1}$ so as to obtain a continuous, monotonic function $\overline{\rho} : [0,1] \to [1,2]$ on the unit interval. And we are done, more or less.

The mapping χ maps the basic open sets of $\overline{\mathbb{B}}$ to open subintervals of the unit interval; more precisely, to subintervals with dyadics at each end. These are precisely the sets $I(m,v)$ defined in eqn 36. They allowed the infinite comb to be built up from unbounded-length but finite strings in $\mathfrak{B}$.

The base of the product topology was mapped to open intervals on the real-number line; and so $\overline{\rho}$ is continuous on the reals. This can also be seen in a different way. The standard measure μ on the reals says that $\mu I(m,v) = 2^{2-v}$. The function G is built from left and right moves. The left moves bump the rank by one, and so always map to sets that are exactly half the size on the real number line. The right moves are given by the leadership function: $\Lambda\Psi_v \subset \bigcup_{h=0}^{\infty} \Psi_{v+1+h}$ for each rank v. The resulting sets are (at a minimum) half the size, and possibly larger, as the union of additional small intervals. This allows a conventional delta-epsilon proof of continuity to go forward: for each epsilon, one can choose a rank v where the basis sets $I(m,v)$ are smaller than epsilon. The function G maps them to other open sets that are strictly smaller. Strictly, because the height of a leader is always a finite number; the union is written as a union over all possible heights, but the union is always finite.

13.3 Proof that roots are correctly bounded

Theorem: The bracket relationship gives roots with $r_\ell < r_f < r_\rho$ being strict inequalities.

Proof: Recursion starts with $\infty \Mapsto 1 \Mapsfrom 0$. The bracket moves are

$$L : (\ell \Mapsto f \Mapsfrom \rho) \mapsto (\ell \Mapsto 2f \Mapsfrom f)$$
$$\mathfrak{R} : (\ell \Mapsto f \Mapsfrom \rho) \mapsto (f \Mapsto \Lambda(f) \Mapsfrom \rho)$$

So $p_1 = \beta^2 - \beta - 1$ and $p_\infty = p_{1/0} = \beta - 1$ and $p_{1/\infty} = \beta - 2$. Must show that bracket bounds are respected in all four cases. The proofs depend on having $p'_n(\beta) > 0$ for all $1 \le \beta \le 2$.

Case 1: Show that $r_{2f} < r_f$. Proof: The L move gives $p_{2f} = \beta(p_f + 1) - 1$. So $1 = r_{2f}(p_f(r_{2f}) + 1)$ but since $r_{2f} > 1$, must have $p_f(r_{2f}) < 0$ and since $p'_n > 0$, conclude that $r_{2f} < r_f$.

Case 2: Show that $r_\ell < r_{2f}$. Proof: For any $\ell \Mapsto f \Mapsfrom \rho$ it is always the case that ρ is in the tree rooted by ℓ. This can be seen by examining the two possibilities arising

from $\infty \mapsto 1 \Leftarrow 0$ and then keeping in mind all trees are isomorphic. The path to ρ from it's root ℓ is $\rho = L^n R \ell$ for some $n \geq 0$. Thus $p_f = \beta^{n+1} p_\ell - 1$ and so $p_{2f} = \beta^{n+2} p_l - 1$. But $p_\ell(r_\ell) = 0$ and so $p_{2f}(r_\ell) = -1 < 0$. Since $p'_n(\beta) > 0$, conclude that $r_\ell < r_{2f}$.

Case 3: Show that $r_f < r_{\Lambda(f)}$. Proof: Proceeds as a modified proof of case 1. For height zero, $\Lambda(f) = 2f + 1$ so $p_{\Lambda(f)} = \beta p_f - 1$. Thus, $1 = r_{\Lambda(f)} p_f(r_{\Lambda(f)})$ so $0 < p_f(r_{\Lambda(f)}) < 1$ and since $p'_n > 0$, conclude that $r_f < r_{\Lambda(f)}$. What about positive heights? For height h, the moves are $L^h \mathfrak{R}$ and so $p_{\Lambda(f)} = \beta^{h+1} p_f - 1$. Again, $1 = (r_{\Lambda(f)})^{h+1} p_f(r_{\Lambda(f)})$ and again $0 < p_f(r_{\Lambda(f)})$ and so again $r_f < r_{\Lambda(f)}$.

Case 4: Show that $r_{\Lambda(f)} < r_\rho$. Proof: This proceeds as a modified version of case 2, with the role of left and right reversed. In this case, ρ roots a tree, and the path to f is $f = R^n L \rho$ for some $n \geq 0$. For $n = 0$, one has $p_f = \beta(p_\rho + 1) - 1$ and so $p_{\Lambda(f)} = \beta^{h+1}(p_f) - 1$. Expanding, $p_{\Lambda(f)} = \beta^{h+2} p_\rho + \beta^{h+2}(\beta - 1) - 1$, and since $p_\rho(r_\rho) = 0$, one concludes that $p_{\Lambda(f)}(r_\rho) > 0$ and therefore $r_{\Lambda(f)} < r_\rho$. For $n > 0$, the proof proceeds the same way, with $p_{\Lambda(f)} = a p_\rho + b$ for some positive numbers a, b. Thus, $p_{\Lambda(f)}(r_\rho) = b > 0$ and again $r_{\Lambda(f)} < r_\rho$.

To conclude: the roots are ordered as $r_\ell < r_f < r_\rho$ with strict inequalities holding. QED.

13.4 Conclusion

The general sketch of closures seems to be complete. It feels a bit rambling. Adding more detail to it feels like it would just clutter up the arguments even more. I don't think there's anything difficult or challenging in there; it feels all very straight-forward.

The proof of bracketing seems like it should have been easy, but is currently incomplete.

14 The involution Jimm

The rollover of the odometer appears to be given by an involution given the name of "Jimm" by Uludağ and Ayral.[45] This is an involutive outer automorphism of the group PGL(2,Z) given by Joan Dyer in 1978.

This is another run-length encoding trick.

15 Galois extensions

The β-golden polynomials defined in the last section form a curious class. One can now embark down a path of arithmetic. For each such polynomial (or analytic series, as the case may be) and given some field K (for example, say, the rationals $\mathbb{Q}$), one can construct the Galois extension $K(\beta) : K$.

For the case of β periodic, there is just one real root φ_n, and it is not rational. This root is an algebraic element of the polynomial ring $\mathbb{Q}[\beta]$ (obviously), and so it generates the (simple) field extension $\mathbb{Q}(\varphi_n)$. The order of the polynomial $p_n(\beta)$ is $\lfloor \log n \rfloor + 1$ and so this is obviously the transcendence degree of the field extension. The elements φ_n^m for $0 \le m \le \lfloor \log n \rfloor$ provide the basis.

For the case of β aperiodic, the vector space is clearly infinite-dimensional.

A few questions:

- What is the automorphism group?

- What is the field of fractions?

16   Islands of Stability as Arnold Tongues

The classical Feigenbaum bifurcation diagram, shown in figure 5, manifests two distinct behaviors: the "islands of stability", in which there are periodic orbits, and the "chaotic regions". These regions alternate and interleave as a parameter λ appearing in the iterated equation $\lambda x(1-x)$ is swept through a range of values. By contrast, the equivalent diagram for the beta-map, shown in figure 2, does not seem to have regions of stability. This is only an illusion: they are there, they are only infinitely thin. This chapter focuses on how to crowbar them open, to finite size.

In the previous chapters, it was demonstrated that there is a countable set of β values, dense in the range $1 \leq \beta \leq 2$, for which orbits are finite and terminate after a fixed number of iterations. Alternately, they can be made periodic, simply by changing a less-than sign to a less-than-or-equals sign. The β values for which orbits are periodic can be placed in correspondence with the periodic orbits of the logistic map; the β values for which orbits are chaotic correspond to the chaotic orbits of the logistic map. The open problem is to demonstrate this correspondence explicitly. This problem is not tackled here; it is just brushed up against.

The circle map $x_{n+1} = x_n + \theta + K \sin 2\pi x_n$ provides one possible mechanism for taking a set of measure zero, and crowbaring it open to a set of finite size. For $K = 0$, this is just the rotation map $x_{n+1} = x_n + \theta$ which has only periodic orbits, when θ is rational, and chaotic orbits, when θ is not rational. As one sweeps θ through a range, the subset of periodic orbits is countable, and is a set of measure zero: thus, the rotation map can be said to be chaotic for almost all θ. Setting K to a non-zero value expands the regions of periodic orbits to finite size, termed Arnold tongues. These are the mode-locking regions that are generically visible in driven oscillator systems. The perturbation by the kick K displaces what would have been chaotic orbits into mode-locked regions. For small K, this perturbation is soft, in that one might say "it shouldn't change things much." But even a whisper of a miniscule K is enough to convert the set of periodic orbits from a measure of zero to a finite measure.

Another possibility is to just crowbar open the periodic regions with a "hard" perturbation, localized at a point. Take the natural saw-tooth shape of the β-map, widen the middle, and insert a slanting downward line, to create a zig-zag. That is, connect the two endpoints in the middle of the beta shift, "widening" it so that it has a finite, not infinite slope, thereby converting the iterated function from a discontinuous to a continuous one. This can be constructed directly: given some "small", real $\varepsilon > 0$, define the piecewise-linear ε-generalization of the map 4 as
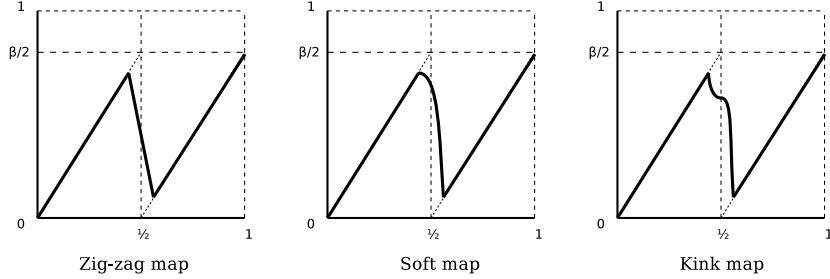
$$T_{\beta,\varepsilon}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \beta\left(\frac{1}{4} - \varepsilon\right)w & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \tag{45}$$

where w is just a handy notation for a downward sloping line:

$$w = \frac{2x - 1}{2\varepsilon}$$

Observe that $w = 1$ when $x = \frac{1}{2} - \varepsilon$ and that $w = -1$ when $x = \frac{1}{2} + \varepsilon$ so that w just

smoothly interpolates between +1 and -1 over the middle interval. The additional factors of $\frac{\beta}{4} - \beta \left(\frac{1}{4} - \varepsilon \right) w$ just serves to insert the downward slope smack into the middle, so that the endpoints join up. The results is the zig-zag map, illustrated in the figure below

| Zig-zag map | Soft map | Kink map |

In the limit of $\varepsilon \to 0$, one regains the earlier beta shift: $\lim_{\varepsilon \to 0} T_{\beta,\varepsilon} = T_\beta$, as the slope of the middle bit becomes infinite. The middle segment is a straight line; it introduces another folding segment into the map. This segment introduces a critical point only when ε is sufficiently large, and β is sufficiently small, so that its slope is less than 45 degrees (is greater than -1). When this occurs, a fixed point appears at $x = 1/2$. A sequence of images for finite ε are shown in figure 45.

The appearance of islands of stability in the Feigenbaum attractor is due to the presence of a fixed point at any parameter value. In order to "surgically add" islands of stability to the beta transform, the middle segment interpolation must also have a critical point at "any" value of ε. To achieve this, consider the curve

$$D_{\beta,\varepsilon}(x) = \begin{cases} \beta x & \text{for } 0 \le x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \beta \left(\frac{1}{4} - \varepsilon \right) g(w) & \text{for } \frac{1}{2} - \varepsilon \le x < \frac{1}{2} + \varepsilon \\ \beta \left(x - \frac{1}{2} \right) & \text{for } \frac{1}{2} + \varepsilon \le x \le 1 \end{cases} \qquad (46)$$

where the straight line has been replaced by a soft shoulder
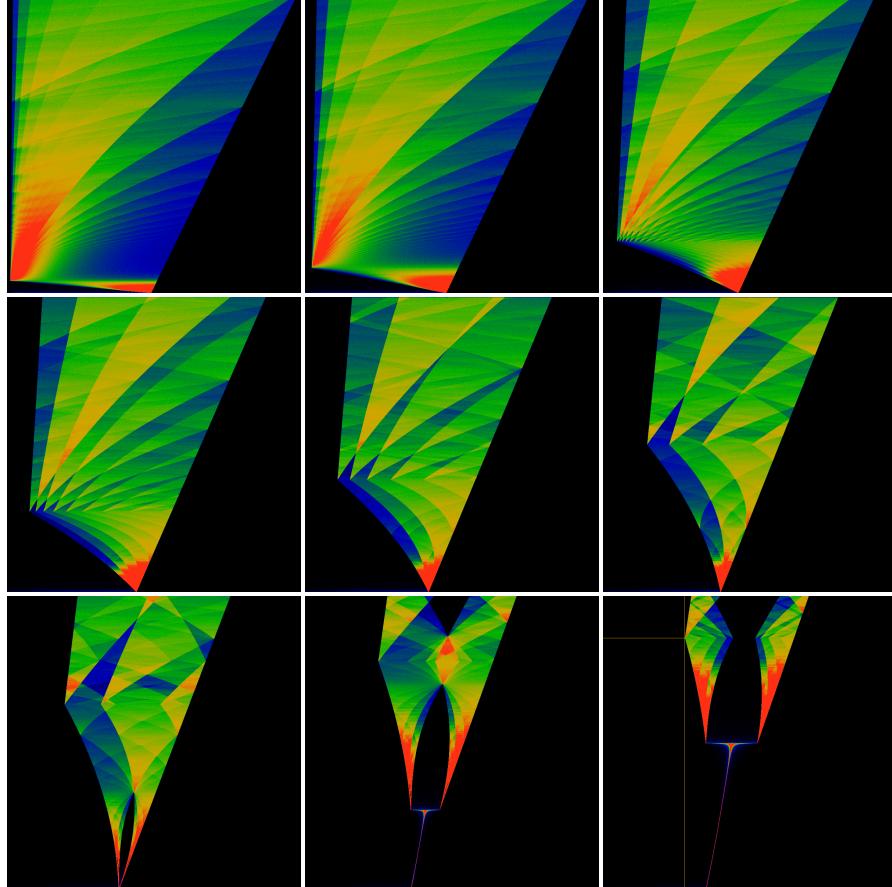
$$g(w) = 1 - 2\cos\frac{\pi}{4}(1 + w)$$

and w is the same as before. This is scaled so that its a drop-in replacement for the straight line: $g \left(\frac{1}{2} - \varepsilon \right) = 1$ and $g \left(\frac{1}{2} + \varepsilon \right) = -1$. A cosine was used to create this soft shoulder, but a parabola would have done just as well. It is illustrated above, with the label "soft map".

This map also interpolates between the left and right arms of the beta transform, forming a single continues curve. The curve is smooth and rounded near $\frac{1}{2} - \varepsilon \lesssim x$, having a slope of zero as x approaches $\frac{1}{2} - \varepsilon$ from above. This introduces a critical point near $\frac{1}{2} - \varepsilon$. Notice that there is a hard corner at $\frac{1}{2} + \varepsilon$. The interpolation is NOT an S-curve! A sequence of images for finite ε are shown in figure 46.
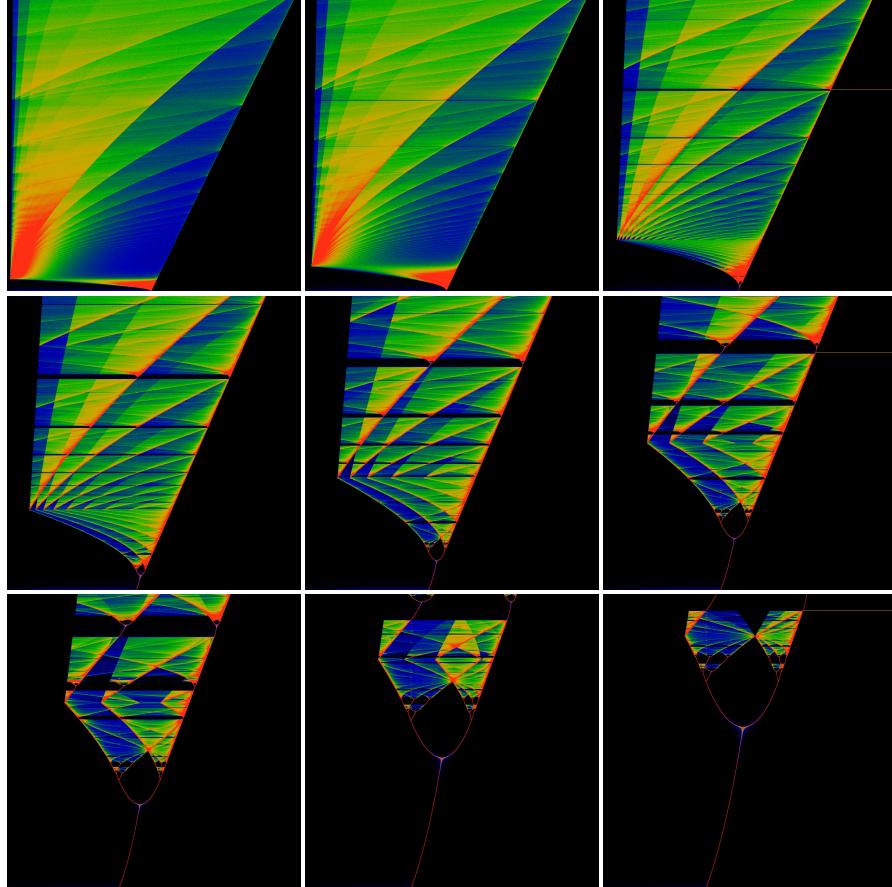
Two more variant maps can be considered. Both replace the center piece with

Figure 45: Z-shaped Map

This illustrates a sequence of iterated maps, obtained from eqn 45. Shown are $\varepsilon = 0.01$, 0.02, 0.04 in the first row, 0.06, 0.08, 0.10 in the second row and 0.12, 0.14, 0.15 in the third row. The image for $\varepsilon = 0$ is, of course, figure 2. The parameter β runs from 1 at the bottom to 2 at the top. Thus, a horizontal slice through the image depicts the invariant measure of the iterated map, black for where the measure is zero, and red where the measure is largest. The sharp corner at the lower-left is located $\beta = (1+2\varepsilon)/(1-2\varepsilon)$ and $x = \varepsilon(1+2\varepsilon)/(1-2\varepsilon)$. A yellow horizontal and vertical line in the last image indicate the location of this corner.

Figure 46: Critical-Z map

This illustrates a sequence of iterated maps, obtained from eqn 46. The sequence of depicted ε values are the same as in figure 45. The top row shows $\varepsilon = 0.01, 0.02, 0.04$, with 0.06, 0.08, 0.10 in the second row and 0.12, 0.14, 0.15 in the bottom row. The image for $\varepsilon = 0$ is, of course, figure 2. The parameter β runs from 1 at the bottom to 2 at the top. Working from bottom to top, one can see islands of stability forming in the $\varepsilon = 0.02$ and 0.04 images. The largest island, one third from the top, corresponds to $\beta = \varphi = 1.618\cdots$ the golden ratio. Moving downwards, the other prominent islands correspond to the "trouble spots" 101, 1001 and 10001, which are the Narayana's Cows number, an unnamed number, and the Silver Ratio, at $\beta = 1.4655\cdots$ and so on. Moving upwards, one can see a faint island at the tribonacci number. Due to the general asymmetry of the map, these islands quickly shift away from these limiting values. For example, the primary island appears to start near $\beta = \delta + (2-\delta)(\varphi - 1)$, where $\delta = (1+2\varepsilon)/(1-2\varepsilon)$. This location is indicated by a horizontal yellow line in the images in the right column. The other islands shift away in a more complicated fashion.

247

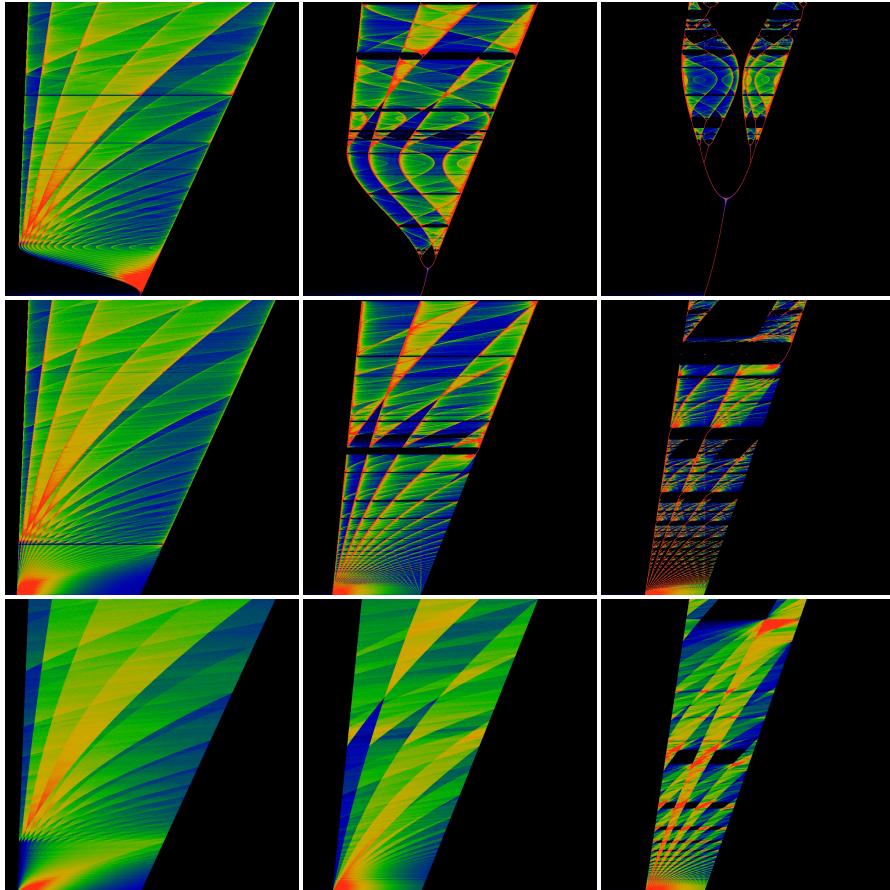symmetrical sinuous S-shaped curves, but in different ways. Consider

$$S_{\beta,\varepsilon,\sigma}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \sigma\beta\left(\frac{1}{4} - \varepsilon\right)\sin\frac{\pi}{2}w & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \tag{47}$$

and

$$H_{\beta,\varepsilon,p,\sigma}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \sigma\beta\left(\frac{1}{4} - \varepsilon\right)\operatorname{sgn}\left(x - \frac{1}{2}\right)|w|^p & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \tag{48}$$

The $S_{\beta,\varepsilon}(x)$ replaces the central segment with a softly-rounded segment, containing two critical points: near $\frac{1}{2} - \varepsilon$ and near $\frac{1}{2} + \varepsilon$, where the curve flattens out to a zero slope. When $\sigma = +1$, the map as a whole is continuous. When $\sigma = -1$, the map consists of three discontinuous pieces. Different values are explored in figure 47.

The $H_{\beta,\varepsilon,p,\sigma}(x)$ replaces the central segment with a segment that has a kink in the middle, when $p > 1$. Note that $H_{\beta,\varepsilon,1,1}(x) = T_{\beta,\varepsilon}(x)$. Here, $\operatorname{sgn}x$ is the sign of x. The general shape of $H_{\beta,\varepsilon,p,\sigma}(x)$ is shown above, labeled as the "kink map". The location of the kink in H is always centered; an off-center kink, as depicted in the figure, is explored below. The bifurcation diagrams for H are illustrated in figure 48.

To summarize: the "trouble spots" aren't "just some periodic orbits" at certain values of β: they are more "fundamental" than that: they indicate the regions where ("phase-locked") periodic orbits can be made to appear. And conversely: bifurcations can only appear here, and not elsewhere! The last sequence of images, shown in figure 48 indicate that the islands of stability need NOT consist of the period-doubling sequences seen in the Feigenbaum map. This is made explicit in figure 49, which shows a zoom by a factor of thirty.

Another interesting visualization is a Poincaré recurrence plot. The islands of stability should manifest as Arnold tongues[46]. These are shown in figures 50 and 51.

To intuitively understand the location of the islands (the location of the Arnold tongues), its easiest to examine a map with a kink in it, whose location is adjustable.

$$H_{\beta,\varepsilon,\alpha,\sigma}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \sigma\beta\left(\frac{1}{4} - \varepsilon\right)h_{\alpha,p} & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases}$$

with

$$h_{\alpha,p}(x) = \begin{cases} \alpha + (1-\alpha)|w|^p & \text{for } x < \frac{1}{2} \\ \alpha - (1+\alpha)|w|^p & \text{for } \frac{1}{2} \leq x \end{cases}$$

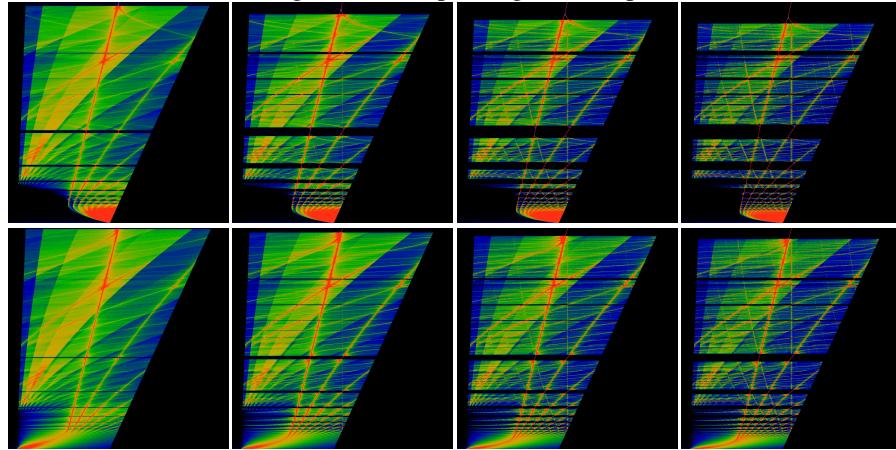As before, $h_{\alpha,p}(x)$ is designed to interpolate appropriately, so that $h_{\alpha,p}\left(\frac{1}{2} - \varepsilon\right) = 1$ and $h_{\alpha,p}\left(\frac{1}{2} + \varepsilon\right) = -1$. The location of the kink is now adjustable: $h_{\alpha,p}\left(\frac{1}{2}\right) = \alpha$. Iterating on this map results in figures that are generically similar to those of figure 48, except that this time, the location of the islands is controllable by the parameter α. Roughly,

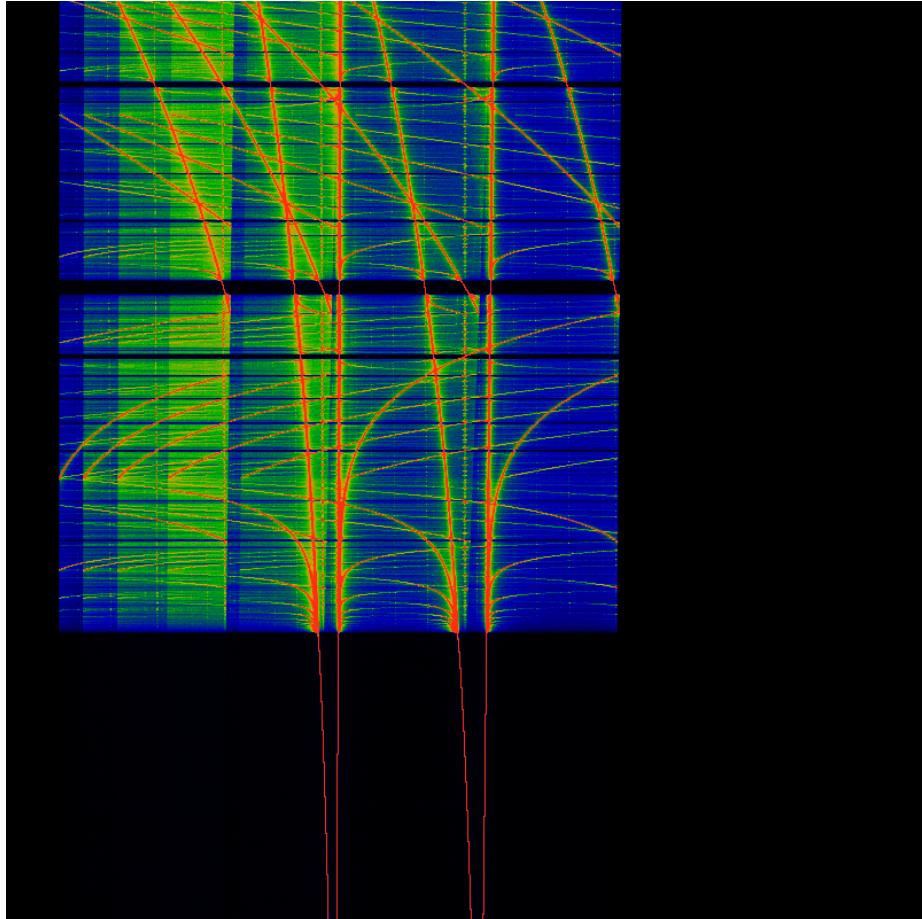Figure 47: Interpolating Sine Map

This illustrates a sequence of iterated maps, obtained from eqn 47. The sequence in the upper row shows $\varepsilon = 0.04$, 0.10 and 0.15; with $\sigma = +1$. The upper row is much like the sequence shown in figure 46, except that its made sinuous, thanks to symmetrical S-shape. The middle row shows the same ε values, but for $\sigma = -1$. The bottom row shows eqn 48 with $p = 1$ and $\sigma = -1$; thus, because $p = 1$ gives a straight-line segment in the middle, this bottom row is directly comparable to the zig-zap map. It should make clear that the islands appear in the middle row due to critical points in the S-curve, and not due to the tripartite map. The lower right diagram exhibits islands, but only because the middle segment has a slope of less than 45 degrees, resulting in a critical point at the middle of the map. As usual, the parameter β runs from 1 at the bottom to 2 at the top.

Figure 48: Interpolating Kink Map
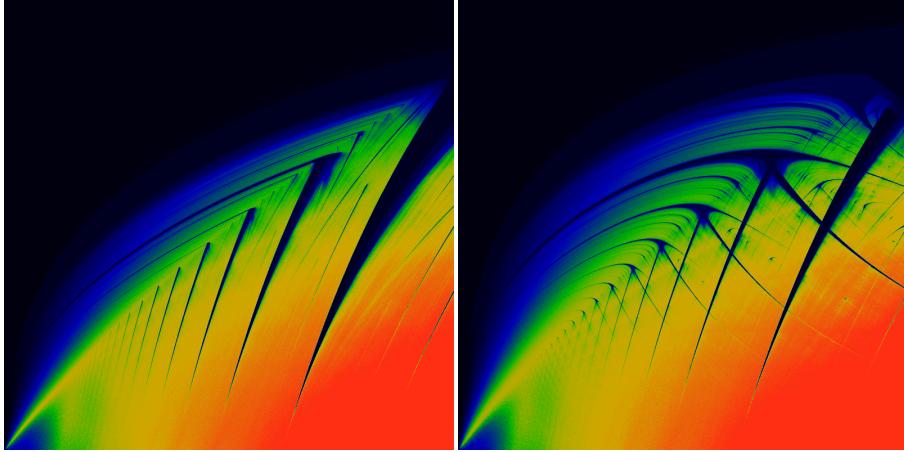
This illustrates a sequence of iterated maps, obtained from eqn 48. All eight images are held at $\varepsilon = 0.04$. The top row has $\sigma = +1$ (and thus the map is continuous) while the bottom row has $\sigma = -1$ (and thus the map has three disconnected branches. Left to right depicts the values $p = 2, 3, 4, 5$. As usual, the parameter β runs from 1 at the bottom to 2 at the top. In all cases, islands appear, and numerous common features are evident. Perhaps most interesting is that the islands do NOT contain period-doubling sequences. The primary sequence of islands, starting from the central largest, proceeding downwards, are located the inverse powers of two, *viz* at $\beta = \sqrt[k]{2}$. Why are the islands located at inverse powers of two, instead or, for example, the golden means? The short answer: it depends on the location of the kink in the map, as explored in the main text.

Figure 49: No Period Doubling
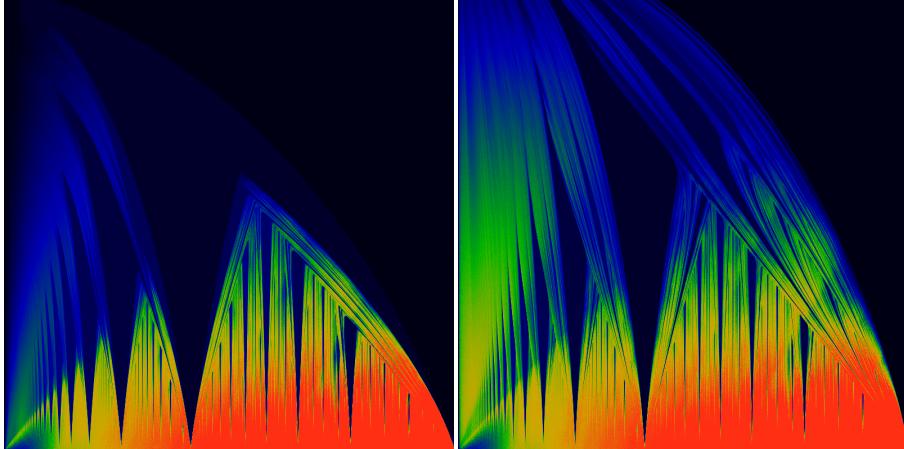
This figure is a zoom, confirming a lack of period doubling in the map $H_{\beta,\varepsilon,p,\sigma}(x)$ of eqn 48. The explored region is $0 \le x \le 1$, viz no zoom in the horizontal direction. Vertically, the image is centered on $\beta = 1.45$, having a total height of $\Delta\beta = 0.015625$. This uses the quintic kink, so $p = 5$ and $\sigma = +1$, making the the continuous variant. The value of $\varepsilon = 0.04$ makes this directly comparable to other images.

Figure 50: Poincaré recurrence

The above visualize the Poincaré recurrence times for the map $D_{\beta,\varepsilon}(x)$ of eqn 46 on the left, and the map $S_{\beta,\varepsilon,1}(x)$ of eqn 47 on the right. In both cases, the parameter β runs from 1 to 2, left to right. The parameter ε runs from 0 to 0.2, bottom to top. The Poincaré recurrence time is obtained by iterating on the maps, and then counting how many iterations it takes to get near an earlier point. The color coding is such that yellow/red indicates large recurrence times; green is intermediate time, blue a short time, and black corresponds to n less than 3 or 4 or so. The vertical black spikes are the Arnold tongues; they correspond to parameter regions which lie in an island of stability. That is, the recurrence time is low, precisely because the the point x is bouncing between a discrete set of values. The yellow/red regions correspond to chaos, where the iterate x is bouncing between all possible values. The largest right-most spike is located at $\beta = \varphi = 1.618\cdots$, with the sequence of spikes to the left located at the other primary golden means (*viz*, $1.3803\cdots$ and the silver mean $1.3247\cdots$ and so on). As noted earlier, the general curve of that spike appears to follow $\beta = \delta + (2-\delta)(\varphi-1)$, where $\delta = (1+2\varepsilon)/(1-2\varepsilon)$. The dramatic swallow-tail shapes in the right-hand image are identical to those that appear in the classic iterated circle map.[46]

Figure 51: Arnold Tongues

The above visualize the Poincaré recurrence times for the map $H_{\beta,\varepsilon,p,\sigma}(x)$ of eqn 48. The parameter β runs from 1 to 2, left to right. The parameter ε runs from 0 to 0.2, bottom to top. The power p is held fixed at $p = 5$. The left image shows $\sigma = -1$; the right shows $\sigma = +1$. The Poincaré recurrence time is obtained by iterating on $H_{\beta,\varepsilon,p,\sigma}(x)$ and counting how many iterations it takes until $\left| x - H^n_{\beta,\varepsilon,p,\sigma}(x) \right| < 0.009$. The shapes depicted are not sensitive to the recurrence delta 0.009; this value is chosen primarily to make the colors prettier. The color coding is such that yellow/red indicates large recurrence times n; green is intermediate time, blue a short time, and black corresponds to n less than 3 or 4 or so. The vertical blue spikes are the Arnold tongues; they correspond to parameter regions which lie in an island of stability. That is, the recurrence time is low, precisely because the the point x is bouncing between a discrete set of values. The yellow/red regions correspond to chaos, where the iterate x is bouncing between all possible values. The central spike is located at $\beta = \sqrt{2}$ with the sequence of spikes to the left located at $\sqrt[k]{2}$ for increasing k. In that sense, the large black region dominating the right side of the figures corresponds to $\beta = 2$. These correspond to the black bands in figure 48.

to first order, the primary series of islands are located at $\sqrt[k]{2/(1-\alpha)}$; as before, these islands do not allow period-doubling to take place.

To get islands with period doubling, one needs to recreate the "soft shoulder" of eqn 46, but at a variable location.

Thus, the above presents a general surgical technique for controlling both the general form of the chaotic regions, the location of the islands of stability, and what appears within the islands.

Conjectures are fun! The above arguments should be sufficient to fully demonstrate that the circle map, which is well-known to exhibit phase locking regions called Arnold tongues, is topologically conjugate to the fattened beta shift $T_{\beta,\varepsilon}$. Or something like that. In a certain sense, this can be argued to be a "complete" solution, via topological conjugacy, of the tent map, the logistic map and the circle map. This is a worthwhile exercise to actually perform, i.e. to give explicit expressions mapping the various regions, as appropriate.

Essentially, the claim is straight-forward: topologically, all chaotic parts of a map correspond to folding (as per Milnor, 1980's on kneading maps), into which one may surgically insert regions that have cycles of finite length. The surgical insertion can occur only at the discontinuities of the kneading map. It almost sounds trivial, expressed this way; but the algebraic articulation of the idea would be worthwhile.

17 Miscellaneous unfinished ideas

An ad-hoc collection of half-finished thoughts.

17.1 Multiplicative Shifts

A multiplicative shift is a shift assembled as an product of functions. The most famous of these is the generating function for integer partitions

$$P(z) = \prod_{n=1}^{\infty} \frac{1}{(1-z^n)}$$

Similarly products occur for the necklace counting functions, most famously the cyclotomic identity

$$\frac{1}{1-\beta z} = \prod_{j=1}^{\infty} \left(\frac{1}{1-z^j} \right)^{M(\beta,j)}$$

where $M(\beta, j)$ the necklace polynomial.

A far more obscure product expresses the Minkowski measure[47], given as

$$?'(x) = \prod_{n=0}^{\infty} \frac{A' \circ A^n(x)}{2}$$

with

$$A(x) = \begin{cases} \frac{x}{1-x} & \text{for } 0 \le x < \frac{1}{2} \\ \frac{2x-1}{x} & \text{for } \frac{1}{2} \le x \le 1 \end{cases}$$

with A' being the derivative of A and A^n being the n'th iterate. The Minkowski measure integrates to the Minkowski Question Mark function $?(y) = \int_0^y ?'(x)\,dx$; it is the prototypical "multi-fractal measure" (although there really is nothing "multi-" about it; the "multi-" prefix stems from a misunderstanding of its multiplicative invariance). The product structure indicates that the Minkowski measure is a Gibbs measure, viz arising from an invariant Hamiltonian on a one-dimensional lattice.

The figure 28 suggests that a similar product can be constructed from the midpoint sequence, namely

$$\prod_{p=0}^{\infty} \frac{4m_p(\beta)}{\beta}$$

for the midpoints $m_p(\beta) = T_\beta^p(m_0)$.

17.2 Midpoints, revisited

The midpoints are defined above as $m_0 = \beta/2$, so that $m_p = T_\beta(m_{p-1}) = T_\beta^p(m_0)$ with $T_\beta(y)$ the beta shift map of eqn 4. Almost all literature uses the beta transform $t_\beta(x)$ of eqn 8 instead. The midpoint sequence and the iterate $t_\beta^p(1)$ are closely related:

$$2m_p \mod 1 = t_\beta^{p+1}(1)$$

Although related, they are not the same. The difference is a sequence of bits:

$$c_p = 2m_p - t_\beta^{p+1} \quad (1)$$

Note that $c_p \in \{0, 1\}$ always. Note that

$$\beta = \sum_{p=0}^{\infty} \frac{c_p}{\beta^p}$$

which is not entirely obvious!

17.3 Rauzy Fractals

Given a polynomial, one has an associated finite matrix, in Hessenberg form, that, iterated upon, generates a sequence. The projection of that sequence to a non-expanding orthogonal plane is a Rauzy fractal. What are the corresponding Rauzy fractals for this situation?

How about the general iterated sequence (e.g. the sequence of midpoints)? Is this space-filling, or not?

17.4 Fourier spectrum

What is the Fourier spectrum of the eigenfunctions?

18   Orthogonal generators

The monomials are used to generate a set of orthonormal eigenfunctions, having eigen-values β^{-n}. Since orthonormality is fixed with respect to the invariant measure, the monomials cannot be used in raw form; they must be orthogonalized. The procedure explored here is basic Gaussian elimination.

Preliminary note about the confusion about measures. The Jacobian is used in the integrals below. That is, the integrals are for

$$\int f(\rho(x))\,dx = \int f(y)\,\frac{dy}{\rho'(\rho^{-1}(y))}$$

after substitution $y = \rho(x)$, so that $\rho'\left(\rho^{-1}(y)\right) = v(y)$ is the Parry–Gelfond measure, and, as the Jacobian of the change of variable, it appears in the denominator, not the numerator. Basically, its the pushforward, not the pullback, so it puts the measure in the denominator.

A function f is normalized with respect to the Gelfond–Parry measure v if

$$\int_0^1 [f(y)]^2\,\frac{dy}{v(y)} = 1$$

Two functions f, g are orthogonal if

$$\int_0^1 f(y)g(y)\,\frac{dy}{v(y)} = 0$$

The Gelfond–Parry measure v is generated from $v_0 = 1$.

Let $f_n(x) = x^n$. Then

$$\int_0^1 [f_0(y)]^2\,\frac{dy}{v(y)} = \int_0^1 \frac{dy}{v(y)} = N_0$$

Likewise, define

$$\int_0^1 f_1(y)f_0(y)\,\frac{dy}{v(y)} = \int_0^1 y\frac{dy}{v(y)} = P_{01}$$

Define $g_1(x) = x - P_{01}/N_0$. It follows that
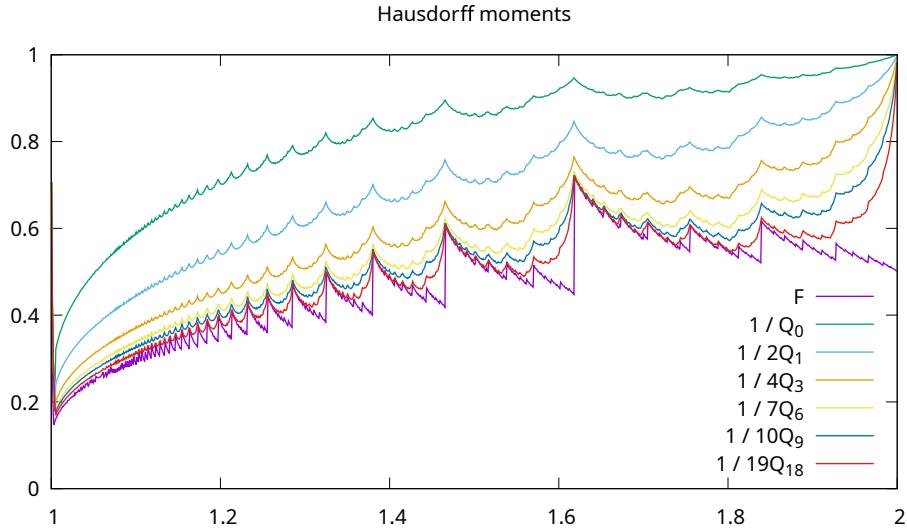
$$\int_0^1 g_1(y)f_0(y)\,\frac{dy}{v(y)} = 0$$

The normalization is obtained by setting

$$\int_0^1 [g_1(y)]^2\,\frac{dy}{v(y)} = N_1$$

and so on down the line, $g_n(x) = f_n(x) - P_{n-1,n}g_{n-1}(x)/N_{n-1}$ so that they are all orthogonal to one-another.

Figure 52: Hausdorff moments

The figure shows the (reciprocal of the) Hausdorff moments Q_n for the Gelfond–Parry measure, as a function of β on the horizontal axis.

18.1 Hausdorff moments

Define

$$Q_n = \int_0^1 y^n \frac{dy}{v(y)}$$

Numerical exploration shows that $nQ_n \to F$ in the $n \to \infty$ limit. Not obvious why this would be. This is shown in figure 52.

18.2 Hamburger moment problem

The Q_n are Hamburger moments, from the Hamburger moment problem. More precisely, this is the Hausdorff moment problem, since the measure is on the unit interval, instead of the whole real-number line.

Given the sequence $F_m^{\{b\}}$ of beta-Fibonacci numbers, there is a corresponding Hankel matrix. The corresponding measure is presumably the Gelfond–Parry measure. TODO: turn the crank and verify that this is so.

The relationship between Hankel matrices, shift operators and Hamburger moments (Hausdorff moments) is well-known. What seems to be missing is an articulation for the beta shift.

Also: See chapter on the Jacobi operator, much later in this text.

19 Midpoint Moments

The derivative of the Gelfond–Parry measure is

$$m_\beta(y) = \sum_{k=0}^{\infty} \frac{\delta(t_k - y)}{\beta^k}$$

where with $\delta(x)$ the Dirac delta, and t_k are the iterated endpoints, $t_k \equiv t_\beta^k(1) = \frac{2}{\beta} T_\beta^k\left(\frac{\beta}{2}\right) = \frac{2}{\beta} m_k$ with $t_\beta(x) = \beta x \mod 1$. It is particularly easy to compute the moments for this function, as the integral becomes a sum:

$$M_n(\beta) = \int_0^1 y^n m_\beta(y) \, dy$$

The zeroth moment is

$$M_0 = \int_0^1 \sum_{k=0}^{\infty} \frac{\delta(t_k - y)}{\beta^k} dy = \sum_{k=0}^{\infty} \frac{1}{\beta^k} = \frac{\beta}{\beta - 1}$$

The first moment is the Gelfond–Parry normalization

$$M_1 = \int_0^1 y \sum_{k=0}^{\infty} \frac{\delta(t_k - y)}{\beta^k} dy = \sum_{k=0}^{\infty} \frac{t_k}{\beta^k} = F$$

In general

$$M_n(\beta) = \int_0^1 y^n \sum_{k=0}^{\infty} \frac{\delta(t_k - y)}{\beta^k} dy = \sum_{n=0}^{\infty} \frac{t_k^n}{\beta^k}$$

Numerics shows that $M_n(1) = n/(n-1)$. Not obvious why this would be. The $M_n(\beta)$ appear to all be discontinuous everywhere, as a function of β.

As $n \to \infty$ one has $t_k^n \to 0$ since $t_k < 1$ for all $k > 0$, except $t_0 = 1$. Conclude that $M_n(\beta) \to 1$ as $n \to \infty$.

19.1 Orthogonal polys

What are the orthogonal polynomials for this measure? These are the polynomials

$$p_n(x) = \sum_{j=0}^{n} a_{nj} x^j$$

with

$$\int_0^1 p_i(x) p_j(x) m_\beta(x) \, dx = \delta_{ij}$$

The first few can be done by hand: Clearly, $p_0(x) = 1/\sqrt{M_0}$ and $p_1(x) = (x - M_1)/\sqrt{N_1}$ with normalization $N_1 = M_2^2 - 2M_1^2 + M_1^2 M_0$.

Writing $a_{nj} = [A]_{nj} = A_{nj}$ for a matrix A, the matrix is upper-triangular, solvable. The polynomials are $p_n = A_{nj} x^j$ and from orthogonality, $\int p_i p_j = \delta_{ij} = A_{ik} A_{jm} M_{k+m}$ it

follows that $I = AMA^T$ where the moments are organized into a Hankel matrix $[M]_{nm} = M_{m+n}$. Since A is solvable, $[A^{-1}]_{nj} = A_{jm}M_{n+m} = [MA^T]_{nj}$ or more directly, $A^{-1} = MA^T$.

Numerically, A is quite nasty, with large entries of alternating sign in each row. Rows alternate sign, too, so that the diagonal is always positive, and rapidly increasing. Seems like the diagonal increases faster than factorial. That is, $a_{nn} \gtrsim \Gamma(n)$.

Numerically, A^{-1} is quite nice. Apparently, the alternating signs result in a nice cancellation, since the entries of M tend to 1 quite rapidly.

The matrix $O = A\sqrt{M}$ is orthogonal: $OO^T = I$. How can $\sqrt{M}$ be computed?

20 Bergman (Hessenberg) Polynomials

Given a matrix operator in Hessenberg form, it can be interpreted as a right-shift on the space of polynomials. Such polynomials form an orthogonal basis for a certain kind of Hilbert space, called the Bergman space. They are studied in applied mathematics, as they are orthogonal over some measure on the complex plane. The Hessenberg operator is a generalization of the better-known case of the Jacobi operator, which has it's own extensive theory, including spectra and scattering, and is important for several exactly solvable non-linear models in physics, including the Toda lattice[48]. The Hessenberg operator presumably has an equally rich theory, but it does not appear to be currently known; the breadth and scope of existing publications is limited.

The general framework for the Hessenberg polynomials is sketched below, including a fast and informal definition of Bergman space (the space on which the polynomials are orthogonal). The Hessenberg matrix is explicitly solvable on the left, and can be explicitly brought into a from that exhibits the right-shift operator. In the general theory, the change of basis from the shift operator to the Hessenberg matrix is known to be the Cholesky decomposition of a moment matrix, and specifically, the moments of the measure on which the polynomials are orthogonal.

There are two Hessenberg operators in this text: the operator $\mathscr{L}_\beta$ in the wavelet basis, and the operator $\mathscr{H}_\beta$ generated from the midpoint orbits. The second is already obviously a shift, and so everything below follows "trivially" from it. The first form is is numerically and analytically difficult. Needless to say, the section below treats the first rather than the second. XXX TODO this should be fixed, as $\mathscr{H}_\beta$ is both simpler and more enlightening overall. Later ...

Working backwards from the beta shift, the first asymptotic term in the measure can be extracted. For $\beta > \varphi$, it appears to be a Dirac delta (point mass) located at $z = 1$ on the complex plane, with a blancmange-like fractal curve giving the weight. For $\beta < \varphi$, it appears to be the derivative of the Dirac delta, with a different blancmange-like fractal curve giving the weight.

20.1 Bergman Space

Given a matrix operator in Hessenberg form, it can be interpreted as a right-shift on the space of polynomials. That is, given an unreduced Hessenberg matrix with matrix entries A_{ij}, one can write a recurrence relation that defines a sequence of polynomials as

$$z p_n(z) = \sum_{k=0}^{n+1} A_{kn} p_k(z) \tag{49}$$

with $p_0(z) = 1$. This relation is easily solvable in closed form, as the recurrence relation terminates in a finite number of steps.

One important property of these polynomials is that the zeros of $p_n(z)$ correspond to the eigenvalues of the $n \times n$ principle submatrix of A. Numeric exploration of these polynomials confirms the previous results on eigenvalues obtained from direct diagonalization: the zeros of the $p_n(z)$ seems to lie mostly near the circle of radius $1/\beta$, distributed uniformly over all angles.

If all of the sub-diagonal entries obey $A_{n+1,n} > 0$, then the polynomials form an orthonormal basis for Bergman space. That is, there exists a domain in the complex plane on which the polynomials provide a basis for a Hilbert space of holomorphic functions on that domain[49, 50, 51]. That is, one has the orthogonality relation

$$\delta_{mn} = \int_D p_m(\bar{z}) \, p_n(z) \, d\mu(z)$$

for some domain $D \subset \mathbb{C}$ of the complex plane, and some (Borel) measure $d\mu$ on that domain.

The matrix A can be interpreted as an operator with a continuous spectrum. To do this, fix a certain, specific value of $z = c$ a constant, and then notice that $\vec{p} = (p_n(z))_{n=0}^{\infty}$ is a vector having the property that $A^T \vec{p} = z\vec{p}$. That is, $\vec{p}$ is a left-eigenvector of A; equivalently, a right-eigenvector of its transpose A^T. Clearly, the spectrum is continuous on the domain D.

The matrix operator A can also be interpreted as a right-shift on Bergman space. To do this, define

$$\mathscr{A}(w,z) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} p_k(w) A_{kn} p_n(\bar{z})$$

Then, given some holomorphic function $f(z)$ decomposed in terms of the polynomials, so that $f(z) = \sum_n a_n p_n(z)$, one has that

$$\begin{aligned} [\mathscr{A}f](w) &= \int \mathscr{A}(w,z) f(z) \, d\mu(z) \\ &= \sum_k \sum_n p_k(w) A_{kn} a_n \\ &= w \sum_n a_n p_n(w) \\ &= wf(w) \end{aligned}$$

That is, given a sequence $(a_0, a_1, a_2, \cdots)$, the Hessenberg matrix acts as a right-shift, mapping it to the sequence $(0, a_0, a_1, \cdots)$.
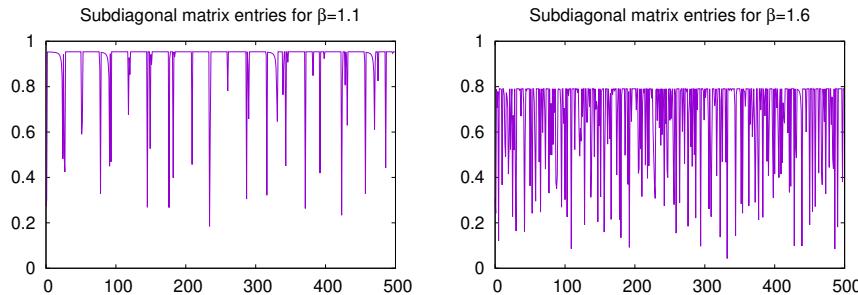
This is perhaps a bit silly, as one could instead just perform the same manipulation without the $f(z)$, by observing that, formally,

$$\mathscr{A}(w,z) = w \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} p_k(w) \, p_n(\bar{z})$$

The above treatment is breezy and "formal", paying no heed to summability, convergence or responding to any questions about what spaces the various vectors may live in. This is as appropriate, since the task here is to discover which spaces are the appropriate ones, when the Hessenberg matrix arises from the beta shift.

Notice that the word "operator" is a bit mis-used, here, as a vague synonym for "infinite-dimensional matrix". Properly, the word "operator" should be reserved for an infinite-dimensional matrix acting on some given space, having general properties that are independent of the basis chosen for that space. So far, that might not be the case here: the infinite-dimensional matrices here might not be bounded operators; they

Figure 53: Sub-diagonal Entries

These charts show the sub-diagonal matrix entries $\langle n+1 | \mathscr{L}_\beta | n \rangle$ for the first $n < 500$. The left graph shows $\beta = 1.1$, the right shows $\beta = 1.6$; other values behave similarly. A scatterplot of the location of the spikes as a function of β does not reveal any structure. That is, except for small n, the location of a spike shows no smooth variation as β is varied smoothly. There does appear to be some structure for small n – some banded sequences – and so perhaps the correct statement is that the system is mixing, as n increases.

might not even be continuous, viz. we have not ruled out the possibility that the space of interest is some Fréchet space or some more general topological vector space. It is well known that operators on such spaces can have "unexpected" discontinuities, unexpected in that they are not seen in ordinary Banach spaces.

At any rate, if polynomials obtained from the beta shift are orthogonal on some domain $D \subset \mathbb{C}$ that is the support of some measure $d\mu$, it is not at all clear what this measure might be. They are certainly not orthogonal on the unit disk, with uniform measure.

Notice also that the above treatment seems to be a special case of a more general principle: when an operator has a continuous spectrum, it can sometimes be interpreted as a right-shift. That is, given some arbitrary operator $\mathscr{H}$, then if one has that $\mathscr{H} f = \lambda f$ and λ takes arbitrary values $\lambda \in D \subset \mathbb{C}$, then $\mathscr{H}$ can be taken to be a right-shift operator, provided that $f = f(\lambda)$ can be decomposed into a set of orthogonal polynomials in λ.

20.2 Beta Bergman Shift

The primary question for this section is whether the β-transform transfer operator, in the Hessenberg basis, can be considered to be a Bergman shift.

To obtain the orthogonal polynomial basis, one must satisfy the constraint that $A_{n+1,n} > 0$ for the matrix elements $A_{kn} = \langle k | \mathscr{L}_\beta | n \rangle$ of eqn 29. Numeric exploration indicates that this is indeed the case, with the sub-diagonal entries all positive (none are zero), and all tend to have the same value, with sporadic exceptions. These are shown in figure 53.

Can one find a domain on the complex plane that would have such Bergman polynomials? The references[49, 51] provide a technique for doing so, provided that the matrix is asymptotically Toeplitz. That is, if the diagonals of A_{ij} have nice limits, that $\lim_{n\to\infty} A_{n-k,n}$ exists for fixed k, then a Jordan arc bounding a domain on the complex plane can be found. The figure 53 indicates that this limit does not exist, in the strict sense: the values bounce away from an obvious limit point indefinitely. Exactly what this implies is unclear. Perhaps it is possible to extend the results of [49, 51] to matrices that are where the diagonals merely have an accumulation point, as opposed to a well-defined limit?

Based on numeric exploration, it appears that the domain is the unit disk. That is, $A^T \vec{p} = z\vec{p}$ holds for $|z| \leq 1$.

20.3   Bergman Alternative

The Bergman polynomials of eqn 49 define an orthonormal basis for some region of the complex plane. For the square-integrable norm, this basis is the basis of a Hilbert space, and specifically, that of a reproducing kernel Hilbert space.

Yet, something funny happens on the unit disk. Let $p_m(z)$ be the polynomials, and for some sequence of coefficients $\{a_n\}$, consider a generic function

$$f(z) = \sum_{k=0}^{\infty} a_k p_k(z)$$

Consider the case where the $\{a_n\}$ are a right-eigenvector of the Hessenberg operator, that is, where

$$\sum_{m=0}^{\infty} A_{km} a_m = \lambda a_k$$

Substituting into the above, one has

$$f(z) = \sum_{k=0}^{\infty} \frac{1}{\lambda} \sum_{m=0}^{\infty} A_{km} a_m p_k(z) = \frac{z}{\lambda} \sum_{m=0}^{\infty} a_m p_m(z) = \frac{zf(z)}{\lambda}$$

There are two alternatives to solving this; either $f(z) = 0$ or $z = \lambda$. Since this is a reproducing kernel Hilbert space, then if $z = \lambda$ is part of the domain of the Bergman space, then one must conclude that $f(z) = 0$ everywhere. That is, right-eigenvalues of A correspond to functions $f(z)$ that are vanishing. To invent a new name, by analogy to the Fredholm alternative, perhaps this can be called the Bergman alternative.

Numerical exploration indicates that, for the matrix elements of eqn, 29, the function $f(z)$ vanishes inside the unit disk $|z| < 1$, and is undefined (infinite) outside of it.

20.4   Left Factorization

XXX TODO: somewhere just above, or maybe only below, I accidentally swapped left and right. Thus, almost everything that is written is correct, except that some of the transposes are in the wrong places. This needs review and correction.

Suppose one is given an (arbitrary) sequence of polynomials $(p_n(z))_{n=0}^{\infty}$, such that the order of p_n is n. Then each individual polynomial can be expanded as or $\beta > \varphi$,

$$p_n(z) = \sum_{k=0}^{n} p_{nk} z^k$$

This defines an infinite matrix $\mathscr{P} = [p_{nk}]$, provided that the coefficients are extended so that $p_{nk} = 0$ whenever $k > n$. This matrix is manifestly lower-triangular. Writing vectors $\vec{z} = (z^n)_{n=0}^{\infty}$ and $\vec{p} = (p_n(z))_{n=0}^{\infty}$ as before, the above is just the matrix equation

$$\vec{p} = \mathscr{P}\vec{z}$$

Consider now the case where the polynomials were constructed from some irreducible Hessenberg matrix A. The earlier observation that A^T is a shift, namely, that $A^T \vec{p} = z\vec{p}$ can now be written as

$$A^T \mathscr{P}\vec{z} = z\mathscr{P}\vec{z} = \mathscr{P}z\vec{z} = \mathscr{P}\mathscr{K}\vec{z}$$

In the above, the z without the vector notation is just a scalar, and thus commutes (trivially) with $\mathscr{P}$. Its eliminated by explicitly making use of the right-shift (Koopman) operator, which, in this basis, is

$$
\mathscr{K} =
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & \ddots \\
0 & 0 & 0 & 0 & \ddots
\end{bmatrix}
$$

Since $\mathscr{P}$ is lower-triangular, it is invertible on the right, that is, the inverse $\mathscr{P}^{-1}$ exists, and so one is left with

$$\mathscr{P}^{-1}A^T \mathscr{P} = \mathscr{K}$$

The irreducibility of A is important, here; nonzero entries on the sub-diagonal are required, else trouble ensues.

Rearranging, this provides an explicit decomposition of A into triangular matrices:

$$A^T = \mathscr{P}\mathscr{K}\mathscr{P}^{-1}$$

Taking the transpose, this gives

$$A = \left[\mathscr{P}^{-1}\right]^T \mathscr{K}^T \mathscr{P}^T$$

with $\mathscr{P}^T$ and $\left[\mathscr{P}^{-1}\right]^T$ both being upper-triangular, and $\mathscr{K}^T$ being the left-shift.

This system is solvable. Given some matrix A in Hessenberg form, the matrix elements of $\mathscr{P}$ can be computed recursively, in a finite number of steps (i.e. in closed form), directly from 49. The explicit expression is

$$A_{n+1,n}p_{n+1,j} = p_{n,j-1} - \sum_{k=0}^{n} A_{kn}p_{kj}$$

The starting conditions are $p_{00} = 1$. To handle the $j = 0$ case in the above, set $p_{n,-1} = 0$.

Because $\mathscr{P}$ is lower triangular, its inverse $\mathscr{P}^{-1} \equiv \mathscr{R} = [r_{kn}]$ can be obtained explicitly. Along the diagonal, one has $r_{nn} = 1/p_{nn}$ while the lower triangular form means $r_{kn} = 0$ for $k < n$. For the remaining entries $m < n$, one has

$$0 = \sum_{k=m}^{n} p_{nk}r_{km}$$

This can be solved in a finite number of iterations on

$$p_{nn}r_{nm} = -\sum_{k=m}^{n-1} p_{nk}r_{km}$$

The above avoids questions of convergence, or any notion of the spaces on which the matrices or operators might act. The norm to be used for $\vec{z}$ and $\vec{p}$ is not specified. This is appropriate at this stage: it is the algebraic manipulations that are interesting, at this point, rather than the spaces on which the matrices/operators might act. One can invent several kinds of norms that might be applicable, but there is no particular reason to believe that $\vec{p}$ might have a finite norm. Likewise, $\mathscr{P}$ may not have a finite norm. For the case of the Hessenberg operator originating with the beta shift operator, it does not; the individual matrix elements p_{nm} increase without bound. That is, $\mathscr{P}$ is an infinite matrix, but it is not clear that it is also an operator. If it is, it is certainly not a compact operator.

Some of the poor behavior can be brought under control by factoring $\mathscr{P} = \mathscr{D}\mathscr{N}$ with $\mathscr{N}$ being unitriangular (all ones on the diagonal) and $\mathscr{D}$ a diagonal matrix, with entries $[\mathscr{D}]_{nk} = p_{nn}\delta_{nk}$. With this factorization, one may then write

$$\mathscr{N}^{-1}A^{T}\mathscr{N} = \mathscr{D}\mathscr{K}\mathscr{D}^{-1}$$

so that $\mathscr{D}\mathscr{K}\mathscr{D}^{-1}$ has off-diagonal matrix entries $\left[\mathscr{D}\mathscr{K}\mathscr{D}^{-1}\right]_{nk} = \delta_{n+1,k}p_{nn}/p_{kk}$. This is a rescaling of the shift $[\mathscr{K}]_{nk} = \delta_{n+1,k}$. The scaling factor is exactly the sub-diagonal of the Hessenberg. That is, $p_{nn}/p_{n+1,n+1} = A_{n+1,n}$. The polynomials $\mathscr{N}\vec{z}$ are monic.
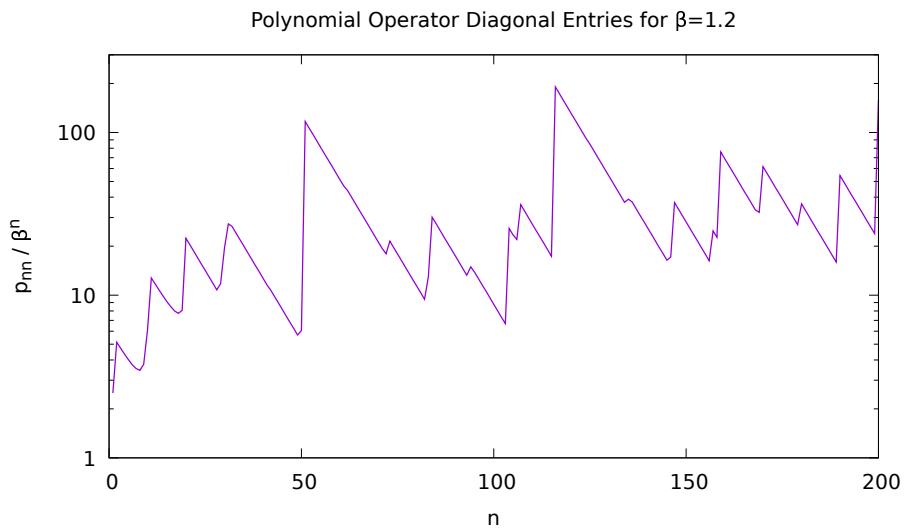
20.5 Beta-transform factoids

An assortment of observations follow, for the case of the beta shift.

First, the matrix entries of $\mathscr{P}$ grow in an unbounded fashion. It appears that $p_{nn} \sim \mathscr{O}(\beta^{n})$; the ratio p_{nn}/β^{n} is depicted in figure 54.

Experimentation reveals two different regimes of behavior, depending on whether or not $\beta < \varphi = \left(1 + \sqrt{5}\right)/2$ the Golden ratio. Exactly why there are two different regimes is unclear. Earlier sections motivated the reason for the appearance of the golden mean; why this shows up dramatically, as it does here, is unclear (to me).

Figure 54: Polynomial Operator Diagonal Entries

This depicts the ratio p_{nn}/β^n of the diagonal matrix entries p_{nn} of the Bergman polynomial matrix operator $\mathscr{P}$ for the beta shift with value $\beta = 1.2$. Other values of β are not dissimilar, although the spikes are pushed more closely together. The height of the spikes seems to be roughly the same, for all β. This is another way of visualizing the same information as in figure 53, as the ratio $p_{nn}/p_{n+1,n+1}$ is just given by the subdiagonal entries $A_{n+1,n}$ of the Hessenberg matrix. In particular, the straight edges correspond to usually-constant values on the subdiagonal.

One such result is that when $\beta < \varphi$, then the sum over columns of the Bergman operator vanishes. That is,

$$\sum_{k=0}^{\infty} p_{nk} = \delta_{n0}$$

This implies that every polynomial $p_n(z)$ has a zero at $z = 1$ (except for $p_0(z) = 1$) when $\beta < \varphi$.

20.6 Decaying Eigenfunctions

The matrix mechanics developed in the previous sections can be used to perform asymptotic expansions that rapidly converge to decaying eigenfunctions. This works most simply for the case of $\varphi < \beta$. TODO Write these down. TODO flesh out. Basically, write a vector $\vec{w}$ with elements $w_n = \omega^n$ for $1 < |\omega|$ so that this is divergent. Then write the formal vector $\vec{a} = \left[\mathscr{P}^T\right]^{-1} \vec{w}$ which is formally divergent, but can be truncated in finite dimensions, and renormalized to be of unit length. Doing so provides an eigenfunction of A. The associated eigenvalue is 1 when $\beta < \varphi$ but is less than 1 when $\varphi < \beta$ (and in fact, the eigenvalue is exactly that depicted in figure 55). TODO graph some of these, explore more thoroughly, address the issues of formal divergence.

20.7 Moment Matrix

When the Hessenberg matrix is derived from measures on the complex plane, it takes the form of $\mathscr{M} = \mathscr{R}\mathscr{R}^T$ with $\mathscr{R} = \mathscr{P}^{-1}$, so that $\mathscr{R}$ is the Cholesky decomposition of $\mathscr{M}$. This matrix is manifestly symmetric: $\mathscr{M} = \mathscr{M}^T$. Direct observation shows that it is almost positive-definite: one finds that $[\mathscr{M}]_{ij} > 0$ for all i, j except for $[\mathscr{M}]_{00} = 0$. This result can be strengthened: when $\beta < \varphi$, then $[\mathscr{M}]_{ij} > 1$ for all i, j except for $[\mathscr{M}]_{00} = 0$ and $[\mathscr{M}]_{0n} = [\mathscr{M}]_{n0} = 1$. But, for $\beta > \varphi$, one finds that $[\mathscr{M}]_{00} = 0$ and $[\mathscr{M}]_{01} = [\mathscr{M}]_{10} = [\mathscr{M}]_{11} = 1$, while all the rest obey $0 < [\mathscr{M}]_{ij} < 1$.
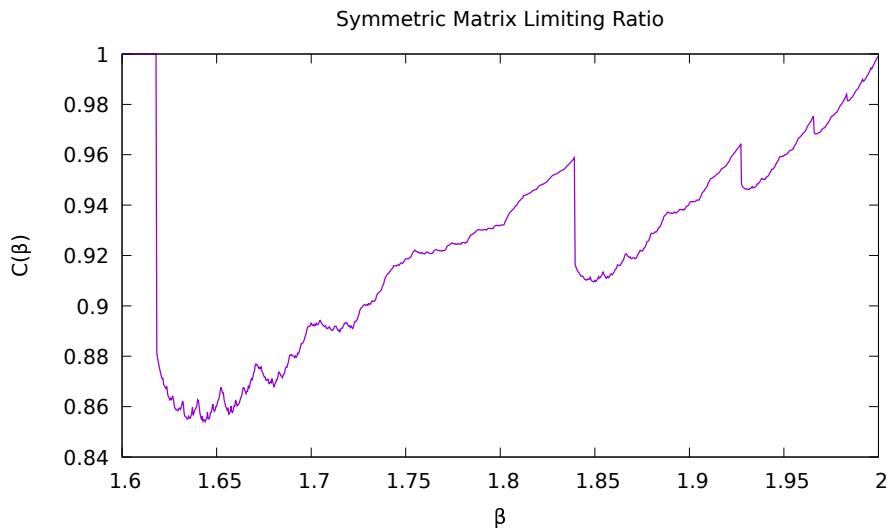
In the standard literature, $\mathscr{M}$ is usually obtained from some moment matrix, viz, for the integral $\int \bar{z}^m z^n d\mu(z)$ for some measure $d\mu(z)$. Might that be the case here? Taking the time to numerically characterize the matrix, one finds that the ratio of successive rows (or columns as its symmetric) very quickly approaches a limit $\lim_{n\to\infty,} [\mathscr{M}]_{nm} / [\mathscr{M}]_{n-1,m} = C(\beta)$ for some constant C that depends only on β but not on m. The limit $C(\beta)$ is graphed in figure 55.

For $\beta < \varphi$, it appears that $\lim_{n\to\infty,} [\mathscr{M}]_{nm} = B(\beta)$ a constant, independent of m. This limiting value $B(\beta)$ is graphed in figure 56.

The asymptotic behavior of the matrix $[\mathscr{M}]_{ij}$ can be obtained as a moment matrix on point sources. A delta function located at $z = C$ for real C has the moments
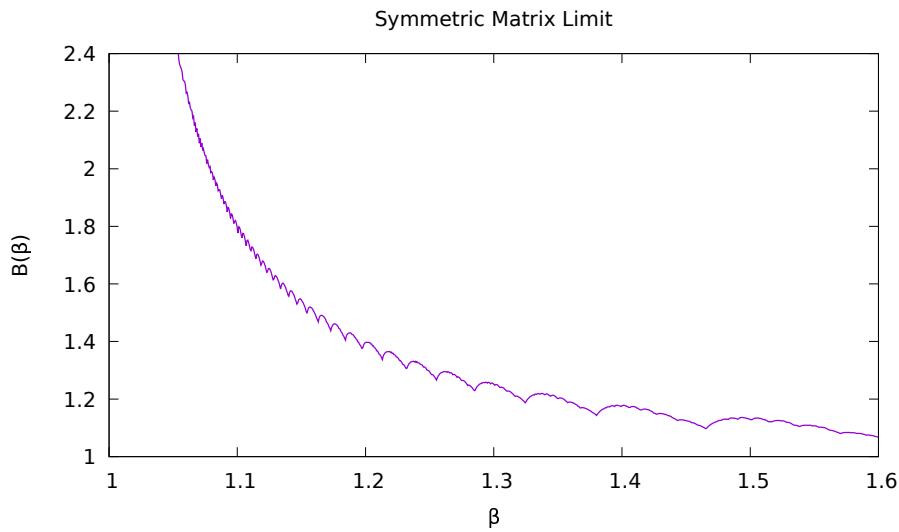
$$C_{mn} = \int \bar{z}^m z^n \delta(z - C) dz$$
$$= \int r^m r^n \delta(r - C) r dr \int \delta(\theta) e^{-im\theta} e^{in\theta} d\theta$$
$$= C^{m+n+1}$$

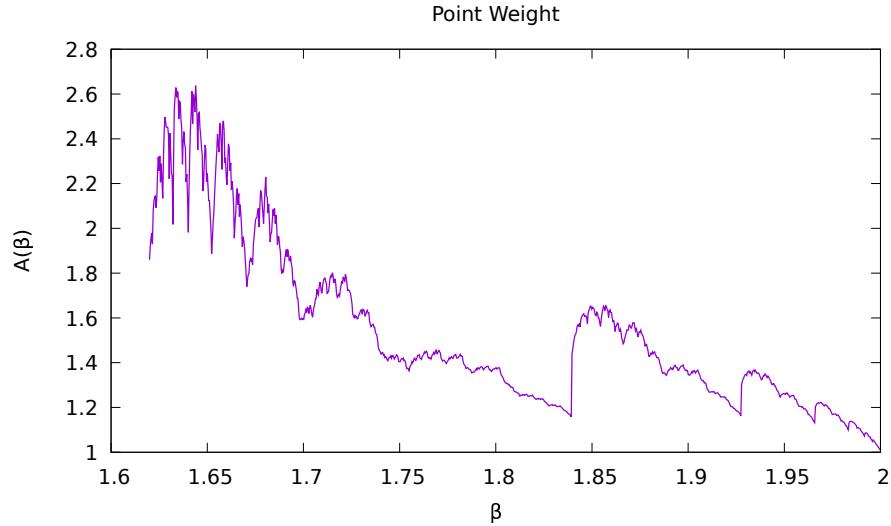Figure 55: Symmetric Matrix Limit Ratio

This figure shows the limit $C(\beta)$ defined in the text. Note that $C(\beta) = 1$ for $\beta < \varphi$. The jump is at about $\beta = 1.83928676\cdots$. Note this is one of the "troublesome midpoints" for the Hessenberg basis expansion, specifically for $T_\beta^3(\beta/2) = 0$ or $\beta/2$. This is one of the first "generalized golden means", the positive real root of $\beta^3 - \beta^2 - \beta - 1 = 0$. The entire fractal structure presumably corresponds to higher iterates p that satisfy $T_\beta^p(\beta/2) = 0$.

Figure 56: Symmetric Matrix Limit

This figure shows the limit $B(\beta)$ defined in the text. The limit is approached fairly quickly for the larger values of β, but convergence proves difficult for $\beta \lesssim 1.1$. The overall shape is that of a hyperbola, but doesn't seem to actually be hyperbolic for either small or large β. The right-most nick in the curve appears to be at $\beta = 1.465571231876768\cdots$, another "generalized golden mean", and the only real root of $\beta^3 - \beta^2 - 1 = 0$; equivalently, the root of $T_\beta^3(\beta/2) = 0$. The remaining nicks are presumably located at $T_\beta^p(\beta/2) = 0$ for higher iterates p.

Figure 57: Point Weight

Point Weight

This figure shows the value of $A(\beta)$ that gives the point weight of the moment matrix. That is, the asymptotic behavior of $\mathcal{M}$ is given by $[\mathcal{M}]_{mn} \to \int \bar{z}^m z^n \rho(z)\, dz$ with the measure given by a point mass $\rho(z) = A(\beta)\delta(z - C(\beta))$. Clearly, there is a strong resemblance to figure 55.

Thus, for $\varphi < \beta$, the asymptotic behavior of $[\mathcal{M}]_{ij}$ is given by the distribution $A(\beta)\delta(z - C(\beta))$. What is $A(\beta)$? This is graphed in figure 57.

What about $\beta < \varphi$? A limiting constant distribution can be obtained from a derivative point mass located at $z = 1$. That is,

$$D_{mn} = \int \bar{z}^m z^n \delta'(z - 1)\, dz$$
$$= \int r^m r^n \delta'(r - 1)\, r\, dr \int \delta(\theta) e^{-im\theta} e^{in\theta}\, d\theta$$
$$= 1$$

so that the asymptotic behavior of $[\mathcal{M}]_{ij}$ for $\beta < \varphi$ is given by the distribution $B(\beta)\delta'(z - 1)$. The prime superscript here means derivative, viz, in colloquial language, $\delta'(z) = d\delta(z)/dz$.

21 The Jacobi Operator

Given a Borel measure on the real number line, one can find a sequence of polynomials that are orthonormal with respect to that measure. These polynomials $p_n(x)$ are coupled together by a three-term recurrence equation

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_n p_n(x) + a_n p_{n-1}(x)$$

with $p_0(x) = 1$ and $p_{-1}(x) = 0$. This recurrence relation can be taken to be an operator, known as the Jacobi operator $\mathcal{J}$, acting on vectors consisting of the polynomials $p(x) = \{p_n(x)\}$ so that

$$[\mathcal{J} p](x) = xp(x)$$

so that p is an eigenvector of $\mathcal{J}$ with eigenvalue x. The two sequences of coefficients $\{a_n\}$ and $\{b_n\}$ form three diagonals of the operator, with $\{a_n\}$ running down the center, and $\{b_n\}$ the two diagonals on either side[48].

Given that the invariant measure for the β-transform, given by eqn 18 and visualized in figure 1 is a Borel measure, it seems reasonable to ask: what is the corresponding Jacobi operator? How can the sequence of polynomials be understood?

Szegő polynomials w.r.t. $d\mu$ are a set of orthogonal polynomials on the unit circle. Applying a Cayley transform gives the Schur functions, obeying a rational recurrence relation solvable via continued fractions. Hmmm.

And then there is Favard's theorem...

21.1 Moments

Construction of the polynomial sequences require moments. Since the invariant measures (and all of the eigenfunctions) are linear combinations of the Hessenberg basis functions, it suffices to compute the moments for these. Since the basis functions are piece-wise constant, and have an explicit expression given by eqn 28, the moments can also be given explicit expression:

$$\int_0^1 x^{n-1} \psi_p(x)\,dx = \frac{C_p}{n}\left[\frac{m_p^n - m_l^n}{m_p - m_l} - \frac{m_u^n - m_p^n}{m_u - m_p}\right]$$

with the midpoint m_p and the lower and upper midpoints $m_l < m_p < m_u$ defined just as before. Clearly, the moments rapidly get small as $n \to \infty$. Likewise, for fixed n, these also rapidly get small as $p \to \infty$.

22 The Multiplication Operator

The difficulties presented in the previous section suggests that studying the multiplication operator might be simpler. Multiplication by β is given by

$$M_\beta (x) = \beta x \tag{50}$$

The corresponding transfer operator is

$$[\mathscr{M}_\beta f](y) = \frac{1}{\beta} f\left(\frac{y}{\beta}\right)$$

The multiplication operator, superficially, in itself, is not terribly interesting; it simply rescales things. It does not generate fractals, at least, not if one confines oneself to real numbers and the canonical topology on the real-number line. If instead one works with the product topology on 2^ω, then the multiplication operator becomes rather complicated and difficult to analyze. In this sense, it is promising: it avoids the overt complexity of the logistic map, the tent map and the beta shift, yet still has a complicated behavior in the product topology. In particular, the multiplication of two numbers appear to involve chaotic dynamics of the carry bit.

22.1 Beta-shift, Revisited

The beta shift of eqn 4 takes a simple form when reinterpreted on bit-strings: it is the concatenation of multiplication, followed by a left-shift. Given a bit-string $(b_n) = 0.b_0 b_1 b_2 \cdots$ denote its left-shift by U given by

$$U(0.b_0 b_1 b_2 \cdots) = 0.b_1 b_2 \cdots$$

which, for real numbers, corresponds to

$$U(x) = \begin{cases} 2x & \text{for } 0 \le x < \frac{1}{2} \\ 2x - 1 & \text{for } \frac{1}{2} \le x \le 1 \end{cases}$$

which is none-other than the Bernoulli shift of eqn 1 with a change of notation. The beta shift is then

$$T_\beta (x) = M_\beta (U(x))$$

so that the iterated beta shift is an alternation between a left-shift and a multiplication. The act of discarding the most significant bit (the MSB) with each left-shift alters the dynamics of iterated multiplication.

This suggests that studying multiplication and the multiplication operator might provide fruitful insight into the beta shift.

22.2 Monomial Eigenfunctions

Some properties of the multiplication operator can be guessed at directly. Obviously, $f = \text{const.}$ is a decaying/growing eigenfunction, depending on whether $\beta > 1$ or not.

That is, one should imagine $f = \text{const.}$ as a uniform distribution of dust; with each iteration, it is spread either farther apart ($\beta > 1$) or bunched closer together ($\beta < 1$).

Clearly, $f(x) = x^n$ is an eigenfunction, with eigenvalue $1/\beta^{n+1}$. If one considers multiplication only to operate on the positive real-number line, then n need not be an integer. In other words, the multiplication operator has a continuous spectrum in this situation.

If the domain of the operator is extended to functions on the non-negative real-number line, then n must be positive, as otherwise $f(0)$ is ill-defined. But if n is positive, then (for $\beta < 1$) the multiplication operator only has eigenvalues greater than one, which is not, in general, very desirable.

If the domain of the multiplication operator is extended to the entire real-number line, then n is forced to be an integer, in order to avoid issues due to multi-valued functions. Extending the domain to the complex plane leads us astray, and so we will not go there.

22.3 A Fractal Eigenfunction

The compressor function is also an eigenfunction. It was previously observed in eqn 15 that

$$\text{cpr}_\beta \left(\frac{x}{\beta} \right) = \frac{1}{2} \text{cpr}_\beta (x)$$

whenever $1 < \beta \le 2$ and $0 \le x < 1$ and so, cpr_β is potentially be an eigenfunction of $\mathcal{M}_\beta$ with eigenvalue $1/2\beta$, provided that it is extended to arguments $1 < x$. This can be done as follows. Define the extended function, valid for $0 \le x < \infty$ and for $1 < \beta \le 2$ as

$$\text{ecpr}_\beta (x) = \begin{cases} \text{cpr}_\beta (x) & \text{if } 0 \le 2x < \beta \\ 2\text{cpr}_\beta \left(\frac{x}{\beta} \right) & \text{if } \beta \le 2x < \beta^2 \\ 4\text{cpr}_\beta \left(\frac{x}{\beta^2} \right) & \text{if } \beta^2 \le 2x < \beta^3 \\ 2^n \text{cpr}_\beta \left(\frac{x}{\beta^n} \right) & \text{if } \beta^n \le 2x < \beta^{n+1} \end{cases}$$

The extension is performed simply by treating the self-similarity as a recurrence relation, which can be iterated to move the argument into a region where the original definition was sufficient. In essence, one applies a right-shift operator to reduce the argument. Since the multiplication operator is odd about $x = 0$, on can trivially extend this to negative x by defining $\text{ecpr}_\beta (-x) = -\text{ecpr}_\beta (x)$.

Note that the original $\text{cpr}_\beta (x)$ also had a translation symmetry: the upper half was equal to the lower half. This translation symmetry has been lost, since after all, multiplication does not preserve translation.

The ecpr function is not square integrable; it does not have an L_p-norm for any p; and this is no surprise, as its hard to imagine how it could be otherwise, for a function to be self-similar under scaling.
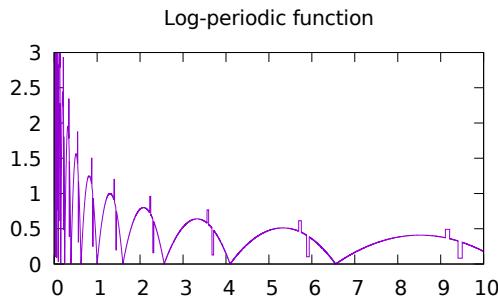
22.4 A Generic log-periodic Eigenfunction

Inspired by the above, its should be clear how to build a generic eigenfunction. Let $g(x)$ be some arbitrary function, defined on the interval $1 \le x < \beta$ (given some fixed $1 < \beta$). Define its extension as

$$g'_w(x) = w^n g\left(\frac{x}{\beta^n}\right) \quad \text{if } \beta^n \le x < \beta^{n+1}$$

This has, by construction, the self-similarity relation $g'_w(\beta x) = w g'_w(x)$ and so is an eigenfunction with eigenvalue w/β:

$$\left[\mathcal{M}_\beta g'_w\right] = \frac{w}{\beta} g'_w$$

This function is merely log-periodic; its not fractal. Perhaps its silly to illustrate this; it should be obvious, but just in case its not, the figure below shows such a function, for $\beta = 1.6$ and $w = 0.8$. It is an eigenfunction of $\mathcal{M}_{1.6}$ with eigenvalue of 1/2.

Log-periodic function

There doesn't seem to be anything particularly interesting with this particular game. There's a simple explanation for this: The multiplication operator is generating a free monoid in one generator (the iteration itself), whereas fractals require at least two generators of self-symmetry. The (usually) free interaction of multiple generators is what forces the fractal to appear.

Note that the cpr_β function constructed above is a special case of this: It's self-similar, but the property that made it interesting, as a fractal, was erased in the construction. As before, note that $g'_w(x^n)$ is an eigenfunction with eigenvalue $1/\beta w^n$ (for integer n).

22.5 Haar Basis Matrix Elements

The Haar basis matrix elements for the beta shift proved to be a bit unwieldy and not terribly useful. The corresponding matrix elements for the multiplication operator have the same general essence, but are slightly simpler and shorter to write down. In all other respects, they still have the same tractability issues.

The multiplication operator $\mathcal{M}_\beta$ has matrix elements in the standard Haar basis:

$$\langle mi | \mathcal{M}_\beta | nj \rangle = \int_{-\infty}^{\infty} h_{mi}(x) \left[\mathcal{M}_\beta h_{nj} \right](x) \, dx$$

$$= \frac{2^{(m+n)/2}}{\beta} \int_{-\infty}^{\infty} h(2^m x - i) h\left(\frac{2^n x}{\beta} - j \right) dx$$

Instead of confining oneself to the unit interval, here it is convenient to consider the entire real-number line, and thus that is the range of the integral. Likewise, i and j an be any integers, positive or negative. As before, matrix elements vanish unless

$$\left[\frac{i}{2^m}, \frac{i+1}{2^m} \right] \cap \left[\frac{\beta j}{2^n}, \frac{\beta(j+1)}{2^n} \right] \neq \emptyset$$

This holds in three cases: where one of the intervals contains an edge transition (left, middle or right) of the other interval, without also containing the other two.

22.6   The Shift and Add algorithm

One can model the multiplication of real numbers with a number of different algorithms applied to bit strings. One of the simplest such algorithms is the shift-and-add algorithm, described here. Its just elementary-school long-form multiplication, applied to the binary expansions of the numbers.

There's a point worth laboring on: a bit string representing a real number is not the same thing as the real number. There are more bit-strings than there are real numbers. Most famously, the two bit strings $0.0111\cdots$ and $0.1000\cdots$ are two obviously distinct bit-strings, but they represent the same real number: one-half. All real numbers of the form $j/2^n$ (the "dyadic rationals") will always have dual representations; all other real numbers have a single, unique representation. These correspond to the "gaps" in the Cantor set, or, equivalently, neighboring infinite branches in the finite binary tree. Bit-strings are not real numbers. They're just a usable model of them. The usability is somewhat limited; its OK for working with individual points, but fails miserably for the topologies: the canonical topology on the reals is sharply different than the product topology on 2^ω.

The goal is to compute the product Kx with $0 \leq K \leq 1$ and $0 \leq x \leq 1$ so that the product is $0 \leq Kx \leq 1$. Both K and x are represented by their binary expansions. Let the binary expansions be

$$x = 0.b_0 b_1 b_2 \cdots = \sum_{n=0}^{\infty} b_n 2^{-n-1}$$

and

$$K = 0.c_0 c_1 c_2 \cdots = \sum_{n=0}^{\infty} c_n 2^{-n-1}$$

where the b_n and c_n are either 0 or 1, always.

Define $s_0 = 0$ and s_{n+1} to be the non-negative integer

$$s_{n+1} = b_n c_0 + b_{n-1} c_1 + \cdots + b_0 c_n = \sum_{k=0}^{n} b_k c_{n-k} \qquad (51)$$

Note that $0 \le s_n \le n$. It is useful to visualize this in terms of the elementary school shifted tabular form:

	0	$c_0 b_0$	$c_0 b_1$	$c_0 b_2$	$c_0 b_3$	$\cdots$
			$c_1 b_0$	$c_1 b_1$	$c_1 b_2$	$\cdots$
				$c_2 b_0$	$c_2 b_1$	$\cdots$
+					$c_3 b_0$	$\cdots$
	s_0	s_1	s_2	s_3	s_4	$\cdots$

This makes clear the shift-and-add form. The value of each individual s_n can be visualized as a stack of blocks. For the special case of $K = 0.111\cdots = 1$ one has that $s_{n+1} = \sum_{k=0}^{n} b_k$, that is, it is simply the total number of one-bits in the first n locations.

The final step is to reduce the the sum series s_n to a bit-string. This is accomplished recursively, by performing a carry operation:

$$d_n = s_n + \left\lfloor \frac{d_{n+1}}{2} \right\rfloor \qquad (52)$$

where $\lfloor d \rfloor = d \mod 1$ denotes the floor of d (the integer part of d). The desired bit sequence is then

$$a_n = d_n \mod 2 \qquad (53)$$

Equivalently, a_n is the remainder, the part of d_n that was not propagated to the next location. Explicitly, is is $a_n = d_n - 2\lfloor d_n/2 \rfloor$. The carry-sum propagation can be imagined as a kind of bulldozer, razing the towers d_n until they are one block high, pushing the razed bits off to the next location. The resulting sequence (a_n) is then the bit-string for the product Kx. That is,

$$Kx = 0.a_0 a_1 a_2 \cdots = \sum_{n=0}^{\infty} a_n 2^{-n-1}$$

The problem with this algorithm is that the relation 52 for the d_n is infinitely recursive, and in general is not guaranteed to terminate. One has to start at $n = \infty$ and move backwards from there. There are two plausible scenarios for computing the a_n in practice. One is to search the n until one finds that spot where $\lfloor d_{N+1}/2 \rfloor = 0$; one can then obtain the a_n for all $n < N$ without issue. The problem here is to find such an N.

The other way to compute is to observe that the iteration is convergent. The recursion 52 only depends on a finite and fixed number of bits "behind it", roughly equal to $\log_2 n$ bits that come after this. As noted earlier, $0 \le s_n \le n$ and likewise, $0 \le d_n \le 2n+1$. To write down d_n, one needs at most $C = 1 + \lfloor \log_2 (2n+1) \rfloor$ bits. This implies that a given d_n can only perturb at most $C - 1$ bits downstream of it. That

is, d_{n-C+1} depends on d_n but d_{n-C} does not. Thus, in order to correctly compute all bits a_k for $0 \leq k \leq n-C$, it is sufficient to set d_n to some arbitrary value (less than $2n+2$) and then iterate (using the correct values for s_k when $k < n$). At the end, discard all d_k and a_k for $n-C < k$, as they are incorrect.

22.7 Tree-view

Points:
1) adding one bit is like shifting the tree over sideways.
2) multiplying by one bit is like shifting the tree down-left.
3) adding a number to itself is like shifting tree up (since its just 2x)

23 Simplified Models of Multiplication

The shift-and-add algorithm is obviously rather complex; can it be replaced by something simpler? The particular question to ask is how much of the chaotic dynamics of the beta shift is due to the propagation of the carry bit, and how much of it is due to other parts of the algorithm? Specifically, the addition of two numbers, which requires a carry bit, can be replaced by a bit-wise XOR of their bit strings: this generates "almost" the same results as addition, when the number of 1-bits in the strings are sparse, but are wrong when 1-bits appear in the same location: the XOR discards the carry bits. Thus, a simplified model of multiplication would the the shift-and-XOR model: it proceeds the same way as shift-and-add, but replaces addition with XOR. What does this look like, and how does the equivalent of the beta shift behave under this operation?

23.1 Shift-and-XOR

The shift-and-XOR algorithm must like the shift-and-add algorithm, except that it drops the carry bits. Starting from the same spot, let $0 \leq K \leq 1$ and $0 \leq x \leq 1$ and represent both by their binary expansions:

$$x = 0.b_0 b_1 b_2 \cdots = \sum_{n=0}^{\infty} b_n 2^{-n-1}$$

and

$$K = 0.c_0 c_1 c_2 \cdots = \sum_{n=0}^{\infty} c_n 2^{-n-1}$$

where the b_n and c_n are either 0 or 1.

Define $s_0 = 0$ and s_{n+1} to be the result of XOR-ing instead of adding the bits.

$$s_{n+1} = b_n c_0 \oplus b_{n-1} c_1 \oplus \cdots \oplus b_0 c_n = \bigoplus_{k=0}^{n} b_k c_{n-k}$$
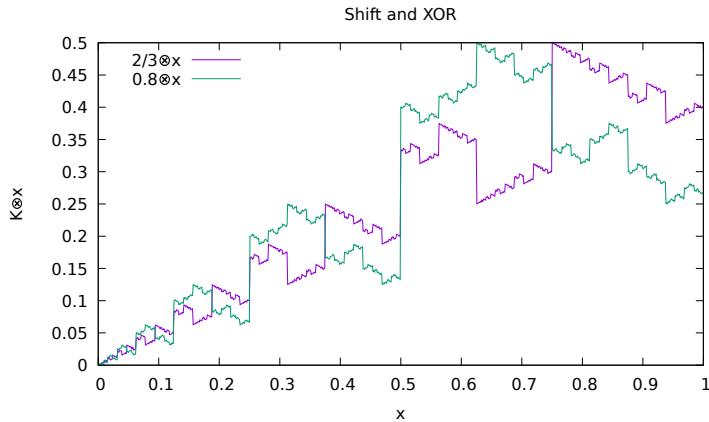
Here, the oplus symbol $\oplus$ denotes the XOR operation. Note that each s_n is either zero or one. Reconstructing a real number from this, one defines

$$K \otimes x = 0.s_0 s_1 s_2 \cdots$$

where the otimes symbol $\otimes$ is pressed into service to indicate the shift-and-XOR product. Note that it is symmetric: $K \otimes x = x \otimes K$ and so behaves at least a bit like ordinary multiplication. Its is not distributive over ordinary addition: $(a+b) \otimes x \neq a \otimes x + b \otimes x$ but it is distributive over XOR: $(a \oplus b) \otimes x = (a \otimes x) \oplus (b \otimes x)$. It is illustrated in figure 58.

The range of the shift-and-XOR operation is fundamentally different from multiplication. First, because the carry bit is dropped, one has that $s_0 = 0$ always, and so that $K \otimes x \leq 1/2$ always, even when both $K \to 1$ and $x \to 1$. Next, for any value of $1/2 < K \leq 1$, the range of $K \otimes x$ runs over the entire interval $[0, 1/2]$ as x runs over the

Figure 58: Shift and XOR Algorithm

This figure shows two functions, $(2/3) \otimes x$ and $(4/5) \otimes x$ as a function of x.

interval $[0,1]$. The measure is not compressed (other than by a factor of 2) , as there is in ordinary multiplication. That is, if $S \subset [0,1]$ is a measurable subset of the unit interval, with measure $\mu(S)$, then one has $\mu(K \otimes S) = \mu(S)/2$. There are several ways to prove this. One formal approach is to consider the correspondence between the natural measure on the reals, and the measure of cylinder sets on the product topology. That is, the Cantor space $\{0,1\}^{\omega}$ is endowed with a natural topology, the product topology. The open sets of this topology are called "cylinder sets". Their measure is uniformly distributed over unit interval, precisely because the Bernoulli shift is ergodic: the one implies the other.

Indeed, the shift-and-XOR algorithm can be best thought of as a formula for shuffling the bit-strings around, without actually altering them: reordering them, not changing them. The intuitive key to this is to observe that subtracting x from 1 just reorders the unit interval, top to bottom, and that this is the same as flipping all zero bits to one, and v.v. That is, $1 - x = x \oplus 0.111 \cdots$.

Another way to see this shuffling is to note that $a \oplus a = 0$ and that $0 \oplus x = x$. Thus, for a fixed value of a, the string x and the string $a \oplus x$ are paired together, in a unique way, so that either can be gotten from the other. The function $a \oplus [0,1] \to [0,1]$ sending $x \mapsto a \oplus x$ is an exchange of these unique pairings of strings. It is not just a bijection, it is an involution. If the strings are given their natural lexicographic sort order, the mapping $x \mapsto a \oplus x$ is just a certain kind of shuffle of the sort order; it neither adds new strings, nor deletes any, nor changes their number. The function is one-to-one and onto. The multiply-and-XOR algorithm is just a repeated sequence of XOR's:

$$K \otimes x = \left(\frac{c_0 x}{2}\right) \oplus \left(\frac{c_1 x}{4}\right) \oplus \left(\frac{c_2 x}{8}\right) \oplus \cdots$$

and so $K \otimes x$ is nothing more than a reshuffling of strings (along with a right-shift equal to the number of leading zero-bits in the binary expansion of K; the right-shift com-

280

mutes with the measure on the product topology.) Thus, $K \otimes x$ preserves the measure on the unit interval (up to a factor of 2^{-n} due to the above-mentioned right-shift). That is, for $1/2 < K \leq 1$, this discussion shows that $\mu(K \otimes S) = \mu(S)/2$.

23.2   Self-similarity

There are several self-similarity properties of the shift-XOR worth noting. It behaves very much like a classic dyadic fractal. Thus, one has that

$$K \otimes \left(\frac{x}{2} \right) = \frac{1}{2}(K \otimes x) = \frac{1}{2} K \otimes x$$

In addition... TODO: illustrate the other symmetry.

23.3   Similarity Transformations

The shift-and-XOR algorithm acts as a permutation on bit-strings. As a result, the XOR-analogs of the beta shift and the tent map become uniformly ergodic, behaving exactly as the Bernoulli shift. The Frobenius-Perron solution to these is just the uniform distribution, which is featureless. All of the structure visible in figures 2 and 4 is entirely due to the dynamics of the carry bit. Effectively, the carry-bit algorithm alters the uniform distribution of the Bernoulli shift (equivalently, the uniform distribution associated with the natural measure on Cantor space.)

Define the XOR-analog of the beta shift as

$$c_\beta(x) = \begin{cases} 2\beta \otimes x & \text{for } 0 \leq x < \frac{1}{2} \\ 2\beta \otimes \left(x - \frac{1}{2} \right) & \text{for } \frac{1}{2} \leq x < 1 \end{cases}$$

The factor of 2 makes up for the fact that shift-XOR effectively drops the top bit; thus the goal is to map each half of the unit interval into the entire interval $[0, 1]$.

Given a fixed β, define $\boxtimes_\beta : [0, 1] \to [0, 1]$ as

$$\boxtimes_\beta(x) = \beta \otimes x$$

As observed previously, $\boxtimes_\beta$ is an automorphism of the unit interval, and more: it is a permutation on Cantor space. Let $b(x)$ be the Bernoulli shift of eqn 1; then one has that $c_\beta = \boxtimes_\beta \circ b$. Taken together, this implies that the ergodic properties of iterating on c_β follow directly from the ergodic properties of the Bernoulli shift; a shuffle, any shuffle on the Cantor set should not alter these ergodic properties.

TODO: similarity transforms on the transfer operator... and the non-alteration of the eigenspectrum, even as the eigenfunctions are altered.

23.4   Multiplication on the Cantor Space

The previous set of results indicates that all of the structure in the bifurcation diagrams of 2 and 4 is entirely due to the dynamics of the propagation of the carry sum. To explore this, the notation needs to be improved on.

The beta shift can be decomposed into multiple distinct stages. First, there is a conversion from the unit interval to the Cantor space; this was defined at the very start, but now we need a less awkward notation for it. Let

$$\pi : 2^\omega \quad \to \quad [0,1]$$
$$0.b_0 b_1 b_2 \cdots \quad \mapsto \quad x$$

be the projection from the Cantor space to the real-number unit interval, given by eqn 2. Note that it is a surjection: dyadic rationals (rationals of the form $m/2^n$) correspond to two distinct bit strings. For example, $1/2$ can be represented as both $0.1000\cdots$ and as $0.0111\cdots$. Cantor space covers the unit interval. Write the inverse mapping as

$$\pi^{-1} : [0,1] \quad \to \quad 2^\omega$$
$$x \quad \mapsto \quad 0.b_0 b_1 b_2 \cdots$$

As a function, it is injective but not surjective. It is usually convenient to ignore this, and to pretend that both π and π^{-1} are bijections, even though they are not. This rarely leads to practical difficulties, as long as one stays conceptually tidy. Better yet, just perform all work on the Cantor space, and project to the unit interval only when needed.

Next, turn to multiplication. This has three parts. First, the summation of the carry bits:

$$S_\beta : 2^\omega \quad \to \quad \mathbb{N}^\omega$$
$$0.b_0 b_1 b_2 \cdots \quad \mapsto \quad (s_0, s_1, s_2, \cdots)$$

where the summation is given by eqn 51. Here, $\mathbb{N}^\omega$ is Baire space, the space of all infinite-length sequences of non-negative integers. In number theory, this would be called the space of arithmetic functions. The second part of multiplication is the propagation of the carry bits. Denote this as

$$C : \mathbb{N}^\omega \quad \to \quad \mathbb{N}^\omega$$
$$(s_0, s_1, s_2, \cdots) \quad \mapsto \quad (d_0, d_1, d_2, \cdots)$$

which is defined in eqn 52. Finally, one extracts the remainder, after propagation:

$$A : \mathbb{N}^\omega \quad \to \quad 2^\omega$$
$$(d_0, d_1, d_2, \cdots) \quad \mapsto \quad (a_0, a_1, a_2, \cdots)$$

which is given by eqn 53. Of the three parts into which we've decomposed multiplication, only the first part is parameterized by K. Thus, multiplication, on Cantor space, can be written as $M_\beta = A \circ C \circ S_\beta$. The shift-and-XOR algorithm omits the propagation of the carry sum. On Cantor space, it is just $\boxtimes_\beta = A \circ S_\beta$: the XOR is just modulo-2 of the carry sum.

To obtain multiplication on the real-number unit interval, we need merely to re-project from Cantor space to the reals. Thus, multiplication, given in eqn 50, decomposes into

$$M_\beta = \pi \circ A \circ C \circ S_\beta \circ \pi^{-1}$$

The beta shift of eqn 4 is then

$$T_\beta = \pi \circ A \circ C \circ S_\beta \circ \pi^{-1} \circ b$$

where b is the Bernoulli shift. To simplify notation, it is convenient to go ahead and provide a symbol for the shift operator:

$$B : 2^\omega \qquad \to \qquad 2^\omega$$
$$(b_0, b_1, b_2, \cdots) \quad \mapsto \quad (b_1, b_2, \cdots)$$

so that $b = \pi \circ B \circ \pi^{-1}$. The corresponding beta shift on the Cantor space is

$$B_\beta = A \circ C \circ S_\beta \circ B$$

which eliminates the pesky projection π. It should be clear that S_β is an injection, the propagation operation C and the remainder A are both surjections.

As noted, the shift-and-XOR algorithm can be written as $\boxtimes_\beta = A \circ S_\beta$; the step where the carry bits are propagated is dropped. The XOR-version of the beta shift is

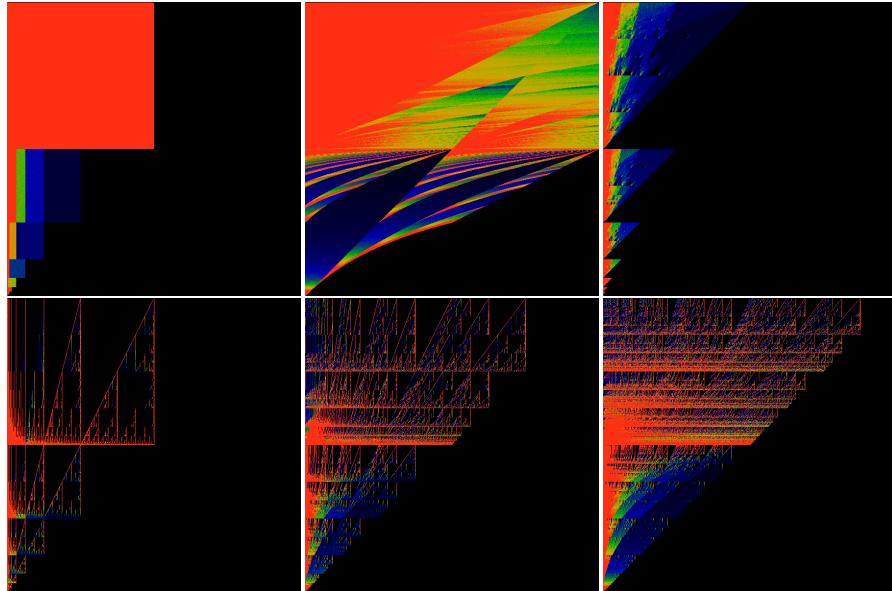$$c_\beta = \boxtimes_\beta \circ B = A \circ S_\beta \circ B$$

Thus, in this new notation, it reaffirms that B is the true source of ergodicity, and that $A \circ S_\beta$ being a permutation does not alter the basic ergodic property of B. All of the structure in the bifurcation diagrams can be blamed on the propagation operator C.

23.5   Propagation games

Pinning the "blame" of complex dynamical structure on the propagation of the carry bits seems to be an open invitation to replace the propagation operator C by just about anything, to see what happens. Figure 59 illustrates some of the things that can happen.

Reviewing the images there makes it clear that although fiddling with the carry bit fundamentally alters point trajectories, it completely fails to open any doors that would provide insight into the structure of the transfer operator. The pictures are pretty, but appear to be meaningless.

Figure 59: Carry-bit propagation

Two triptychs of different carry-bit behaviors. Define $F : \mathbb{N}^\omega \to \mathbb{N}^\omega$ by $F = f \times f \times f \times \cdots$ and then iterate on $A \circ C \circ F \circ S_\beta \circ B$. For $f(n) = n$ one obtains, of course, the standard beta shift of figure 2. The top-left image shows $f(n) = n \mod 2$, which is the same as iterating on the shift-XOR function c_β. Here, β runs from 0 at the bottom, to 2 at the top; x runs from 0 to 1, left to right. The uniform red square simply indicates that the iteration is completely independent of β when $1 < \beta \le 2$: it is fully uniform and ergodic in the same way that the Bernoulli shift is. The top-middle image shows $f(n) = n + 1$, that is, pretending that there is one carry bit too many. The top-right shows $f(n) = \max(0, n - 1)$, that is, having one carry-bit too few.

The bottom three shows a progression of $f(n) = \max(n, 1)$, $f(n) = \max(n, 2)$ and $f(n) = \max(n, 3)$, allowing more and more carry bits to propagate. In the limit, this becomes figure 2 once again. Except for the top-left image, the rest seem pointlessly goofy.

284

24 Sci-fi day-dreaming

This section provides two day-dreams inspired by this material. They are just that: daydreams. If you don't like fictional daydreaming, you won't like the material here. Sorry about that.

24.1 Limits to computation

There are many limits to computation. One limit is the speed of light. In current generation CPU chips, clock rates in the vicinity of 3 gigahertz$= 3 \times 10^9$ cycles per second. By comparison, the speed of light in a vacuum is about 3×10^8 meters per second. Dividing, one finds that light can travel about $3 \times 10^8 / 3 \times 10^9 = 10^{-1}$ meters, or about four inches: a bit bigger than the actual physical dimensions of a chip (typically around half-an-inch on a side), but not by much. Of course, the speed of light in a metal conductor is lower – about half the speed in a vacuum. And transistors are small – more than twenty-thousand times smaller. So, measured in terms of the size of the transistor, the speed of light is about ten or twenty transistor-widths per clock-cycle. So, OK, its still fast, at that length scale. But not really all that fast. The point here is that the speed of light is a potential limit to the speed of computation, and it is not all that far away.

In this setting, one can imagine the situation where the speed of propagating the carry bit during multiplication becomes a limiting factor. The above work hints at a somewhat mind-boggling idea: can multiplication be effectively parallelized by working with transfer operators instead? That is, the multiplication of two numbers corresponds to point-wise particle dynamics: a discrete particle following a chaotic path through a complex numerical computation. By contrast, the transfer operator describes how a distribution propagates through a computation: it effectively performs "an infinite number" of multiplications at the same time, in parallel. That is, rather than asking how single values propagate, one could, and perhaps should, ask how distributions propagate – parallelize multiplication (for example) to an "infinite" degree. It is this rather ridiculous idea that suggests that the above explorations are not purely abstract, but have a potentially practical application. As I suggested – its a bit of science-fiction day-dreaming at this point. But it does hint at an alternate model of computation.

Variants of this model have already been explored, for decades. For example, Crutchfeld defined "geometric state machines" as generalizations of finite state machines, where, instead of having a finite matrix (a "transition matrix") act on a finite vector (the "state vector"), one instead considers operators acting on homogeneous spaces – that is, applying a sequence of such operators on homogeneous space. The most famous and celebrated such space would the $\mathbb{CP}^n$ – complex projective space, with the operators that act on it being the the unitary ones: $U(n)$ – such a system defining the n-qubit quantum state machine. Distributions on $\mathbb{CP}^n$ are mixed states – and the idea of quantum computing is to evolve such states through a set of operations.

Note that ALL quantum computing can be understood to be nothing more than a Crutchfeld machine. For whatever reason, this is almost never stated outright. This is unfortunate, since the Crutchfeld machine is quite easy to understand, whereas "quantum computing" seems to be mired in mystery, at least for the uninitiated.

The point here is that computation, by means of the time-like evolution of distributional densities, is already being explored, but in a rather different context than the one explored here. Here, it seems like we are bowled over by the complexities of a seemingly much simpler system.

24.2 Wave function collapse

There is also a different, bizarrely hypothetical way in which all of this apparatus could manifest itself. Currently, in order to avoid the rather severe issues associated with the concept of quantum-mechanical wave-function collapse, the (vast?) majority of practicing physicists believe in the many-worlds hypothesis. Clearly, this belief is entirely correct for microscopic systems, isolated from the usual thermodynamic hustle and bustle (chlorophyll, rhodopsin and the magnetically sensitive cryptochromes notwithstanding). But it seems to fly in the face of daily experience, where we are aware of just one reality. One of my favorite hypotheses is that this is the result of the (rapid) decay of macroscopic quantum states down to a probability of zero. The mechanism is presumably that of decaying subshift measures. Penrose argues that this has something to do with gravity; but we can go one better: the natural setting for shift spaces are hyperbolic spaces, as that is where there is enough room to "fit everything" in a uniform way consistent with a metric. Curiously, the world we live in – Minkowski space, is hyperbolic. This suggests that the Many Worlds interpretation is exactly right, as long as one truly is in Minkowski space, but that gravitation, which essentially bends or distorts it, squeezes down the room available for multiple quantum states, effectively forcing the collapse in this way.

Put another way: the standard treatment for quantum field theory is the Feynman functional integral; it can be viewed as an integral over all possible paths that a "particle" might take. The daydream is to equate a specific path with the idea of point-dynamics in an iterated function. As long as one considers only points, and their movement, one can be completely unaware of either the invariant measure, or of the decaying eigenstates of the shift operator. In a standard QFT textbook, all equations appear microscopically time-reversible. There's almost no idea of a measure, except for the $\exp -i\hbar S$ in the Feynman integral. The incorporation of gravity into this is famously difficult. The daydream here is that gravity manifests itself as eigenfunctions that live off of the shell of unitary evolution.

There is some practical hope of bringing this daydream to fruition: the theory of subshifts has seen dramatic advances over the last few decades, getting increasingly abstract, and gaining a firm footing in very general settings: viz not just in metric spaces, but even in more general topological vector spaces, and specifically in stereotype spaces, where most of the devices used in analysis can be exercised in reasonably safe manner. The point here is that most of QFT can be formulated using these more-or-less conventional tools and notations. The trick is to locate and extract those parts that renormalize to zero, not unlike some of the formally divergent sums explored above, which can none-the-less be regulated and made to give reasonable answers. Or at least, that's the daydream. Clearly, got far to got before it can be reality.

25 Conclusion

What, exactly, is the point of analytic mathematics, especially in the computational age? Can't one just get a fast computer, iterate on the logistic map, and find out everything there is to find? Well, of course, yes, and no: these questions can be taken as either silly or as deeply philosophical, and it is worth the effort to understand them and address them properly.

First, lets dispose of some obvious mis-perceptions. If one carefully scrutinizes figure 1, one will find signs of a slight unevenness in the horizontal bars. These are numerical artifacts due to statistical under-sampling: they smooth out and fade away with additional sampling of the iterated equations. There is a way to obtain this same figure, far more rapidly, and without this particular form of numerical noise: one can instead iterate on equation 19. This suggests one philosophical answer: the goal of mathematics is to find faster ways of computing things; to discover better algorithms.

A uniting theme between this, and the other text that I have written on fractal issues, is that they are all explorations of the structure of the Cantor set, the structure of the space of infinite sequences of symbols, and the structure of the continuum. That is, we know the continuum in two different ways: one way is by means of the natural topology on the real number line; the other is the product topology on the space of binary strings. The former is suggested by the physical universe that we actually live in: a continuum with spatial extents. The latter is suggested by the notion of time and repetition: the making of choices naturally leads to a tree structure; tree structures necessarily embed in hyperbolic spaces; the Minkowski space that we live in is hyperbolic, and this is why, every day, as time passes on, we get to make new choices precisely because the amount of room for possibilities is ever-increasing as time flows forward.

The idea of analytic combinatorics takes on a whole new meaning in the computational age. Historically, the ability to provide an "exact solution" in the form of an analytic series has been highly prized; the ultimate achievement in many cases. Being able to expression a solution in terms of the addition and multiplication of real numbers is very comforting. Every school student eventually comes to feel that arithmetic on the real numbers is very natural and normal. It's more than that: Cartesian space is smooth and uniform, and all of differential geometry and topology are founded on notions of smoothness.

The inner workings of computers expose (or hide!) a different truth. The most efficient algorithm for computing $sine(x)$ is not to sum the analytic series. Arbitrary precision numerical libraries open the rift further: neither addition nor multiplication are simple or easy. Both operations have a variety of different algorithms that have different run-times, different amounts of memory usage. In the effort to minimize space and time usage, some of these algorithms have grown quite complex. The root cause of the complexity is bewildering: it is the use of the binary digit expansion to represent a real number. Computers use the Cantor space $\{0,1\}^\omega$ or at least a subset thereof, under the covers.

Different representations of the real numbers potentially offer different algorithms and performance profiles. One could represent reals by rationals, but then several other issues arise. One is that the rationals are not evenly distributed across the real number line: rationals with small denominators cluster about in a fractal fashion. This is easily

exhibited by considering continued fractions. As a result, one promptly gets stuck in a quagmire of trying to understand what a "uniform distribution" should be. Binary expansions are more "obviously" uniform. A more basic issue is that, if working with rationals, one must somehow accomplish the addition or multiplication of two integers. To accomplish this, one has to represent the integers as sequences of bits, which only takes us back to where we started. There is no computational oracle that automatically knows the sum or product of integers: it has to be computed.

Compare this situation to that of iterated functions and fractals. At first impression, these seem pathological in almost every respect: differentiable nowhere, unbounded and nonuniform: somehow they feel like the quintessential opposite of the analytic series, of the smoothness of Cartesian space, of the smoothness of addition and multiplication. The place where these two worlds come together is that both are attempts to approach countable infinity, and both are attempts to harness the first uncountable infinity. The real number number is an infinite string of binary digits. The analytic series is an infinite sum. The iterated function is recursively infinite. The historic labor of finding "exact solutions" to problems can perhaps be better views as the discovery of correspondences between finite structures ("the problem to be solved") and infinite structures ("the solution").

The situation here is more easily illustrated in a different domain. The hypergeometric series was presented and studied by Gauss; then Kummer, Pfaff and Euler observed various identities yoking together different series. By the 1950's, thousands of relations were known, along with some algorithms that can enumerate infinite series of relations. The current situation is that there is no known algorithm that can enumerate all such relations; there is no systematic way to classify them. There is an interplay between infinite series and algorithmic relationships between them. Stated a different way: hypergeometric series have a class of self-similarities, and the identities relating them are expressions of that self-similarity. What is that class of self-similarities? For the hypergeometric series, it remains unknown.

For Cantor space, that place where we represent real numbers, the situation is much better. The Cantor space itself has the structure of an infinite binary tree; the tree and it's subtrees are obviously self-similar; the class of similarities is described by the dyadic monoid. The dyadic monoid embeds naturally into the modular group; this in turn is a gateway to vast tracts of modern mathematics. The recursive aspects, the shadow that the Cantor space seems to leave behind everywhere appears to be "explained" by Ornstein theory.

Yet, the picture remains incomplete. The β-transform provides a simple, silly model for multiplying two real numbers together: β and x. The "extra complication" of taking mod 1 after multiplication just reveals how complex multiplication really is. After all, mod 1 is just the subtraction of 1; how hard can that be? Moving in one direction: the fastest, most-efficient-possible algorithm for multiplying two numbers is not known. Moving in another direction, the simple iterated maps, shown in figures 2, 4 and 5 are obviously not only self-similar, but also are surely topologically conjugate to one-another, and in all cases are presumably described by the dyadic monoid; likewise the Mandelbrot set and it's exterior. Yet the details remain obscure.

The meta-question is: what is the correct framework by which one can best understand the interplay between symmetries, infinite series, infinite recursion and al-

gorithms? Until modern times, mathematical practice has reified addition and multiplication into oracular operations that magically obtain "the right answer". Modern computers have put a lie to this: the theory of numerical methods has made clear that addition and multiplication are necessarily algorithmic operations performed on finite truncations of infinite series. What other algorithms are hiding nearby, and what is their relationship to analytic series?

26 Bibliography

The references below provide a bibliography that attempts to touch on all the different ways in which the beta transform and beta expansions have been studied. Search engines exist to help you find the things you don't know, and want to find out more about.

References

[1] Georges de Rham, "On Some Curves Defined by Functional Equations (1957)", in *Classics on Fractals*, edited by Gerald A. Edgar, Addison-Wesley, 1993, pp. 285–298.

[2] A. Rényi, "Representations for real numbers and their ergodic properties", *Acta Math Acad Sci Hungary*, 8, 1957, pp. 477–493.

[3] W. Parry, "On the β-expansion of real numbers", *Acta Math Acad Sci Hungary*, 11, 1960, pp. 401–416.

[4] Karma Dajani, et al., "The natural extension of the beta-transformation", *Acta Math Hungary*, 73, 1996, pp. 97–109, URL https://www.researchgate.net/publication/2257842.

[5] A.O. Gel'fond, "A common property of number systems", *Izvestiya Akad Nauk SSSR Seriya Matematicheskaya*, 23, 1959, pp. 809–814, URL http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=im&paperid=3814&option_lang=eng.

[6] Nikita Sidorov, "Almost every number has a continuum of β-expansions.", *The American Mathematical Monthly*, 110, 2003, pp. 838–842, URL http://www.maths.manchester.ac.uk/nikita/amm.pdf.

[7] Martijn de Vries and Vilmos Komornik, "Unique Expansions of Real Numbers", *ArXiv*, arXiv:math/0609708, 2006, URL https://www.esi.ac.at/static/esiprpr/esi1810.pdf.

[8] Vaughn Climenhaga and Daniel J. Thompson, "Intrinsic ergodicity beyond specification: beta-shifts, S-gap shifts, and their factors", *Israel Journal of Mathematics*, 2010, URL https://arxiv.org/abs/1011.2780.

[9] P. Erdös and V. Komornik, "Developments in non-integer bases", *Acta Math Hungar*, 79, 1998, pp. 57–83.

[10] Nikita Sidorov, "Universal β-expansions", *Arxiv*, 2002, URL https://arxiv.org/abs/math/0209247v1.

[11] Boris Adamczewski, et al., "Rational numbers with purely periodic β-expansion", *Bull London Math Soc*, 42, 2010, pp. 538–552, URL http://adamczewski.perso.math.cnrs.fr/AFSS_BLMS.pdf.

[12] K. Schmidt, "On periodic expansions of Pisot numbers and Salem numbers", *Bull London Math Soc*, 12, 1980, pp. 269–278.

[13] F. Blanchard, "Beta-expansions and Symbolic Dynamics", *Tehoretical Comp Sci*, 65, 1989, pp. 131–141.

[14] Bruno Henrique Prazeres de Melo e Maia, "An equivalent system for studying periodic points of the beta-transformation for a Pisot or a Salem number", , 2007, URL http://repositorio.ual.pt/bitstream/11144/471/1/bmaia_thesis.pdf.

[15] Shigeki Akiyama, "Finiteness and Periodicity of Beta Expansions - Number Theoretical and Dynamical Open Problems", , 2009, URL http://math.tsukuba.ac.jp/~akiyama/papers/proc/BetaFiniteCIRM.pdf.

[16] Louis-Sébastien Guimond, et al., "Arithmetics on beta-expansions", *Acta Arithmetica*, 112, 2001, pp. 23–40, URL https://www.researchgate.net/profile/Edita_Pelantova/publication/259299735_Arithmetics_of_beta-expansions/links/5434e32a0cf294006f736e7c/Arithmetics-of-beta-expansions.pdf.

[17] Bernat Julien, "Arithmetics in β-numeration", *Discrete Mathematics and Theoretical Computer Science*, 9, 2006, URL http://www.iecl.univ-lorraine.fr/~Julien.Bernat/arithbetanum.pdf.

[18] M. Hbaib and Y. Laabidi, "Arithmetics in the set of beta-polynomials", *Int J Open Problems Compt Math*, 6, 2013, URL http://www.i-csrs.org/Volumes/ijopcm/vol.6/vol.6.3.1.pdf.

[19] W. P. Thurston, "Groups, tilings and finite state automata", *AMS Colloquium Lectures*, 1989.

[20] Valérie Berthé and Anne Siegel, "Tilings associated with beta-numeration and substitutions", *Integers: Electronc Journal of Combinatorial Number Theory*, 5, 2005.

[21] Sh. Ito and H. Rao, "Purely periodic β-expansions with Pisot unit base", *Proc Amer Math Soc*, 133, 2005, pp. 953–964.

[22] Shigeki Akiyama, "Beta expansion and self-similar tilings", , 2017, URL http://cloud.crm2.uhp-nancy.fr/pdf/Manila2017/Akiyama.pdf.

[23] Valérie Berthé and Anne Siegel, "Purely periodic β-expansions in the Pisot non-unit case", *ArXiv*, arXiv:math/0407282, 2004, URL https://hal.archives-ouvertes.fr/hal-00002208/document.

[24] Jakob Grue Simonsen, "Beta-Shifts, their Languages, and Computability", *Theory of Computing Systems*, 48, 2011, pp. 297–318, URL http://www.diku.dk/~simonsen/papers/j12.pdf.

[25] Leopold Flatto, et al., "The Zeta Function of the Beta Transformation", *Ergodic Theory and Dynamical Systems*, 14, 1994, pp. 237–266.

[26] Daniel J. Thompson, "Irregular sets and conditional variational principles in dynamical systems", , 2010, URL https://people.math.osu.edu/thompson.2455/thesis_thompson.pdf.

[27] L. Barreira and B. Saussol, "Variational Principles and Mised Multifractal Spectra", *Transactions of the American Mathematical Society*, 353, 2001, pp. 3919–3944, URL http://www.math.univ-brest.fr/perso/benoit.saussol/art/mixed.pdf.

[28] Nishant Chandgotia, et al., "One dimensional Markov random fields, Markov chains and Topological Markov fields", *arXiv*, arXiv:1112.4240 [math.DS], 2011, URL https://arxiv.org/abs/1112.4240.

[29] Lyndsey Clark, "The beta-transform with a hole", *arXiv*, arXiv:1412.6384 [math.DS], 2014, URL https://arxiv.org/abs/1412.6384.

[30] Hans-Peter Blatt, et al., "On a generalization of Jentzsch's theorem", *J Approx Theory*, 159, 2009, pp. 26–38, URL https://core.ac.uk/download/pdf/210808388.pdf.

[31] Dean J. Driebe, *Fully Chaotic Maps and Broken Time Symmetry*, Kluwer Academic Publishers, 1999.

[32] Linas Vepstas, "The Bernoulli Map", , 2004, URL https://www.linas.org/math/bernoulli.pdf, self-published on personal website.

[33] Linas Vepstas, "The Minkowski Question Mark and the Modular Group PSL(2,Z)", , 2004, URL https://www.linas.org/math/chap-minkowski.pdf, self-published on personal website.

[34] Linas Vepstas, "Symmetries of Period-Doubling Maps", , 2004, URL https://www.linas.org/math/chap-takagi.pdf, self-published on personal website.

[35] Wikipedia, "Quotient space", , 2018, URL https://en.wikipedia.org/wiki/Quotient_space_(topology).

[36] Kevin Hare, et al., "Three Series for the Generalized Golden Mean", *Arxiv*, 2014, URL https://arxiv.org/abs/1401.6200.

[37] A.P. Stakhov, "The Generalized Principle of the Golden Section and its applications in mathematics, science, and engineering", *Chaos, Solitons and Fractals*, 26, 2005, pp. 263–289, URL http://www.student.oulu.fi/~taneliha/Phi6/1/The%20Generalized%20Principle%20of%20the%20Golden%20Section%20and%20its%20applications%20in%20mathematics,%20science,%20and%20engineering.pdf.

[38] Solomon W. Golomb, "Irreducible polynomials, synchronizing codes, primitive necklaces and cyclotomic algebra", *Procedings Conf Combinatorial Mathematics and its Applicaltions*, 1969, pp. 358–370.

[39] Christophe Reutenauer, *Free Lie Algebras*, London Mathematical Society Monographs, 1993.

[40] S. Duzhin and D. Pasechnik, "Automorphisms of necklaces and sandpile groups", *ArXiv*, arXiv:math/1304.2563, 2013, URL https://arxiv.org/pdf/1304.2563.

[41] Kevin Cattell, et al., "Fast Algorithms to Generate Necklaces, Unlabeled Necklaces, and Irreducible Polynomials over GF(2)", *Journal of Algorithms*, 37, 2000, pp. 267–282, URL http://www.cis.uoguelph.ca/~sawada/pub.html].

[42] Shagnik Das, "A brief note on estimates of binomial coefficients", , 2015, URL http://discretemath.imp.fu-berlin.de/DMII-2015-16/binomials.pdf, university course notes.

[43] Mahdi S. Hosseini, et al., "On the Closed Form Expression of Elementary Symmetric Polynomials and the Inverse of Vandermonde Matrix", *arXiv*, arXiv:1909:08155 [math.NA], 2019, URL https://arxiv.org/abs/1909.08155.

[44] E. A. Rawashdeh, "A Simple Method for Finding the Inverse Matrix of the Vandermonde Matrix", *Matematicki Vesnik*, 71, 2019, pp. 207–213, URL http://www.vesnik.math.rs/vol/mv19303.pdf.

[45] A. Muhammed Uludağ and Hakan Ayral, "Jimm, a Fundamental Involution", *arXiv*, arXiv:1501.03787 [math.NT], 2015, URL https://arxiv.org/abs/1501.03787.

[46] Wikipedia, "Arnold tongue", , 2006, URL https://en.wikipedia.org/wiki/Arnold_tongue.

[47] Linas Vepstas, "On the Minkowski Measure", *ArXiv*, arXiv:0810.1265, 2008, URL http://arxiv.org/abs/0810.1265.

[48] Gerald Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*, American Mathematical Society, 2000, URL https://www.mat.univie.ac.at/~gerald/ftp/book-jac/jacop.pdf.

[49] Edward B. Saff and Nikos Stylianopoulos, "Asymptotics for Hessenberg matrices for the Bergman shift operator on Jordan region", *arXiv*, arXiv:1205.4183 [math.CV], 2012, URL https://arxiv.org/abs/1205.4183.

[50] E. Torrano V. Tomeo, "Two applications of the subnormality of the Hessenberg matrix related to general orthogonal polynomials", *Linear Algebra and its Applications*, Volume 435, Issue 9, Pages 2314-2320, 2011, URL http://oa.upm.es/id/eprint/8725/contents.

[51] Carmen Escribano, et al., "The Hessenberg matrix and the Riemann mapping", *arXiv*, arXiv:1107.603 [math.SP], 2011, URL https://arxiv.org/abs/1107.6036.

Appendix

A lot of notation has been introduced, and can be difficult to track. This is an abbreviated listing of the definitions.

The β-transformation   The β-transformation is $t_\beta(x) = \beta x \mod 1$

The beta shift   The β-shift is defined in 4

$$T_\beta(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} \\ \beta\left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases}$$

The relation to the β-transformation is given in 9 as

$$T_\beta^n(x) = \frac{\beta}{2} t_\beta^n\left(\frac{2x}{\beta}\right)$$

for all positive integers n.

The beta bitsequence   The beta shift generates a bitsequence, defined in 5 as

$$k_n(x) = \begin{cases} 0 & \text{if } 0 \leq T_\beta^n(x) < \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leq T_\beta^n(x) \leq 1 \end{cases}$$

Equivalently, written with the Heaviside Θ, this is

$$k_n(x) = \Theta\left(T_\beta^n(x) - \frac{1}{2}\right)$$

Substituting in the β-transformation,

$$k_n(x) = \Theta\left(\beta t_\beta^n\left(\frac{2x}{\beta}\right) - 1\right)$$

The β-expansion   The bit-sequence defines a unique base-β expansion, given in 6 as

$$x = \frac{1}{2} \sum_{n=0}^{\infty} \frac{k_n}{\beta^n}$$

The beta shift acts on this series as $T(k_0 k_1 k_2 \cdots) = k_1 k_2 \cdots$

The midpoint/endpoint orbit   The iterated endpoint of the β-transformation is the same as the iterated mid-point of the beta shift. It is

$$t_n \equiv t_\beta^n(1) = \frac{2}{\beta} T_\beta^n\left(\frac{\beta}{2}\right)$$

The iterated mid-point is written as $m_0 = \beta/2$ and $m_n = T_\beta^n(\beta/2)$.

Gelfond–Parry digit sequence   The Gelfond–Parry digit sequence ε_n, defined for a fixed β and a real number $0 \leq y \leq 1$, is given in 11 is

$$\varepsilon_n(y) = \begin{cases} 1 & \text{if } y \leq t_\beta^n(1) \\ 0 & \text{otherwise} \end{cases}$$

Using the Heaviside step function, this can be written as $\varepsilon_n(y) = \Theta(t_n - y)$. This is rescaled to the beta shift in 23

$$d_n(x) = \varepsilon_n\left(\frac{2x}{\beta}\right) = \Theta\left(\frac{\beta}{2}t_n - x\right) = \Theta\left(T^n\left(\frac{\beta}{2}\right) - x\right) = \begin{cases} 1 & \text{if } x < T^n\left(\frac{\beta}{2}\right) \\ 0 & \text{otherwise} \end{cases}$$

This digit sequence runs in the "opposite direction" from the beta bitsequence. The transfer operator is a pushforward; the Koopman operator is a pullback.

Gelfond-Parry measure   The Gelfond-Parry measure, defined for a fixed β and a real number $0 \leq y \leq 1$, is given in 10

$$\nu_\beta(y) = \frac{1}{F} \sum_{n=0}^{\infty} \frac{\varepsilon_n(y)}{\beta^n}$$

with normalization

$$F = \sum_{n=0}^{\infty} \frac{t_\beta^n(1)}{\beta^n}$$

Characteristic bit-sequence   Every value of β has a characteristic bit-sequence, obtained from iterating the mid-point $m_0 = \beta/2$.

$$b_n = \Theta\left(m_n - \frac{1}{2}\right) = \Theta\left(T_\beta^n\left(\frac{\beta}{2}\right) - \frac{1}{2}\right) = d_n\left(\frac{1}{2}\right) = k_n\left(\frac{\beta}{2}\right) = \varepsilon_n\left(\frac{1}{\beta}\right) = \Theta(\beta t_k - 1)$$

Note that $b_0 = 1$ always.

Stopping algo   The bracketing constraint given in 32 is a boolean function (returning true/false)

$$\theta_n(\rho) = \begin{cases} \Theta\left(r_{n/2} - \rho\right) \cdot \theta_{n/2}(\rho) & \text{for } n \text{ even} \\ \theta_{\lfloor n/2 \rfloor}(\rho) & \text{for } n \text{ odd} \end{cases}$$

where r_n is the positive real root $p_n(r_n) = 0$. The above defintion is written in tail-recursive form, and so can be expressed as a loop:

```
m_prev := n
m := ⌊n/2⌋
while (0 < m)
     m_prev is even and r_m < ρ then fail
     m_prev := m
     m := ⌊m/2⌋
if arrived here, then success
```

Polynomials   The self-describing polynomials are

$$p_n(z) = z^{k+1}\left(1 - b_0 z^{-1} - b_1 z^{-2} - \cdots - b_k z^{-k-1}\right)$$

These have the property that if β generated the characteristic bitsequence $\{b_k\}$, then $p_n(\beta) = 0$. For infinite-length bitsequences, define the holomorphic function

$$q^{\{b\}}(\zeta) = 1 - \sum_{j=0}^{\infty} b_j \zeta^{j+1}$$

which recovers the polynomial when $\zeta = 1/z$.

That's all, folks!