The Takagi Function and Related Functions

Jeff Lagarias,
University of Michigan
Ann Arbor, MI, USA

(December 14, 2010)

Functions in Number Theory and Their Probabilistic Aspects,

(RIMS, Kyoto University, Dec. 2010)

Topics Covered

- Part I. Introduction and History
- Part II. Number Theory
- Part III. Probability Theory
- Part IV. Analysis
- Part V. Rational Values of Takagi Function
- Part VI. Level Sets of Takagi Function

Credits

- J. C. Lagarias and Z. Maddock , Level Sets of the Takagi Function: Local Level Sets, `arXiv:1009.0855`

- J. C. Lagarias and Z. Maddock , Level Sets of the Takagi Function: Generic Level Sets, `arXiv:1011.3183`

- Work partially supported by NSF grant DMS-0801029.

Part I. Introduction and History

- Definition The distance to nearest integer function

$$\ll x \gg = dist(x, \mathbb{Z})$$

- The map $T(x) = 2 \ll x \gg$ is sometimes called the symmetric tent map, when restricted to $[0, 1]$.

The Takagi Function

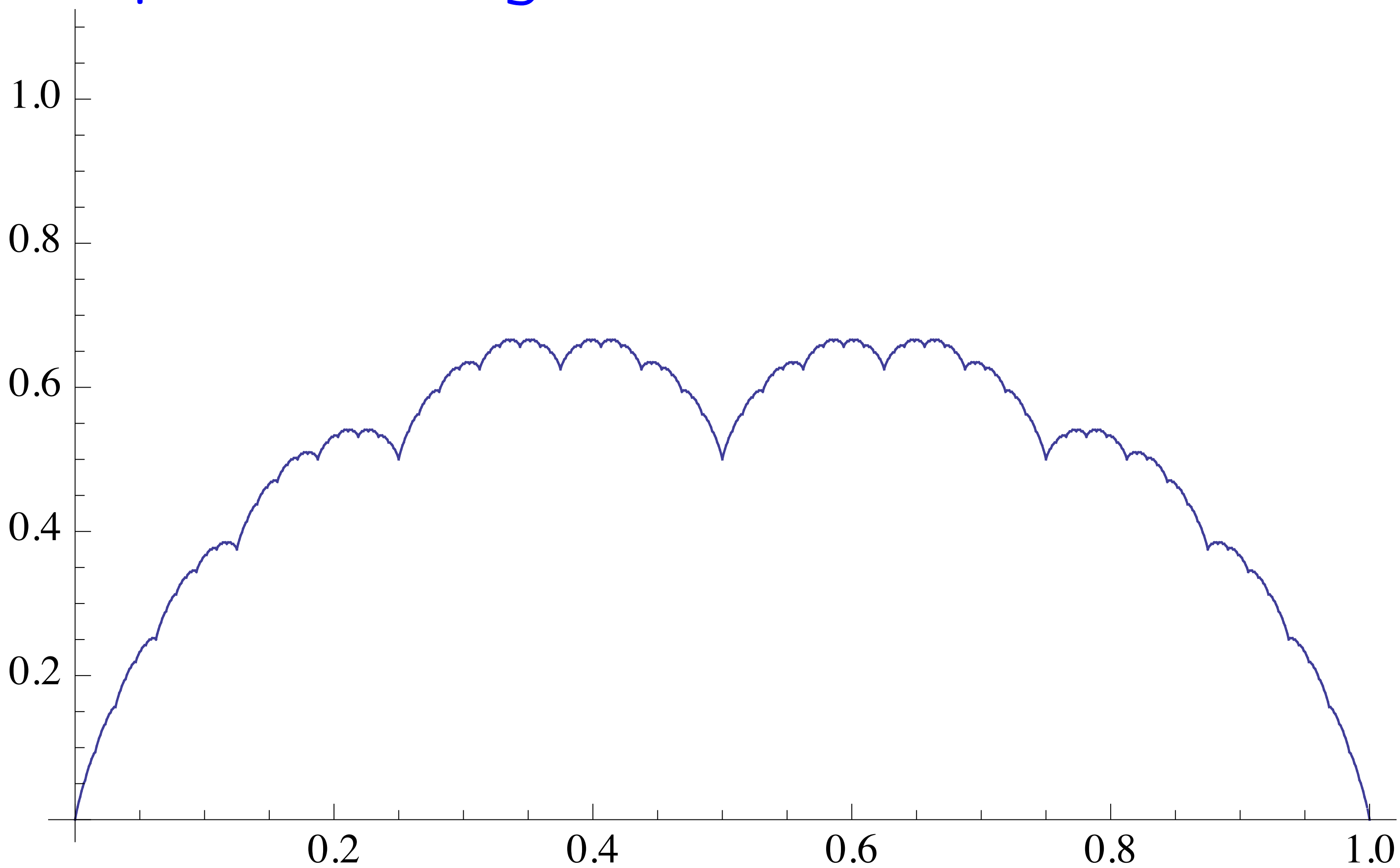
- The Takagi Function $\tau(x) : [0,1] \to [0,1]$ is

$$\tau(x) = \sum_{j=0}^{\infty} \frac{1}{2^j} \ll 2^j x \gg$$

- This function was introduced by Teiji Takagi in 1903.

- Motivated by Weierstrass nondifferentiable function. (Visit to Germany 1897-1901.)

Graph of Takagi Function



Main Property: Everywhere Non-differentiability

- Theorem (Takagi 1903) The function $\tau(x)$ is continuous on $[0,1]$ and has no derivative at each point $x \in [0,1]$ on either side.

- Base 10 variant function independently discovered by van der Waerden (1930), same theorem.

- Takagi function also rediscovered by de Rham (1956).

Generalizations

- For $g(x)$ periodic of period one, and $a, b > 1$, set

$$F_{a,b,g}(x) := \sum_{j=0}^{\infty} \frac{1}{a^j} g(b^j x)$$

- This class includes Weierstrass nondifferentiable function. Properties of functions depend sensitively on a, b and the function $g(x)$.

- Smooth function example (Hata and Yamaguti (1984))

$$F(x) = \sum_{j=0}^{\infty} \frac{1}{4^j} \ll 2^j x \gg = 2x(1 - x).$$

Recursive Construction

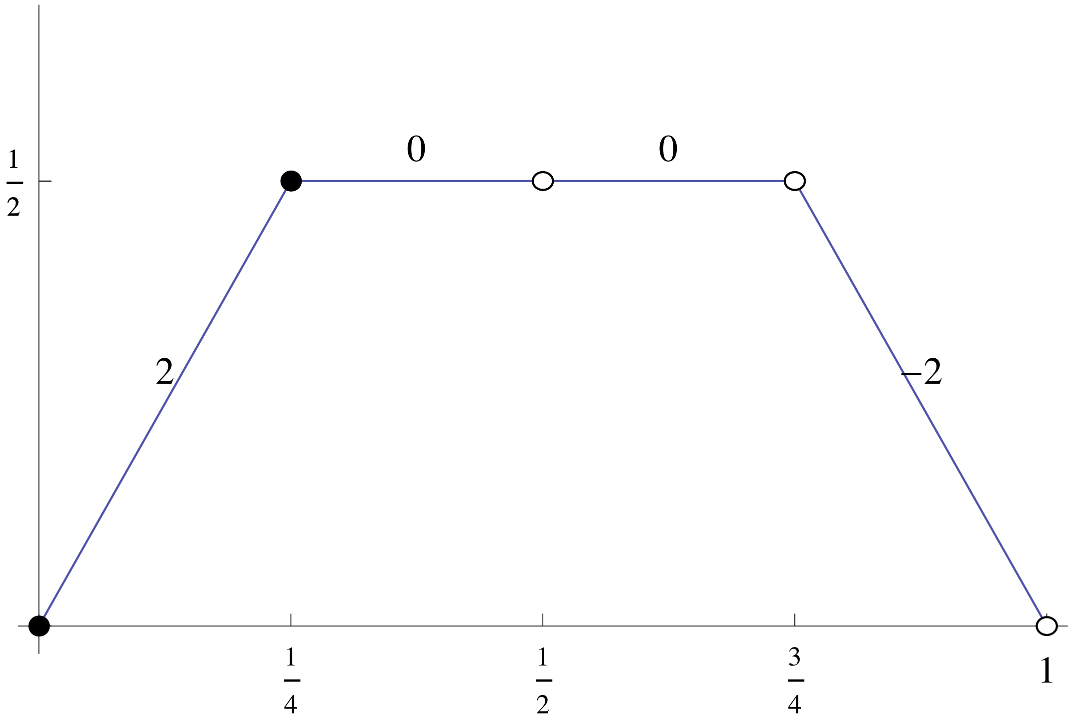
- The n-th approximant

$$\tau_n(x) := \sum_{j=0}^{n} \frac{1}{2^j} \ll 2^j x \gg$$

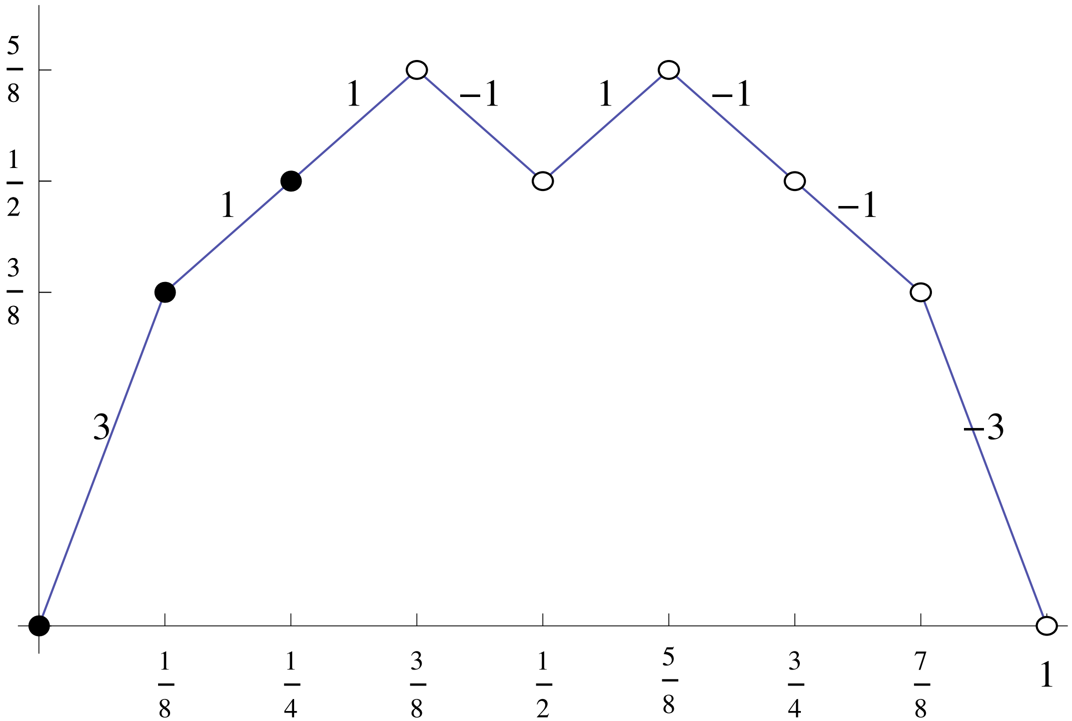
- This is a piecewise linear function, with breaks at the dyadic integers $\frac{k}{2^n}$, $\;1 \le k \le 2^n - 1$.

- All segments have integer slopes, in range between $-n$ and $+n$. The maximal slope $+n$ is attained in $[0, \frac{1}{2^n}]$ and the minimal slope $-n$ in $[1 - \frac{1}{2^n}, 1]$.

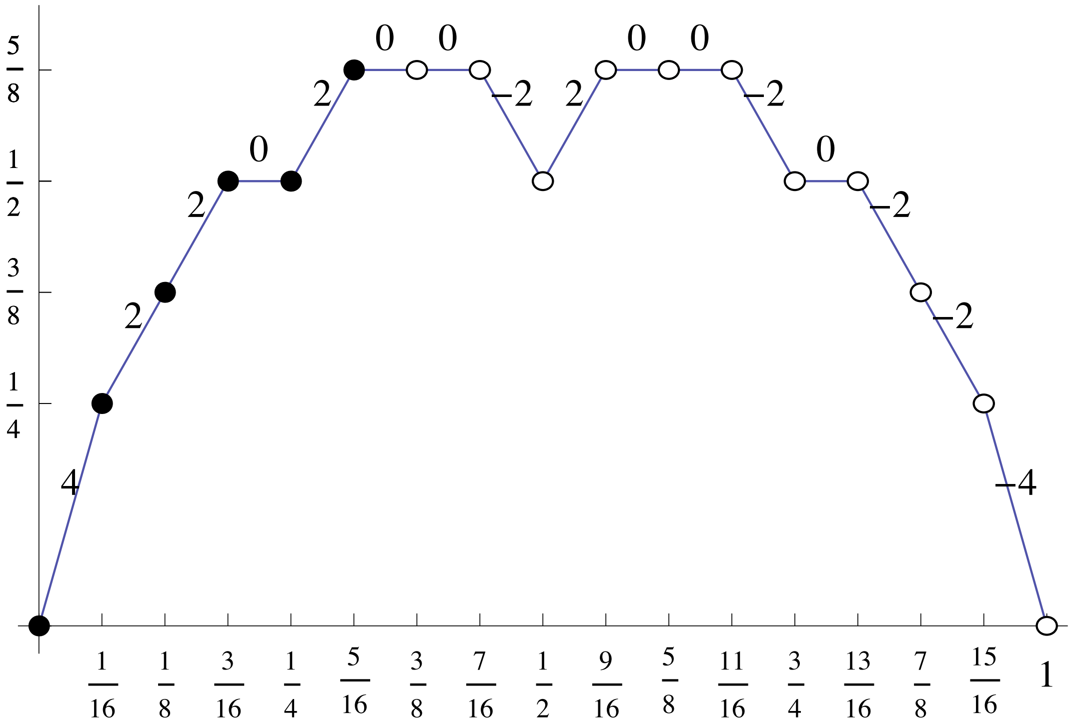
Takagi Approximants-τ_2



Takagi Approximants-τ_3



Takagi Approximants-τ_4



Properties of Approximants

- The n-th approximant

$$\tau_n(x) := \sum_{j=0}^{n} \frac{1}{2^j} \ll 2^j x \gg$$

agrees with $\tau(x)$ at all dyadic rationals $\frac{k}{2^n}$.
These values then freeze, i.e. $\tau_n(\frac{k}{2^n}) = \tau_{n+j}(\frac{k}{2^n})$.

- The approximants are nondecreasing at each step, They approximate Takagi function $\tau(x)$ from below.

Symmetry

- Local symmetry

$$\tau_n(x) = \tau_n(1-x).$$

- Thus

$$\tau(x) = \tau(1-x).$$

Functional Equations

- Fact. The Takagi function, satisfies, for $0 \le x \le 1$, two functional equations:

$$\begin{aligned} \tau(\frac{x}{2}) &= \frac{1}{2}\tau(x) + \frac{1}{2}x \\ \tau(\frac{x+1}{2}) &= \frac{1}{2}\tau(x) + \frac{1}{2}(1-x). \end{aligned}$$

- These are a kind of dilation equation, relating function on two different scales.

Takagi Function Formula

- Takagi's Formula (1903): Let $x \in [0, 1]$ have binary expansion

$$x = .b_1 b_2 b_3 ... = \sum_{j=1}^{\infty} \frac{b_j}{2^j}$$

Then

$$\tau(x) = \sum_{n=1}^{\infty} \frac{l_n(x)}{2^n}.$$

with

$$\begin{aligned} l_n(x) &= b_1 + b_2 + \cdots + b_{n-1} && \text{if digit} \quad b_n = 0. \\ &= n - 1 - (b_1 + b_2 + \cdots + b_{n-1}) && \text{if digit} \quad b_n = 1. \end{aligned}$$

Fourier Series

- Theorem. The Takagi function $\tau(x)$ has period 1, and is an even function. It has Fourier series

$$\tau(x) := \sum_{n=0}^{\infty} c_n e^{2\pi i n x}$$

in which

$$c_0 = \int_0^1 \tau(x)dx = \frac{1}{2}$$

and for $n > 0$ there holds

$$c_n = c_{-n} = \frac{1}{2^{m+1}(2k+1)^2} \cdot \frac{1}{\pi^2}, \quad \text{where} \quad n = 2^m(2k+1).$$

Takagi Function as a Boundary Value

- Theorem. Let $\{c_n : n \in \mathbb{Z}\}$ be the Fourier coefficients of the Takagi function, and define the power series
$$f(z) = \frac{1}{2}c_0 + \sum_{n=1}^{\infty} c_n z^n.$$
It converges on unit disk and defines a continuous function on the boundary of the unit disk,
$$f(e^{2\pi i\theta}) = \frac{1}{2}(T(\theta) + iU(\theta))$$
in which $T(\theta) = \tau(\theta)$ is the Takagi function, and $U(\theta)$ is a function which we call the conjugate Takagi function.

- Open Problem. Study properties of $U(x)$.

History

- The Takagi function $\tau(x)$ has been extensively studied in all sorts of ways, during its 100 year history, often in more general contexts.

- It has some surprising connections with number theory and (less surprising) with probability theory.

- It has showed up as a "toy model" in study of chaotic dynamics, as a fractal, and it has connections with wavelets.

Part II. Number Theory: Counting Binary Digits

- Consider the integers $1, 2, 3, \ldots$ represented in binary notation. Let $S_2(N)$ denote the sum of the binary digits of $0, 1, \ldots, N-1$, i.e. it counts the total number of $1's$ in these expansions.

- Bellman and Shapiro (1940) showed $S_2(N) \sim \frac{1}{2} N \log_2 N$. Mirsky (1949) showed $S_2(N) \sim \frac{1}{2} N \log_2 N + O(N)$.

- Trollope (1968) showed $S_2(N) = \frac{1}{2} N \log_2 N + N\, E_2(N)$ where $E_2(N)$ is an oscillatory function. He gave an exact combinatorial formula for $N E_2(N)$ involving the Takagi function.

Counting Binary Digits-2

- Delange (1975) gave an elegant reformulation and sharpening of Trollope's result...

- Theorem. (Delange 1975) There is a continuous function $F(x)$ of period 1 such that, for all integer N,

$$\frac{1}{N} S_2(N) = \frac{1}{2} \log_2 N + F(\log_2 N).$$

  Here

$$F(x) = \frac{1}{2}(1 - \{x\}) - 2^{-\{x\}} \tau(2^{\{x\}-1})$$

  where $\tau(x)$ is the Takagi function, and $\{x\} := x - [x]$.

Counting Binary Digits-3

- The function $F(x) \leq 0$, with $F(0) = 0$.

- The function $F(x)$ has an explicit Fourier expansion whose coefficients involve the values of the Riemann zeta function on the line $Re(s) = 0$, at $\zeta(\frac{2k\pi i}{\log 2})$, $k \in \mathbb{Z}$.

Counting Binary Digits-4

- Flajolet, Grabner, Kirchenhofer, Prodinger and Tichy (1994) gave a direct proof of Delange's theorem using Dirichlet series and Mellin transforms.

- Identity 1. Let $e_2(n)$ sum the binary digits in n. Then

$$\sum_{n=1}^{\infty} \frac{e_2(n)}{n^s} = 2^{-s}(1 - 2^{-s})^{-1}\zeta(s).$$

Counting Binary Digits-5

- Identity 2: Special case of Perron's Formula. Let
$$H(x) := \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \frac{\zeta(s)}{2^s-1} x^s \frac{ds}{s(s-1)}.$$
Then for integer N have an exact formula
$$H(N) = \frac{1}{N} S_2(N) - \frac{N-1}{2}.$$

- Proof. Shift the contour to $Re(s) = -\frac{1}{4}$. Pick up contributions of a double pole at $s = 0$ and simple poles at $s = \frac{2\pi i k}{\log 2}$, $k \in \mathbb{Z}, k \neq 0$. Miracle occurs: The shifted contour integral vanishes for all integer values $x = N$. (It is a kind of step function, and does not vanish identically.)

Part III. Probability Theory: Singular Functions

- Łomnicki and Ulam (1934) constructed singular functions as solutions to various functional equations.

- Draw binary digits of a number, at random:

  0 with probability α

  1 with probability $1 - \alpha$.

  Let $L_\alpha(x)$ be the cumulative distribution function of resulting distribution μ_α. Call this the Lebesgue function with parameter α.

Singular Functions-2

- These functions satisfy the functional equations ($0 \le x \le 1$),

$$\begin{aligned} L_\alpha(\frac{x}{2}) &= \alpha\, L_\alpha(x), \\ L_\alpha(\frac{x+1}{2}) &= \alpha + (1-\alpha) L_\alpha(x). \end{aligned}$$

- Claim. The measure $\mu_\alpha(x) = dL_\alpha(x)$ is a (singular) measure supported on a set of Hausdorff dimension

$$H_2(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha).$$

(binary entropy function)

Singular Functions-3

- Salem (1943) determined the Fourier series of $L_\alpha(x)$. He obtained it using

$$\int_0^1 e^{2\pi i t x} dL_\alpha(x) = \prod_{k=1}^{\infty} \left(\alpha + (1-\alpha) e^{\frac{2\pi i t}{2^k}} \right).$$

- Product formulas like this occur in wavelet theory (solutions of dilation equations), see Daubechies and Lagarias (1991), (1992).

Singular Functions-4

- Theorem (Hata and Yamaguti 1984) For fixed x the Lebesgue function $L_\alpha(x)$ extends in the variable α to an analytic function on the lens-shaped region

$$\{\alpha \in \mathbb{C} : \ |\alpha| < 1 \ \text{ and } \ |1 - \alpha| < 1\}.$$

The Takagi function appears as:

$$2\,\tau(x) = \frac{d}{d\alpha} L_\alpha(x) \,|_{\alpha=\frac{1}{2}}$$

- Hata and Yamaguti, Japan J. Applied Math. **1** (1984), 183-199. A very interesting paper!

Open Problem: Invariant Measure

- Observation The absolutely continuous measure $\mu_T := 2\tau(x)dx$ is a probability measure on $[0, 1]$. Call it the Takagi measure.

- General Query. Are there any interesting maps of the interval $f : [0, 1] \to [0, 1]$ for which the Takagi measure $\mu_T(x)$ is an invariant measure?

Part IV. Analysis: Fluctuation Properties

- The Takagi function oscillates rapidly. It is an analysis problem to understand the size of its fluctuations on various scales.

- These problems have been completely answered, as follows...

Fluctuation Properties: Single Fixed Scale

- The maximal oscillations at scale h are of order: $h \log_2 \frac{1}{h}$.

- Proposition. For all $0 < h < 1$ the Takagi function satisfies

$$|\tau(x+h) - \tau(x)| \leq 2\,h\,\log_2 \frac{1}{h}.$$

- This bound is sharp within a multiplicative factor of 2. Kôno (1987) showed that as $h \to 0$ the constant goes to 1.

Maximal Asymptotic Fluctuation Size

- The asymptotic maximal fluctuations at scale $h \to 0$ are of order: $h\sqrt{2\log_2 \frac{1}{h} \log\log\log_2 \frac{1}{h}}$ in the following sense.

- Theorem (Kôno 1987) Let $\sigma_l(h) = \sqrt{\log_2 \frac{1}{h}}$. Then for all $x \in (0,1)$,

$$\limsup_{h\to 0^+} \frac{\tau(x+h)-\tau(x)}{h\,\sigma_l(h)\sqrt{2\log\log \sigma_l(h)}} = 1,$$

and

$$\liminf_{h\to 0^+} \frac{\tau(x+h)-\tau(x)}{h\,\sigma_l(h)\sqrt{2\log\log \sigma_l(h)}} = -1.$$

Average Scaled Fluctuation Size

- Average Fluctuation size at scale h is Gaussian, proportional to $h\sqrt{\log_2 \frac{1}{h}}$.

- Theorem (Gamkrelidze 1990) Let $\sigma_l(h) = \sqrt{\log_2 \frac{1}{h}}$. Then for each real y,

$$\lim_{h \to 0^+} Meas\,\{x : \frac{\tau(x+h) - \tau(x)}{h\,\sigma_l(h)} \le y\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{1}{2}t^2} dt.$$

- Kôno's result on maximum asymptotic fluctuation size is analogous to the law of the iterated logarithm.

Part V. Rational Values

- Easy Fact.

  (1) The Takagi function maps dyadic rational numbers $\frac{k}{2^n}$ to dyadic rational numbers $\tau(\frac{k}{2^n}) = \frac{k'}{2^{n'}}$, where $n' \leq n$.

  (2) The Takagi function maps rational numbers $r = \frac{p}{q}$ to rational numbers $\tau(r) = \frac{p'}{q'}$. Here the denominator of $\tau(r)$ may sometimes be larger than that of r.

- Next formulate three (hard?) unsolved problems...

Rational Values: Pre-Image Problems

- Problem 1. Determine whether a rational r' has some rational preimage r with $\tau(r) = r'$.

- Problem 2. Determine which rationals r' have an uncountable level set $L(r')$.

- Problem 2 was raised by Donald Knuth (2004) in Volume 4 of the Art of Mathematical Programming (Fascicle 3, Problem 83 in 7.2.1.3) He says:
  "WARNING: This problem can be addictive."

Rational Values: Iteration Problems

- Problems 3 and 4. Determine the behavior of $\tau(x)$ under iteration, restricted to dyadic rational numbers, resp. all rational numbers.

- Remarks. (1) For dyadic rationals the denominators are nonincreasing, so all iterates go into periodic orbits. But figuring out orbit structure could be an interesting problem.

  (2) For general rational numbers it is not clear what happens. Denominators could potentially increase to $+\infty$. (Any invariant measure will be supported in $[\frac{1}{2}, \frac{2}{3}]$.)

Part VI. Level Sets of the Takagi Function

- Definition. The level set $L(y) = \{x : \quad \tau(x) = y\}$.

- Problem. How large are the level sets of the Takagi function?

- Quantitative Problem. determine exact count if finite; determine Hausdorff dimension if infinite.

- Answer depends on sampling method: Could choose random x-values (abscissas) or random y-values (ordinates)

Size of Level Sets: Cardinality

- Fact. There exist levels y such that $L(y)$ is finite, countable, or uncountable.

- $L(\frac{1}{5})$ is finite, containing two elements.
  Knuth (2005) showed that $L(\frac{1}{5}) = \{\frac{3459}{87040}, \frac{83581}{87040}\}$.

- $L(\frac{1}{2})$ is countably infinite.

- $L(\frac{2}{3})$ is uncountably infinite.
  Baba (1984) observed this holds...because...

Size of Level Sets: Hausdorff Dimension

- Theorem (Baba 1984) The set $L(\frac{2}{3})$ has Hausdorff dimension $\frac{1}{2}$.

- This result followed up by...

- Theorem (Maddock 2010) All level sets $L(y)$ have Hausdorff dimension at most 0.699.

- Conjecture (Maddock 2010) All level sets $L(y)$ have Hausdorff dimension at most $\frac{1}{2}$.

Local Level Sets-1

- Approach to understand level sets: break them into local level sets, which are easier to understand.

- The local level set containing x is described completely in terms of the binary expansion of $x = \sum_{n \geq 1} b_n 2^{-n}$.

- Definition. The deficient digit function $D_n(x)$ counts the excess of 0's over 1's in the first n digits of the binary expansion of x.

- For random x the values $(D_1(x), D_2(x), D_3(x), ...)$ are sums of a simple random walk, taking steps $+1$ or -1.

Local Level Sets-2

- Given x, look at all the breakpoint values

$$0 = c_0 < c_1 < c_2 < \dots$$

  where $D_{c_j}(x) = 0$, i.e. values n where the random walk returns to the origin. Call this set the breakpoint set $Z(x)$.

- The binary expansion of x is broken into blocks of digits with position $c_j < n \le c_{j+1}$. The flip operation exchanges digits 0 and 1 inside a block.

- Definition. The local level set L_x^{loc} consists of all numbers $x' \sim x$ by a (finite or infinite) set of flip operations. All numbers in L_x^{loc} have the same breakpoint set $Z(x) = Z(x')$.

Propeties of Local Level Sets

- Property 1. L_x^{loc} is a closed set.

- Property 2. L_x^{loc} is either a finite set of cardinality $2^{Z(x)}$, if there are finitely many blocks in $Z(x)$, or is a Cantor set if there are infinitely many blocks in $Z(x)$.

- Property 3. Each level set partitions into a disjoint union of local level sets.

Level Sets-Abscissa Viewpoint

- Problem. Draw a random point x uniformly with respect to Lebesgue measure. How large is the level set $L(\tau(x))$?

- Partial Answer. At least as large as the local level set L_x^{loc}.

- Theorem A. For a randomly drawn point x, with probability one the local level set L_x^{loc} is an uncountable (Cantor) set , of Hausdorff dimension 0.

Proof of Theorem A

(1) With probability one, the set of breakpoints $Z(x)$ is infinite: the one-dimensional random walk $D_n(x)$ returns to the origin infinitely often almost surely. This makes L_x^{loc} a Cantor set.

(2) With probability one, the expected time for the random walk $D_n(x)$ to return to the origin is infinite. This "implies" that L_x^{loc} has Hausdorff dimension 0.

Number of Local Level Sets per Level

- Fact. The number of local level sets on a level can take an arbitrary integer value and also can be countably infinite.

- Theorem. There are a dense set of levels y such that $L(y)$ contains a countably infinite number of local level sets.

- Known such levels all have y a dyadic rational, including $y = \frac{1}{2}$.

Level Sets-Abscissa View

- Problem.What is the average size of full level set $L(\tau(x))$ where x is picked at random?

- This problem is unsolved. Expect the same answer as Theorem A: Most $L(\tau(x))$ uncountable of Hausdorff dim. 0.

- Difficulty. The mysterious problem is to understand how many local level sets there are on a given level, when (abscissa) x is picked at random.

Expected Number of Local Level Sets: Ordinate View

- We are able to estimate the number of local level sets when the ordinate y is picked at random:

- Theorem B. The expected number of local level sets for an (ordinate) y drawn uniformly from $0 \leq y \leq 2/3$ is finite. This number is exactly $3/2$.

Level Sets-Ordinate view

- We can compute the expected size of a level set $L(y)$ for a random (ordinate) level y...

- Theorem C.
  (1) (Buczolich (2008)) The expected size of a level set $L(y)$ for y drawn at random (Lebesgue measure) is finite.

  (2) The expected number of elements in a level set $L(y)$ for y drawn at random (Lebesgue measure) is infinite.

- Our proof of (1) differs from the proof of Buczolich.
  It gives extra information, namely (2).

Local Level Sets: Size Paradox?

- Ordinate View: Level sets $L(y)$ are finite with probability 1.

- Abscissa View: Level sets $L(\tau(x))$ are uncountably infinite with probability 1.

- Reconciliation Mechanism: x-values preferentially select level sets that are "large".

Reconciling Size of Local Level Sets

- Theorem D. The set Big of levels y such that the level set $L(y)$ has positive Hausdorff dimension, is itself a set of full Hausdorff dimension 1.

- Proof Idea. Explicit construction of local level sets giving distinct y values having Hausdorff dimension $> 1 - \epsilon$, for any given $\epsilon > 0$.

Approach to Results

- We study the left hand endpoints of local level sets...

- Definition. The deficient digit set Ω^L is the set of left-hand endpoints of all local level sets.

- Fact. The set Ω^L consists of all real numbers x whose binary expansions have at least as many 0's as 1's after n steps. That is, all $D_n(x) \geq 0$.
  (There is a unique choice of flips to achieve this.)

Approach to Results-cont'd.

- Key point. Ω^L keeps track of all local level sets. It is a closed set obtained by removing a countable set of open intervals from $[0, 1]$. It has has a Cantor set structure.

- Theorem E. Ω^L has Lebesgue measure 0 and has Hausdorff dimension 1.

- This holds because...

Proof of Theorem E

- Heuristic Argument: Count the number of allowable strings in expansions in Ω^L. There are about $n^{-3/2}2^n$ strings of length n. The fact that $\sum n^{-3/2} < \infty$ implies Lebesgue measure 0. The fact that allowed number exceeds $2^{(1-\epsilon)n}$ "implies" deficient digit set Ω^L has Hausdorff dimension 1.

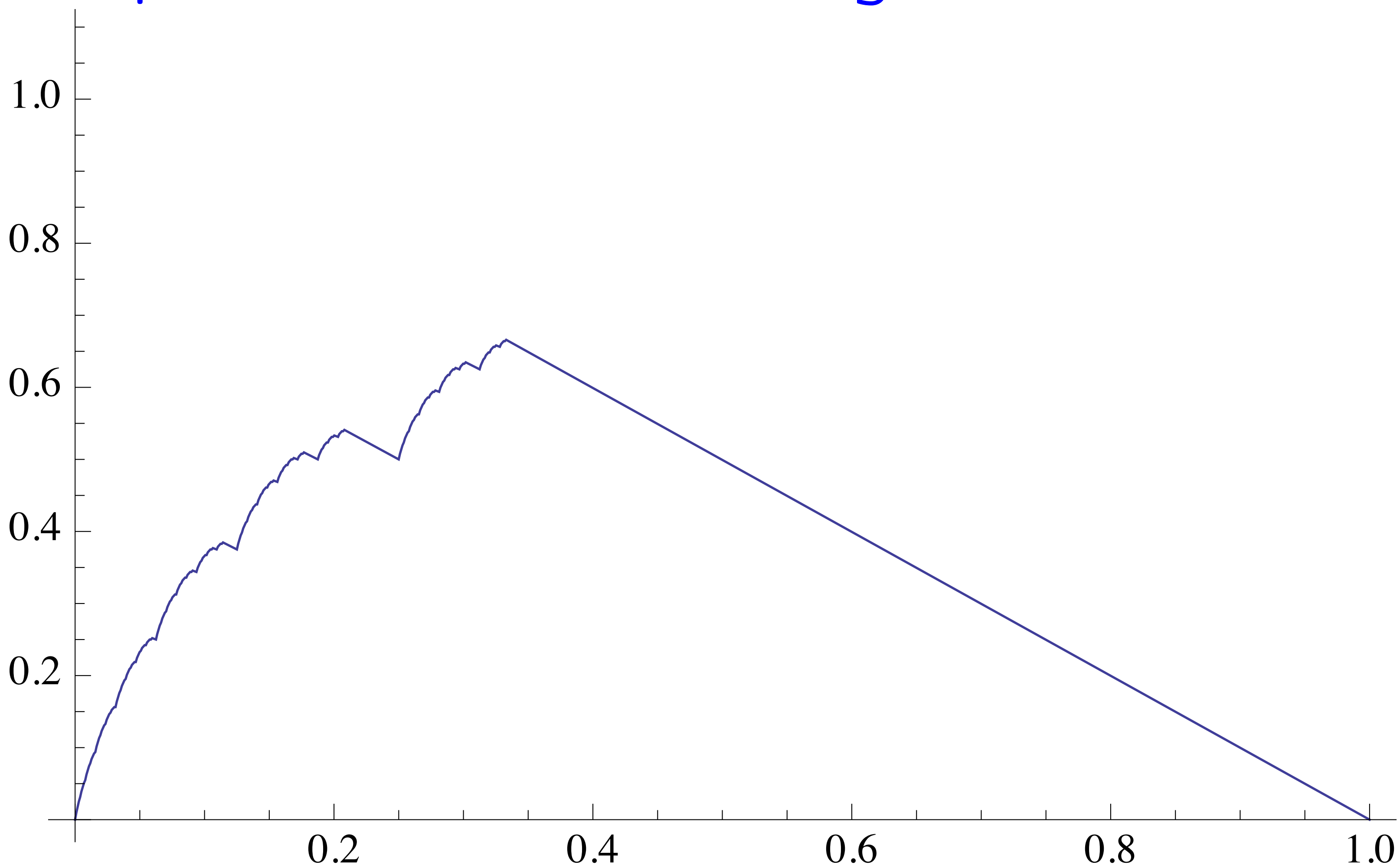
Flattened Takagi Function

- Restrict the Takagi function to Ω^L. On every open interval that was removed to construct Ω^L, linearly interpolate this function between the two endpoints.

- Call the resulting function $\tau^L(x)$ the flattened Takagi function.

- Amazing Fact. (Or Trivial Fact.) All the linear interpolations have slope -1.

Graph of Flattened Takagi Function



Flattened Takagi Function-2

- Claim. The flattened Takagi function has much less oscillation than the Takagi function. Namely...

- Theorem F.
  (1) The flattened Takagi function $\tau^L(x)$ is a function of bounded variation. That is, it is the sum of an increasing function (means: nondecreasing) and a decreasing function (means: nonincreasing).
  (This is called: Jordan decomposition of BV function.)

  (2) $\tau^L(x)$has total variation $V_0^1(\tau^L) = 2$.

- This theorem follows from...

Takagi Singular Function

- Theorem. (1) The flattened Takagi function has a minimal Jordan decomposition

$$\tau^L(x) = \tau^S(x) + (-x),$$

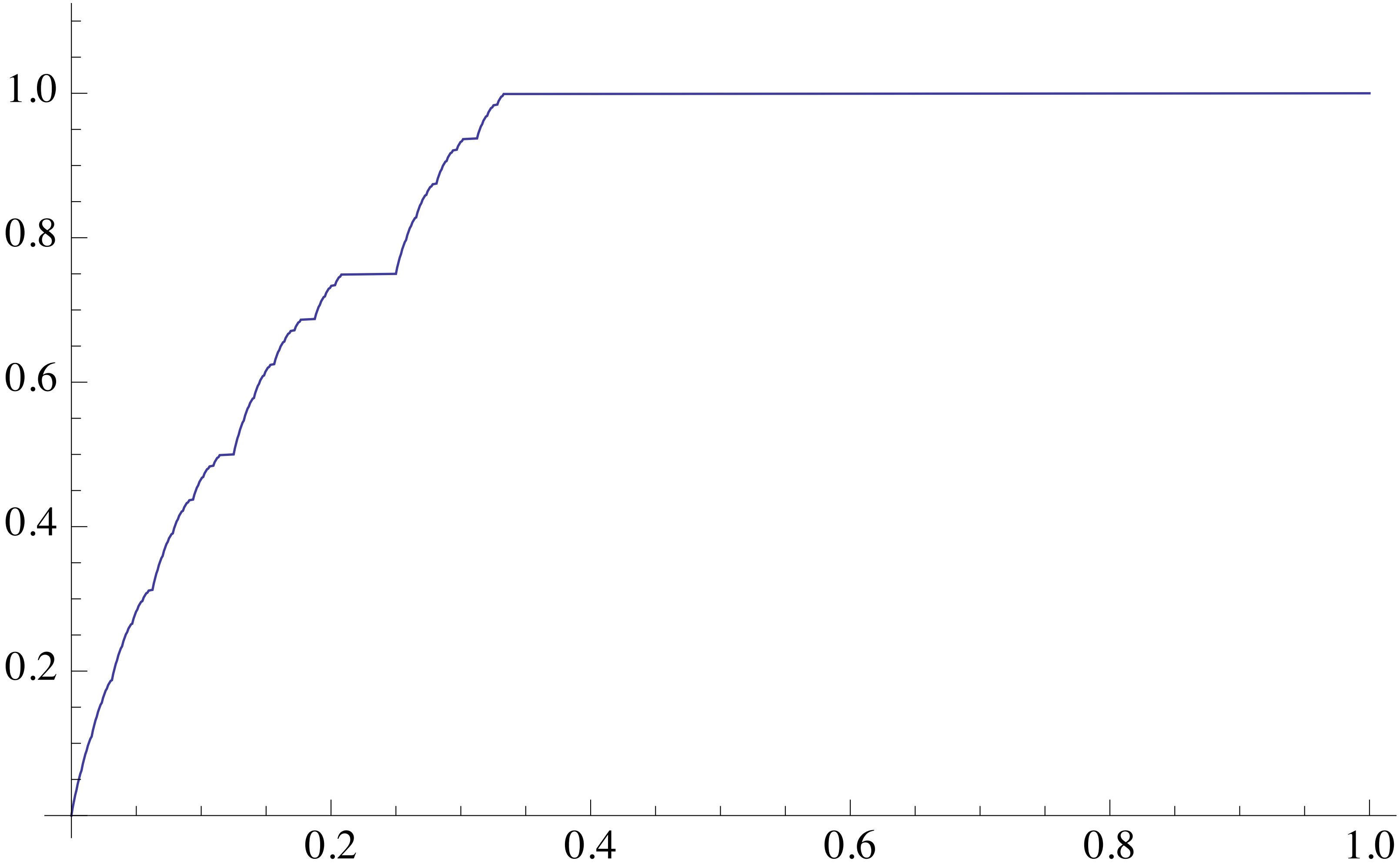
in which the function

$$\tau^S(x) := \tau^L(x) + x$$

is nondecreasing, and the function $-x$ is strictly decreasing.

(2) The function $\tau^L(x)$ is a nondecreasing singular continuous function; it has derivative 0 off the set Ω^L. Call it the Takagi singular function.

Graph of Takagi Singular Function



Takagi Singular Function

- The Takagi singular function is the integral of a singular measure:

$$\tau^S(x) = \int_0^x d\mu^S(t)$$

  Call μ^S the Takagi singular measure. It is supported on Ω^L.

- The Takagi singular measure is obviously not translation-invariant. But it satisfies various functional equations coming from those of the Takagi function. It is possible to compute with it.

Proof of Theorem B

- Compute expected value of number of local level sets at random level y using the co-Area formula for BV-functions, applied to $\tau^L(x)$.

- This counts the expected number of points of the function on each level. Exactly half of these endpoints correspond to left hand endpoint of a level set. End up with answer $\frac{3}{2}$.

Proof of Theorem C

(1) Compute the Takagi singular measure of various subsets of Ω^L, those for which the breakpoint set $Z(x)$ takes a finite value $m \geq 1$. Show that summing over $1 \leq m < \infty$ accounts for all of the Takagi singular measure. This shows that, on Ω^L only, drawing x with respect to Takagi singular measure, the number of points of $\tau^L(x) = \tau(x)$ is finite on a full measure set of Ω^L. Then carry this over to Lebesgue measure on ordinates.

(2) Explicit computation of average value shows that these subsets also shows that the expected number of points is infinite. QED

Concluding Remarks.

- Found interesting new internal structures: Takagi singular measure. Relation to random walks.

- Raised various open problems: Study the Conjugate Takagi function $U(\theta)$. Study rational levels. Study Takagi function as dynamical system (map of interval $[0, 1]$).

The End