Discretization of expanding maps and percolation on a graph.

Oscar E. Lanford III and Paul Philipp Flockermann
Mathematics Department
ETH-Zurich

Talk by OEL at ESI on 4 June 2008

Starting question: Does "iterating a (smooth) map on the computer" reliably reflect the true ergodic properties of that map?

Iterating on the computer: First discretize, then iterate the discretization.

The discretized map is qualitatively very different from the original smooth one. For example, every orbit is eventually periodic.

Which properties of the smooth map should one expect to persist under discretization? And what can one hope to prove?

Paul Philipp Flockermann, *Discretizations of expanding maps*, Diss., Mathematische Wissenschaften ETH Zürich, Nr. 14448, 2002, 89pp.

P. P. Flockermann and O. E. Lanford, *Discretization of expanding maps,* in preparation. A preliminary version – not for general distribution, read at your own risk – is available at
http://www.math.ethz.ch/~lanford/9973/paper.pdf

The slides for this lecture are available at
http://www.math.ethz.ch/~lanford/9973/slides.pdf

For my views on the underlying philosophy, see my paper *Some informal remarks on the orbit structure of discrete approximations to chaotic maps.* Experimental Mathematics **7:4** (1998) pp. 317-324.)

We study **discretization of expanding circle endomorphisms of degree** 2

- very good ergodicity properties (unique absolutely continuous invariant probability measure ρ_f.)
- about as simple as possible without being trivial

$\check{f}$ will denote a smooth mapping $\mathbf{R} \to \mathbf{R}$ with

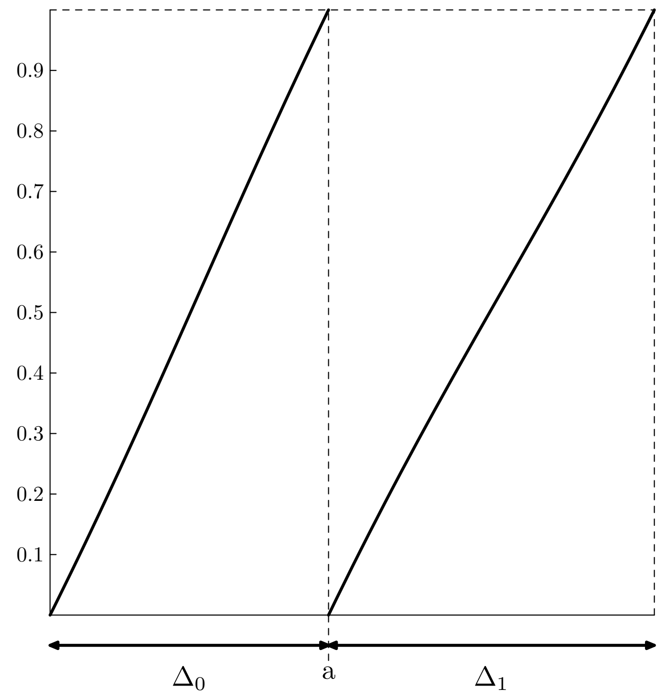
- $\check{f}(x+1) = \check{f}(x) + 2$ ($2 =$ *degree*)
- $\check{f}'(x) \geq \alpha^{-1}$ everywhere, $\alpha < 1$ (*expansive*)

Such an $\check{f}$ induces by passage to quotients a smooth 2-to-1 mapping

$$f : \mathbf{T} \to \mathbf{T} \quad \mathbf{T} := \mathbf{R}/\mathbf{Z}$$

Convenient simplification: $\check{f}(0) = 0$.

"Identify" $\mathbb{T}$ with $[0, 1)$. In this representation f has a discontinuity where its value passes through 1. It has two smooth globally defined smooth contractive "inverse branches" f_0^{-1} and f_1^{-1}



$f(x) = 2x + \frac{1}{4\pi}\sin^2(\pi x) \bmod 1$

Discretization: Put down a grid Γ_N of N evenly spaced points in $\mathbf{T}$:

$$\Gamma_N := \left\{ \frac{j}{N} : 0 \leq j < N \right\}$$

($\mathbf{T}$: *continuum state space*, Γ_N: *discret(ized) state space.* Elements of Γ_n are *grid points*)

Discretized $\mathbf{f}$: Apply (the exact continuum) f to the grid point x and round the result to the nearest grid point.

$$f_N : \Gamma_N \rightarrow \Gamma_N$$

This model for discretization is only semi-realistic:

- Γ_N evenly spaced – not like machine floating point numbers which are evenly spaced "within octaves" (makes f_N locally injective)

- only final result of computing $f(x)$ is rounded – no "rounding of intermediate quantities"

Which N? Study a *limiting regime* with $N \to \infty$.

Discretized f is **not** a "random perturbation of f" Probability comes into the story – as it should – by considering not a single grid point but an ensemble of them with macroscopic fluctuations $\to 0$ in the $N \to \infty$ limit.

Orientation: what might one hope to prove? One possibility: Let δ_N denote normalized counting measure on Γ_N, but regarded as a probability measure on $\mathbf{T}$. $f_N^m\,\delta_N$ denotes the "push-forward" of δ_N under the m-th iterate of f_N.

Conjecture: *For generic f, $f_N^m\,\delta_N \to \rho_f$, (weak-$*$) as $N, m \to \infty$ with $\log N \ll m \ll \sqrt{N}$. Here ρ_f denotes the unique absolutely continuous probability measure invariant under f.*

So far as I know, no one has *any idea at all* about how to prove such a statement. Compare with the following

Easy theorem:

$$\lim_{m\to\infty} \lim_{N\to\infty} f_N^m\,\delta_N = \rho_f.$$

Disclaimer:

- **Our results:** $N \to \infty$, m fixed.
- **Would like to understand:** $m, N \to \infty$, $m/logN$ large.

$1/N$: grid spacing, m: number of iterations.

$m/logN$ large means that happenings on the scale of the grid can be magnified by the intrinsic expansivity of the map to the macroscopic scale.

Notation: For $x \in \mathbf{R}$:

- $\mathsf{round}(x)$ denotes the nearest integer to x (round up if x half odd integral), and

- $\mathsf{frac}(x)$ denotes $x - \mathsf{round}(x)$,

so

$$x = \mathsf{round}(x) + \mathsf{frac}(x), \quad \mathsf{round}(x) \in \mathbf{Z}, \quad -1/2 \leq \mathsf{frac}(x) < 1/2.$$

Then:

$$f_N(x) = \frac{1}{N}\mathsf{round}(N \cdot f(x)) \quad \text{(only for } x \in \Gamma_N\text{)}$$

Object of study: Local structure of $f_N^m\,\delta_N$ for $N \to \infty$, m fixed.

Main result: There is a well-defined limit determined by a non-trivial probabilistic set-up (percolation on a tree)

Consequences:

$$\delta_N\left(\mathsf{support}(f_N^m\,\delta_N)\right) \to 0 \quad \text{as } N, m \to \infty.$$
$$\delta_N\left(\{\text{periodic points of } f_N\}\right) \to 0 \quad \text{as } N \to \infty$$

Observation:

$$f_N^m\,\delta_N(\{y\}) = \frac{1}{N} \cdot \left| f_N^{-m}\{y\} \right|$$

We study: Preimages of points of Γ_N under iterates of f_N *(discrete preimages.)*

Continuum preimages. The smooth mapping f is precisely 2-to-1; there is an a, $0 < a < 1$, so that f maps each of $\Delta_0 := [0, a)$ and $\Delta_1 := [a, 1)$ bijectively onto all of $[0, 1)$. Thus

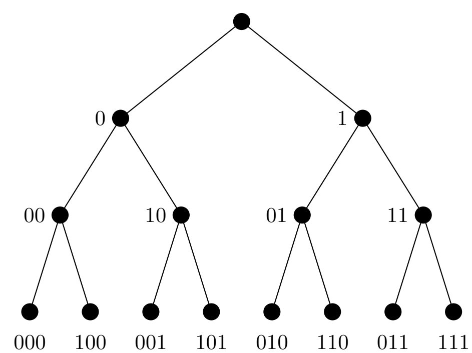
every point $\bar{y} \in [0, 1)$ has exactly two preimages, one ($\bar{y}_0$, left preimage) in Δ_0 and another ($\bar{y}_1$, right preimage) in Δ_1.

Each of these has in turn 2 preimages, and so on, so

every point $\bar{y} \in [0, 1)$ has exactly 2^m preimages under f^m. These can be labelled $\bar{y}_{i_1,\ldots i_m}$, $i_1, \ldots i_m \in \{0, 1\}$ (preimage of type $i_1 \ldots i_m$); the rule is

$$\bar{y}_{i_1\ldots i_m} \in \Delta_{i_1} \quad \textit{and} \quad f(\bar{y}_{i_1\ldots i_m}) = \bar{y}_{i_2\ldots i_m}$$

We think of the set of all preimages, of all orders, as arranged in a tree:



Note: The order of the points in a row is not the same as their order on the circle.

Discrete primages. Surprisingly, the situation for preimages under the discretized map is totally different.

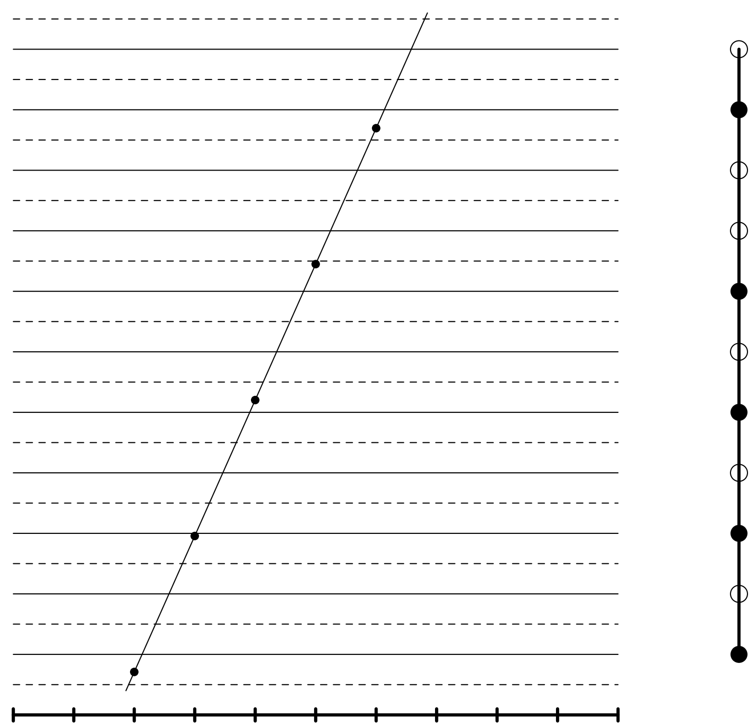
Observation: *If g is any map of a finite set X to itself, then each point of X has – on the average – exactly one preimage under g.*

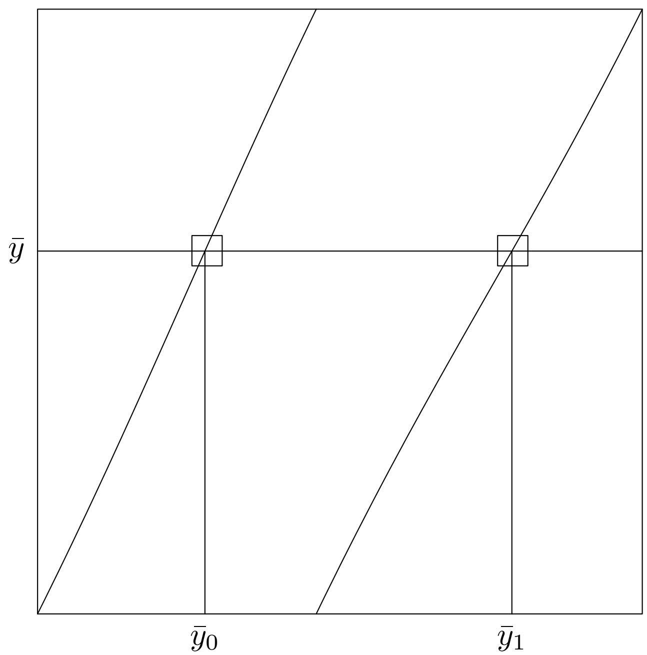
Proof. The preimages of the various points are disjoint, and their union is all of X. Hence, the sum of the cardinalities of the preimages is exactly the cardinality of X.

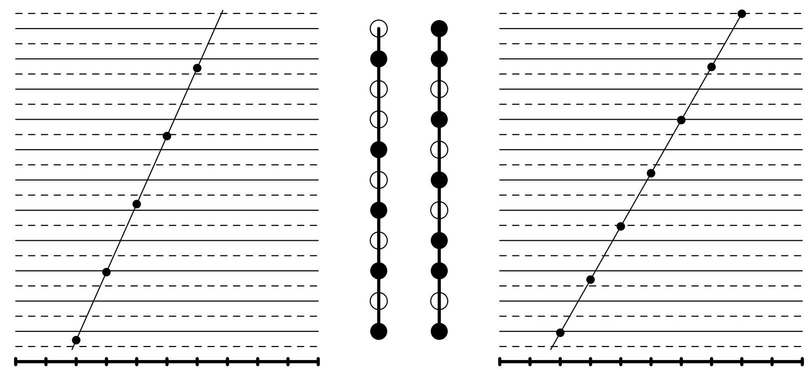
This applies to f_N, but also to its iterates: No matter how fine the discretization is, and no matter how large m is, each point has on the average *one* preimage under f_N^m. Some of the continuum preimages "disappear" under discretization.

Orientation. With our hypotheses, a grid point y has *at most one* preimage under f_N in Δ_0 (y_0, *left discrete preimage*) and at most one in Δ_1 (y_1, *right discrete preimage*) Similarly: At most one *discrete* $i_1 \ldots i_m$ *preimage,* (written $y_{i_1 \ldots i_m}$, but need not exist.)

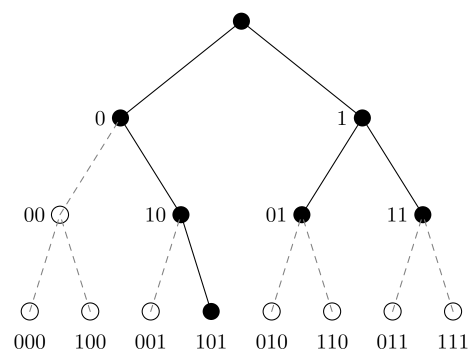
Where do the continuum preimages go? f is strictly expansive, so moving x over one grid spacing moves $f(x)$ up by strictly more than one grid spacing. For large enough j, as x moves over j grid spacings, $f(x)$ must move up by at least $j + 1$, so the rounded values must skip over at least one of the possible image grid points.




$\bar{y}$
$\bar{y}_0$
$\bar{y}_1$



Observation: The set of discrete preimages which do exist for a given $y \in \Gamma_N$ forms a *subtree* of the continuum preimage tree. (If y has no discrete left preimage, it can't possibly have a discrete preimage of type 00 or 10.) Speak of the *tree of discrete preimages of* y. Better: describe the existing preimages by giving their *combinatorial types (labels)* $i_1 \ldots i_m$.



As the grid point y varies, the set of combinatorial types of the existing discrete preimages changes. We have a function:

$$y \mapsto T(y): \quad \Gamma_n \to \{\texttt{subtrees}\}$$

Technically: Probably best to cut off these trees at some fixed height m_0 (which can be arbitrarily large)

Ensembles and limiting regime.
We want to study $y \mapsto T(y)$ for gridpoints y near to some given continuum $\bar{y}$, as $N \to \infty$.

Pick for each N a "tolerance" $\eta_N > 0$ so that

$$\left.\begin{array}{r} \eta_N \to 0 \\ N \cdot \eta_N \to \infty \end{array}\right\} \quad \text{as } N \to \infty.$$

Make a probability measure μ_N on Γ_N by giving equal weight to all grid points y with $|y - \bar{y}| < \eta_N$ and zero weight to the others.

Study: Distribution of the tree-valued random variable $T(\,.\,)$ with respect to μ_N in the limit $N \to \infty$.

Next step: Will show concretely how to compute limiting values for some representative probabilities.

Weyl's equidistribution theorem. *Let* $\lambda_1, \ldots \lambda_m$ *be real numbers such that*

$$1, \lambda_1, \ldots, \lambda_m \quad \textit{is linearly independent over } \mathbf{Q}$$

Then

$$\{(j \cdot \lambda_1, \ldots j \cdot \lambda_m) : j = 0, 1, \ldots\}$$

is uniformly distributed modulo $\mathbf{Z}^m$.

With apologies, I want to remind you explicitly what "uniformly distributed modulo $\mathbf{Z}^m$" means. Let π denote the canonical projection $\mathbf{R}^m \to \mathbf{T}^m$. The assertion is as follows: Let J_1, J_2, ... be any sequence of positive integers with $J_n \to \infty$. For each n, let ν_n be normalized counting measure on the finite subset

$$\{\pi(j \cdot \lambda_1, \ldots, j \cdot \lambda_m) : 0 \leq j < J_n\}$$

of $\mathbf{T}^m$ Then ν_n converges in the weak-$*$ topology to normalized Lebesgue measure as $n \to \infty$.

Two small twists:

1. If a_n is an arbitrary sequence in $\mathbf{R}^m$, then $T_{a_n}\nu_n$ (ν_n translated by a_n) again converges weak-$*$ to normalized Lebesgue measure.

2. The canonical projection $\pi : \mathbf{R} \to \mathbf{T}$ can be replaced by the fractional part map described above. We get a sequence of normalized counting measures which converges to Lebesgue measure on the cube $[-1/2, 1/2]^m$. Again, an arbitrary sequence of translations – applied before taking fractional parts – doesn't change things.

First example: Probability of existence of a discrete left preimage.

The setup: We are fixing the continuum localization point $\bar{y}$; $\bar{y}$ will be subject to conditions which we will impose as they arise. Grid points y have the form j/N. We write either j or y as convenient. Say a grid point is *allowed* if $|y - \bar{y}| < \eta_N$, i.e., if it carries a non-zero weight in our ensemble. Allowed grid points correspond to j's with $J_0(N) \leq j < J_1(N)$, where $J_1(N) - J_0(N) \approx 2 \cdot N \cdot \delta_N \to \infty$.

Expository fiction: f_0^{-1} is exactly affine on the set of allowed grid points:

$$f_0^{-1}(y) = \alpha_0 y + \beta_0$$

Let $y = j/N$ be an allowed grid point. Then

$$
\begin{aligned}
& y \text{ has a discrete left preimage} \\
\Leftrightarrow \quad & \exists\, x \in \Gamma_N \cap \Delta_0 \text{ with } y - 1/2N \le f(x) < y + 1/2N \\
\Leftrightarrow \quad & f_0^{-1}[y - 1/2N, y + 1/2N) \text{ contains a grid point } k/N = x \\
\Leftrightarrow \quad & \exists\, k \in \mathbf{Z} : \alpha_0 \cdot j + N\beta_0 - \alpha_0/2 \le k < \alpha_0 \cdot j + N\beta_0 + \alpha_0/2 \\
\Leftrightarrow \quad & \exists\, k \in \mathbf{Z} : k - \alpha_0/2 < \alpha_0 j + N\beta_0 \le k + \alpha_0/2 \\
\Leftrightarrow \quad & \mathrm{frac}(\alpha_0 \cdot j + N\beta_0) \in (-\alpha_0/2, \alpha_0/2] =: I_0
\end{aligned}
$$

If α_0 is irrational then, by the simplest case of Weyl's theorem, $\{\mathrm{frac}(\alpha_0 \cdot j + N\beta_0) \ : \ j$ allowed for $N\ \}$ becomes uniformly distributed over $[-1/2, 1/2]$ as $N \to \infty$, so the fraction of allowed j's for which j/N has a discrete left preimage – i.e., the probability of having a discrete left preimage – goes to the length α_0 of the interval I_0 as $N \to \infty$.

Summary: If $\alpha_0 = 1/f'(\bar{y}_0)$ is irrational, then the probability of having a left discrete preimage converges as $N \to \infty$ to α_0.

Actually, this much can be proved more easily, and without the irrationality assumption.

A similar argument – making again a local linearity assumption that $f_1^{-1}(y) = \alpha_1 \cdot y + \beta_1$ for relevant y's – shows that

$y = j/N$ has a discrete right preimage if and only if

$$\text{frac}(\alpha_1 \cdot j + N\beta_1) \in (-\alpha_1/2, +\alpha_1/2] =: I_1$$

and hence

y has both left and right discrete preimages if and only if

$$(\text{frac}(\alpha_0 \cdot j + N\beta_0), \text{frac}(\alpha_1 \cdot j + N\beta_1)) \in I_0 \times I_1$$

By Weyl's theorem with $m = 2$: If

$$\{1, \alpha_0, \alpha_1\} \quad \text{is linearly independent over } \mathbf{Q},$$

then

$$\{(\mathsf{frac}(\alpha_0 \cdot j + N\beta_0), \mathsf{frac}(\alpha_1 \cdot j + N\beta_1)) : \ j \text{ allowed for } N\}$$

becomes uniformly distributed over the square $[-1/2, 1/2] \times [-1/2, 1/2]$ as $N \to \infty$. Hence, under the irrationality assumption above, the probability of having both left and right discrete preimages goes to $\alpha_0 \cdot \alpha_1$. In particular: The events $\{$discrete left preimage exists$\}$ and $\{$discrete right preimage exists$\}$ are asymptotically independent.

Second example: Probability of existence of discrete 10 preimage. $y = j/N$ has a discrete 10 preimage if and only if

1. y has a discrete left preimage $y_0 = j_0/N$ and

2. y_0 has a discrete right preimage

1. holds if and only if

$$\theta_0 := \mathrm{frac}(j \cdot \alpha_0 + N\beta_0) \in I_0 := (-\alpha_0/2, +\alpha_0/2]$$

and, further,

$$j_0 = \mathrm{round}(j \cdot \alpha_0 + N\beta_0) = j \cdot \alpha_0 + N\beta_0 - \theta_0.$$

Assuming – as usual – $f_1^{-1}(y) = \alpha_{10} y + \beta_{10}$ near $\bar{y}_0$: $y_0 = j_0/N$ has a discrete right preimage if and only if

$$\mathrm{frac}(\alpha_{10} \cdot j_0 + N\beta_{10}) \in I_{10} := (-\alpha_{10}/2, +\alpha_{10}/2].$$

The expression on the left can be written as

$$\mathrm{frac}(\alpha_{10} \cdot (\alpha_0 j + N\beta_0 - \theta_0) + N\beta_{10}) = \mathrm{frac}(A_{10} j + NB_{10} - \alpha_{10}\theta_0)$$

with $A_{10} := \alpha_0 \cdot \alpha_{10}$. Putting

$$\theta_{10} := \mathrm{frac}(A_{10} j + NB_{10}),$$

we can combine the two conditions for the existence of y_{10} into

$$(\theta_0, \theta_{10} - \alpha_{10}\theta_0) \in I_0 \times I_{10},$$

Summarizing:

$$\theta_0(j) := \mathrm{frac}(\alpha_0 \cdot j + N\beta_0), \quad \theta_{10}(j) := \mathrm{frac}(A_{10} \cdot j + NB_{10}).$$

j/N has a discrete 10 preimage if and only if the *vertical shear*

$$(\phi_1, \phi_2) \mapsto (\phi_1, \phi_2 - \alpha_0\, \phi_1)$$

sends (θ_0, θ_{10}) into the rectangle $I_0 \times I_{10}$.

The preimage of the rectangle under the shear is a *parallelogram* P_{10}; since the shear is area-preserving

$$\mathrm{area}(P_{10}) = \mathrm{area}(I_0 \times I_{10}) = \alpha_0 \times \alpha_{10}.$$

Irrationality: Assume $\{1, \alpha_0, A_{10}\}$ linearly independent over $\mathbf{Q}$. By Weyl's theorem, the $(\theta_0(j), \theta_{10}(j))$'s become uniformly distributed over the square $[-1/2, 1/2] \times [-1/2, 1/2]$.

Hence the fraction of allowed j's for which this pair lands in P_{10} converges as $N \to \infty$ to the area $\alpha_0 \cdot \alpha_{10}$ of P_{10}.

Summary: Under the irrationality assumption, the probability of existence of a discrete 10 preimage $\to \alpha_0 \cdot \alpha_{10}$ as $N \to \infty$.

Notation. For arbitrary $i_1 \dots i_m$, we write

$$\alpha_{i_1 \dots i_m} := \left(f'(\bar{y}_{i_1 \dots i_m})\right)^{-1}, \quad A_{i_1 \dots i_m} := \alpha_{i_1 \dots i_m} \cdot \alpha_{i_2 \dots i_m} \cdots \alpha_{i_m}.$$

Comprehensive irrationality assumption:

$$\{A_{i_1 \dots i_m} : m = 0, 1, 2, \dots\}$$

is linearly independent over $\mathbf{Q}$.

Proposition. *Under the comprehensive irrationality assumption, the distribution of the "random discrete preimage tree" converges as* $N \to \infty$*. The proof gives an explicit characterization of the limit.*

Orientation on the irrationality assumption. This is a condition on both f and the continuum localization point $\bar{y}$. For given f it is *never* satisfied for all $\bar{y}$. We show

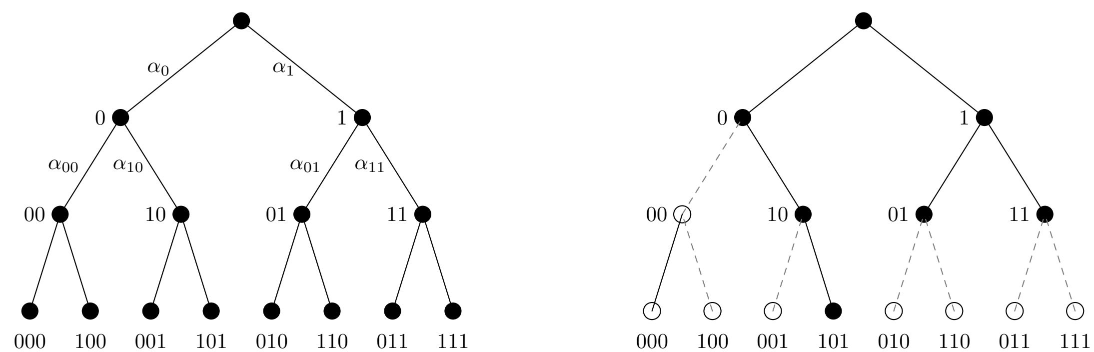
Proposition. *The set of f's such that the comprehensive irrationality assumption holds for* all but countably many $\bar{y}$'s *is residual in the $\mathcal{C}^r$ topology for all $2 \leq r \leq \infty$.*

The proof is complicated and ugly.

Describing the limiting distribution on trees.

Make a bond-percolation system on the (trivial, binary) tree of all continuum preimages.

- Label the bonds by their lower ends.
- Bond $i_1 \ldots i_m$ is present with probability $\alpha_{i_1 \ldots i_m}$.
- Different bonds are independent.
- The connected component of the root is a subtree: *tree-valued random variable on the probability space of bond configurations.*
- Distribution of this random variable is the $N \to \infty$ limit of the finite-N preimage tree distribution.



α_0
α_1
0
1
α_{00}
α_{10}
α_{01}
α_{11}
00
10
01
11
000
100
001
101
010
110
011
111
0
1
00
10
01
11
000
100
001
101
010
110
011
111

Alternative description.

Describe a preimage tree by splitting it into order-m components:

$$T = T_1 \cup T_2 \cup \ldots \cup T_m \cup \ldots; \quad T_m \subset \{0,1\}^m$$

(Many exclusions: If $i_2, \ldots i_m \notin T_{m-1}$, then neither $0i_2 \ldots i_m$ nor $1i_1 \ldots i_m$ is in T_m.)

In the limit: $T_1, T_2, \ldots T_m, \ldots$ becomes a Markov chain (state space varying with m) with transition probability ...

Consequence 1.

$$\langle |T_m| \rangle_N \xrightarrow[N\to\infty]{} ((P_f)^m 1)(\bar{y}) \xrightarrow[m\to\infty]{} \rho_f(\bar{y})$$

where $|\,.\,|$ denotes cardinality, $\langle\,.\,\rangle_N$ the mean value in the ensemble for N, P_f the *Ruelle-Perron-Frobenius operator* for f, and ρ_f the absolutely continuous probability density invariant under f.

In other words: The average number of discrete preimages of order m converges, as first $N \to \infty$, then $m \to \infty$, to $\rho_f(\bar{y})$.

Remark. This drops out of the above results, but a simple direct proof – not using the comprehensive irrationality assumption – can also be given.

Consequence 2. *Let P_k^m denote the probability with respect to the limiting distribution that $|T_m| = k$. Then*

$$\lim_{m\to\infty} P_k^m \to 0 \quad \text{for } k = 1, 2, \ldots.$$

Proof. Using the formula for transition probabilities of the process (T_m) – which I didn't give – it is easy to find an $\epsilon_k > 0$ so that the conditional probability

$$Pr\{|T_{m+1}| = 0 : T_m\} \geq \epsilon_k \quad \text{for } |T_m| = k.$$

Then

$$P_0^{m+1} \geq P_0^m + \epsilon_k \cdot P_k^m.$$

Hence,

$$\sum_m P_k^m \leq \frac{1}{\epsilon_k}.$$

Since, by the preceding result,

$$\langle |T_m| \rangle = \sum_{k=0}^{\infty} k \cdot P_k^m \xrightarrow[m\to\infty]{} \rho_f(\bar{y}) \neq 0, \infty,$$

$P_0^m \to 1$ as $m \to \infty$ – most grid points near $\bar{y}$ have no high-order discrete preimages at all – but a few of the remaining ones have enough of them so that the average number remains of order 1 as $m \to \infty$.

Among other things, the above says that, for m large, most grid points near $\bar{y}$ have no discrete preimages of order m at all *provided that the comprehensive irrationality condition is satisfied by f at $\bar{y}$.*

If f is such that the comprehensive irrationality condition holds for almost all $\bar{y}$, we can integrate over $\bar{y}$ and use the dominated convergence theorem to show that that most points of Γ_N have no m-th order discrete preimages at all, i.e., that $(f_N)^m\delta_N$ is concentrated on a small subset of Γ_N.

There are many intriguing questions about the percolation picture which we have been unable to answer; here is one:

Question: How does the probability of having at least one discrete preimage of order m behave as $m \to \infty$?

There is a very suggestive hint: If f is affine on each of Δ_0 and Δ_1, the percolation picture reduces to a classical *critical branching process,* and there is a wonderful trick – using a generating function – for analyzing these. One result: The probability in question is asymptotic to c/m for a computable constant c. It seems very likely that something like this remains true in the general situation we are studying.

Numerical experiments with $N = 2^{52}$ (IEEE double precision numbers) show a rough proportionality to $1/m$ for values of m up to several hundred, i.e., arguably into the range $m/\log N$ “large”.