

Spectrum of the Beta Transformation

Linas Vepštas

December 2017 (Current version: September 2020)*

linasvepstas@gmail.com

Abstract

The beta transformation is the iterated map $\beta x \bmod 1$. 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 transform 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 spectrum of the Ruelle–Frobenius–Perron operator of the beta transformation is examined. Explicit expressions for a set of bounded eigenfunctions is given; they compose a discrete spectrum accumulating on a circle of radius $1/\beta$ in the complex plane.

Periodic orbits of the beta transform give eigenvalues that are roots of polynomials. These form an curious class; they are “quasi-cyclotomic” and can be counted with the necklace-counting function. They do not have any obvious relation to either Lyndon words or to polynomials over the field \mathbb{F}_2 , despite both of those also being countable by the necklace function. The positive real roots include the Golden and silver ratios, the Pisot numbers, the n-bonacci (tribonacci, tetranacci, etc.) numbers and yoke all of these into a regular structure.

Periodic orbits of the beta transform appear to be in one-to-one correspondence with the “islands of stability” of the Feigenbaum map. The beta transformation has no such “islands”, but this is apparently because they are degenerate or collapsed down to a point. The Feigenbaum map opens up at exactly these points, making the islands manifest.

This text assumes very little mathematical sophistication on the part of the reader, and should be approachable by those with minimal or no prior experience in ergodic theory. Most of the development is casual. As a side effect, the introductory sections are perhaps a fair bit longer than strictly needed to present the new results.

*This text is an extract of a diary of research notes, first published in December 2017, with major updates published in February 2018 and December 2018. The current version expunges all the unfinished parts and the boring parts and the speculative parts, and prettifies the rest into a presentable format. MSC2020 Classification 37E05, 37G10. Submitted to Transactions of the AMS on 4 October 2020.

1 Introduction

The beta transformation is the map

$$t_\beta(x) = \beta x \mod 1$$

iterated on the real unit interval $[0, 1]$ and β taken to be a real positive constant. For the special case of $\beta = 2$, this is the Bernoulli process, the ergodic process of random coin flips. This case is solvable in more ways than one; it is very well-explored, and is a canonical model of ergodic behavior. The beta transformation is less-well studied; a full review of results is presented at the end of this introduction.

There are two common ways to study iterated functions: in terms of point dynamics and orbits (“where does the point x go, when iterated?”) and in terms of distributions (“how does a scattered dust of points evolve over iteration?”). The approach via distributions is arguably more powerful and constraining, as the “dust” is a continuum, a measurable set, rather than a handful of points “of measure zero”.

The action of an iterated map on a distribution is described in terms of the pushforward of the map. Informally, in the present case, the pushforward of the beta transform, acting on a distribution $\rho : [0, 1] \rightarrow \mathbb{R}$ can be written as

$$[\mathcal{L}_\beta \rho](x) = \sum_{y \in t^{-1}(x)} \frac{\rho(y)}{|df/dx|_y}$$

where the denominator is understood as the Jacobian determinant. The operator \mathcal{L}_β is a linear operator, acting on some space of functions ρ . Different spaces result in different properties for \mathcal{L}_β ; commonly explored are the polynomials, and the square-integrable functions, and, in certain special cases, the analysis can be extended to highly singular, continuous-nowhere functions. The pushforward operator is generically called the “transfer operator” or the “Ruelle–Frobenius–Perron operator”.

The largest eigenvalue is conventionally called the Frobenius–Perron eigenvalue; for the beta transform, as for many other iterated maps, it is equal to one. Solutions to $\mathcal{L}_\beta \mu = \mu$ define an invariant measure μ on the unit interval. It is invariant in the usual sense of dynamical systems: one has $(\mu \circ t_\beta^{-1})(x) = \mu(x)$ for almost all x . The proper formal statement of this is to invoke the sigma-algebra \mathcal{B} of Borel sets on the unit interval. The inverse of the beta transform is then a map $t_\beta^{-1} : \mathcal{B} \rightarrow \mathcal{B}$ while the measure $\mu : \mathcal{B} \rightarrow \mathbb{R}$ is a sigma-additive measure. The pushforward nature of the transfer operator is now more readily visible: $\mathcal{L}_\beta \rho = \rho \circ t_\beta^{-1}$. One is generally interested in the spectrum of \mathcal{L}_β . In the current case, the spectrum is simple enough that it can be described by eigenfunctions $\mathcal{L}_\beta \rho = \lambda \rho$ of this operator. In general, the precise spectrum depends on how the class of mappings $\rho : \mathcal{B} \rightarrow \mathbb{C}$ or $\rho : \mathcal{B} \rightarrow \mathbb{R}$ might be restricted: *e.g.* by limiting it to functions ρ that can be interpreted as being polynomials, or square-integrable, *etc.* on the unit interval.

This text will present a complete description of the spectrum in terms of finite, bounded functions on the unit interval. It is described by the zeros of a certain analytic function. For special values of β corresponding to the generalized Fibonacci numbers, the analytic function becomes a polynomial 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). These are quasi-cyclotomic, in that their roots are approximately evenly distributed in an approximately circular fashion on the complex plane. This class of polynomials has numerous interesting properties, but seems to never have been studied before.

But first, the remainder of this introduction: it provides a very basic review of some elementary properties of the beta transform, starting with the full shift, aka the Bernoulli shift. The section is completed with a survey of other research results.

1.1 Bernoulli shift

The Bernoulli shift (*aka* the bit-shift map, the dyadic transform or the full shift) is an iterated map on the unit interval, given by iteration of the function

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

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

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

(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.) Of course, the symbolic dynamics recreates the initial real number:

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

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

$$x = 0.b_0b_1b_2\cdots$$

The Bernoulli shift has many interesting properties, connecting it to the Cantor set and a large class of self-similar fractals. It is also connected to the modular group and elliptic functions (thence to number theory and Riemann surfaces). It is also the primary example of an ergodic system, and, in a sense, has far-reaching implications for ergodic systems via Ornstein theory. These various connections are far too numerous to review here, beyond these brief mentions.

The current task is to attempt to see how many of these properties still hold in the more general case of the β -transformation.

1.2 Shift space

A “shift space” is formally defined to be a set of infinite sequences of a set of N letters (or symbols), together with a shift operator T that takes each sequence, and lops off the left-most symbol. For the Bernoulli shift, there are $N = 2$ letters, taken from the set $\{0, 1\}$. For the Bernoulli shift, one is typically interested in the set of *all* possible infinite sequences: this is the “full shift”. One writes $\{0, 1\}^\omega$ for this shift space, ω denoting countable infinity. For the Bernoulli shift, the map $b(x)$ is the shift operator: it just lops off the left-most symbol.

In general, a shift space does not have to include every possible sequence of symbols; it does, however, by definition, have to be shift-invariant. That is, given some set S of infinite sequences of N symbols, the set S is a shift space if and only if, by lopping off the leading symbol of each string, one regains S again. That is, a shift space S must obey

$$TS = S$$

For example, $S = \{000\cdots, 111\cdots\}$ contains only two elements: the string of all zeros, and the string of all ones; lopping off the leading digit just returns S again. In general, shift spaces may contain a finite, or a countable, or an uncountable number of elements. In general, one defines the “full shift” as the space N^ω of all possible strings of N symbols. Subsets that are shift spaces are called “subshifts”.

The words “symbolic dynamics” also deserve some mention: given one specific sequence $x \in S$ out of the shift space S , one can ponder “where it goes to”, as one lops off a symbol at a time. This gives the “symbolic dynamics” or the “point dynamics” of the sequence. The “orbit” is defined as the set $\{T^m x \mid \text{integer } m \geq 0\}$ – that is, the set of all places that x goes to. There are several possibilities: one is that x is a fixed point, so that $Tx = x$. Another is that x is a repeating sequence of symbols, in which case iteration repeats as well: $T^m x = x$ holds whenever the repeat length is m ; this is a periodic orbit.

Most importantly, when there are non-periodic orbits, then there are uncountably many of them. That is, the number of periodic orbits is always countable: there are at most N starting points for an orbit, and they are of at most length m . This includes the “eventually periodic” orbits: those that begin with some finite initial sequence of “random” moves, eventually settling into a periodic recurrence. This cannot be done for the non-periodic orbits: they cannot be placed in lexicographic order; as the Cantor slash argument applies. The non-periodic orbits are always uncountable. In this sense, the non-periodic orbits are “more important”, in that they make up the bulk of all orbits. There are situations in which the periodic orbits can be said to “govern” the dynamics of the aperiodic ones; this is termed “coherent structure”. However such concepts are advanced, and are not obviously applicable to the current situation.

For the most part, the goal of this text is to describe what happens to the uncountable number of non-periodic orbits. The technique to do this is not to describe them individually, but instead as collections, distributions, on the unit interval.

1.3 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

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

This map, together with similar maps, is illustrated in figure 5 below.

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

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

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

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

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

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

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

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

That is, 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 . This is explored in greater detail in a later section.

1.4 Associated polynomial

The iterated shift can also be written as a finite sum. This is noted here; it will be useful in later sections. 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

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

The general form is then:

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

Since the k_m depend on both β and on x , and are piece-wise discontinuous functions of x , this is not a true polynomial; however, this polynomial-like representation will be useful, later.

1.5 Density Visualizations

The “long-term dynamics” of the β -shift can be visualized by means of a “bifurcation diagram”. Scare quotes are used, as these are imprecise terms; greater precision will be provided shortly. The idea of a bifurcation diagram was made famous by the Feigenbaum map (shown in figure 4), and so 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 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, as follows. One divides the unit interval 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 $j/N \leq b_\beta^n(x) < (j+1)/N$, then increment the count for the j 'th bin. After a total of M iterations, let $c(j; M)$ be the count in the 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:

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

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:

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

¹Formal 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,[1] weakly mixing[2] and weakly Bernoulli.[3]

Numerically, this integration can be achieved by randomly sampling a large number of starting points. Observe that, 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 .

1.6 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 3. Its is worth contemplating the similarities between this, and the corresponding beta shift diagram. Clearly, there are a number of shared features.

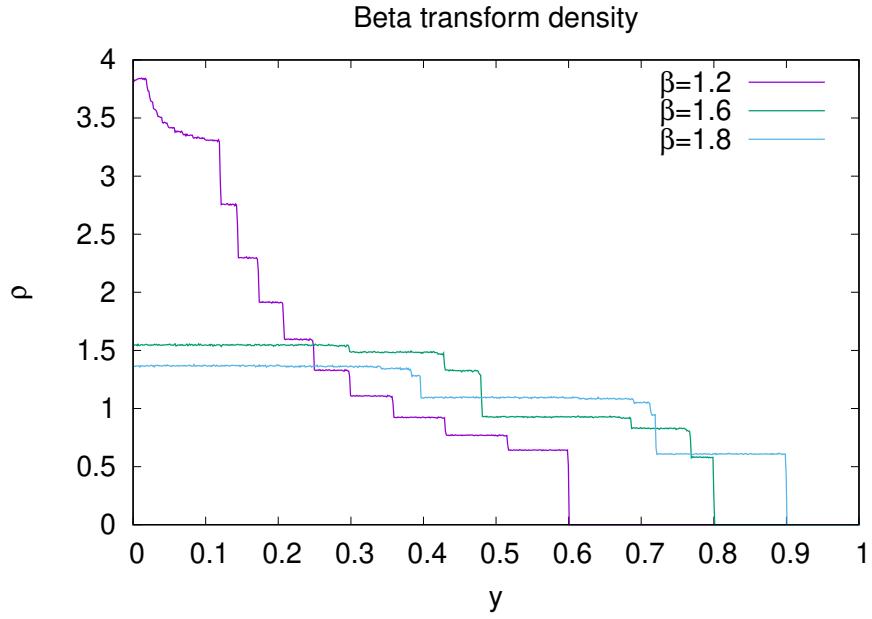
1.7 Logistic Map

The logistic map is related to the tent map, and is given by iteration of the function

$$f_\beta(x) = 2\beta x(1-x)$$

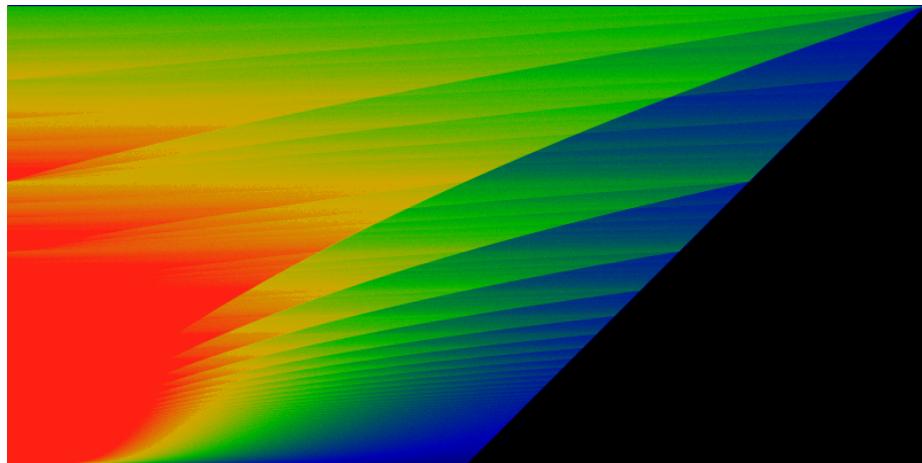
It essentially replaces the triangle forming the tent map with a parabola of the same height. That is, the function is defined here so that the the same value of β corresponds to the same height for all three maps. Although the heights of the iterators have been aligned so that they match, each exhibits 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

Figure 1: Beta-shift Density Distribution



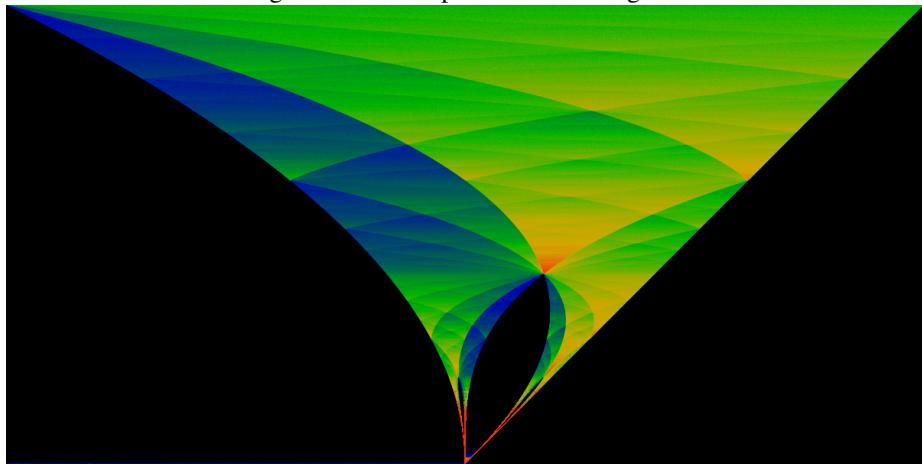
The above figure shows three different density distributions, for $\rho_{1.2}(y)$, $\rho_{1.6}(y)$ and $\rho_{1.8}(y)$, calculated numerically. These are obtained by histogramming a large number of point trajectories, as described in the text. The small quantities of jitter are due to a finite number 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 tiny 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.

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. This means that one cannot apply simple-minded guess-work to discover the overall diagram shown here.

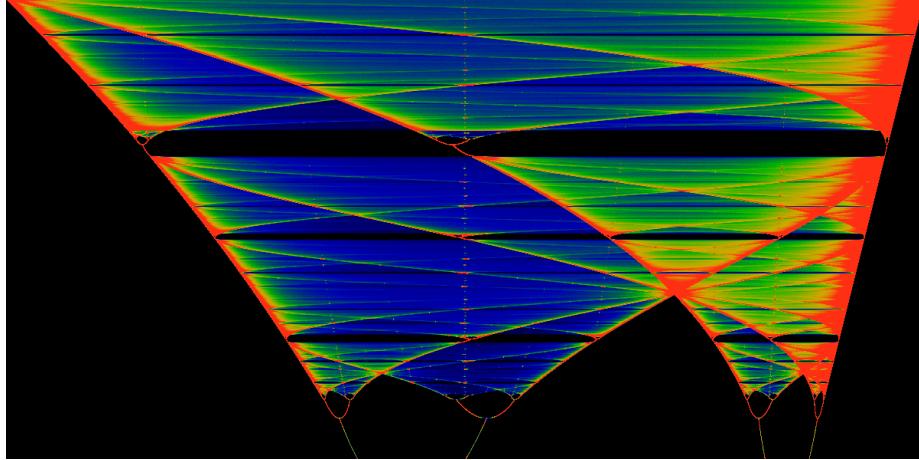
Figure 3: Tent Map Bifurcation Diagram



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 $0 \leq x \leq \beta(1 - \beta/2)$ are never visited more than a finite number of times.

Figure 4: 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 corresponding diagram for the beta shift.

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 4. 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, 3 and 4 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: there is a way to make all three of these maps be “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.

1.8 Beta Transformation

After exactly one iteration of the beta shift, all initial points $\beta/2 \leq x \leq 1$ are swept up into the domain $0 \leq x < \beta/2$, and never leave. Likewise, the range of the iterated beta-shift is $0 \leq x < \beta/2$. 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

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

This can be written more compactly as $t_\beta(x) = \beta x \bmod 1$. 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:

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

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 doing this is to retain contact with the β -shift being a subshift of the full shift.

After a single iteration of the tent map, a similar situation applies. After one iteration, all initial points $\beta/2 \leq x \leq 1$ are swept up into the domain $0 \leq x < \beta/2$. After a finite number of iterations, all points $0 < x \leq \beta(1 - \beta/2)$ 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

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

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

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

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”:

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

The six different maps under consideration here are depicted in figure 5. It is interesting to compare three of the bifurcation diagrams, side-by-side. These are shown in figure 6.

1.9 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.

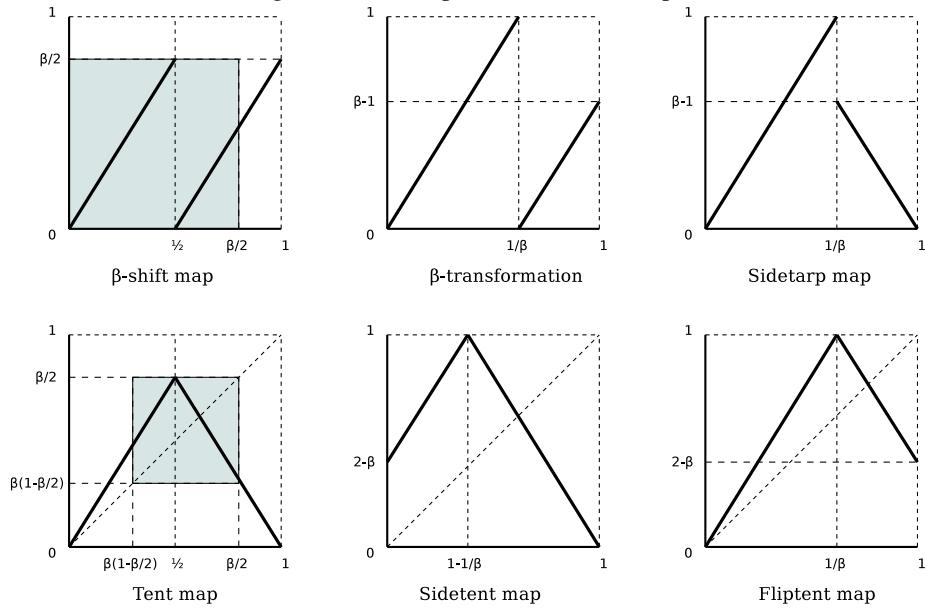
The Cantor space $\{0, 1\}^\omega$ can be given a topology, the product topology. The open sets of this topology are called “cylinder sets”. These are the infinite strings in three symbols: a finite number of 0 and 1 symbols, and an infinite number of * symbols, the latter meaning “don’t care”. Set union 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: $\bar{0} = 1$, $\bar{1} = 0$ and $\bar{*} = *$. 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.

Denote the Borel algebra by \mathcal{B} . A shift is now a map $T : \mathcal{B} \rightarrow \mathcal{B}$ that lops off the leading symbol of a given cylinder set. This 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 $f : \mathcal{B} \rightarrow \mathbb{R}$ 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 $\mu : \mathcal{B} \rightarrow \mathbb{R}$ that are positive, and are “compatible” with the sigma algebra, in that $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A \cap B = \emptyset$ and (for product-space measures) that $\mu(A \cap B) = \mu(A)\mu(B)$ for all $A, B \in \mathcal{B}$. 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

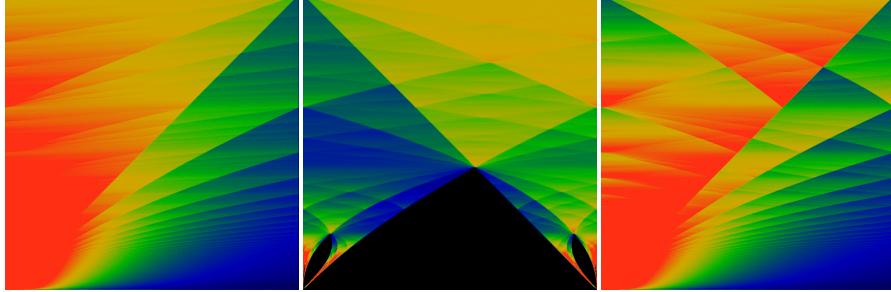
Figure 5: Iterated piece-wise linear maps



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 fixed points and periodic orbits, essentially because the diagonal intersects the map. The top three maps have no periodic orbits, and 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 rearrangements of the same basic elements.

Figure 6: Beta transform and Side-tent



The left figure shows the bifurcation diagram for the β -transform, as it is normally defined as the $\beta x \bmod 1$ 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 3, 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.

$p^m(1-p)^n$. 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 \mathcal{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 $0***\dots$ is mapped to the interval $[0, 1/2]$ and $1***\dots$ 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.

Although the map of eqn 3 is not an isomorphism, it is more-or-less safe for most practical purposes to treat it as such. For any closed interval, there is a unique corresponding cylinder set. The only trouble arises at a countable number of points: all dyadic rationals have two expansions, not one, and so there, the map is not invertible. However, the points all have a measure of zero in both cases, and there are only a countable number of such points, and the mapping always takes points to points. Thus, the inverse exists for almost all members of the algebra and is unique whenever an interval has positive 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 \mathcal{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 $\mu(T^{-1}(\sigma)) = \mu(\sigma)$ 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 \mathcal{L}_T on the space \mathcal{F} of all functions $f : \mathcal{B} \rightarrow \mathbb{R}$. It is defined as $\mathcal{L}_T : \mathcal{F} \rightarrow \mathcal{F}$ by setting $\mathcal{L}_T : f \mapsto f \circ T^{-1}$. It is obviously linear, in that $\mathcal{L}_T(af + bg) = a\mathcal{L}_T(f) + b\mathcal{L}_T(g)$. 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 \mathcal{F} .

The canonical example is again the Bernoulli shift. For this, we invoke the inverse of the mapping of eqn 3 so that $f : [0, 1] \rightarrow \mathbb{R}$ is a function defined on the unit interval, instead of $f : \mathcal{B} \rightarrow \mathbb{R}$. When \mathcal{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 \mathcal{L}_T is discrete. It consists of the Bernoulli polynomials $B_n(x)$ corresponding to an eigenvalue of 2^{-n} . That is, $\mathcal{L}_T B_n = 2^{-n} B_n$. Note that $B_0(x) = 1$ is the invariant measure on the full shift. For the space of square-integrable functions $f : [0, 1] \rightarrow \mathbb{R}$, the spectrum of \mathcal{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 $\mathcal{L}_T ?' = ?'$ with $?$ the Minkowski question mark function, and $?'$ its 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 \mathcal{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 \mathcal{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 measureability 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.

1.10 Beta Transformation Literature Review and References

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

An explicit expression for the invariant measure was independently obtained by

A.O. Gel'fond[4] and by W. Parry[2], as a summation of step functions

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

where ε_n is the digit sequence

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

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

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

Similar 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![5] This signals that the beta shift behaves rather differently from the Cantor set in its 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.[6]

The β -transformation has been shown to have the same ergodicity properties as the Bernoulli shift.[3] The fact that the beta shift, and its subshifts are all ergodic is established by Climenhaga and Thompson.[7]

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 Erdős and Komornik[8]. Its is shown by N. Sidorov that almost every β -expansion is universal.[9] Conversely, there are some values of β for which rational numbers have purely periodic β -expansions;[10] all such numbers are Pisot numbers.[11]

The symbolic dynamics of the beta-transformation was analyzed by F. Blanchard[12]. A characterization of the periodic points are given by Bruno Maia[13]. A discussion of various open problems with respect to the beta expansion is given by Akiyama.[14]

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.[15, 16, 17]

Certain values of β - generally, the Pisot numbers, generate fractal tilings,[18, 19, 20, 10, 14] which are generalizations of the Rauzy fractal. An overview, with common terminology and definitions is provided by Akiyama.[21] The tilings, sometimes called (generalized) Rauzy fractals, can be thought of as living in a direct product of Euclidean and p -adic spaces.[22]

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.[23]

The zeta function, and a lap-counting function, are given by Lagarias[24]. The Hausdorff dimension, the topological entropy and general notions of topological pressure arising from conditional variational principles is given by Daniel Thompson[25]. A proper background on this topic is given by Barreira and Saussol[26].

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.

2 Transfer operators

The discovery and study of invariant measures, as well as of decaying states can be approached via the transfer operator, or, properly named, the Ruelle-Frobenius-Perron operator. This is an operator that captures the behavior of a distribution under the action of a map. The invariant measure is an eigenstate of this operator; indeed, it provides a formal definition for what it means to be invariant under the action of the map.

Given an iterated map $g : [0, 1] \rightarrow [0, 1]$ on the unit interval, the transfer operator defines how distributions are acted on by this map. It is defined as

$$[\mathcal{L}_g f](y) = \sum_{x=g^{-1}(y)} \frac{f(x)}{|g'(x)|}$$

The left adjoint of the transfer operator is the composition operator (Koopman operator). This is defined as

$$[\mathcal{C}_g f](y) = f(g(y))$$

The Koopman operator is adjoint, in the sense that $\mathcal{L}_g \mathcal{C}_g = 1$ but that, in general, $\mathcal{C}_g \mathcal{L}_g \neq 1$.

2.1 The β -shift Transfer Operator

The transfer operator for the β -shift and the β -transform are only superficially different. We'll give both of them here, as it is sometimes useful to reference one or the other, depending on the situation. Superficially, they seem to have a different form, but are in the end just scaled versions of one-another. The transfer operator the beta-shift map $T_\beta(x)$ is

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

or, written more compactly

$$[\mathcal{L}_\beta f](y) = \frac{1}{\beta} \left[f\left(\frac{y}{\beta}\right) + f\left(\frac{y}{\beta} + \frac{1}{2}\right) \right] \Theta\left(\frac{\beta}{2} - y\right) \quad (11)$$

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

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

The density distributions graphed in figure 1 are those functions satisfying

$$[\mathcal{L}_\beta \rho](y) = \rho(y) \quad (12)$$

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

$$\rho(y) = \frac{1}{\beta} \left[\rho\left(\frac{y}{\beta}\right) + \rho\left(\frac{y}{\beta} + \frac{1}{\beta}\right) \right] \Theta\left(\frac{\beta}{2} - y\right) \quad (13)$$

Likewise, the Gelfond-Parry measure of eqn 10 satisfies

$$[\mathcal{M}_\beta v_\beta](y) = v_\beta(y)$$

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

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

$$[\mathcal{L}_\beta \rho](y) = \lambda \rho(y)$$

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 any of the Banach spaces, and in particular, the space of square-integrable functions. Particularly interesting is the space of almost-smooth functions, those having discontinuities at only a countable number of locations, but otherwise being infinitely differentiable. Although the discussion so far implicitly conditions one to restrict oneself to real-valued functions, and to consider only real-valued eigenvalues, this is perhaps too sharp a restriction. It will become clear in the following chapters that even the most basic form of \mathcal{L}_β has a complex-valued spectrum. At any rate, it should be obvious that, whatever the choice of function space, one must have that $\rho(y) = 0$ whenever $\beta < 2y$. This turns out to be a rather harsh condition.

At least one basic fact is known: for at least some kinds of function spaces, the RFP eigenfunction is given by Gel'fond and Parry, as shown in eqn 10. More precisely, it is just the rescaled form $\rho(x) = v(2x/\beta)$ for $x < \beta/2$ and zero otherwise. Changing vocabulary, this is sometimes called the “invariant measure”; as it describes a measure on the unit interval. That is, for the space of all possible measures on the unit interval, the Gel'fond-Parry measure is one of the eigenfunctions of the transfer operator. Some caution is advised here: for the special case of $\beta = 2$, that is, the Bernoulli shift, one has as an eigenfunction the Minkowski measure[27]; it has eigenvalue 1, but is otherwise

quite pathological: it is continuous nowhere, zero on the rationals, and divergent on the rest (i.e. on a “fat” Cantor set). There’s no particular reason to think that this holds only for $\beta = 2$; measures can be, in general, very unusual functions.

A very minor simplification can be achieved with a change of variable. Let $y = \frac{\beta}{2} - \varepsilon$. Then the eigenequation becomes

$$\lambda \beta \rho \left(\frac{\beta}{2} - \varepsilon \right) = \rho \left(\frac{1}{2} - \frac{\varepsilon}{\beta} \right) + \rho \left(1 - \frac{\varepsilon}{\beta} \right)$$

The second term vanishes whenever $\beta/2 < 1 - \varepsilon/\beta$ or $\varepsilon < \beta(1 - \beta/2)$ and so one has the simpler recurrence relation

$$\lambda \rho(y) = \frac{1}{\beta} \rho \left(\frac{y}{\beta} \right) \quad (14)$$

whenever $\beta(\beta - 1) < 2y \leq \beta$.

The equations 13 and 14 can be treated as recurrence relations, defining the $\lambda = 1$ eigenstate. Recursing on these gives exactly the densities shown in figure 1. Computationally, these are much, much cheaper to compute, at least for β much larger than 1, although convergence issues present themselves as β approaches 1. The resulting density may be called the Ruelle-Frobenius-Perron eigenstate; because it can be quickly computed, it provides an alternative view of figure 1, free of stochastic sampling noise.

2.2 The Gelfond-Parry measure

This measure is given by eqn 10 and is given by Gelfond[4] and by Parry[2]. 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 is used to construct general eigenfunctions.

The easiest route seems to be direct verification that eqn 10 is correct, as opposed to a derivation of eqn 10 from first principles. This 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 $t(x) \equiv t_\beta(x) = \beta x \bmod 1$ 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

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

where it is more convenient to write $\varepsilon_n(y) = \Theta(t_n - y)$. The transfer operator \mathcal{M} for the beta-transformation is slightly more convenient to work with than \mathcal{L} for this particular case. It is given by

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

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

$$\begin{aligned} P &= v\left(\frac{y}{\beta}\right) + v\left(\frac{y+1}{\beta}\right)\Theta(\beta - 1 - y) = \sum_{n=0}^{\infty} \frac{1}{\beta^n} \left(\Theta\left(t_n - \frac{y}{\beta}\right) + \Theta\left(t_n - \frac{y+1}{\beta}\right)\Theta(\beta - 1 - y) \right) \\ &= \sum_{n=0}^{\infty} \frac{1}{\beta^n} (\Theta(\beta t_n - y) + \Theta(\beta t_n - 1 - y)) \end{aligned}$$

since $\Theta(ax) = \Theta(x)$ for all constants a , and since $t_n \leq 1$ for all n . These terms simplify, depending on whether t_n is small or large. Explicitly, one has

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

and so

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

These can be collapsed by noting that

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

and so

$$\begin{aligned} P &= \sum_{n=0}^{\infty} \frac{1}{\beta^n} (\Theta(1 - \beta t_n) \Theta(t_{n+1} - y) + \Theta(\beta t_n - 1) (1 + \Theta(t_{n+1} - y))) \\ &= \sum_{n=0}^{\infty} \frac{1}{\beta^n} (\Theta(t_{n+1} - y) [\Theta(1 - \beta t_n) + \Theta(\beta t_n - 1)] + \Theta(\beta t_n - 1)) \\ &= \sum_{n=0}^{\infty} \frac{1}{\beta^n} (\Theta(t_{n+1} - y) + \Theta(\beta t_n - 1)) \\ &= \beta v(y) - \beta + \sum_{n=0}^{\infty} \frac{\Theta(\beta t_n - 1)}{\beta^n} \\ &= \beta v(y) \end{aligned}$$

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 $\mathcal{M}v = v$ as claimed.

2.3 Analytic Gelfond-Parry function

The steps performed above can be repeated verbatim for a “rotated” or “coherent” function

$$v_{\beta;z}(y) = \sum_{n=0}^{\infty} z^n \frac{\Theta(t_n - y)}{\beta^n} \tag{15}$$

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

$$\begin{aligned} [\mathcal{M}v_{\beta;z}] (y) &= \frac{v_{\beta;z}(y)}{z} - \frac{1}{z} + \sum_{n=0}^{\infty} z^n \frac{\Theta(\beta t_n - 1)}{\beta^{n+1}} \\ &= \frac{v_{\beta;z}(y)}{z} + C(\beta; z) \end{aligned}$$

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

$$v_{\beta;z}(x) = \sum_{n=0}^{\infty} \frac{d_n(x)}{\beta^n} z^n \quad (16)$$

where the $d_n(x)$ are exactly the same digits as defined by Parry, just rescaled for the beta-shift. That is,

$$d_n(x) = \varepsilon_n\left(\frac{2x}{\beta}\right) = \Theta\left(\frac{\beta}{2}t_n - x\right) = \Theta\left(T^n\left(\frac{\beta}{2}\right) - x\right) = \begin{cases} 1 & \text{if } x < T^n\left(\frac{\beta}{2}\right) \\ 0 & \text{otherwise} \end{cases}$$

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:

$$[\mathcal{L}_{\beta}v_{\beta;z}] (x) = \frac{1}{z} v_{\beta;z}(x) + C(\beta; z)$$

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 $x < \beta/2$ and is zero otherwise!

The interesting limit is where $|z| \rightarrow \beta$ and so its convenient to re-express C in terms of $\zeta = z/\beta$. With some rearrangements, one obtains

$$E(\beta; \zeta) \equiv \zeta \beta C(\beta; \zeta \beta) = -1 + \zeta \sum_{n=0}^{\infty} \zeta^n d_n\left(\frac{1}{2}\right) \quad (17)$$

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.

2.4 Analytic ergodics

The constant term can be independently derived through a different set of manipulations. Explicitly plugging in eqn 16 into the transfer operator yields

$$\begin{aligned} C(\beta; z) &= [\mathcal{L}_\beta v_{\beta;z}] (y) - \frac{v_{\beta;z}(y)}{z} \\ &= \sum_{n=0}^{\infty} \frac{z^n}{\beta^n} \left[\frac{1}{\beta} \left[d_n \left(\frac{y}{\beta} \right) + d_n \left(\frac{y}{\beta} + \frac{1}{2} \right) \right] \Theta \left(\frac{\beta}{2} - y \right) - \frac{1}{z} d_n(y) \right] \end{aligned}$$

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 numerical 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. 17 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 tempting to give this some silly name, such as the “fundamental theorem of analytic ergodics”.

Let's be exceptionally clear about the meaning: 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 the point, its best to compare side-by-side. Copying equation 5, we had

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

while

$$d_n(x) = \begin{cases} 1 & \text{if } x < T^n\left(\frac{\beta}{2}\right) = T^{n+1}\left(\frac{1}{2}\right) \\ 0 & \text{otherwise} \end{cases}$$

The iterations are running in opposite directions; this is as appropriate, since the transfer operator was a push-forward.

It is useful here to return to the language of cylinder sets, as opposed to point dynamics. Recall that the Borel algebra \mathcal{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 $\mathcal{S} \subset \mathcal{B}$ together with a map $T : \mathcal{S} \rightarrow \mathcal{S}$ that lops off the leading symbol of a given cylinder set but otherwise preserves the subshift: $T\mathcal{S} = \mathcal{S}$. The inverted map T^{-1} is a pushforward, in that it defines the transfer operator, a linear operator $\mathcal{L}_T : \mathcal{F} \rightarrow \mathcal{F}$ on the space \mathcal{F} of all functions $f : \mathcal{S} \rightarrow \mathbb{R}$; explicitly, it is given by $\mathcal{L}_T : f \mapsto f \circ T^{-1}$. 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 point dynamics is that one cannot meaningfully write $T^{-n}(x)$ for a real number, a point x , at least, not without sever 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 $d_n : \mathcal{S} \rightarrow [0, 1]$. Let $\mu : \mathcal{B} \rightarrow [0, 1]$ 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 $A \in \mathcal{S}$, 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 $T^{-1} : \mathcal{S} \rightarrow \mathcal{S}$, and otherwise yield eqn 17.

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 $A \in \mathcal{S}$ 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 \bar{D} .

The pushforward action is then

$$\begin{aligned} v(t^{-1}(A)) &= \sum_{n=0}^{\infty} \zeta^n \mu(\Delta_n \cap t^{-1}(A)) \\ &= \sum_{n=0}^{\infty} \zeta^n [\mu(\Delta_n \cap D \cap t^{-1}(A)) + \mu(\Delta_n \cap \bar{D} \cap t^{-1}(A))] \end{aligned}$$

Two distinct cases emerge. When $t^n(1) < 1/\beta$ then one has that $\Delta_n \cap \bar{D} = \emptyset$. Thus, the second term can be written as

$$\begin{aligned} \mu(\Delta_n \cap \bar{D} \cap t^{-1}(A)) &= \Theta\left(t_n - \frac{1}{\beta}\right) \mu(\Delta_n \cap \bar{D} \cap t^{-1}(A)) \\ &= \Theta\left(t_n - \frac{1}{\beta}\right) \frac{1}{\beta} \mu(\Delta_{n+1} \cap A) \end{aligned}$$

where the second line follows from the first by linearity, and that \bar{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:

$$\begin{aligned} \Theta\left(t_n - \frac{1}{\beta}\right) \mu(\Delta_n \cap D \cap t^{-1}(A)) &= \Theta\left(t_n - \frac{1}{\beta}\right) \mu(D \cap t^{-1}(A)) \\ &= \Theta\left(t_n - \frac{1}{\beta}\right) \frac{1}{\beta} \mu(A) \end{aligned}$$

while

$$\Theta\left(\frac{1}{\beta} - t_n\right) \mu(\Delta_n \cap D \cap t^{-1}(A)) = \Theta\left(\frac{1}{\beta} - t_n\right) \frac{1}{\beta} \mu(\Delta_{n+1} \cap A)$$

Reassembling these pieces and making use of $\Delta_0 \cap A = A$ one gets

$$\begin{aligned} v(t^{-1}(A)) &= \sum_{n=0}^{\infty} \frac{\zeta^n}{\beta} \left[\mu(A) \Theta\left(t_n - \frac{1}{\beta}\right) + \mu(\Delta_{n+1} \cap A) \right] \\ &= \frac{1}{z} v(A) - \frac{\mu(A)}{z} + \mu(A) \sum_{n=0}^{\infty} \frac{\zeta^n}{\beta} \Theta\left(t_n - \frac{1}{\beta}\right) \\ &= \frac{1}{z} v(A) + \frac{\mu(A)}{z} E(\beta; z) \end{aligned}$$

with the constant term as before, in eqn 17:

$$\begin{aligned} E(\beta; z) &= -1 + \zeta \sum_{n=0}^{\infty} \zeta^n \Theta\left(t_n - \frac{1}{\beta}\right) \\ &= -1 + \zeta \sum_{n=0}^{\infty} \zeta^n d_n\left(\frac{1}{2}\right) \end{aligned}$$

As before, one has for $z = 1$ that $E(\beta; 1) = 0$ and so $v \circ t^{-1} = v$ 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 $E(\beta; z)$.

2.5 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 $|\zeta| \rightarrow 1$. 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 $d_n(\frac{1}{2})$ is periodic, and thus $E(\beta; z)$ 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.

β	z	$ z $	$1/z$
1.8	-1.591567859	1.591567859	-0.6283112558
1.8	-1.1962384 +i 1.216022231	1.705783215	-0.4111213835 - i 0.4179206604
1.8	0.9919147363 +i 1.446092984	1.753590535	0.3225655308 - i 0.4702619429
1.6	-1.063651387 +i 1.008959895	1.466067646	-0.4948701876 - i 0.4694246429
1.4	0.550836432 +i 1.178171082	1.300579822	0.3256481633 - i 0.6965211931
1.2	0.9578845659 +i 0.6073301155	1.134192537	0.7446284155 - i -0.4721187476

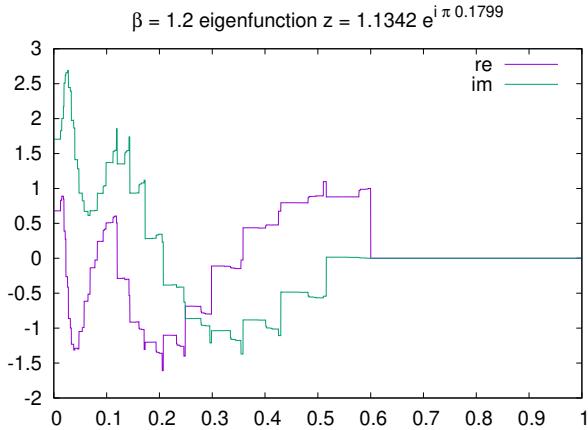
These are not particularly meaningful numbers; they just give a flavor for some locations of eigenvalues. Given a zero, the corresponding eigenfunction is also very easily computed. A typical eigenfunction is shown in figure 7; this is for the zero listed in the last row of the table above. Although 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 $|\zeta| \rightarrow 1$, 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 16. The lower bound is $\lambda = 1/\beta$; if there are eigenfunctions with smaller eigenvalue, they cannot be written in the form of eqn 16. It is possible that these exist: certainly, that is the suggestion from the special case of $\beta = 2$, where there is a discrete spectrum $\lambda = 2^{-n}$ of real-analytic eigenfunctions, given by the Bernoulli polynomials, and a continuous spectrum filling the unit disk, consisting of square-integrable eigenfunctions. The $\beta \rightarrow 2$ limit of eqn 16 does not approach either of these cases. The description seems incomplete.

3 Periodic 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

Figure 7: Typical Eigenfunction



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:

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

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:

$$T_{\beta}^{<}(x) = \begin{cases} \beta x & \text{for } 0 < x < \frac{1}{2} \\ \beta(x - \frac{1}{2}) & \text{for } \frac{1}{2} < x < 1 \end{cases}$$

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 $T_{\beta}^{\leq}(x)$.

All three variants can be considered together, so that the “true” beta shift is taken as the quotient space or identification space[28] 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 non-trivial 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 long-standing problem of exactly how to properly describe the topological conjugacy between different kinds of iterated maps.

3.1 The β -generalized Golden Ratio

The first periodic orbit can be found when $\beta = \varphi = (1 + \sqrt{5})/2$ the Golden Ratio. In this situation, one has that $m_0 = \varphi/2$ and $m_1 = 1/2$. At this location, further iteration breaks down. That is, $m_2 = T_\varphi(m_1)$ can either be taken to be $m_2 = 0$ or $m_2 = m_0$. For lack of a better term, this can be called a “trouble spot”.

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 $m_p = m_0$. For $p = 3$, 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 8. 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 similar suggestion by Dajani[3], numbers of this kind 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; one that is, is given by Hare *et al.*[29] 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[30] 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[13]. 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.

3.2 Counting Orbits

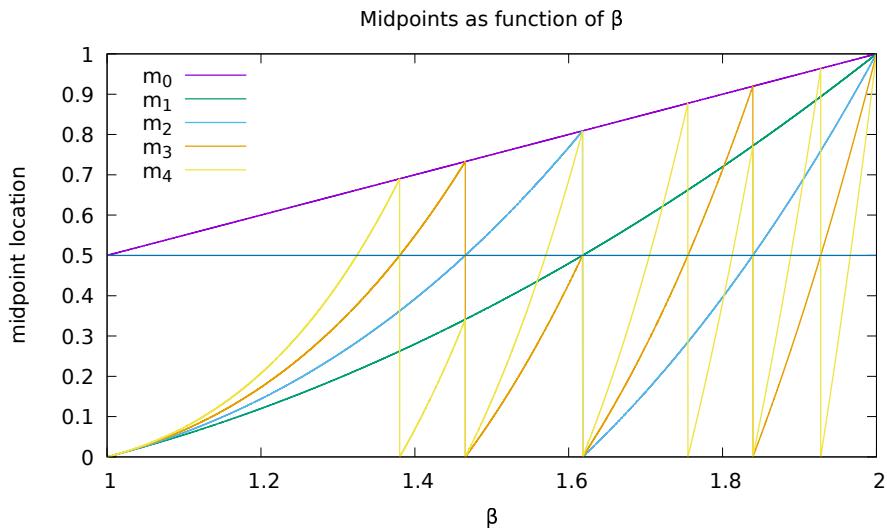
How many trouble spots are there? The table below shows the count M_n of the number of “new” trouble spots, as a function of the midpoint index n .

n	2	3	4	5	6	7	8	9	10	11	12
M_n	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 trouble spots are the positive real roots of polynomials of the form

$$p_{\{b_k\}}(\beta) = \beta^n - \beta^{n-1} - b_1 \beta^{n-2} - b_2 \beta^{n-3} - \cdots - 1 = 0$$

Figure 8: Location of Midpoints



This rather busy image illustrates the location of the first five midpoints, m_0, m_1, \dots, m_4 as a function of β . The locations of the discontinuities are termed “trouble spots”; the first trouble spot occurring for m_2 at $\beta = \varphi$. The midpoint m_3 has two new trouble spots at $\beta = 1.465\dots$ and $\beta = 1.839\dots$; the trouble spot at $\beta = \varphi$ being disallowed, as it already lead to a termination of midpoint iteration. The midpoint m_4 has three new trouble-spots. That is, trouble spots arise on a pruned binary tree: one generally expects new spots to occur on the left and right of the immediate predecessor, unless they’re already been encountered before.

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.

Although the digits b_k must be zero or one, this definition of irreducibility, plus the counting, suggests some relationship to the irreducible polynomials over the field \mathbb{F}_2 , as that is what the definition of OEIS A001037 counts. Yet the relationship, if any, is quite unclear.²

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

$$\frac{t}{1-t} = \sum_{n=1}^{\infty} n M_n \frac{t^n}{1-t^n}$$

which has a radius of convergence $|t| < 1/2$. For large n , the asymptotic behavior can be trivially deduced from the defining sum:

$$M_n = \frac{2^n}{n} - \mathcal{O}\left(\frac{2^{n/2}}{n}\right)$$

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.[31, 32] A number of other results exist, including [33, 34]. At any rate, a closer study of the polynomials seems to be called for.

²A hypothesis presented in a later section suggests that each orbit should be thought of as a Galois group, with the length of the orbit corresponding to the number of elements in the Galois group. It seems that this might explain much of the structure.

3.3 β -Golden Polynomials

The “trouble spots” occur whenever the k ’th iterate $m_k = T_\beta^k(m_0)$ lands on the starting midpoint $m_k = m_0$. 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 , and of the form $p_{\{b_k\}}(\beta)$. These must be arranged in the manner such that $p_{\{b_k\}}(\beta) = 0$ at each discontinuity, as illustrated in figure 8. This limits the polynomials that can appear; it limits the possible coefficients $\{b_k\}$; not all bit-sequences appear.

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 its real root is bracketed by the roots of the earlier iterates. Adopting an ordinal numbering, $p_n(\beta)$ must have the form

$$p_n(\beta) = \begin{cases} \beta(p_{n/2}(\beta) + 1) - 1 & \text{for } n \text{ even} \\ \beta p_{(n-1)/2}(\beta) - 1 & \text{for } n \text{ odd} \end{cases}$$

The roots must be bracketed 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 relationship is rather awkwardly expressed in the following pseudo-code. Here, r_n is the root $p_n(r_n) = 0$. The polynomial p_n is included in the list if it is the case that this pseudo-code does not fail:

```

mprev := n
m := ⌊n/2⌋
while (0 < m)
    mprev is even and rm < rn then fail
    mprev := m
    m := ⌊m/2⌋

```

The above is a rather awkward way of stating that roots must be bracketed by pairs of previous roots. It can perhaps be more easily understood by studying the location of the discontinuities in figure 8: new discontinuities at higher orders must occur before earlier ones.

Thus, 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 $1.324717957244746\dots$, 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.

The polynomial $p_5 = \beta^4 - \beta^3 - \beta - 1$ is excluded because it has $\varphi = 1.618\dots$ as a root, which was previously observed by p_1 . The polynomial $p_9 = \beta^5 - \beta^4 - \beta - 1$ is excluded because its root, $r_9 = 1.497094048762796\dots$ is greater than its predecessor r_2 ; the recursive algorithm does not compare to r_4 . Note that p_9 is relatively prime to the earlier polynomials, so irreducibility is not a sufficient criterion.

There are other ways to start the table, and to index the polynomials. The given indexing has the property that $2n+1$, taken as binary, encodes the coefficients of the

polynomial. The order of the polynomial is $\lceil \log_2(2n+1) \rceil$. The index n itself encodes the orbit of the midpoint. That is, writing $n = b_0b_1b_2\cdots b_p$ for binary digits b_k , then $T_\beta^k(\beta/2) < 1/2$ if and only if $b_k = 0$. Note that $b_0 = 1$ always corresponds to $1/2 < \beta/2$ always. By convention, the last digit is always 1, also.

order	$p_n(\beta)$	n	binary	root
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	1.7548776662466924 \cdots
	$\beta^4 - \beta^3 - \beta^2 - \beta - 1$	7	1111	1.9275619754829252 \cdots
5	$\beta^5 - \beta^4 - 1$	8	10001	1.324717957244746 \cdots
	$\beta^5 - \beta^4 - \beta^2 - 1$	10	10101	1.5701473121960547 \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 polynomials 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 is not currently known to OEIS.

order	sequence
5	8,10,12,13,14,15
6	16,20,24,25,26,28,29,30,31
7	32,36,40,42,44,48,49,50,52,53,54,56,57,58,59,60,61,62,63

Although there are as many of these polynomials as there are Lyndon words, there is no obvious way to write a bijection between the two. It is almost possible to do so by writing $2n$ in binary, and then reversing the order of the bits, left-to-right. One almost gets the Lyndon words in the correct order, except “in the middle”: so, for example, in the table above, one can get the Lyndon order by exchanging 10 with 12, and 13 with 14. But the table above cannot be re-ordered: the given ordering encodes the orbit of the midpoint. Apparently, although a given orbit can be cyclically rotated to obtain a Lyndon word, the initial segment of the orbit is not a Lyndon word itself.

Although the necklace polynomial also describes the number of elements of length n in a Hall set, there is no obvious correspondence. That is, if one writes down the Hall set in two generators, one does not obtain the binary strings given in the table above.

Questions that present themselves include:

- Is there a generating function for the sequence of allowed values of n ? What is it?

- How long is the initial segment of each periodic orbit, before the orbit attains its Lyndon word form? What are the values of β where the initial orbits are not in Lyndon form?

3.4 Distribution of β -Golden Roots

The location and distribution of the roots can be visualized in several ways. One direct technique is via the normalization of the Parry-Gelfond measure. The function in eqn 10 or more generally 15 can be integrated in a straight-forward 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 9. Each discontinuity corresponds to the real root of one of the polynomials. The first few are labelled by the integer labels from the previous table. The labels follow an obvious doubling sequence. Each doubling sequence has a leader. The leaders appear to form a sequence 1,3,7,10,13,15,25,29,31,... The leaders are easily obtained: they are what's left of the earlier polynomial-labelling sequence, after knocking out power-of-two sequences. This sequence is not currently known to OEIS. The leaders above 1 follow a simple pattern: they are always at $2n + 1$ of the previous dominant spike. So, for example, the first leader above 6 is 13. The leaders below 1 are given by $n(2n + 1)$, so for example, the leader above 16 is $16 \times 33 = 528$.

It seems natural to assume that the real roots have some distribution. This seems not to be the case. Figure 10 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 $k \rightarrow \infty$, almost all $p_n(\beta)$ have roots that approach 2. Although the roots appear to be dense in $1 < \beta < 2$, essentially all of the weight of that density is at 2. Since the roots are countable, the density clearly becomes very thin.

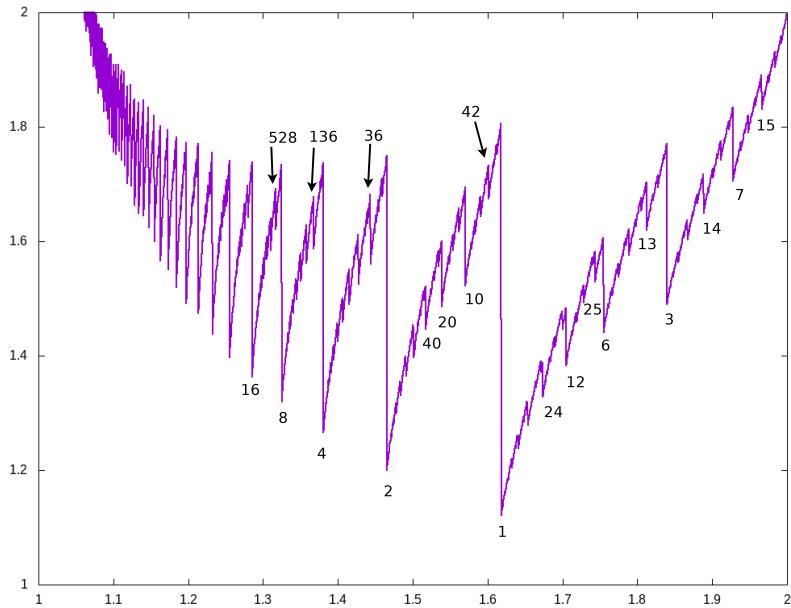
The local distribution of roots can be sensed from the figure 11, which visualizes the distance between neighboring roots.

3.5 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.

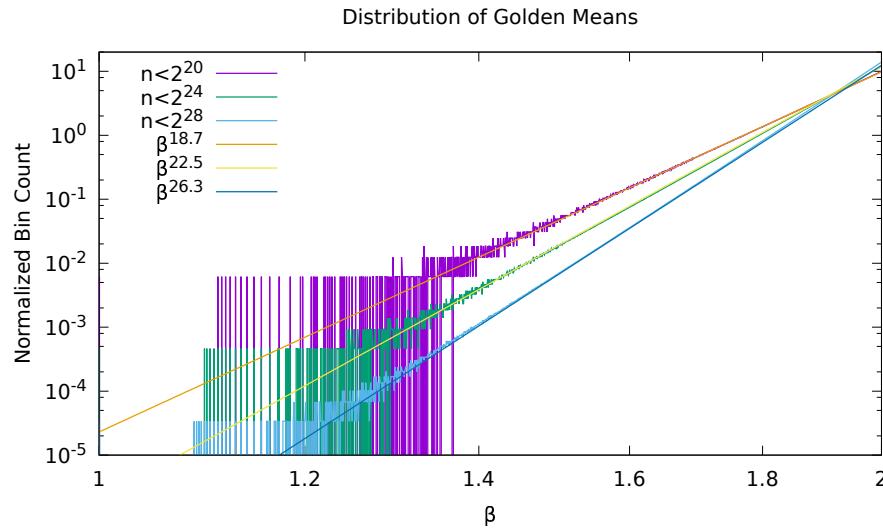
These numerical results can be argued heuristically: just divide the polynomial by

Figure 9: Normalization Integral



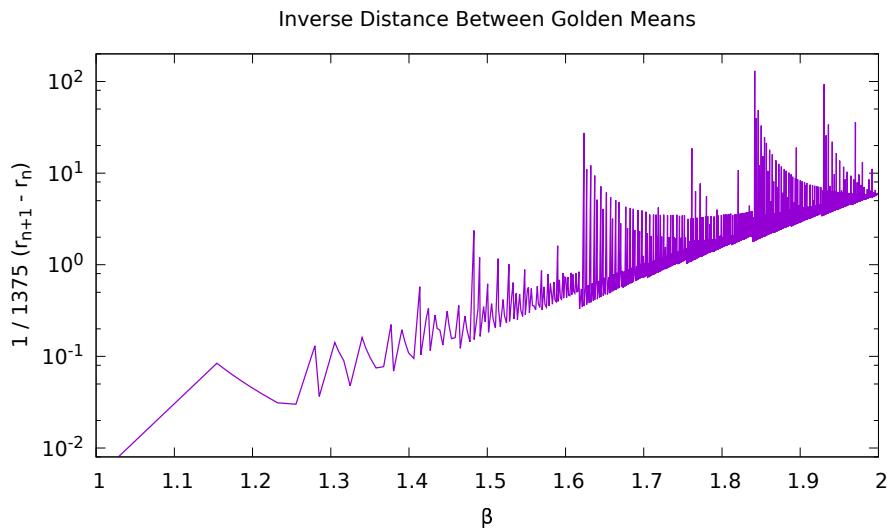
This figure shows the integral $I(\beta) = \sum_{n=0}^{\infty} \beta^{-n} T^n \left(\frac{\beta}{2} \right)$ with $1 < \beta \leq 2$ running along the horizontal axis. Each discontinuity corresponds to the location of a real root of one of the β -Golden polynomials. Some of these are manually labelled by integers, corresponding to the polynomial labels from the previous polynomial table.

Figure 10: 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 $k \rightarrow \infty$ limit accumulates all the measure at $\beta = 2$.

Figure 11: 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.

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 $n = b_0 b_1 b_2 \cdots b_p$ corresponding to a terminating orbit. Dividing by z^{k+1} gives a series

$$1 - \frac{1}{z} - \frac{b_1}{z^2} - \frac{b_2}{z^3} - \dots$$

Clearly, this can have a zero only when $|z| < 2$ as otherwise, the terms get too small too quickly.

3.6 β -Golden β -Fibonacci Sequences

It is well known that the golden ratio occurs as limit of the ratio of adjacent Fibonacci numbers:

$$\varphi = \lim_{m \rightarrow \infty} \frac{F_m}{F_{m-1}}$$

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

$$\alpha_n = \lim_{m \rightarrow \infty} \frac{F_m^{(n)}}{F_{m-1}^{(n)}}$$

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 $\{b\} = b_0, b_1, \dots, b_k$, 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

$$\alpha^{\{b\}} = \lim_{m \rightarrow \infty} \frac{F_m^{\{b\}}}{F_{m-1}^{\{b\}}}$$

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,

$$\sum_{m=0}^{\infty} z^m F_m^{\{b\}} = \frac{z^k}{1 - b_0 z - b_1 z^2 - \cdots - b_k z^{k+1}} = \frac{1}{z p_n(\frac{1}{z})}$$

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 $\{b_0, 2b_1, 3b_2, \dots, (k+1)b_k\}$. 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, $\{b_0, 2b_1, 3b_2, \dots, (k+1)b_k\} = \{1 \cdot 1, 2 \cdot 1, 3 \cdot 0, 4 \cdot 0, 5 \cdot 1\} = \{1, 2, 5\}$. 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	$\varphi = \frac{1+\sqrt{5}}{2} = 1.618\dots$	golden ratio	Fibonacci A000045
2	101	1.465571231876768\dots	OEIS A092526	Narayana A000930
3	111	1.839286755214161\dots	tribonacci A058265	tribonacci A000073
4	1001	1.380277569097613\dots	2nd pisot A086106	A003269
6	1101	1.754877666246692\dots	OEIS A109134	A060945
7	1111	1.927561975482925\dots	tetranacci A086088	tetranacci A000078
8	10001	1.324717957244746\dots	silver A060006	A003520
10	10101	1.570147312196054\dots	pisot A293506	A060961
12	11001	1.704902776041646\dots		A079971
13	11011	1.812403619268042\dots		A079976
14	11101	1.888518845484414\dots		A079975
15	11111	1.965948236645485\dots	pentanacci A103814	A001591

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, there are various known relations for the “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 roots of 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[29]:

$$\frac{1}{\alpha_k} = \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{\infty} \frac{1}{j} \binom{j(k+1)}{j-1} \frac{1}{2^{j(k+1)}}$$

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

$$r_3 = \frac{1}{3} \left(1 + \sqrt[3]{19 + 3\sqrt{33}} + \sqrt[3]{19 - 3\sqrt{33}} \right) \simeq 1.839 \dots$$

The silver number (plastic number) r_8 is given by

$$r_8 = \frac{1}{6} \left(\sqrt[3]{108 + 12\sqrt{69}} + \sqrt[3]{108 - 12\sqrt{69}} \right) \simeq 1.324 \dots$$

The Narayana's cows number r_2 is given by

$$r_2 = \frac{1}{6} \sqrt[3]{116 + 12\sqrt{93}} + \frac{2}{3\sqrt[3]{116 + 12\sqrt{93}}} + \frac{1}{3} \simeq 1.645 \dots$$

The root r_6 is related to the silver number r_8 as $r_8 = r_6(r_6 - 1)$ and is given by

$$r_6 = \frac{1}{6} \sqrt[3]{108 + 12\sqrt{69}} + \frac{2}{\left(\sqrt[3]{108 + 12\sqrt{69}} \right)^2} \simeq 1.754 \dots$$

Do the other roots have comparable expressions? To obtain them, is it sufficient to articulate the theory of “complex multiplication” on elliptic curves? The appearance of only the cube and square roots is certainly suggestive of an underlying process of points on elliptic curves.

3.7 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 a $(k+1) \times (k+1)$ matrix in lower-Hessenberg form:

$$B = \begin{bmatrix} b_0 & 1 & 0 & 0 & \cdots & 0 \\ b_1 & 0 & 1 & 0 & \cdots & 0 \\ b_2 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ b_{k-1} & 0 & 0 & 0 & \cdots & 1 \\ b_k & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (18)$$

That is, the m 'th element of the sequence is obtained from the m 'th iterate B^m . That such iteration results in integer sequences has long been observed in the theory of continued fractions. It's useful to work an explicit example. For the golden ratio, one has

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

and the iterates are

$$B^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, B^3 = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}, B^4 = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}, B^n = \begin{bmatrix} F_n & F_{n-1} \\ F_{n-1} & F_{n-2} \end{bmatrix}$$

with F_n being the n 'th Fibonacci number, as usual. For the general case, one has that

$$B^m = \begin{bmatrix} F_m^{\{b\}} & F_{m-1}^{\{b\}} & F_{m-2}^{\{b\}} & \cdots & F_{m-k+1}^{\{b\}} & F_{m-k}^{\{b\}} \\ F_{m-1}^{\{b\}} & F_{m-2}^{\{b\}} & F_{m-3}^{\{b\}} & \cdots & F_{m-k}^{\{b\}} & F_{m-k-1}^{\{b\}} \\ F_{m-2}^{\{b\}} & F_{m-3}^{\{b\}} & & \cdots & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{m-k+1}^{\{b\}} & & \cdots & & & \\ F_{m-k}^{\{b\}} & & \cdots & & & F_{m-2k}^{\{b\}} \end{bmatrix}$$

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 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.

3.8 Equivalent labels for orbits

At this point, it should be clear that there are several equivalent ways of labeling the expressions under consideration. These are recapped here. Proofs are omitted; they are straight-forward.

3.8.1 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 Parry sequence; namely $T_\beta^p(\beta/2) = (\beta/2)t_\beta^p(1)$, recalling here the notation of eqn 8 and 10.

3.8.2 Orbit encoding

The midpoint generates a unique sequence of bits $\{b_0, b_1, \dots, b_k, \dots\}$ 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 .

3.8.3 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 .

3.8.4 Polynomial numbering

If the orbit is periodic, then there exists a polynomial $p_n(z) = z^{k+1} - b_0 z^k - b_1 z^{k-1} - \dots - b_{k-1} z - 1$ with $k = 1 + \lfloor \log_2 n \rfloor$ being the length of the orbit. The positive real root of $p_n(z)$ is β . 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 $2n + 1 = \sum_{j=0}^k 2^j b_j$.

3.8.5 Integer sequences

If the orbit is periodic, 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 $\{b\} = b_0, b_1, \dots, b_k$, and with the value of β .

3.8.6 Shift matrix

If the orbit is periodic, then the finite bit sequence $\{b\} = b_0, b_1, \dots, b_k$ defines a lower-Hessenberg “golden shift” matrix B , as shown in eqn 18.

3.8.7 Summary

To summarize: any one of these: the integer n , the polynomial $p_n(x)$, 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. Specifying one determines the others; they all uniquely map to one-another. They are all equivalent labels. Fashionably abusing notation, $n \equiv p_n(x) \equiv \{b\} \equiv F_m^{\{b\}} \equiv m_p \equiv w(\beta) \equiv \beta \equiv B$.

An explicit expression relating the orbit encoding and the orbit can be read off directly from eqn 7. Plugging in,

$$m_p = T_\beta^{p+1} \left(\frac{\beta}{2} \right) = \frac{\beta}{2} \left[\beta^{p+1} - \sum_{j=0}^p b_j \beta^{p-j} \right] \quad (19)$$

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 periodic orbits. Three of these can be extended to non-periodic 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-periodic orbits. Indeed, there are too many: the non-periodic orbits are uncountable.

3.9 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:

$$\beta = \lim_{m \rightarrow \infty} \frac{F_m^{\{b\}}}{F_{m-1}^{\{b\}}}$$

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 periodic 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 (non-recurring) bit-sequence $\{b\} = 10101001010010100000100\dots$. The generated integer sequence is $F_m^{\{b\}} = 1, 1, 1, 2, 3, 5, 8, 12, 20, 32, 51, 82, 130, 209, 335, 535, \dots$ 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 $535/335 = 1.597\dots$ and the previous is $335/209 = 1.603\dots$ and the ratio of successive elements eventually converges to 1.6. By comparison, the Fibonacci sequence is generated by the bit-string 10101010... 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 periodic or not. This dual-representation of periodic 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 $1.000\dots = 0.999\dots$ as an example. So likewise here, one can toggle between finite and infinite-periodic strings. So, given a finite string $\{b\} = b_0, b_1, \dots, b_{k-1}, b_k$

which has, by definition, $b_k = 1$, create a new finite string that is twice as long: $\{b'\} = b_0, b_1, \dots, b_{k-1}, 0, b_0, b_1, \dots, b_k$ 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.

3.10 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

$$p_n(z) = z^{k+1} \left(1 - b_0 z^{-1} - b_1 z^{-2} - \dots - b_k z^{-k-1} \right)$$

Set $\zeta = 1/z$ to get

$$\zeta^{k+1} p_n \left(\frac{1}{\zeta} \right) = 1 - b_0 \zeta - b_1 \zeta^2 - \dots - b_k \zeta^{k+1}$$

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 17, 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.

3.11 β -Hessenberg operator

Extending the golden shift matrix B of eqn 18 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 $\det(B - \lambda I) = p_n(\lambda) = 0$; 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 $\mathcal{H}_\beta = B^T/\beta$ is in upper-Hessenberg form. Rescaled in this way, it now seems safe to declare, by fiat, that the operator \mathcal{B}_β is the correct extension of the matrix B to infinite dimensions. Just to be explicit: given the bit-sequence $\{b\}$, the operator \mathcal{H}_β has the matrix elements

$$\begin{aligned}\langle 0 | \mathcal{H}_\beta | j \rangle &= \frac{b_j}{\beta} \\ \langle j+1 | \mathcal{H}_\beta | j \rangle &= \frac{1}{\beta}\end{aligned}$$

with all other entries being zero. This is clearly in upper-Hessenberg form, with the subdiagonal providing the shift.

The invariant measure is then the Frobenius-Perron eigenvector σ solving $\mathcal{H}_\beta \sigma = \sigma$. It is easy to write down σ explicitly: $\sigma = (1, \beta^{-1}, \beta^{-2}, \dots)$, that is, $\sigma_j = \beta^{-j}$. This is obviously so: the subdiagonal entries of \mathcal{H}_β act as a shift on σ and the top row is just

$$1 = \sum_{j=0}^{\infty} \langle 0 | \mathcal{H}_\beta | j \rangle \sigma_j = \sum_{j=0}^{\infty} b_j \beta^{-j-1} = 1 - q^{\{b\}} \left(\frac{1}{\beta} \right) = 1$$

Although \mathcal{H}_β is not triangular, (and thus not “solvable”), it does have a dramatically simple form. It also re-affirms the Ansatz 16 for the eigenfunctions. To be explicit: if v is a vector satisfying $\mathcal{H}_\beta v = \lambda v$, 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, $[\mathcal{L}_\beta v](x) = \lambda v(x)$, or, explicitly:

$$\frac{1}{\beta} \left[v\left(\frac{x}{\beta}\right) + v\left(\frac{x}{\beta} + \frac{1}{2}\right) \right] \Theta\left(\frac{\beta}{2} - x\right) = \lambda v(x) \quad (20)$$

which is just eqn 11. So, for $\lambda = 1$, this is just $v = \sigma$ which is just eqn 16 for $z = 1$, the invariant measure, as always. But it also says more: the *only* solutions to $\mathcal{H}_\beta v = \lambda v$ are necessarily of the form $v = (1, (\lambda\beta)^{-1}, (\lambda\beta)^{-2}, \dots)$, because the subdiagonal forces this shift. To satisfy the top row of \mathcal{H}_β , one must have that

$$\lambda = \sum_{j=0}^{\infty} \langle 0 | \mathcal{H}_\beta | j \rangle v_j = \frac{1}{\beta} \sum_{j=0}^{\infty} \frac{b_j}{(\lambda\beta)^j} = \lambda \left(1 - q^{\{b\}} \left(\frac{1}{\lambda\beta} \right) \right) = \lambda$$

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: periodic orbits have an associated shift matrix B ; this extends naturally to a shift operator \mathcal{H}_β for non-periodic 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 states³. 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 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 20 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 20 holds for any x .

3.12 Eigenfunctions from periodic orbits

To recap: eigenstates of the transfer operator correspond with the zeros of $q^{\{b\}}(\zeta)$, or, more precisely, the zeros for which $|\zeta| \leq 1$. 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...

3.12.1 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

$$q^{\{1010101\cdots\}}(\zeta) = 1 - \zeta - \zeta^3 - \zeta^5 - \cdots = (1 - \zeta - \zeta^2) / (1 - \zeta^2)$$

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

$$v(x) = \sum_{m=0}^{\infty} \frac{d_m(x)}{\varphi^m} = \begin{cases} \varphi & \text{for } 0 \leq x < \frac{1}{2} \\ 1 & \text{for } \frac{1}{2} \leq x < \varphi \\ 0 & \text{for } \varphi \leq x \end{cases}$$

There is a negative real zero at $\zeta = -\varphi$, but the eigenfunction summation is not convergent here.

³The 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.

3.12.2 Case n=2

The $n = 2$ case has the finite bit-string $\{b\} = 101$ and the periodic bit-string $\{b\} = 1001001\dots$. The corresponding finite beta-polynomial is $q^{\{101\}}(\zeta) = 1 - \zeta - \zeta^3$; the infinite series is

$$q^{\{1001\dots\}}(\zeta) = 1 - \zeta - \zeta^4 - \zeta^7 - \dots = (1 - \zeta - \zeta^3) / (1 - \zeta^3)$$

which has a positive real zero at $\zeta = 1/\beta = 0.6823\dots$ and three poles on the unit circle. The FP eigenvalue provides $\beta = 1.4655\dots$. The invariant measure is

$$v(x) = \sum_{m=0}^{\infty} \frac{d_m(x)}{\beta^m} = \begin{cases} \frac{\beta}{\beta-1} & \text{for } 0 \leq x < T\left(\frac{\beta}{2}\right) \\ \frac{1}{\beta-1} & \text{for } T\left(\frac{\beta}{2}\right) \leq x < \frac{1}{2} \\ \frac{1/\beta}{\beta-1} & \text{for } \frac{1}{2} \leq x < \beta \\ 0 & \text{for } \beta \leq x \end{cases}$$

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 $\beta = 1/\beta(\beta - 1)$.

3.12.3 Case n=3

The $n = 3$ case has the finite bit-string $\{b\} = 111$ and the periodic bit-string $\{b\} = 1101101\dots$. The corresponding finite beta-polynomial is $q^{\{111\}}(\zeta) = 1 - \zeta - \zeta^2 - \zeta^3$; the infinite series is

$$q^{\{110110\dots\}}(\zeta) = 1 - \zeta - \zeta^2 - \zeta^4 - \dots = (1 - \zeta - \zeta^2 - \zeta^3) / (1 - \zeta^3)$$

which has a positive real zero at $\zeta = 1/\beta = 0.5436\dots$ and three poles on the unit circle. The FP eigenvalue gives $\beta = 1.8392\dots$. The invariant measure is

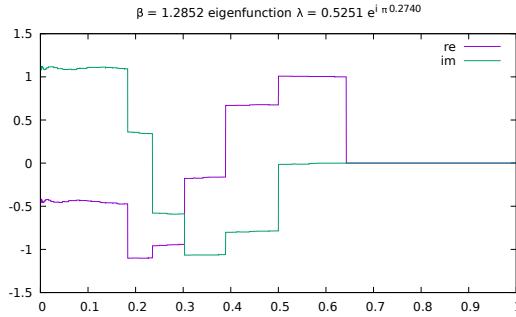
$$v(x) = \begin{cases} \frac{\beta}{\beta-1} & \text{for } 0 \leq x < \frac{1}{2} \\ \beta & \text{for } \frac{1}{2} \leq x < T\left(\frac{\beta}{2}\right) \\ \frac{1}{\beta-1} & \text{for } T\left(\frac{\beta}{2}\right) \leq x < \beta \\ 0 & \text{for } \beta \leq x \end{cases}$$

3.12.4 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 bit-strings are those previously listed in tables; they are $\{b\} = 1001$, $\{b\} = 1101$ and $\{b\} = 1111$. The infinite series is $q^{\{b\dots\}}(\zeta) = q^{\{b\}}(\zeta) / (1 - \zeta^4)$. 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.

3.12.5 Case n=16

The $n = 16$ polynomial is the first one to have complex zeros inside the unit disk. The finite bit-string 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 $\zeta = 0.965709509\cdots \exp \pm i\pi 0.2740452363\cdots$ which corresponds to eigenvalues are $\lambda = 0.525107\cdots \pm i0.611100\cdots = 0.805718\cdots \exp \pm i\pi 0.274045\cdots$. The corresponding eigenfunction is shown immediately below.



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.

3.12.6 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 $T^m(\frac{1}{2})$ for $m = \{0, \dots, k-1\}$ (so that $T^0(\frac{1}{2}) = \frac{1}{2}$ and $T^1(\frac{1}{2}) = \frac{\beta}{2}$, and so on).
- The left-most 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.

3.13 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

$$p_2(x) = x^3 - x - 1 = (x - r)(x^2 + xp_0(r) + rp_0(r))$$

while

$$p_3(x) = x^3 - x^2 - x - 1 = (x - r)(x^2 + xp_0(r) + p_1(r))$$

Continuing in this way, there are three order-4 polynomials. They factor as

$$p_7(x) = x^4 - x^3 - x^2 - x - 1 = (x - r)(x^3 + x^2p_0(r) + xp_1(r) + p_3(r))$$

and

$$p_6(x) = x^4 - x^3 - x^2 - 1 = (x - r)(x^3 + x^2p_0(r) + xp_1(r) + rp_1(r))$$

and (noting that there is no p_5 that occurs in the series)

$$p_4(x) = x^4 - x^3 - 1 = (x - r)(x^3 + x^2p_0(r) + xrp_0(r) + r^2p_0(r))$$

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

$$p_2(x) = q_{101}(x) = (x - r)(x^2 + xq_1 + q_{01})$$

where, by definition, $q_{01}(x) \equiv rq_1(x)$. Next,

$$p_3(x) = q_{111}(x) = (x - r)(x^2 + xq_1 + q_{11})$$

is the second factorization of order 3. For order 4, one has

$$p_4(x) = q_{1001}(x) = (x - r)(x^3 + x^2q_1 + xq_{01} + q_{001})$$

where, this time, $q_{001}(x) = xq_{01}(x) = x^2q_1(x)$. Continuing,

$$p_6(x) = q_{1101}(x) = (x-r)(x^3 + x^2q_1 + xq_{11} + q_{011})$$

where, by definition, $q_{011}(x) \equiv xq_{11}(x)$. Finally,

$$p_7(x) = q_{1111}(x) = (x-r)(x^3 + x^2q_1 + xq_{11} + q_{111})$$

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

$$\begin{aligned} p_{12}(x) &= q_{11001}(x) = (x-r)(x^4 + x^3p_0(r) + x^2p_1(r) + xrp_1(r) + r^2p_1(r)) \\ &= (x-r)(x^4 + x^3q_1 + x^2q_{11} + xq_{011} + q_{0011}) \end{aligned}$$

The general pattern should now be clear. Given one of the admissible bit sequences $b_0b_1b_2 \dots b_{k-1}b_k$ and recalling that $b_k = 1$ always, (and that $b_0 = 1$ always) one has

$$p_n(z) = q_{b_0b_1b_2 \dots b_{k-1}b_k}(z) = z^{k+1} - b_0z^k - b_1z^{k-1} - \dots - b_{k-1}z - 1$$

which has the factorization, with bits reversed:

$$q_{b_0b_1b_2 \dots b_{k-1}b_k}(z) = (z-r)\left(z^k + z^{k-1}q_{b_0} + z^{k-2}q_{b_1b_0} + z^{k-3}q_{b_2b_1b_0} + \dots + q_{b_{k-1}b_{k-2} \dots b_1b_0}\right)$$

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

$$q_{b_0b_1b_2 \dots b_{k-1}b_k}(z) = (z-r)(z-s)\left(z^{k-1} + (s+q_{b_0})z^{k-2} + (s^2 + sq_{b_0} + q_{b_0b_1})z^{k-3} + \dots\right)$$

The coefficient of the next term being $s^3 + s^2q_{b_0} + sq_{b_0b_1} + q_{b_0b_1b_2}$ and so on. From this point one, 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 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 re-arrangement of the polynomial labels. A generic polynomial factorization looks like the below, for comparison. If

$$p(x) = x^{n+1} + c_0x^n + c_1x^{n-1} + \dots + c_n$$

and if r is a root of $p(x)$ viz $p(r) = 0$ then

$$\begin{aligned} p(x) &= (x-r)(x^n + (r+c_0)x^{n-1} + (r^2 + c_0r + c_1)x^{n-2} + \dots) \\ &= (x-r)(x^n + a_0x^{n-1} + a_1x^{n-2} + \dots) \end{aligned}$$

with

$$a_k = r^{k+1} + \sum_{j=0}^k c_j r^{k-j}$$

There are some notable values occurring in the factorization. These are shown in the table below:

o	n	bin	root r	q polynomial	OEIS	root of
2	1	11	$\varphi = 1.618 \cdot$	$q_1 = 0.618 \dots$		
3	2	101	$1.465571 \cdot$	$q_1 = 0.46557123187676 \dots$	A088559	$q^3 + 2q^2 + q - 1$
				$q_{01} = 0.68232780382801 \dots$	A263719	$q^3 + q - 1$
	3	111	$1.839286 \cdot$	$q_1 = 0.83928675521416 \dots$ $q_{11} = 0.54368901269207 \dots$	A192918	$q^3 - 2q^2 - 2$ $q^3 + q^2 + q - 1$
4	4	1001	$1.380277 \cdot$	$q_1 = r - 1$ $q_{01} = 0.52488859865640 \dots$ $q_{001} = 0.72449195900051 \dots$	A072223	$q^4 - 2q^2 - q + 1$
				$q_1 = r - 1$ $q_{11} = 0.32471795724474 \dots$	A075778	$q^3 + q^2 - 1$ silver - 1
				$q_{011} = 0.56984029099805 \dots$		
	6	1101	$1.7548776 \cdot$	$q_1 = r - 1$ $q_{11} = 0.78793319384471 \dots$		
				$q_{111} = 0.51879006367588 \dots$		

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.

3.14 Beta odometer

The periodic orbits can be assigned a label, in the style of continued fractions, shown in figure 12. The goal of the label is to provide a unique identifier for each discontinuity on the graph; yet one that is easy to read off by visual examination. Recall that each discontinuity corresponds to a periodic orbit, and so these labels apply to the periodic orbits.

This labeling scheme is hopefully apparent. The largest discontinuity is given a label of [0]. It is the leader of a period-doubling sequence, given labels [1], [2], and so on. The largest discontinuity between $[n]$ and $[n-1]$ is given the label $[n,0]$ and it leads a sequence $[n,1]$, $[n,2]$, and so on. At the right-most end of the graph, at $\beta = 2$, the appropriate label appears to be $[-1]$. With this observation, the labeling is completely regular. Each interval brackets a sequence of discontinuities, ad infinitum.

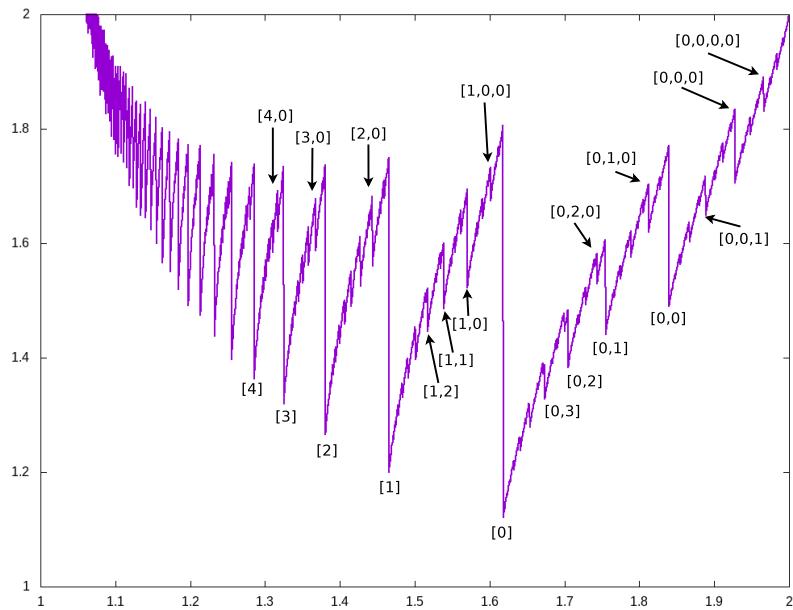
The bracketing relationship can be written as

$$[m_1, m_2, \dots, m_{k-1}] \Rightarrow [m_1, m_2, \dots, m_k] \Leftarrow [m_1, m_2, \dots, m_p - 1]$$

with where p is the largest integer $1 < p \leq k-1$ for which $m_p \neq 0$. The notation here is that each label is brackets by a pair of shorter labels, on the left, and on the right. Each label corresponds to a specific β value; write $\beta = \beta([m_1, m_2, \dots, m_k])$ for this value. The bracketing relationship states that

$$\beta([m_1, m_2, \dots, m_{k-1}]) < \beta([m_1, m_2, \dots, m_k]) < \beta([m_1, m_2, \dots, m_p - 1])$$

Figure 12: Discontinuity Labels



This figure shows the integral $I(\beta) = \sum_{n=0}^{\infty} \beta^{-n} T^n\left(\frac{\beta}{2}\right)$ with $1 < \beta \leq 2$ 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 “continued-fraction”-style labelling.

with the inequalities being strict. Each interval is of decreasing size, limiting to zero. One can thus conclude that infinite sequences $[m_1, m_2, \dots]$ are unique, and correspond one-to-one with the bracketed β value. That is, there is a bijection between the non-periodic β and the infinite length labels.

There is a slight ambiguity for the labeling of the periodic β values. Apparently, one can make the identification

$$[-1] = [0, 0, \dots]$$

in that $\beta = 2$ appears as the limit of ever-lengthening sequences $[0, 0, \dots, 0]$. Closer examination reveals that each discontinuity can be labeled this way. That is, each finite-length label has a corresponding infinite-length label:

$$[m_1, m_2, \dots, m_k] = [m_1, m_2, \dots, m_k + 1, -1] = [m_1, m_2, \dots, m_k + 1, 0, 0, \dots]$$

From this, we can easily conclude that there is a bijection between each infinite length of strings of (non-negative) integers, and β values; they are in one-to-one correspondence. (It is not hard to see that every possible sequence occurs once and only once.)

As each such β corresponds to a periodic orbit, it is the one real root of one of the beta polynomials p_n for some n ; equivalently, it is the convergent of one of the β -Fibonacci sequences, for that same n . That value is given by a recursive formula:

$$n = [m_1, m_2, \dots, m_k] = 2^{m_k + m_1 + m_2 + \dots + m_{k-2}} (2[m_1, m_2, \dots, m_{k-1}] + 1)$$

which is anchored by $[m] = 2^m$. These are the integer labels appears on figure 9.

The sequences can be generated in proper order via a simple recursive algorithm, given here.

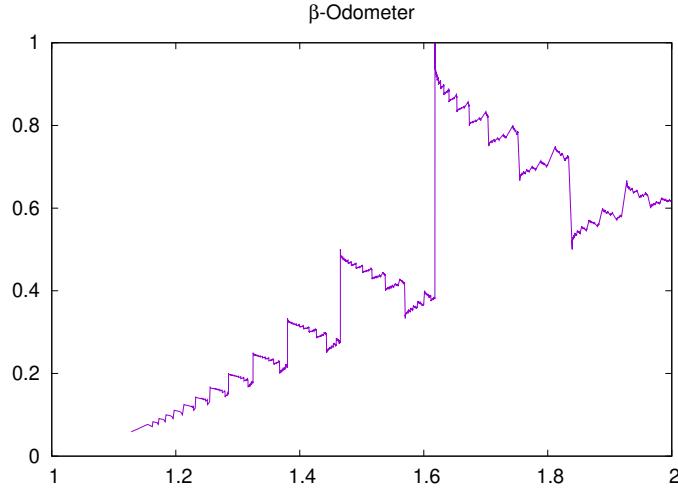
```
# Print sequences in descending beta order
# The 'sequence' is the list [m_1, m_2, ... , m_k]
# The 'maxlength' parameter limits recursion to short sequences (max k)
# The 'maxdepth' parameter limits the largest value of each m_j
define iterate (sequence, maxlength, maxdepth)

    # First, iterate to maximum length ,
    # appending a zero to the end of the sequence .
    if (length(sequence) < maxlength)
        iterate (append (sequence, 0), maxlength, maxdepth)

    # Output the resulting sequence .
    print(sequence)

    # Next, increment the tail element of the sequence
    # and recurse .
    if (tail(sequence) < maxdepth)
        tail(sequence) += 1
        iterate (sequence, maxlength, maxdepth)
```

Figure 13: Beta odometer



This visualizes the β -odometer in terms of continued fractions. Any given sequence $[m_1, m_2, \dots, m_k]$ 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.

The above algorithm will generate polynomials in descending β -order, that is, in an order such that the β of each sequence is less than the previous one.

The sequences $[m_1, m_2, \dots]$ can be re-interpreted as continued fractions, albeit off-by-one. That is, for each such sequence, there is a real number

$$x = x([m_1, m_2, \dots]) = \frac{1}{m_1 + 1 + \frac{1}{m_2 + 1 + \dots}}$$

With some care, one can construct a bijection between real numbers and continued fractions. There is no particular difficulty for non-terminating sequences: these correspond to irrationals, and the correspondence is unique; again a bijection. Difficulties show only for the rationals, where any given rational has several inequivalent continued fraction expansions. The ambiguity can be avoided by fiat: one could, for example, take the shortest possible continued fraction expansion.

Combining these two interpretations gives a bijection from the unit interval back to itself. This is terms the “beta odometer” and is illustrated in figure 13.

The use of the word “odometer” is perhaps a bit cryptic; a quick review is in order. An odometer of characteristic p is defined as the set of all infinite-length integer sequences (a_0, a_1, \dots) 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

$$T(a_0, a_1, \dots) \mapsto (0, \dots, 0, a_k + 1, a_{k+1}, \dots) \text{ if } (a_0, a_1, \dots) = (p-1, \dots, p-1, a_k, a_{k+1}, \dots)$$

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, \dots) 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. Odometers are bijections of the unit interval into itself.

The β -odometer can be taken to be an odometer of characteristic zero. The rationale for this opaque claim can be clarified by looking at how it “rolls over”. The prototypical roll-over occurs at the juncture of $\beta = \varphi = 1.618\dots$. One observes the sequence

$$\begin{aligned} [0, 1] &= \beta = 1.75488\dots \\ [0, 2] &= \beta = 1.70490\dots \\ &\vdots \\ [0, 16] &= \beta = 1.61816\dots \\ [0, N] &= \beta > \varphi \text{ for } N \rightarrow \infty \\ [0] = [1, -1] &= [1, 0, 0, \dots] = \beta = \varphi \\ [1, 0, 0, \dots, 1, \dots] &= \beta < \varphi \end{aligned}$$

That is, the sequence $[0, N]$ for $N \rightarrow \infty$ rolls over to $[1, 0, 0, \dots]$. 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

$$[1, N] = \frac{1}{2 + \frac{1}{N+1}} = 0.5 - \varepsilon$$

which rolls over to

$$[0, 0, 0, \dots, 1, \dots] = \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}} = 0.5 + \varepsilon$$

which is, in a sense, backwards: the sequence $[1, N]$ for $N \rightarrow \infty$ does not roll over to $[2, 0, 0, \dots]$ as one might hope or guess or expect, but instead to $[0, 0, 0, \dots]$.

The p -adic odometers are interval maps; they re-arrange 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 $(\frac{1}{2}, 1)$ 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 $(\frac{1}{n}, \frac{1}{n-1})$. By self-symmetry, the sub-intervals will re-arrange as well.

A precise expression for the self-symmetry is easy to write in symbolic form. There are two, actually; one is a shift:

$$g : [m_1, m_2, \dots] \mapsto [m_1 + 1, m_2, \dots]$$

and the other is a reflection.

$$r : [m_1, m_2, \dots] \mapsto [1, m_1, m_2, \dots]$$

When the sequences are interpreted as continued fractions, they have a ready expression:

$$c(r([m_1, m_2, \dots])) = c([1, m_1, m_2, \dots]) = 1 - c([m_1, m_2, \dots])$$

that is, $r(x) = 1 - x$; thus a reflection. The shift acting on the continued fractions is likewise a Möbius transformation: $g(x) = x/(x+1)$. By contrast, there is no obvious simple expression for g and r in terms of β for the odometer.

Together, these two generate the dyadic monoid. That is, a general expression $\gamma = g^a \circ r \circ g^b \circ \dots$ is still a self-symmetry of the odometer. Such elements can be put in one-to-one correspondence with the infinite binary tree, i.e. the dyadic monoid.

4 Islands of Stability as Arnold Tongues

The trouble-spots, the eventual fixed-points of the map, can be placed in one-to-one accordance with the “islands of stability” seen in the iterated logistic map. They are, in essence, locations where periodic orbits “could be pasted”, or where they “naturally would appear”, if the map supported periodic attractors. That is, the beta shift only supports a single attractor, of period-one at $x = 0$; there is no “room” for anything more. This is analogous, in a way, to the phase locked loop, at zero coupling constant. At finite coupling strength, these “trouble spots” expand out as Arnold tongues, to have a finite size, visible on the Feigenbaum diagram for the logistic map as regions where period-doubling is occurring.

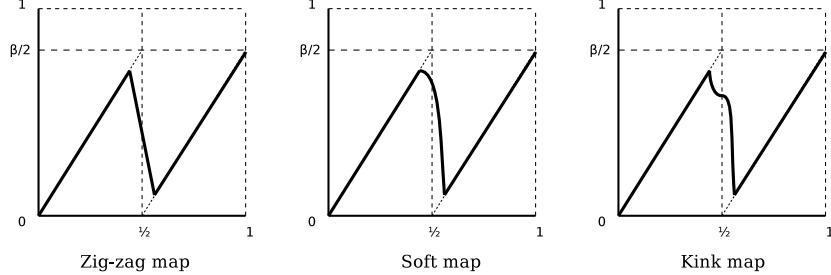
The idea here can be illustrated explicitly. Basically, take the natural saw-tooth shape of the map, widen the middle, and insert a slanting downward line, to create a zig-zag. That is, connect the two endpoints in the middle of the beta shift, “widening” it so that it has a finite, not infinite slope, 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

$$T_{\beta,\varepsilon}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \beta \left(\frac{1}{4} - \varepsilon\right) w & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta \left(x - \frac{1}{2}\right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \quad (21)$$

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



In the limit of $\varepsilon \rightarrow 0$, one regains the earlier beta shift: $\lim_{\varepsilon \rightarrow 0} T_{\beta,\varepsilon} = T_\beta$, 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 14.

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

$$D_{\beta,\varepsilon}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \beta \left(\frac{1}{4} - \varepsilon \right) g(w) & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta \left(x - \frac{1}{2} \right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \quad (22)$$

where the straight line has been replaced by a soft shoulder

$$g(w) = 1 - 2 \cos \frac{\pi}{4} (1 + w)$$

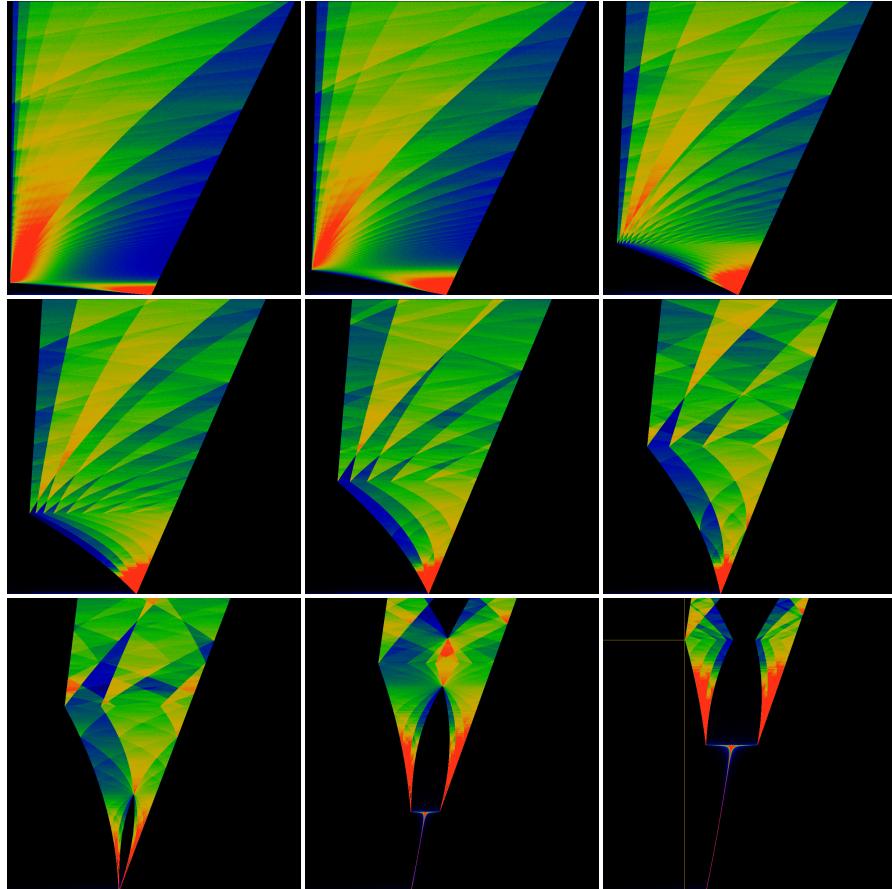
w is the same as before. This is scaled so that it’s a drop-in replacement for the straight line: $g(\frac{1}{2} - \varepsilon) = 1$ and $g(\frac{1}{2} + \varepsilon) = -1$. A cosine was used to create this soft shoulder, but a parabola would have done just as well. It is illustrated above, with the label “soft map”.

This map also interpolates between the left and right arms of the beta transform, forming a single continuous 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 15.

Two more variant maps can be considered. Both replace the center piece with symmetrical sinuous S-shaped curves, but in different ways. Consider

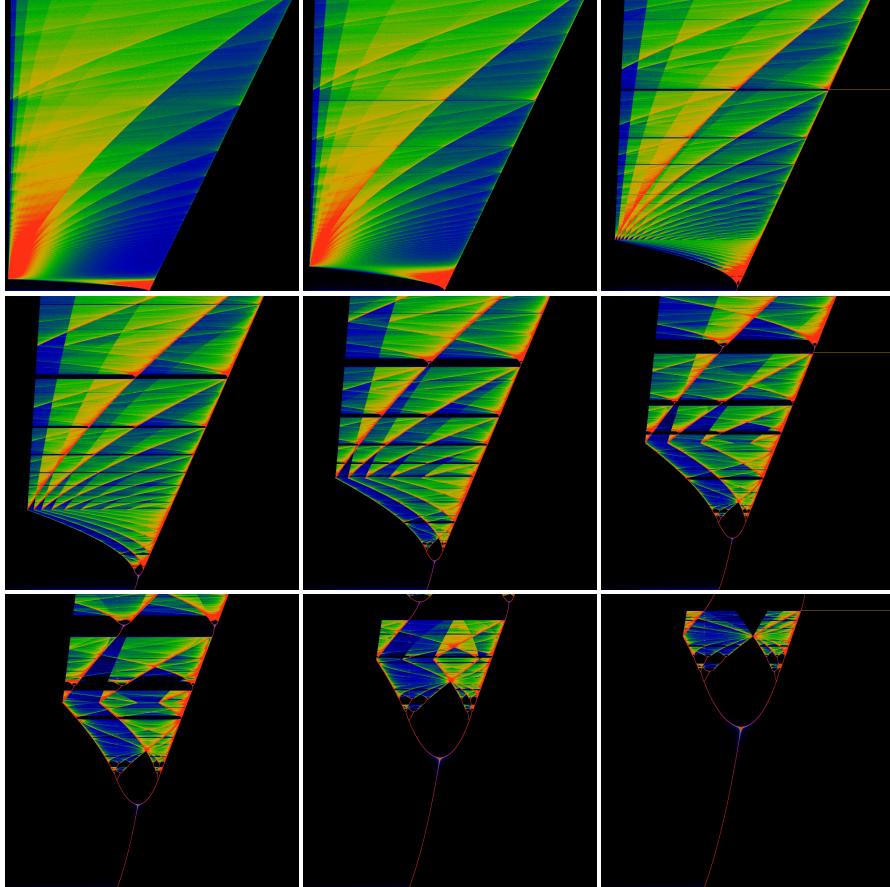
$$S_{\beta,\varepsilon,\sigma}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \sigma \beta \left(\frac{1}{4} - \varepsilon \right) \sin \frac{\pi}{2} w & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta \left(x - \frac{1}{2} \right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \quad (23)$$

Figure 14: Z-shaped Map



This illustrates a sequence of iterated maps, obtained from eqn 21. 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 15: Critical-Z map



This illustrates a sequence of iterated maps, obtained from eqn 22. The sequence of depicted ε values are the same as in figure 14. 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 1001, 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.

and

$$H_{\beta,\varepsilon,p,\sigma}(x) = \begin{cases} \beta x & \text{for } 0 \leq x < \frac{1}{2} - \varepsilon \\ \frac{\beta}{4} - \sigma \beta \left(\frac{1}{4} - \varepsilon \right) \operatorname{sgn}(x - \frac{1}{2}) |w|^p & \text{for } \frac{1}{2} - \varepsilon \leq x < \frac{1}{2} + \varepsilon \\ \beta \left(x - \frac{1}{2} \right) & \text{for } \frac{1}{2} + \varepsilon \leq x \leq 1 \end{cases} \quad (24)$$

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 16.

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 17.

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 17 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 18, 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[35]. These are shown in figures 19 and 20.

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

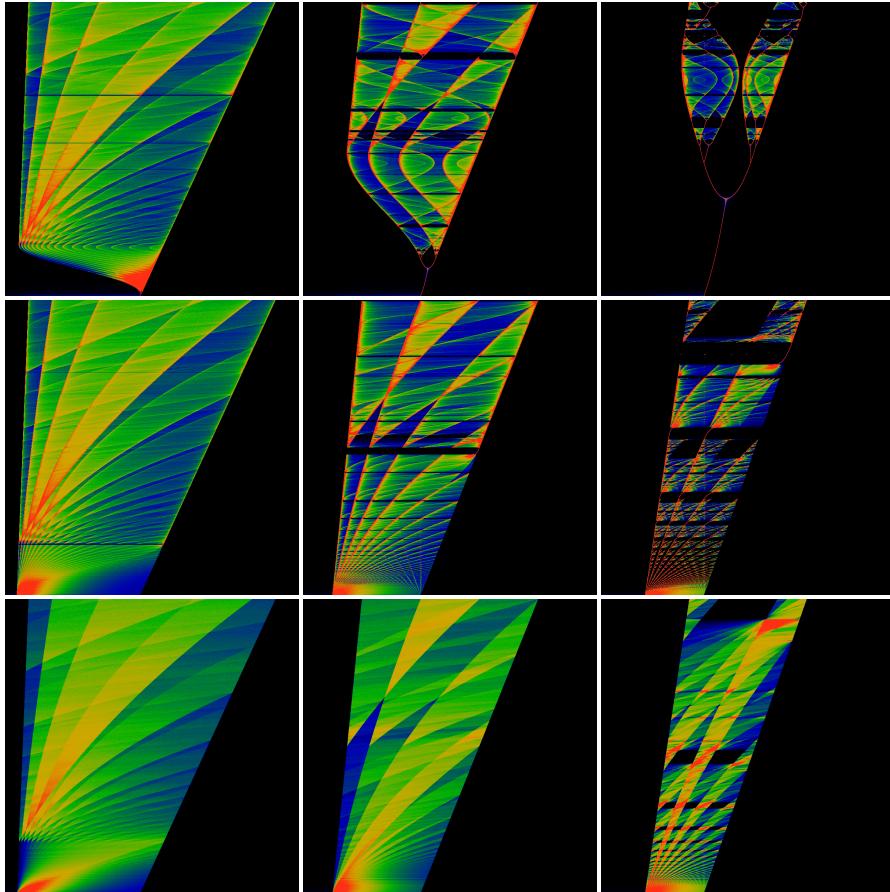
$$h_{\alpha,p}(x) = \begin{cases} \alpha + (1 - \alpha) |w|^p & \text{for } x < \frac{1}{2} \\ \alpha - (1 + \alpha) |w|^p & \text{for } \frac{1}{2} \leq x \end{cases}$$

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 17, except that this time, the location of the islands is controllable by the parameter α . Roughly, 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 re-create the “soft shoulder” of eqn 22, but at a variable location.

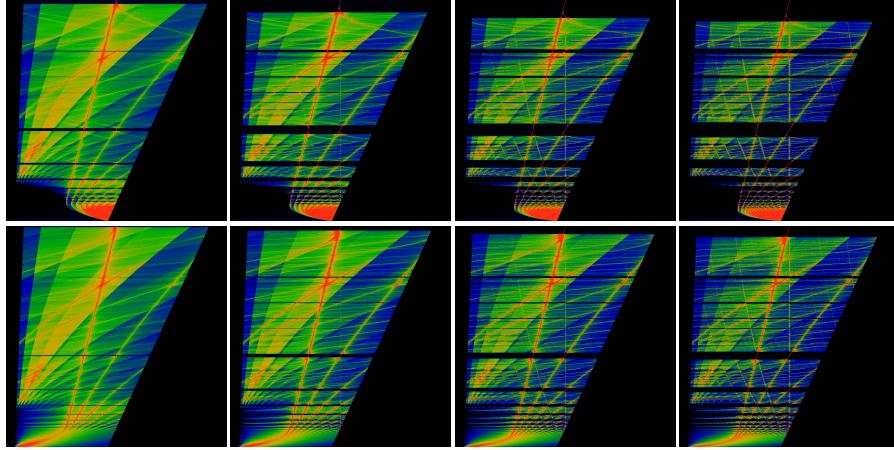
Thus, the above presents a general surgical technique for controlling both the general form of the chaotic regions, the location of the islands of stability, and what appears within the islands.

Figure 16: Interpolating Sine Map



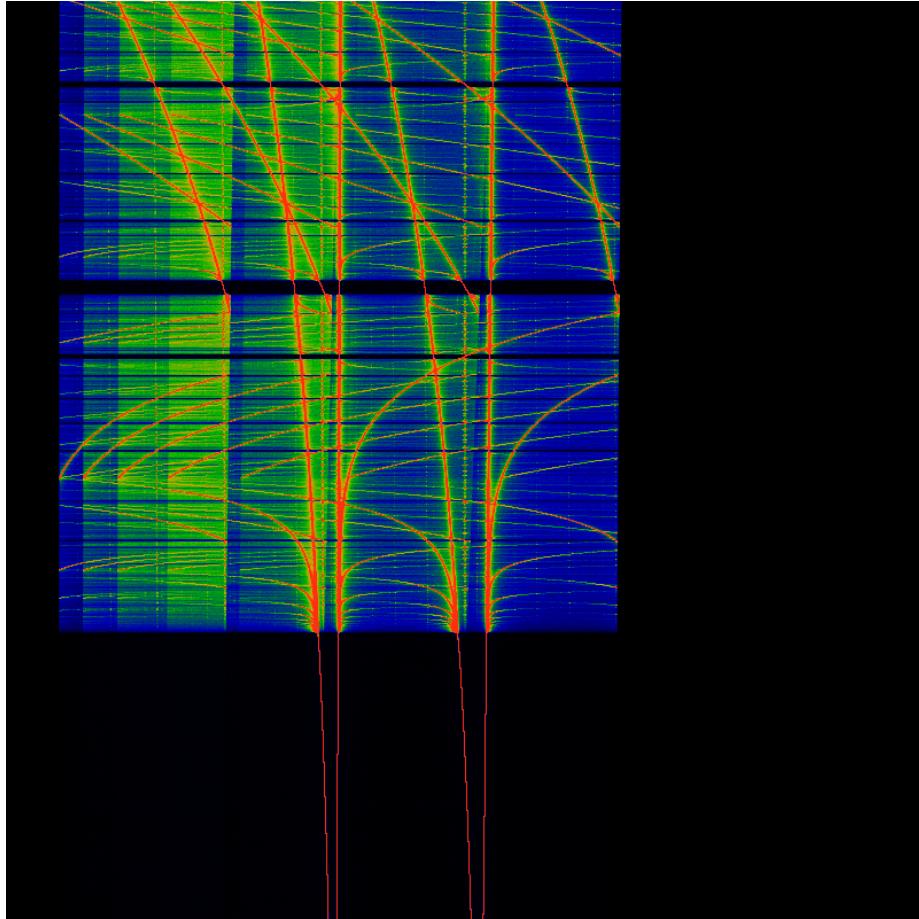
This illustrates a sequence of iterated maps, obtained from eqn 23. The sequence in the upper row shows $\epsilon = 0.04, 0.10$ and 0.15 ; with $\sigma = +1$. The upper row is much like the sequence shown in figure 15, except that its made sinuous, thanks to symmetrical S-shape. The middle row shows the same ϵ values, but for $\sigma = -1$. The bottom row shows eqn 24 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-zag 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 17: Interpolating Kink Map



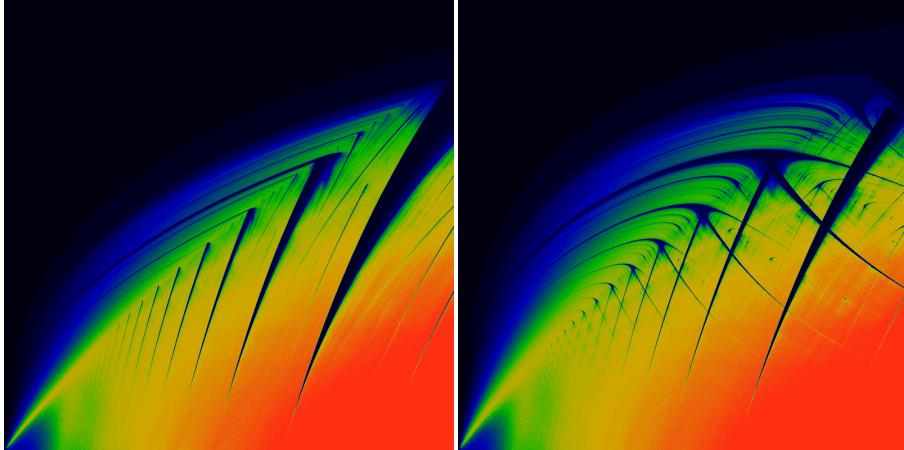
This illustrates a sequence of iterated maps, obtained from eqn 24. All eight images are held at $\varepsilon = 0.04$. The top row has $\sigma = +1$ (and thus the map is continuous) while the bottom row has $\sigma = -1$ (and thus the map has three disconnected branches). Left to right depicts the values $p = 2, 3, 4, 5$. 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 18: No Period Doubling



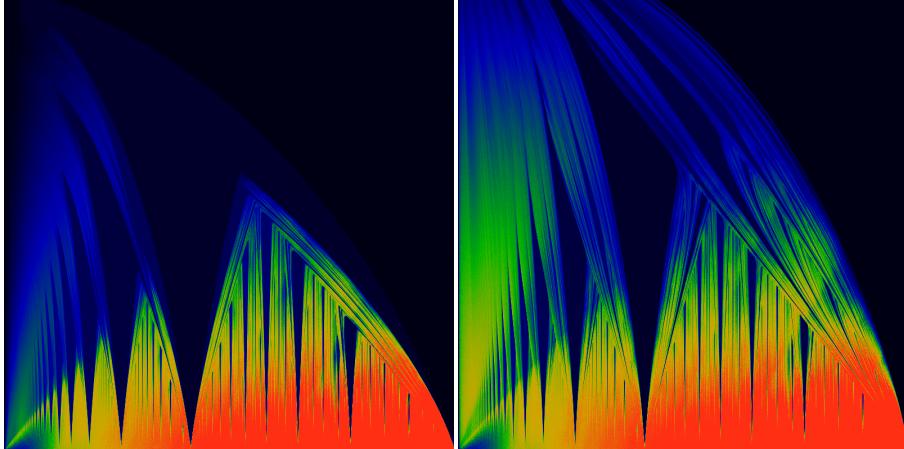
This figure is a zoom, confirming a lack of period doubling in the map $H_{\beta,\varepsilon,p,\sigma}(x)$ of eqn 24. The explored region is $0 \leq x \leq 1$, viz no zoom in the horizontal direction. Vertically, the image is centered on $\beta = 1.45$, having a total height of $\Delta\beta = 0.015625$. This uses the quintic kink, so $p = 5$ and $\sigma = +1$, making the continuous variant. The value of $\varepsilon = 0.04$ makes this directly comparable to other images.

Figure 19: Poincaré recurrence



The above visualize the Poincaré recurrence times for the map $D_{\beta,\varepsilon}(x)$ of eqn 22 on the left, and the map $S_{\beta,\varepsilon,1}(x)$ of eqn 23 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 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 $\beta = \varphi = 1.618\dots$, with the sequence of spikes to the left located at the other primary golden means (*viz.*, $1.3803\dots$ and the silver mean $1.3247\dots$ 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.[35]

Figure 20: Arnold Tongues



The above visualize the Poincaré recurrence times for the map $H_{\beta,\epsilon,p,\sigma}(x)$ of eqn 24. 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 $H_{\beta,\epsilon,p,\sigma}(x)$ and counting how many iterations it takes until $|x - H_{\beta,\epsilon,p,\sigma}^n(x)| < 0.009$. 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 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 17.

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.

5 Conclusion

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 express 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 $\sin(x)$ 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 non-uniform: 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 its 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, 3 and 4 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 its 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 algorithms? 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?

References

- [1] A. Rényi, “Representations for real numbers and their ergodic properties”, *Acta Math Acad Sci Hungary*, 8, 1957, pp. 477–493.
- [2] W. Parry, “On the β -expansion of real numbers”, *Acta Math Acad Sci Hungary*, 11, 1960, pp. 401–416.
- [3] 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>.
- [4] 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.
- [5] 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>.
- [6] 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>.
- [7] 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>.
- [8] P. Erdős and V. Komornik, “Developments in non-integer bases”, *Acta Math Hungar*, 79, 1998, pp. 57–83.
- [9] Nikita Sidorov, “Universal β -expansions”, *Arxiv*, 2002, URL <https://arxiv.org/abs/math/0209247v1>.
- [10] 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.
- [11] K. Schmidt, “On periodic expansions of Pisot numbers and Salem numbers”, *Bull London Math Soc*, 12, 1980, pp. 269–278.
- [12] F. Blanchard, “Beta-expansions and Symbolic Dynamics”, *Tehoretical Comp Sci*, 65, 1989, pp. 131–141.
- [13] 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.

- [14] 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>.
- [15] 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.
- [16] 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>.
- [17] 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>.
- [18] W. P. Thurston, “Groups, tilings and finite state automata”, *AMS Colloquium Lectures*, 1989.
- [19] Valérie Berthé and Anne Siegel, “Tilings associated with beta-numeration and substitutions”, *Integers: Electronic Journal of Combinatorial Number Theory*, 5, 2005.
- [20] Sh. Ito and H. Rao, “Purely periodic β -expansions with Pisot unit base”, *Proc Amer Math Soc*, 133, 2005, pp. 953–964.
- [21] Shigeki Akiyama, “Beta expansion and self-similar tilings”, , 2017, URL <http://cloud.crm2.uhp-nancy.fr/pdf/Manila2017/Akiyama.pdf>.
- [22] 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>.
- [23] 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>.
- [24] Leopold Flatto, et al., “The Zeta Function of the Beta Transformation”, *Ergodic Theory and Dynamical Systems*, 14, 1994, pp. 237–266.
- [25] Daniel J. Thompson, “Irregular sets and conditional variational principles in dynamical systems”, , 2010, URL https://people.math.osu.edu/thompson.2455/thesis_thompson.pdf.
- [26] L. Barreira and B. Saussol, “Variational Principles and Mixed 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>.

- [27] Linas Vepstas, “On the Minkowski Measure”, *ArXiv*, arXiv:0810.1265, 2008, URL <http://arxiv.org/abs/0810.1265>.
- [28] Wikipedia, “Quotient space”, , 2018, URL [https://en.wikipedia.org/wiki/Quotient_space_\(topology\)](https://en.wikipedia.org/wiki/Quotient_space_(topology)).
- [29] Kevin Hare, et al., “Three Series for the Generalized Golden Mean”, *Arxiv*, 2014, URL <https://arxiv.org/abs/1401.6200>.
- [30] 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>.
- [31] Solomon W. Golomb, “Irreducible polynomials, synchronizing codes, primitive necklaces and cyclotomic algebra”, *Proceedings Conf Combinatorial Mathematics and its Applications*, 1969, pp. 358–370.
- [32] Christophe Reutenauer, *Free Lie Algebras*, London Mathematical Society Monographs, 1993.
- [33] S. Duzhin and D. Pasechnik, “Automorphisms of necklaces and sandpile groups”, *ArXiv*, arXiv:math/1304.2563, 2013, URL <https://arxiv.org/pdf/1304.2563.pdf>.
- [34] 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>.
- [35] Wikipedia, “Arnold tongue”, , 2006, URL https://en.wikipedia.org/wiki/Arnold_tongue.