Pieter C. Allaart, Department of Mathematics, University of North Texas, 1155 Union Circle #311430, Denton, TX 76203-5017, U.S.A.   email: `allaart@unt.edu`

Kiko Kawamura, Department of Mathematics, University of North Texas, 1155 Union Circle #311430, Denton, TX 76203-5017, U.S.A.   email: `kiko@unt.edu`

THE TAKAGI FUNCTION: A SURVEY

Abstract

This paper sketches the history of the Takagi function T and surveys known properties of T, including its nowhere-differentiability, modulus of continuity, graphical properties and level sets. Several generalizations of the Takagi function, in as far as they are based on the "tent map", are also discussed. The final section reviews a number of applications of the Takagi function to various areas of mathematics, including number theory, combinatorics and classical real analysis.

1   Introduction.

More than a century has passed since Takagi [76] published his simple example of a continuous but nowhere differentiable function, yet Takagi's function – as it is now commonly referred to despite repeated rediscovery by mathematicians in the West – continues to inspire, fascinate and puzzle researchers as never before. For this reason, and also because we have noticed that many aspects of the Takagi function continue to be rediscovered with alarming frequency, we feel the time has come for a comprehensive review of the literature. Our goal is not only to give an overview of the history and known characteristics of the function, but also to discuss some of the fascinating applications it has found –

---

some quite recently! – in such diverse areas of mathematics as number theory, combinatorics, and analysis. We also include a section on generalizations and variations of the Takagi function. In view of the overwhelming amount of literature, however, we have chosen to limit ourselves to functions based on the "tent map". In particular, this paper shall not make more than a passing mention of the Weierstrass function and is not intended as a general overview of continuous nowhere-differentiable functions. We thank Prof. Paul Humke for encouraging us to write this survey, and for issuing periodic cheerful reminders.
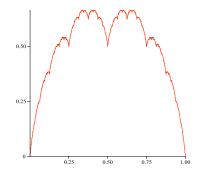
1.1 Early history

Figure 1: Graph of the Takagi function

Takagi's function is indeed simple: in modern notation, it is defined by

$$T(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} \phi(2^n x), \tag{1.1}$$

where $\phi(x) = \text{dist}(x, \mathbb{Z})$, the distance from x to the nearest integer. The graph of T is shown in Figure 1. Takagi himself expressed his function differently, and this is perhaps one reason (in combination with Japan's isolation at the beginning of the twentieth century) why it was largely overlooked in the West. Unlike for the more famous Weierstrass function, it is easy to show that T has at no point a finite derivative; we include the short proof due to Billingsley [17] in Section 2.1. However, it does possess an infinite derivative at many points, and for this reason Knopp [40], in his 1918 review of the rapidly growing body of "strange" functions, did not consider it truely nowhere differentiable. Knopp outlined his own geometric method for producing functions which have

no derivative, finite or infinite, at any point. The example most similar to the Takagi function is

$$f(x) = \sum_{n=0}^{\infty} a^n \phi(b^n x),$$

where $0 < a < 1$, b is an integer and $ab > 4$. Knopp's general construction also includes the Weierstrass function and Faber's example [26] as special cases.

A variant of Takagi's function, using the base 10, was discovered in 1930 by Van der Waerden, who credits Dr. A. Heyting for sending in the proof of its nondifferentiability, in response to a problem in the publication *Wiskundige Opgaven* of the Dutch Mathematical Society. Three years later, Hildebrandt [33] showed that one can more simply use the base 2, thus rediscovering Takagi's original example. An editorial note affixed to Hildebrandt's paper solicited answers to the "interesting and probably not too difficult" question at which set of points $T(x)$ has an infinite derivative. Surprisingly, it would take 77 years for this natural question to be answered correctly, perhaps because the answer is not all that easy to guess; see Section 2.1 below.

In 1939 the Takagi function was rediscovered by Tambs-Lyche [77], who was also inspired by Van der Waerden's paper and motivated by a desire to give "an example easy to understand for beginning students of analysis". Tambs-Lyche defined it by a formula different from both (1.1) and Takagi's original definition (see Section 4), and stated without proof its equivalence to (1.1). Tambs-Lyche was also the first to publish a graph of Takagi's function, hand-drawn but remarkably accurate.

Apparently unaware of Hildebrandt's and Tambs-Lyche's notes, de Rham [63] rediscovered Takagi's function once more in 1957. His main contribution, however, was to identify T as a member of a more general class of functions which are solutions to a certain family of functional equations; in today's language, de Rham observed that the Takagi function is self-affine. His paper soon inspired Kahane [34] to determine the points of global and local extremum of T, and to modify the definition (1.1) in order to create functions with a prescribed modulus of continuity.

By the 1960s, the Takagi-van der Waerden function was sufficiently well known that it could be used as the key element in solutions to other problems, both in classical real analysis and in number theory. Lipinski [53] used it in his elegant characterization of zero sets of continuous nowhere differentiable functions, after Schubert [67] had started the investigation. And Trollope [78] observed that the Takagi function was the missing piece of the puzzle in the binary digital sum problem; his proof was simplified and extended further by Delange [25]. These applications are described in detail in Sections 8.1 and 8.3, respectively.

1.2   The Takagi function comes home

Interest in the Takagi-van der Waerden function spiked after 1980, with two more or less independent streams of publications. In the West, Billingsley [17] drew new attention to the function with his short note in the *American Mathematical Monthly*, providing perhaps the most lucid proof of the function's nowhere differentiability. His argument was modified by Cater [22] to show that T does not even possess a finite one-sided derivative anywhere, and Shidfar and Sabetfakhri [69] proved that T is Lipschitz of every order $\alpha < 1$. A sharper result was obtained by Mauldin and Williams [59], who investigated a much larger class of functions defined by infinite series and showed that the Takagi function is "convex Lipschitz" of order $h \log(1/h)$. Anderson and Pitt [12] slightly improved on this by showing that

$$T(x + h) - T(x) = O(h \log(1/|h|)), \quad \text{as } h \to 0,$$

and this estimate is the best possible. As a result, the Hausdorff dimension of the graph of T is one.

   Meanwhile, the Takagi function had become popular in the country of its birth, due to the influential paper by Hata and Yamaguti [31]. Besides finally restoring credit to its original inventor, these authors did much to elevate Takagi's function beyond the realm of recreational mathematics, by pointing out its connection with chaotic dynamical systems and proving a beautiful relationship between T and Lebesgue's singular function (also called Salem's function or the Riesz-Nagy function). Hata and Yamaguti also replaced the factor $1/2^n$ in (1.1) by an arbitrary constant c_n, calling their new family of functions the Takagi class. Kôno [44] characterized completely the differentiability properties of members of the Takagi class – there are three qualitatively different cases – and proved several other results about these functions, most of them of a probabilistic nature. Gamkrelidze [29] later applied Kôno's methods to obtain a Central Limit Theorem-type result for the small-scale oscillations of T. In the same year as Hata and Yamaguti's paper, Baba [13] calculated the maxima of the general Takagi-van der Waerden function (with arbitrary base $r \geq 2$), after Martynov [57] had rediscovered Kahane's result about the maximum of T. Tsujii [79] constructed a Takagi-like function of two variables, and Yamaguchi et al. [82] viewed the graph of T as the invariant repeller of a dynamical system.

1.3   Recent work

In the last two decades, the literature on the Takagi function and related topics seems to have grown exponentially. Papers from this period can be loosely

classified into three categories: papers about the Takagi function itself, papers dealing primarily with applications, and papers discussing various generalizations and variations. Some papers fit more than one category. It is impossible to describe each individual contribution in this introduction. We limit ourselves here to succinct groupings of papers by topic, referring to later sections for the details.

1. *Papers about the Takagi function itself.* These can be further divided as follows. Kairies et al. [36] and Kairies [35] characterize T by its functional equations. Other papers, such as Brown and Kozlowski [19], Abbott et al. [1] and Watanabe [81], focus on various local continuity properties. The infinite derivatives of T are dealt with in Krüppel [45, 47] and Allaart and Kawamura [8]. There is also a sustained effort ongoing to understand the complicated level set structure of T; see Buczolich [21], Maddock [55], Lagarias and Maddock [50, 51], Allaart [5, 6], and de Amo et al. [9]. A richly illustrated expository article by Martynov [58] gives step-by-step explanations of the main characteristics of T, aimed at undergraduate students.

2. *Papers concerned with applications.* A number of authors have extended Trollope's result about binary digital sums in various directions. The papers most closely related to the Takagi function are Okada et al. [61], Kobayashi [42] and Krüppel [46]. Recently, Házy and Páles [32] and Boros [18] found Takagi's function to be the extremal case in the theory of approximately mid-convex functions. Their work was elaborated on by Tabor and Tabor [74, 75] and Makó and Páles [56], and this last paper includes many further references. Allaart [4] reduces the crucial inequality in the above papers to a simple inequality for binary digital sums, thus linking the two applications. Takagi's function also arises naturally as the limit in certain counting problems in graph theory; see Frankl et al. [28], Knuth [41] and Guu [30]. It even has been used in an equivalent statement of the Riemann hypothesis; see Balasubramanian et al. [14].

3. *Papers about generalizations and variations.* There are literally hundreds of papers about generalizations of the Takagi function. Many replace the tent map ϕ with a more general bounded "base" function; some also introduce random phase shifts. We will not discuss such functions here, but limit ourselves to generalizations based on the tent map. The most direct extension, a subcollection of the Takagi class, is the family of functions $f_\alpha(x) = \sum_{n=0}^{\infty} \alpha^n \phi(2^n x)$ where $0 < \alpha < 1$. They were studied for $\alpha > 1/2$ by Ledrappier [52], who computed their Hausdorff dimension, and for $\alpha < 1/2$ by Tabor and Tabor [74, 75] and Allaart [4]. Other extensions allow the

"tents" at each stage of the construction to be flipped up or down individually. This way one obtains for instance the Gray Takagi function of Kobayashi [42] or the function T^3 of Kawamura [38]. Anderson and Pitt [12], Abbott et al. [1] and Allaart [3] investigate general properties of this larger class of functions. Sekiguchi and Shiota [68], generalizing the work of Hata and Yamaguti, obtained another family of continuous functions, which were examined more closely by Allaart and Kawamura [7]. A version of the Takagi function with random signs is studied in Allaart [2], and Kawamura [39] considers the composition of T with a singular function. Finally, Sumi [73] introduces a complex version of the Takagi function in connection with random dynamics in the complex plane.

1.4   Organization of this paper

This survey is organized as follows. Section 2 focuses on analytic aspects of the Takagi function. We give Billingsley's proof of the nowhere-differentiability of T and characterize the set of points where T has an infinite derivative. In Section 2.2, we treat the Hölder continuity of T and explain the work of Abbott et al. [1] regarding slow points. This is followed by a more detailed examination of the modulus of continuity of T.

Section 3 deals with the graph of T. We first discuss the global and local extrema of T. Then we point out the partial self-similarity of the graph and illustrate how to use this to prove a specific theorem, namely that the graph of T has σ-finite linear measure.

In Section 4 we give a number of different expressions for $T(x)$, show how these can be derived from one another, and explain how they have been used to prove various aspects of the Takagi function.

Section 5 gives functional equations for T, presents T as the unique bounded solution of a system of infinitely many difference equations, and discusses the connection of T with Lebesgue's singular function.

Section 6 is devoted to the level sets of T. This area of research is currently very active: Nearly all the results in this section were found in the last five years or so. We outline a proof, based on the partial self-similarity ideas of Section 3, of the fact that almost all level sets of T are finite, and give an overview of other known facts concerning the level sets. The section ends with a list of open problems.

Section 7 gives an overview of some of the generalizations of the Takagi function and other related functions. This includes the general Takagi-van der Waerden functions, the Takagi class, and the Zygmund spaces Λ_d^*, λ_d^* and $\Lambda_{d,1}^*$, all of which are in some sense fairly direct extensions of the Takagi function. This section too concludes with a list of open problems.

Section 8 deals with applications, and is divided into four parts. Subsection 8.1 presents Trollope's formula for the sum of binary digits of the first N positive integers and discusses several related results. In Subsection 8.2 we treat applications of the Takagi function to the problem of finding the minimum shadow size in uniform hypergraphs and to the edge-discrete isoperimetric problem on the n-cube. Subsection 8.3 deals with applications in classical real analysis and consists of two parts: one on the use of T in Lipinski's characterization of zero sets of continuous nowhere differentiable functions, and one on the role of T and its generalizations in approximate convexity problems. Finally, Subsection 8.4 explains the connection between Takagi's function and the Riemann hypothesis.

We have not attempted to give equal coverage to all the players in this arena. The things we have chosen to emphasize reflect our interests and expertise, not the importance or quality of the cited works.

While we were preparing this article, we learned that Jeffrey Lagarias [49] was working on a survey paper of his own. The two surveys evolved for the most part independently, and while there is inevitably a considerable degree of overlap, the two surveys emphasize different things. For example, we treat in detail the differentiability aspects and fine structure of the graph of T, and discuss various generalizations and applications in considerable detail. (Hence the length of the last two sections of this paper.) Lagarias, on the other hand, focuses on connections of the Takagi function with several areas of analysis, including wavelets, complex power series, and dynamical systems. In view of this, we feel that our survey and that of Lagarias complement each other quite well.

1.5   Frequently used notation

We collect here some notation that will be used regularly throughout this paper. First, for definiteness, we let $\mathbb{N}$ denote the set of natural numbers, and $\mathbb{Z}_+$ the set of nonnegative integers. Most important is the binary expansion of a point $x \in [0, 1)$, which we denote by

$$x = \sum_{n=1}^{\infty} \frac{\varepsilon_n}{2^n} = 0.\varepsilon_1\varepsilon_2\ldots\varepsilon_n\ldots, \qquad \varepsilon_n \in \{0, 1\}.$$

For dyadic rational x (i.e. x of the form $x = k/2^m$ with $k \in \mathbb{Z}_+$ and $m \in \mathbb{N}$) we choose the representation ending in all zeros. When necessary, to avoid confusion, we write $\varepsilon_n(x)$ instead of ε_n.

Let $I_n := I_n(x)$ be the number of ones, and $O_n := O_n(x)$ the number of zeros, among the first n binary digits of x, and let $D_n := D_n(x) := O_n(x) - I_n(x)$. Thus, we have

$$I_n = \sum_{k=1}^{n} \varepsilon_n, \qquad O_n = n - I_n,$$

and

$$D_n = \sum_{k=1}^{n} (1 - 2\varepsilon_k) = \sum_{k=1}^{n} (-1)^{\varepsilon_k}.$$

If the limit

$$d_1(x) := \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k, \tag{1.2}$$

exists, we call $d_1(x)$ the *density* (or long-run frequency) of the digit "1" in the binary expansion of x. In that case, the number

$$d_0(x) := 1 - d_1(x)$$

is the density of the digit "0".

The orthogonal projections onto the x- and y-axes will be denoted by π_X and π_Y, respectively.

By $\mathcal{H}^\alpha$ we will denote α-dimensional Hausdorff measure, and by $\dim_H A$, the Hausdorff dimension of a set A. For a function f, $\dim_H(f)$ denotes the Hausdorff dimension of the graph of f.

2   Analytic properties.

In this section we focus on analytic aspects of the Takagi function, including infinite derivatives, Hölder continuity and slow points. We begin with a short proof of the function's nowhere-differentiability.

2.1   Derivatives, or lack thereof

Takagi himself gave a proof of the fact that T has nowhere a finite derivative [76], as did Hildebrandt [33] and de Rham [63]. Van der Waerden's simple proof for the base 10 case, however, does not immediately transfer to the case of base 2. While all published proofs of non-differentiability follow more or less the same logic, the one by Billingsley [17] is arguably the most natural, and that is the one we present here.

Theorem 2.1 (Takagi). *The Takagi function T does not possess a finite derivative at any point.*

PROOF. (Billingsley) Put $\phi_k(x) = 2^{-k}\phi(2^k x)$ for $k = 0, 1, \ldots$. Fix a point x, and, for each $n \in \mathbb{N}$, let u_n and v_n be the dyadic rationals of order n with $v_n - u_n = 2^{-n}$ and $u_n \leq x < v_n$. Then

$$\frac{T(v_n) - T(u_n)}{v_n - u_n} = \sum_{k=0}^{n-1} \frac{\phi_k(v_n) - \phi_k(u_n)}{v_n - u_n},$$

since $\phi_k(u_n) = \phi_k(v_n) = 0$ for all $k \geq n$. But for $k < n$, ϕ_k is linear on $[u_n, v_n]$ with slope $\phi_k^+(x)$, the right-hand derivative of ϕ_k at x. Thus,

$$\frac{T(v_n) - T(u_n)}{v_n - u_n} = \sum_{k=0}^{n-1} \phi_k^+(x).$$

Since $\phi_k^+(x) = \pm 1$ for each k, this last sum cannot converge to a finite limit. Hence, T does not have a finite derivative at x.   □

Billingsley's argument was modified by Cater [22] to show that T does not have a finite *one-sided* derivative anywhere. The above proof makes it plausible, however, that there exist points with $T'(x) = \pm\infty$. An Editor's note affixed to Hildebrandt's paper asked readers to characterize the set of such points. The call was answered three years later by Begle and Ayres [15], who claimed that $T'(x) = \infty$ if $D_n(x) \to \infty$, and $T'(x) = -\infty$ if $D_n(x) \to -\infty$. This is certainly believable at first sight: If we agree that for dyadic rational points we choose the binary expansion ending in all zeros, then the last equation in the above proof can be written as

$$\frac{T(v_n) - T(u_n)}{v_n - u_n} = D_n(x). \tag{2.1}$$

The limit of the left-hand side as $n \to \infty$, if it exists, is called the *dyadic derivative* of T at x. Convergence of the slopes in (2.1) to $\pm\infty$ is necessary in order that $T'(x) = \pm\infty$, but it is of course, *a priori*, not sufficient. (In fact, there are examples of nowhere differentiable functions for which the dyadic derivative exists almost everywhere [12, Example 3.3].) Begle and Ayres assumed that for fixed n, the slope $D_n(x)$ cannot jump by more than ± 2 as one moves from one dyadic interval into the next. But this is already false for $n = 4$, as $D_4(x) = -2$ for $7/16 \leq x < 1/2$, and $D_4(x) = 2$ for $1/2 \leq x < 9/16$.

The paper by Begle and Ayres appears to have been forgotten soon after its publication, as was Hildebrandt's note. In any case, there is no evidence

in the literature that the mistake was ever noticed – until a few years ago, that is. We learned of Begle and Ayres' work from an historical survey by Prof. H. Okamoto [62]. Knowing that Prof. M. Krüppel [45] had recently written about the improper derivatives of the Takagi function, we sent him a courtesy notification. Krüppel's stunning reply was that, while he did not know about Begle and Ayres, the result could not possibly be true, as his own paper contained a counterexample! (Curiously, Lemma 7.4 of Anderson and Pitt [12] implies the same incorrect statement. In their case, however, the culprit appears to be a typographical error.)

We present Krüppel's example here in somewhat simplified form. Let $x = \sum_{n=1}^{\infty} 2^{-a_n}$, where $a_n = 4^n$. Then certainly $D_n(x) \to \infty$. A well-known formula for $T(x)$ at dyadic rational points is

$$T\left(\frac{k}{2^m}\right) = \frac{1}{2^m} \sum_{j=0}^{k-1} (m - 2s_j), \tag{2.2}$$

where s_j is the number of ones in the binary representation of the integer j. (There are several ways to derive this formula; see Section 4.) For given m, let k be the integer such that $k/2^m < x < (k+1)/2^m$. Then the secant slopes over the dyadic intervals $[k/2^m, (k+1)/2^m]$ containing x indeed tend to $+\infty$ in view of (2.1). However, if we put $m = a_{n+1} - 1$, then $s_k = n$, $s_{k-1} = n + a_{n+1} - a_n - 2$, and $s_{k-2} = n + a_{n+1} - a_n - 3$. Thus, (2.2) yields

$$2^m \left[T\left(\frac{k+1}{2^m}\right) - T\left(\frac{k-2}{2^m}\right) \right] = 3m - 2s_k - 2s_{k-1} - 2s_{k-2}$$

$$= 4a_n - a_{n+1} - 6n + 7$$

$$\to -\infty,$$

as $n \to \infty$. Since the intervals $[(k-2)/2^m, (k+1)/2^m]$ also contain x, it follows that T cannot have an infinite derivative at x.

Intrigued by these developments, the present authors and Prof. Krüppel independently set out to find the correct answer. The result:

Theorem 2.2 (Allaart and Kawamura, Krüppel). *Let $x \in (0,1)$ be non-dyadic, and write*

$$x = \sum_{n=1}^{\infty} 2^{-a_n},$$

where $\{a_n\}$ is a strictly increasing sequence of positive integers. Then $T'(x) = \infty$ if and only if

$$a_{n+1} - 2a_n + 2n - \log_2(a_{n+1} - a_n) \to -\infty. \tag{2.3}$$

By the symmetry of the Takagi function, $T'(x) = -\infty$ if and only if $T'(1 - x) = \infty$. It is easy to see from the definition (1.1) that if x is a dyadic rational, then $T'_+(x) = +\infty$ and $T'_-(x) = -\infty$, where T'_+ and T'_- denote the right- and left-hand derivatives of T, respectively. Combined with these facts, Theorem 2.2 gives a complete characterization of the infinite derivatives of T.

In fact, the condition (2.3) is necessary in order that $T'_-(x) = +\infty$. For $T'_+(x) = +\infty$, it is sufficient that $a_n - 2n \to \infty$, and this can be seen to be equivalent to the Begle and Ayres condition that $D_n(x) \to \infty$. Allaart and Kawamura [8] give several examples illustrating the condition (2.3). For instance, the condition holds for $a_n = 3n$; for any increasing polynomial of degree 2 or higher; and for any exponential sequence $a_n = \lfloor \alpha^n \rfloor$ with $1 < \alpha < 2$. On the other hand, it fails whenever $\limsup_{n\to\infty} a_{n+1}/a_n > 2$. The logarithmic term in (2.3) is sometimes a difference maker: The sequence $a_n = 2^n$ does not satisfy (2.3); neither does $a_n = 2^n + n$. But $a_n = 2^n + (1+\varepsilon)n$ satisfies (2.3) for any $\varepsilon > 0$.

Theorem 2.2 implies that, if the density $d_1(x)$ exists and lies strictly between 0 and 1/2, then $T'(x) = \infty$. By symmetry, $T'(x) = -\infty$ if $1/2 < d_1(x) < 1$. As a result, the sets $\{x : T'(x) = \infty\}$ and $\{x : T'(x) = -\infty\}$ have Hausdorff dimension 1.

2.2 Continuity properties

Since T is nowhere differentiable, it is certainly not Lipschitz. However, Shidfar and Sabetfakhri [69] showed that T is Hölder continuous of any order $\alpha < 1$. That is, for each $0 < \alpha < 1$, there is a constant C_α such that

$$|T(x) - T(y)| \le C_\alpha |x - y|^\alpha,$$

for all x and y in $[0, 1]$. This result prompted an interesting question. A theorem of Marcinkiewicz says that for every Lipschitz function f on $[0, 1]$ there is a C^1 function g which agrees with f outside a set of arbitrarily small measure, and Brown and Koslowski [19] wondered if the Lipschitz requirement in this theorem can be replaced with the weaker condition that f be Hölder continuous of any order $\alpha < 1$. They show that this is not so: the Takagi function provides a counterexample, since for any set M of positive Lebesgue measure, the set of difference quotients

$$\left\{ \frac{T(x) - T(y)}{x - y} : x \in M, y \in M \text{ and } x \ne y \right\}$$

is unbounded.

While the Takagi function is not Lipschitz, it does satisfy a local Lipschitz condition at each "slow point". Abbott et al. [1] call a point x in $[0,1]$ a *slow point* with constant K (K a positive integer) if $|D_n(x)| \leq K$ for every n. They show that there is a uniform bound $P = P(K) > 0$ such that for each slow point x with constant K and for each $y \in [0,1]$, $|T(y) - T(x)| \leq P|y - x|$. They also manage to compute the Hausdorff dimension of the set of slow points with constant K; it is $1 + \log_2 r$, where $r = \cos(\pi/(2(K+1)))$. The paper by Abbott et al. also contains results for more general functions based on the tent map; we will return to it in Section 7.

The result of Shidfar and Sabetfakhri was sharpened by Anderson and Pitt [12], who showed that not only T, but every function f in the so-called Zygmund space Λ_d^* is Lipschitz of order $\theta(y) = y \log(1/y)$; that is to say, there is a constant M such that, for all x and y with $y > 0$ sufficiently small,

$$|f(x + y) - f(x)| \leq My \log(1/y).$$

From this, one can deduce that the graph of T has Hausdorff dimension 1, a result first obtained by Mauldin and Williams [59] on which we will elaborate in the next section.

More precise estimates on the oscillations of T were obtained by Kôno [44]. He describes both the "worst-case" behavior and the "typical" size (in the Lebesgue sense) of the oscillations. Let

$$\sigma_u(h) = \log_2(1/h) \qquad \text{and} \qquad \sigma_l(h) = \sqrt{\log_2(1/h)}, \qquad h > 0.$$

Theorem 2.3 (Kôno 1987). *The oscillations of the Takagi function satisfy*

$$\limsup_{|x-y| \to 0} \frac{T(x) - T(y)}{(x - y)\sigma_u(|x - y|)} = 1 = -\liminf_{|x-y| \to 0} \frac{T(x) - T(y)}{(x - y)\sigma_u(|x - y|)}.$$

The extremal case of the above theorem is rare – at most points x the oscillations are of a smaller order.

Theorem 2.4 (Kôno 1987). *For almost every $x \in [0,1]$, we have*

$$\limsup_{h \to 0} \frac{T(x + h) - T(x)}{h\sigma_l(|h|)\sqrt{2 \log \log \sigma_l(|h|)}} = 1 = -\liminf_{h \to 0} \frac{T(x + h) - T(x)}{h\sigma_l(|h|)\sqrt{2 \log \log \sigma_l(|h|)}}.$$

Kôno proves the last theorem by developing $T(x)$ in terms of Rademacher functions and applying the law of the iterated logarithm. Note that for *fixed* $x \in [0,1]$, Theorem 2.3 implies

$$-1 \leq \liminf_{h \to 0} \frac{T(x + h) - T(x)}{h \log_2(1/|h|)} \leq \limsup_{h \to 0} \frac{T(x + h) - T(x)}{h \log_2(1/|h|)} \leq 1. \qquad (2.4)$$

Within these bounds, various kinds of behavior are possible. Krüppel [45] shows that if x is dyadic rational, then

$$\lim_{h \to 0} \frac{T(x+h) - T(x)}{|h| \log_2(1/|h|)} = 1.$$

Allaart and Kawamura [8] characterize for which points x the limit

$$\lim_{h \to 0} \frac{T(x+h) - T(x)}{h \log_2(1/|h|)} \tag{2.5}$$

exists. This requires the following definition.

Definition 2.5. Let $x \in [0,1]$ be non-dyadic, and let $\{a_n\}$ and $\{b_n\}$ be the (unique) strictly increasing sequences of positive integers such that

$$x = \sum_{n=1}^{\infty} 2^{-a_n}, \qquad 1 - x = \sum_{n=1}^{\infty} 2^{-b_n}.$$

We say x is *density-regular* if $d_1(x)$ exists and one of the following holds:

(a) $0 < d_1(x) < 1$; or
(b) $d_1(x) = 0$ and $a_{n+1}/a_n \to 1$; or
(c) $d_1(x) = 1$ and $b_{n+1}/b_n \to 1$.

Theorem 2.6 (Allaart and Kawamura 2010)**.** *Let x be non-dyadic. The limit in* (2.5) *exists if and only if x is density-regular, in which case the limit is equal to $d_0(x) - d_1(x)$.*

One can also consider the probability distribution of $T(x+h) - T(x)$ for small h when x is chosen at random from $[0,1]$. Gamkrelidze [29] adapts Kôno's approach to give the following Central Limit Theorem-type result:

$$\lim_{h \downarrow 0} \lambda \left(\left\{ x : \frac{T(x+h) - T(x)}{h \sqrt{\log_2(1/h)}} \leq y \right\} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt,$$

where λ denotes Lebesgue measure.

3 Graphical properties.

Figure 1 shows the graph of the Takagi function, restricted to the interval $[0,1]$. A number of features quickly jump out. The graph is symmetric about the line $x = 1/2$, and it has cusps and local minima at the dyadic rational

points. An important aspect of the graph is that it has two $1/4$-scale copies of itself at its top; this is due to the self-affine nature of T, and leads quickly to the fact, observed by many an author, that the absolute maximum of T is attained at uncountably many points. More specifically, we have

Theorem 3.1 (Kahane 1959 [34]). *The maximum value of T is $2/3$. The set $\mathcal{M}$ of points where T attains the maximum value is a perfect set of Hausdorff dimension $1/2$, and consists of all the points x with binary expansion satisfying $\varepsilon_{2n-1} + \varepsilon_{2n} = 1$ for each n.*

PROOF. The simplest way to see this is to rewrite (1.1) as

$$T(x) = \sum_{n=0}^{\infty} \frac{1}{4^n} \phi_1(4^n x),$$

where ϕ_1 is the "table-top" function $\phi_1(x) = \phi(x) + (1/2)\phi(2x)$. Let

$$T_n(x) := \sum_{k=0}^{n-1} \frac{1}{2^k} \phi(2^k x), \tag{3.1}$$

and note that $T_{2n}(x) = \sum_{k=0}^{n-1} 4^{-k} \phi_1(4^k x)$, for $n \in \mathbb{N}$. Figure 2 shows the graphs of T_2 and T_4. One sees by induction that the maximum value of T_{2n} is

$$\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} + \cdots + \frac{1}{2}\left(\frac{1}{4}\right)^{n-1},$$

and hence,

$$M := \max\{T(x) : x \in [0,1]\} = \sum_{k=0}^{\infty} \frac{1}{2}\left(\frac{1}{4}\right)^k = \frac{2}{3}.$$

If $x \in [0,1]$, then $T(x)$ achieves this maximum value of $2/3$ if and only if x lies in the middle half of each quarternary interval to which it belongs – in other words, if the quarternary expansion of x contains only 1's and 2's. In terms of the binary expansion $x = \sum_{n=1}^{\infty} \varepsilon_n/2^n$ of x, it means that $\varepsilon_{2n-1} + \varepsilon_{2n} = 1$ for each n. Thus, the set $\mathcal{M} := \{x \in [0,1] : T(x) = M\}$ is a Cantor-like set constructed by removing at each step the two outside fourths of each remaining quarternary interval. As a result, $\dim_H \mathcal{M} = \log 2/\log 4 = 1/2$. $\qquad\square$

(Kahane did not show that the dimension of $\mathcal{M}$ is $1/2$, but he easily could have: the technique for calculating the dimensions of generalized Cantor sets was by 1959 well established.)
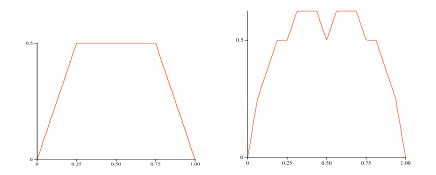
Figure 2: The functions $T_2(x) = \phi_1(x)$ (left) and $T_4(x) = \phi_1(x) + (1/4)\phi_1(4x)$ (right)

3.1   Humps and local extrema

The appearance of smaller-scale similar copies of the graph of T is not limited to the central part of the graph; it happens everywhere. We introduce two definitions and a lemma to make this precise. Let

$$\mathcal{G}_T := \{(x, T(x)) : 0 \le x \le 1\}$$

denote the graph of T over the unit interval $[0, 1]$. The term 'balanced' in the following definition is taken from Lagarias and Maddock [50].

Definition 3.2. A dyadic rational of the form $x = 0.\varepsilon_1\varepsilon_2 \ldots \varepsilon_{2m}$ is called *balanced* if $D_{2m}(x) = 0$. If $D_j(x) = 0$ for exactly n indices $1 \le j \le 2m$, we say x is a balanced dyadic rational of *generation n*. By convention, we consider $x = 0$ to be a balanced dyadic rational of generation 0. The set of all balanced dyadic rationals is denoted by $\mathcal{B}$. For each $n \in \mathbb{Z}_+$, the set of balanced dyadic rationals of generation n is denoted by $\mathcal{B}_n$. Thus, $\mathcal{B} = \bigcup_{n=0}^{\infty} \mathcal{B}_n$.

Lemma 3.3. *Let $m \in \mathbb{N}$, and let $x_0 = k/2^{2m} = 0.\varepsilon_1\varepsilon_2 \ldots \varepsilon_{2m}$ be a balanced dyadic rational. Then for $x \in [k/2^{2m}, (k+1)/2^{2m}]$ we have*

$$T(x) = T(x_0) + \frac{1}{2^{2m}} T\left(2^{2m}(x - x_0)\right).$$

In other words, the part of the graph of T above the interval $[k/2^{2m}, (k + 1)/2^{2m}]$ is a similar copy of the full graph $\mathcal{G}_T$, reduced by a factor $1/2^{2m}$ and shifted up by $T(x_0)$.
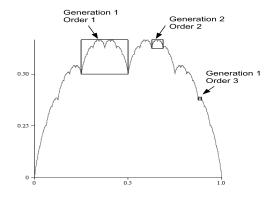
Figure 3: The "humps" $H(1/4)$, $H(5/8)$ and $H(7/8)$, enclosed in rectangles from left to right. Note that in binary, $1/4 = 0.01$, $5/8 = 0.1010$, and $7/8 = 0.111000$. Only the first of these, $H(1/4)$, is a leading hump.

PROOF. This follows immediately from the definition (1.1), since the slope of T_{2m} over the interval $[k/2^{2m}, (k+1)/2^{2m}]$ is equal to $D_{2m}(x_0) = 0$, and $T(x_0) = T_{2m}(x_0)$. □

Definition 3.4. For a balanced dyadic rational $x_0 = k/2^{2m}$ as in Lemma 3.3, let $H(x_0)$ denote the portion of the graph of T restricted to the interval $[k/2^{2m}, (k+1)/2^{2m}]$. By Lemma 3.3, $H(x_0)$ is a similar copy of the full graph $\mathcal{G}_T$; we call it a *hump*. Its height is $\frac{2}{3}(\frac{1}{4})^m$, and we call m its *order*. By the *generation* of the hump $H(x_0)$ we mean the generation of the balanced dyadic rational x_0. A hump of generation 1 will be called a *first-generation hump*. By convention, the graph $\mathcal{G}_T$ itself is a hump of generation 0. If $D_j(x_0) \geq 0$ for every $j \leq 2m$, we call $H(x_0)$ a *leading hump*. See Figure 3 for an illustration of these concepts.

The Takagi function T takes on a local maximum value at a point $x = \sum_{n=0}^{\infty} \varepsilon_n/2^n$ precisely when the point $(x, T(x))$ is located at the top of some hump. This is the case if and only if for some $m \in \mathbb{N}$,

$$\varepsilon_1 + \cdots + \varepsilon_{2m} = m, \quad \text{and} \quad \varepsilon_{2n-1} + \varepsilon_{2n} = 1 \quad \text{for each } n > m.$$

The first part of the above condition ensures that $(x, T(x))$ lies on a hump of order m; the second part implies that it lies at the top of that hump. In

particular, the points of local maximum of T lie dense in $[0, 1]$. This result too is due to Kahane [34]. On the other hand, since $T'_+(x) = \infty$ and $T'_-(x) = -\infty$ at each dyadic x, T has a local minimum value at every dyadic rational point x. Kahane shows that there are no local minima at non-dyadic points.

3.2   Humps and Hausdorff measure

It is often necessary to count the humps of a given order and/or generation, and this counting involves the *Catalan numbers*

$$C_n := \frac{1}{n+1} \binom{2n}{n}, \qquad n = 0, 1, 2, \ldots.$$

Lemma 3.5. *Let $m \in \mathbb{N}$.*

   (i) There are $\binom{2m}{m}$ humps of order m.

   (ii) There are C_m leading humps of order m.

   (iii) There are $2C_{m-1}$ first-generation humps of order m.

This lemma is helpful in the study of the level sets of T (see Section 6). Another use is the following. Mauldin and Williams [59] first showed that the graph of T has Hausdorff dimension one, but remarked that they did not know whether it has σ-finite linear measure. Anderson and Pitt [12] showed that the answer is affirmative, not only for the Takagi function but for a much wider class of functions (the so-called Zygmund space Λ_d^*). Odani [60] explicitly decomposed the graph of T into countably many sets of finite linear measure, as follows.

Let S denote the set of points (x, y) on $\mathcal{G}_T$ which belong to humps of infinitely many generations, and for $n = 0, 1, 2, \ldots$, let E_n denote the set of points which belong to a hump of generation n, but not to a hump of generation $n + 1$. Then

$$\mathcal{G}_T = S \cup E_0 \cup E_1 \cup E_2 \cup \ldots.$$

Since E_0 is the graph of T with all the first-generation humps removed, it is intuitively clear (and can be made rigorous) that the restriction of T to $\pi_X(E_0)$ is monotone increasing on $[0, 1/2]$, and monotone decreasing on $[1/2, 1]$. Hence, E_0 has finite linear measure. Next, E_1 consists of countably many copies of E_0, one inside each first-generation hump. For each $m \in \mathbb{N}$, there are $2C_{m-1}$ copies with contraction ratio $1/4^m$ by Lemma 3.5. Thus,

$$\mathcal{H}^1(E_1) = \sum_{m=1}^{\infty} 2C_{m-1} \left(\frac{1}{4}\right)^m \mathcal{H}^1(E_0) = \frac{1}{2} \sum_{n=0}^{\infty} C_n \left(\frac{1}{4}\right)^n \mathcal{H}^1(E_0) = \mathcal{H}^1(E_0),$$

where we have used the well-known fact that $\sum_{n=0}^{\infty} C_n(1/4)^n = 2$. Inductively, this argument can be continued to show that $\mathcal{H}^1(E_n) < \infty$ for each n. It remains to verify that $\mathcal{H}^1(S) < \infty$. Order the first-generation humps $H_1, H_2, \ldots$ in some arbitrary manner, and for $i \in \mathbb{N}$, let Φ_i be the similarity map which maps $\mathcal{G}_T$ onto H_i. Then it is easy to check that

$$S = \bigcup_{i=1}^{\infty} \Phi_i(S).$$

The open set condition is satisfied (take $(0, 1) \times \mathbb{R}$, say). As above, the contraction ratios of the Φ_i sum to 1, so Moran's equation gives $\dim_H S = 1$. An easy exercise (it is clear which coverings to use) shows that $\mathcal{H}^1(S) < \infty$. Thus, the graph of T has σ-finite linear measure. (In fact, $\mathcal{H}^1(S) > 0$ as well, since $\pi_X(S)$ has full Lebesgue measure.)

Essentially the same construction is given by Buczolich [21, Theorem 9]. He shows additionally that the set S is an "irregular" 1-set, meaning that it intersects every continuously differentiable curve in a set of $\mathcal{H}^1$-measure zero. In [20, Theorem 10], Buczolich shows also that the Takagi function is "micro self-similar", in the sense that the graph of T itself is a micro tangent set of T at almost every point $x \in [0, 1]$.

4   Alternative representations of $T(x)$.

While (1.1) is arguably the simplest and certainly the most common expression for $T(x)$, many other representations occur in the literature, and most have some unique advantage in proving certain things about the Takagi function or its generalizations.

1. *Dynamical systems view.* To begin, put $\psi(x) := 2\phi(x)$ for $x \in [0, 1]$, and note that ψ is a special case of a "tent map", which maps $[0, 1]$ onto itself. It is easy to see that we can write (1.1) as

$$T(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} \psi^{(n)}(x), \qquad x \in [0, 1], \tag{4.1}$$

where $\psi^{(n)}$ denotes n-fold iteration of ψ. Since $\sum_{n=1}^{\infty} 1/2^n = 1$, (4.1) represents $T(x)$ as a weighted average of the iterates of x under the chaotic dynamical system ψ.

2. *Takagi's definition.* Most authors define the Takagi function by either (1.1) or (4.1), but it should be pointed out that Takagi himself defined $T(x)$

differently. For $n \in \mathbb{N}$, let a_n denote the number of binary digits among $\{\varepsilon_1, \ldots, \varepsilon_{n-1}\}$ that are different from ε_n. In other words, $a_n = O_n(x)$ if $\varepsilon_n = 1$, and $a_n = I_n(x)$ if $\varepsilon_n = 0$. Takagi defined $T(x)$ by

$$T(x) = \sum_{n=1}^{\infty} \frac{a_n}{2^n}. \tag{4.2}$$

To see that (1.1) and (4.2) are equivalent, express $\phi(x)$ in terms of the binary expansion of x by

$$\phi(x) = \sum_{k=1}^{\infty} \frac{\varepsilon_k(1 - \varepsilon_1) + (1 - \varepsilon_k)\varepsilon_1}{2^k}.$$

This generalizes to

$$\frac{\phi(2^n x)}{2^n} = \sum_{j=1}^{\infty} \frac{\varepsilon_{n+j}(1 - \varepsilon_{n+1}) + (1 - \varepsilon_{n+j})\varepsilon_{n+1}}{2^{n+j}}. \tag{4.3}$$

We can similarly write $a_n = \varepsilon_n O_n(x) + (1 - \varepsilon_n)I_n(x)$. Inserting (4.3) into (1.1) and interchanging summations it is now easy to obtain (4.2). (This is essentially the proof given by Lagarias and Maddock [50, Lemma 2.1].)

Kôno [44], and later Gamkrelidze [29], used a form similar to (4.2) (expressing a_n in terms of the Rademacher functions $X_n(x) = (-1)^{\varepsilon_n}$) to investigate probabilistic properties of the graph of T.

3. *Tambs-Lyche's definition.* In 1939, Tambs-Lyche [77] gave the following expression for $T(x)$. Write

$$x = \sum_{j=1}^{\infty} 2^{-l_j},$$

where $\{l_j\}$ is a strictly increasing sequence of integers (the sum being finite if x is dyadic). Then

$$T(x) = \sum_{j=1}^{\infty} \frac{l_j - 2(j - 1)}{2^{l_j}}. \tag{4.4}$$

This formula is useful for approximating solutions to the equation $T(x) = y$, as shown in [6, Section 4.2].

Tambs-Lyche actually *defined* his function by the summation in the right hand side of (4.4), and stated without proof that it is equivalent to (1.1). Tambs-Lyche's formula too has been rediscovered many times. In her thesis

and subsequent publication [38], Kawamura used a relationship with Lebesgue's singular function (explained in the next section) to obtain the expression

$$T(x) = \sum_{n=1}^{\infty} \varepsilon_n(x) \frac{O_n(x) - I_n(x) + 2}{2^n} = \sum_{n=1}^{\infty} \varepsilon_n(x) \frac{n - 2(I_n(x) - 1)}{2^n}, \quad (4.5)$$

which is clearly equivalent to (4.4). De Amo and Fernández-Sánchez [11] derive (4.4) explicitly from (1.1), while Kuroda [48] deduces it directly from Takagi's definition (4.2). Here we give a short proof using (1.1). Recall the definition of T_n from (3.1), and note that T_n is piecewise linear with slope $D_n(x)$ at all points not of the form $j/2^n$. Moreover, $T(j/2^n) = T_n(j/2^n)$. Thus, if $0 < l \le m$ and $j \in \mathbb{Z}_+$, we have

$$T\left(\frac{j}{2^l} + \frac{1}{2^m}\right) - T\left(\frac{j}{2^l}\right) = \frac{1}{2^m} D_m\left(\frac{j}{2^l}\right) = \frac{m - 2s_j}{2^m}, \quad (4.6)$$

where s_j is the number of 1's in the binary expansion of j. This immediately gives $T(1/2^m) = m/2^m$, and a straightforward induction argument yields

$$T\left(\sum_{j=1}^n 2^{-l_j}\right) = \sum_{j=1}^n \frac{l_j - 2(j-1)}{2^{l_j}}$$

for all $n \in \mathbb{N}$ and integers $1 \le l_1 < l_2 < \cdots < l_n$. The continuity of T gives (4.4).

As a by-product of (4.6) (putting $l = m$ and summing over j), we obtain the formula given by Krüppel [45]:

$$T\left(\frac{k}{2^m}\right) = \frac{1}{2^m} \sum_{j=0}^{k-1} (m - 2s_j), \quad (4.7)$$

which was used in Section 2.1.

4. *Random walk definition.* Lagarias and Maddock [50, Section 2] use (4.2) to express $T(x)$ in terms of the sequence $\{D_n(x)\}$ as follows:

$$T(x) = \frac{1}{2} - \frac{1}{4} \sum_{n=1}^{\infty} (-1)^{\varepsilon_{n+1}} \frac{D_n(x)}{2^n}. \quad (4.8)$$

This formula is useful in the study of level sets, because one easily infers from it the following important fact.

Lemma 4.1. *If* $|D_n(x)| = |D_n(x')|$ *for every* n, *then* $T(x) = T(x')$.

Note that $\{D_n(x)\}_n$ is a symmetric simple random walk when x is chosen at random in $[0, 1]$. Thus, (4.8) expresses T as a functional of a random walk.

5. *Fourier series.* While the definition (1.1) gives $T(x)$ directly as a Schauder series, it is also relatively easy to compute the Fourier series for $T(x)$. Hata and Yamaguti [31] point out that

$$\phi(x) = \frac{1}{4} - \frac{2}{\pi^2} \sum_{k \in \mathbb{N},\ k\ \text{odd}} \frac{\cos 2\pi k x}{k^2},$$

and this results in the Fourier series

$$T(x) = \frac{1}{2} - \frac{2}{\pi^2} \sum_{m=1}^{\infty} A_m \cos 2\pi m x,$$

where $A_m = (2^n k^2)^{-1}$ if $m = 2^n k$, with k odd. The Fourier coefficients A_m satisfy $1/m^2 \leq A_m \leq 1/m$. In particular, the Fourier series of $T(x)$ is non-lacunary, in contrast to the Weierstrass function which is defined as a lacunary Fourier series.

5   Functional and difference equations.

De Rham [63] was the first to point out that the Takagi function on $[0, 1]$ satisfies the functional equation

$$f(x) = \begin{cases} (1/2)f(2x) + x, & 0 \leq x \leq 1/2, \\ (1/2)f(2x - 1) + (1 - x), & 1/2 \leq x \leq 1. \end{cases} \tag{5.1}$$

Kairies, Darsow and Frank [36] observed (5.1) and proved the following results:

1. Any function $f : [0, 1] \to \mathbb{R}$ satisfying (5.1) is nowhere differentiable, and coincides with T on the dyadic rationals.

2. If $f : [0, 1] \to \mathbb{R}$ satisfies (5.1) and is bounded, then $f = T$.

3. A recursive relation for the moments $M_n = \int_0^1 x^n f(x)\, dx$ of any function f satisfying (5.1) is given by

$$M_0 = 1/2, \qquad M_n = \frac{1}{(n+1)(n+2)} + \frac{1}{2(2^{n+1} - 1)} \sum_{k=0}^{n-1} \binom{n}{k} M_k.$$

(The paper [36] was inadvertently printed before the page proofs were received; a list of corrections is given in [24].) Later, extending the results of [36], Kairies [35] gave a list of seven functional equations satisfied by $T(x)$, and investigated which subsets of these equations imply that a bounded function f satisfying them must in fact be the Takagi function.

In 1984, Hata and Yamaguti [31] started a new direction by regarding the Takagi function and related functions as solutions of discrete boundary value problems. This was quite natural, since (1.1) gives $T(x)$ directly as a Schauder series, and from the Schauder expansion of a function one quickly obtains an infinite system of difference equations which the function satisfies. Using these difference equations, Hata and Yamaguti showed that the Takagi function is closely related to another special function, namely Lebesgue's singular function, which plays an important role in many areas of probability. Kawamura [38] later adopted their approach and found a close relationship between other nowhere differentiable functions, singular functions, and self-similar sets in the plane.

First, we briefly recall Schauder expansions. Whereas a function's Fourier expansion uses trigometric functions, the Schauder expansion uses the "tent" functions

$$S_{n,i}(x) = \begin{cases} 2\phi(2^n x), & \text{if } \frac{i}{2^n} \leq x \leq \frac{i+1}{2^n} \\ 0, & \text{otherwise}, \end{cases}$$

for $0 \leq i \leq 2^n - 1$ and $n \in \mathbb{Z}_+$. Thus, the graph of $S_{n,i}$ is the regular isosceles triangle of unit height whose base is the interval $[i/2^n, (i+1)/2^n]$.

It is well known that every continuous function $f : [0,1] \to \mathbb{R}$ which vanishes at 0 and 1 has a unique Schauder expansion of the form

$$f(x) = \sum_{n=0}^{\infty} \sum_{i=0}^{2^n-1} a_{n,i} S_{n,i}(x), \tag{5.2}$$

where

$$a_{n,i} = f\left(\frac{2i+1}{2^{n+1}}\right) - \frac{1}{2}\left\{f\left(\frac{i}{2^n}\right) + f\left(\frac{i+1}{2^n}\right)\right\}.$$

Applying this to the Takagi function, we immediately obtain

Theorem 5.1 (Hata-Yamaguti, 1983). *The Takagi function $T(x)$ is the unique continuous solution of the discrete boundary value problem*

$$T\left(\frac{2i+1}{2^{n+1}}\right) - \frac{1}{2}\left\{T\left(\frac{i}{2^n}\right) + T\left(\frac{i+1}{2^n}\right)\right\} = \frac{1}{2^{n+1}}, \tag{5.3}$$

where $0 \leq i \leq 2^n - 1$, $n \in \mathbb{Z}_+$, and the boundary conditions are $T(0) = T(1) = 0$.

Next, recall Lebesgue's singular function. Imagine flipping an unfair coin with probability $r \in (0, 1)$ of heads and probability $1 - r$ of tails. Note that $r \neq 1/2$. Let the binary expansion of $t \in [0, 1]$: $t = \sum_{n=1}^{\infty} \omega_n / 2^n$ be determined by flipping the coin infinitely many times. More precisely, $\omega_n = 0$ if the n-th toss is heads and $\omega_n = 1$ if it is tails. We define *Lebesgue's singular function* $L_r(x)$ as the distribution function of t:

$$L_r(x) := Prob\{t \leq x\}, \qquad 0 \leq x \leq 1.$$

With the function L_r is associated a probability measure μ_r on $[0, 1]$, called the *binomial measure*, under which the binary digits of a number $t \in [0, 1]$ are independent, taking the values 0 and 1 with probabilities r and $1 - r$, respectively.

It is well-known that $L_r(x)$ is strictly increasing, but its derivative is zero almost everywhere. De Rham [64] showed that $L_r(x)$ is the unique continuous solution of the functional equation

$$L_r(x) = \begin{cases} rL_r(2x), & 0 \leq x \leq \frac{1}{2}, \\ (1 - r)L_r(2x - 1) + r, & \frac{1}{2} \leq x \leq 1. \end{cases} \qquad (5.4)$$

In particular, the graph of L_r is self-affine. Hata and Yamaguti showed that $L_r(x)$ is also the unique continuous solution of the following discrete boundary value problem:

$$L_r\left(\frac{2i + 1}{2^{n+1}}\right) = (1 - r)L_r\left(\frac{i}{2^n}\right) + rL_r\left(\frac{i + 1}{2^n}\right), \qquad (5.5)$$

where $0 \leq i \leq 2^n - 1$ and $n \in \mathbb{Z}_+$. The boundary conditions are $L_r(0) = 0$ and $L_r(1) = 1$. From (5.3) and (5.5), they proved the important and useful relationship

$$\frac{1}{2} \frac{\partial}{\partial r} L_r(x)\Big|_{r=1/2} = T(x). \qquad (5.6)$$

This identity can also be obtained from the following expression for $L_r(x)$, due to Lomnicki and Ulam [54]:

$$L_r(x) = \frac{r}{1 - r} \sum_{n=1}^{\infty} \varepsilon_n r^{n - I_n} (1 - r)^{I_n}. \qquad (5.7)$$

Differentiating this with respect to r and setting $r = 1/2$ gives the right hand side of (4.5), and hence we have (5.6). (In [38] the reverse approach is taken,

and (4.5) is derived from (5.7) using (5.6).) We note that (5.6) also leads to a very short proof of (4.7): It is easy to see that

$$L_r\left(\frac{k}{2^m}\right) = \sum_{j=0}^{k-1} r^{m-s_j}(1-r)^{s_j},$$

where s_j is the number of 1's in the binary expansion of j. Differentiating gives (4.7).

The functional equations (5.1) and (5.4) are both special cases of the general family of functional equations studied by de Rham [64]. De Rham considers the Takagi function and Lebesgue's singular function in separate papers, and does not appear to have noticed the relationship (5.6).

5.1 Evaluating $T(x)$ for rational x

One particular use of the functional equation (5.1) is to the exact evaluation of $T(x)$ for rational x. As noted by Knuth [41, p. 32, p. 103], $T(x)$ is rational whenever x is, and by applying (5.1) repeatedly one obtains a system of linear equations which is easily solved. We give the details here, and also examine the number of iterations required to compute $T(x)$.

Let $x = p/q$, where $p, q \in \mathbb{N}$ with $gcd(p,q) = 1$. Assume first that $p < q/2$. Then, by (5.1) and the symmetry of T, we can write $T(p/q) = \frac{1}{2}T(p'/q)+(p/q)$, where $p' = \min\{2p, q - 2p\}$. If q is even, the fraction p'/q simplifies. If q is odd, then $gcd(p',q) = 1$ again. These ideas lead to the following two-stage algorithm for evaluating $T(p/q)$:

Step 1. Let $q = 2^m q'$, with q' odd, and assume $gcd(p,q) = 1$. Put $q_0 := q$, $p_0 := \min\{p, q-p\}$, and $q_j := q_{j-1}/2$, $p_j := \min\{p_{j-1}, q_j - p_{j-1}\}$, $j = 1, \ldots, m$. Let $p' := p_m$. Then, after applying the functional equation m times, putting the results together and simplifying, we obtain

$$T\left(\frac{p}{q}\right) = \frac{1}{2^m}T\left(\frac{p'}{q'}\right) + \frac{1}{q}\sum_{j=0}^{m-1} p_j.$$

So it remains to compute $T(p/q)$ for odd q.

Step 2. Let q be odd and $gcd(p,q) = 1$. Put $p_0 := \min\{p, q-p\}$, and $p_j := \min\{2p_{j-1}, q - 2p_{j-1}\}$, $j = 1, 2, \ldots$. Note that $T(p_j/q) = \frac{1}{2}T(p_{j+1}/q)+(p_j/q)$ for each $j \geq 0$. Since $p_j \equiv \pm 2p_{j-1} \pmod q$, we have $p_j \equiv \pm 2^j p_0 \pmod q$ for each j, so there will be some positive integer j such that $p_j = p_0$. The smallest such j is the number $k := \min\{j \in \mathbb{N} : 2^j \equiv \pm 1 \pmod q\}$. Such a j

always exists by Euler's theorem, and $k \leq \varphi(q)$, where the Euler function $\varphi(q)$ denotes the number of integers in $\{1, 2, \ldots, q-1\}$ relatively prime to q. Let $\operatorname{ord}_q(2)$ denote the order of 2 in the group of units of $\mathbb{Z}/q\mathbb{Z}$. That is, $\operatorname{ord}_q(2)$ is the smallest positive integer n such that $2^n \equiv 1 \pmod{q}$. It is an elementary exercise in number theory to show that

$$k = \begin{cases} \frac{1}{2} \operatorname{ord}_q(2), & \text{if } q | 2^j + 1 \text{ for some } j \in \mathbb{N} \\ \operatorname{ord}_q(2), & \text{otherwise.} \end{cases}$$

It now takes precisely k iterations of the functional equation to express $T(p/q)$ in terms of itself. Solving for $T(p/q)$ and simplifying, we eventually obtain

$$T\left(\frac{p}{q}\right) = \frac{1}{q(2^k - 1)} \sum_{j=0}^{k-1} 2^{k-j} p_j.$$

The inverse problem is also interesting: given a rational y in the range of T, is there a rational point $x \in [0,1]$ such that $T(x) = y$? This question is still open. Knuth [41, p. 103] gives an algorithm which produces solutions of the equation $T(x) = y$ for many rational y. This involves starting with an initial value v and walking along a particular directed graph, updating v by some fixed arithmetic operation at each node. Which node may be visited next depends not only on the graph but also on whether the transition will keep the value of v within certain bounds. The algorithm terminates when one visits a node for the second time with the same value of v. It is, however, not known whether the algorithm always terminates. A different method which gives solutions for many (but not all) rational y is given by Allaart [6, Section 4.2].

6   Level sets.

In this section we consider the level sets

$$L(y) := \{x \in [0,1] : T(x) = y\}, \qquad y \in \mathbb{R}.$$

Of course, $L(y) = \emptyset$ if $y \notin [0, 2/3]$. The simplest level set is $L(0) = \{0, 1\}$. At the other extreme, in view of Theorem 3.1, we have that $L(2/3)$ is an uncountable (Cantor) set of dimension $1/2$. In general, $L(y)$ can be finite, countably infinite or uncountable, and which of these three possibilities is the most common depends on the precise mathematical meaning assigned the word "most". If $L(y)$ is finite, it must have an even number of points by the symmetry of the graph of T, but any even positive number is possible. If $L(y)$

is uncountable, its Hausdorff dimension can be zero or strictly positive, but never more than $1/2$.

Level sets are partitioned into easier to understand pieces called *local level sets*. Each local level set is either finite or a Cantor set, and the members of a local level set are easily obtained from one another by certain combinatorial operations ("block flips") on their binary expansions. A level set can consist of finitely many, countably infinitely many, or uncountably many local level sets. As with the cardinalities of the level sets, which of these three is the most common depends on how one defines "most". Many questions about the level sets of T remain open.

6.1   Finite or infinite?

The level sets of the Takagi function and related functions were first considered by Anderson and Pitt [12]. Their Theorem 7.3, which applies to a large class of functions, implies that $L(y)$ is countable for almost every $y \in [0, \frac{2}{3}]$. This result was improved recently by Buczolich [21], who focused on the Takagi function itself and concluded the following.

Theorem 6.1 (Buczolich 2008)**.** *For almost every ordinate y, $L(y)$ is finite.*

SKETCH OF PROOF. (Allaart 2011). We sketch a proof here that is slightly different from the original proof by Buczolich, and which uses the notion of humps and leading humps defined in Section 3. For the full details, see [5, Section 3]. Observe first that Lemma 4.1 immediately implies the following.

Lemma 6.2. *For every hump H there is a leading hump H' of the same order and generation as H, such that $\pi_Y(H) = \pi_Y(H')$. On the other hand, for every leading hump H' there are only finitely many humps H such that $\pi_Y(H) = \pi_Y(H')$.*

Next, define a set

$$X^* := [0,1] \backslash \bigcup_{x_0 \in \mathcal{B}_1} I(x_0).$$

In other words, X^* is obtained by removing all the dyadic closed intervals above which the graph of T has a first-generation hump. The importance of X^* is made clear by the next lemma, which we state here without proof.

Lemma 6.3. *The Takagi function T maps X^* onto $[0, \frac{1}{2}]$. Moreover, T is strictly increasing on $X^* \cap [0, \frac{1}{2})$.*

Lemmas 6.2 and 6.3 can be used to prove the crucial fact that $L(y)$ is finite whenever the horizontal line l_y at level y intersects only finitely many leading

humps. Using this fact, the proof of the theorem can be completed as follows. If y is chosen at random from $[0, \frac{2}{3}]$ and H is a leading hump of order m, the probability that the line l_y intersects H is $\left(\frac{1}{4}\right)^m$. Letting $\mathcal{H}'$ denote the set of all leading humps, this gives

$$\sum_{H \in \mathcal{H}'} \mathrm{P}(y \in \pi_Y(H)) = \sum_{m=0}^{\infty} C_m \left(\frac{1}{4}\right)^m < \infty,$$

since there are C_m leading humps of order m by Lemma 3.5, and $C_m \sim 4^m/(m^{3/2}\sqrt{\pi})$. Thus, by the Borel-Cantelli lemma, the probability that l_y intersects infinitely many leading humps is zero. Therefore, $L(y)$ is finite with probability 1. $\qquad\square$

Despite the above result, the *average* cardinality of the level sets of T is infinite. That is,

$$\int_0^{2/3} |L(y)| = \infty.$$

This was shown by Lagarias and Maddock [51]. An alternative proof based on Lemma 6.3 is given by Allaart [5].

Whereas the above results are probabilistic in nature, a quite different picture emerges when one views the level sets of T from the perspective of Baire category. Define the sets

$$S_\infty^{co} := \{y \in [0, \tfrac{2}{3}] : L(y) \text{ is countably infinite}\},$$
$$S_\infty^{uc} := \{y \in [0, \tfrac{2}{3}] : L(y) \text{ is uncountably infinite}\}.$$

Theorem 6.4 (Allaart 2011). *The set S_∞^{uc} has the decomposition $S_\infty^{uc} = E \cup M$, where E is a dense G_δ set, and M is a countable set disjoint from E which consists exactly of the local maximum ordinate values of T. As a result, the set $\{y \in [0, \tfrac{2}{3}] : L(y)$ is countable$\}$ is of the first category.*

For the proof and a more explicit description of the set E, see Allaart [5, Section 4]. There it is also shown that S_∞^{uc} does not contain any dyadic rational ordinates y. Note that the residual set S_∞^{uc} has Lebesgue measure zero by Theorem 6.1. However, it has full Hausdorff dimension one; see Lagarias and Maddock [51].

The set S_∞^{co} is rather more difficult to describe. It contains the images of all dyadic rational abscissas x in $[0, 1]$, so it is dense in $[0, \frac{2}{3}]$. But it is not known whether it contains more.

Closely related to the level sets of T is the *occupation measure* defined by $\mu_T(A) = \lambda(\{x \in [0, 1] : T(x) \in A\})$, for Borel sets $A \subset \mathbb{R}$. Buczolich

[21] shows that μ_T is singular with respect to Lebesgue measure. This is witnessed by the set S from Section 3.2: It is relatively straightforward to show that $|L(y)| = \infty$ when $y \in \pi_Y(S)$, so Theorem 6.1 implies $\lambda(\pi_Y(S)) = 0$. On the other hand, $\pi_X(S)$ has full measure in $[0, 1]$, because almost every $x \in [0, 1]$ has the property that $D_n(x) = 0$ for infinitely many n. Consequently, $\mu_T(\pi_Y(S)) = \lambda(\pi_X(S)) = 1$.

6.2   Cardinalities of finite level sets

Since almost all level sets are finite in the Lebesgue sense, it is natural to ask what these finite cardinalities can be. By the symmetry of T and the fact that $L(T(\frac{1}{2})) = L(\frac{1}{2})$ is countably infinite, the number of points in each finite level set must be even. Allaart [6] shows that, vice versa, every positive even number occurs. In fact, we have the following. Let

$$S_{2n} := \{y : |L(y)| = 2n\}, \qquad n \in \mathbb{N}.$$

Theorem 6.5 (Allaart 2011)**.** *For each $n \in \mathbb{N}$, S_{2n} is uncountable but nowhere dense.*

It is not known whether S_{2n} has positive Lebesgue measure for each n. The argument used in the proof of Theorem 6.5 unfortunately does not give enough to prove this stronger statement. As a partial result in this direction, however, Allaart [6] shows that $\lambda(S_{2n}) > 0$ whenever n is either a power of 2, or the sum or difference of two powers of 2. And for the specific case of S_2, fairly tight bounds on its Lebesgue measure can be given which show, not surprisingly perhaps, that 2 is the most common finite cardinality.

Theorem 6.6 (Allaart 2011)**.** *The Lebesgue measure $\lambda(S_2)$ of S_2 satisfies*

$$\frac{5}{12} < \lambda(S_2) < \frac{35}{72}.$$

Since the graph of T has height $2/3$, this implies that if an ordinate y is chosen at random in the range of T, the probability that $L(y)$ contains exactly two points lies between 62.5% and 72.9%. The proof of Theorem 6.6 is based on a simple counting argument which uses the fact, from Lemma 3.5, that the graph of T contains exactly C_{m-1} first-generation leading humps. See [6] for the details.

Which specific ordinates y satisfy $|L(y)| = 2$? It is clear from the graph that y must be less than $\frac{1}{2}$. Allaart [6] gives two sufficient conditions, the first condition being directly in terms of the binary expansion of y.

Theorem 6.7 (Allaart 2011). *Let $0 < y < \frac{1}{2}$ such that y is not a dyadic rational, and suppose the binary expansion of y does not contain a string of three consecutive 0's anywhere after the occurrence of its first 1. Then $|L(y)| = 2$.*

Thus, for instance, the set S_2 includes the points $y = 1/3 = 0.(01)^\infty$, $y = 1/7 = 0.(001)^\infty$, and $y = 1/40 = 0.0^5(1100)^\infty$.

The second condition, which is neither weaker nor stronger than the first, involves the orbit of y under iteration of the map

$$\Psi(x) = \begin{cases} 0, & \text{if } y = 0 \\ 4\left(y - \frac{k}{2^k}\right), & \text{if } \frac{k}{2^k} \le y < \frac{k-1}{2^{k-1}}, \quad k = 3, 4, \dots. \end{cases}$$

Note that Ψ maps $[0, \frac{1}{2})$ onto itself. Let Ψ^n denote the nth iterate of Ψ, with $\Psi^0(y) := y$. For $n = 0, 1, 2, \dots$, let k_n be that number $k \ge 3$ for which $k/2^k \le \Psi^n(y) < (k-1)/2^{k-1}$, or put $k_n = \infty$ if $\Psi^n(y) = 0$.

Theorem 6.8 (Allaart 2011). *Let $0 \le y < \frac{1}{2}$. If $k_{n+1} \le 2k_n$ for every n, then $|L(y)| = 2$.*

This gives many more examples. For instance, the condition of the theorem obviously holds for the fixed points of Ψ, which are $y_k^* := k/(3 \cdot 2^{k-2})$, $k \ge 4$. Perhaps unexpectedly, this shows that infinitely many dyadic rational ordinates belong to S_2, the first three being $1/8$, $3/2^7$ and $1/2^8$. More generally, of course, many of the periodic points of Ψ satisfy $k_{n+1} \le 2k_n$ and are therefore in S_2. For instance, $y = 1/11$ does not satisfy the "no 3 zeros" condition of Theorem 6.7, but it has $(k_n)_{n \ge 0} = (7, 6, 5, 5, 6, 5, 5, 4, 4, 4)^\infty$, and since $7 \le 2 \cdot 4$, Theorem 6.8 yields $1/11 \in S_2$.

Allaart [6] actually gives a slightly weaker condition than the one given in Theorem 6.8, which includes lower order terms, and also gives an accompanying necessary condition in terms of the sequence (k_n). Together these conditions cover most cases, but still leave a small gap.

Theorems 6.7 and 6.8 are corollaries to the following, precise but somewhat abstract, characterization of membership in S_2.

Theorem 6.9 (Allaart 2011). *Let $0 < y < \frac{1}{2}$. Then $|L(y)| = 2$ if and only if*

$$\Phi(\Psi^n(y)) > \frac{2}{3} \qquad \text{for all } n \ge 0,$$

where

$$\Phi(y) := \begin{cases} 0, & \text{if } y = 0, \\ 4^k(y - t_k), & \text{if } \frac{k}{2^k} \le y < \frac{k-1}{2^{k-1}}, \quad k = 3, 4, \dots. \end{cases}$$

6.3   Hausdorff dimension

Recall from Section 3 that the maximum value of T is $2/3$, and $L(2/3)$ is a Cantor set of Hausdorff dimension $1/2$. One might ask whether there exist any level sets with Hausdorff dimension strictly greater than $1/2$. This question was first addressed by Maddock [55], who proved that the intersection of the graph of T with *any* line of integer slope has Hausdorff dimension at most 0.668. In particular, 0.668 is an upper bound for the dimensions of the level sets. Maddock himself conjectured that the real maximum is $1/2$. The issue was finally settled by de Amo et al. [9].

Theorem 6.10 (de Amo et al. 2011). *For each ordinate y, the box-counting dimension of $L(y)$ is at most $1/2$, and hence, $\dim_H L(y) \leq 1/2$.*

The proof, which is surprisingly elementary, makes good use of the self-affinity of the graph of T and uses a cleverly devised induction argument.

A direct consequence of Theorem 6.1 is that almost all level sets (in the Lebesgue sense) have Hausdorff dimension zero. On the other hand, Lagarias and Maddock [51] have shown that the set

$$\{y : \dim_H L(y) > 0\}$$

has full Hausdorff dimension one. This is accomplished by putting a sequence of subsets of the range $[0, \frac{2}{3}]$ in one-to-one bi-Lipschitz correspondence with certain well-behaved subsets of the domain $[0, 1]$ whose Hausdorff dimension is easy to calculate and gets arbitrarily close to 1.

It is not known exactly which numbers occur as the Hausdorff dimension of some level set of T.

6.4   Local level sets

Lagarias and Maddock [50, 51] introduce the concept of a local level set of the Takagi function. They first define an equivalence relation on $[0, 1]$ by

$$x \sim x' \quad \overset{\text{def}}{\iff} \quad |D_j(x)| = |D_j(x')| \text{ for each } j \in \mathbb{N}. \qquad (6.1)$$

The *local level set* containing x is defined by

$$L_x^{loc} := \{x' : x' \sim x\}.$$

Note that by Lemma 4.1, $x \sim x'$ implies $T(x) = T(x')$, so each local level set is contained in some level set. Lagarias and Maddock point out that each

local level set is either finite or a Cantor set. Members of the same local level set can be obtained from one another by simple operations on their binary expansions, called "block flips" in [50]. This works as follows. Let $Z(x) = \{n \geq 0 : D_n(x) = 0\} \cup \{\infty\}$. For any two elements $k, l \in Z(x)$ with $k < l \leq \infty$, form the point x' with binary expansion $x' = \sum_{n=1}^{\infty} 2^{-n} \varepsilon_n'$ by setting $\varepsilon_n' = \varepsilon_n$ if $n \leq k$ or $n > l$, and $\varepsilon_n' = 1 - \varepsilon_n$ if $k < n \leq l$. Then $|D_n(x')| = |D_n(x)|$ for each n, and so $x' \in L_x^{loc}$. Every element of L_x^{loc} can be obtained from x by at most countably many operations of this type.

One of the results in [50] concerns the average number of local level sets contained in a level set chosen at random. Let $N^{loc}(y)$ denote the number of local level sets contained in $L(y)$.

Remark 6.11. Lagarias and Maddock [50] define L_x^{loc} slightly differently, effectively viewing local level sets as subsets of the Cantor space $\{0,1\}^{\mathbb{N}}$. However, this distinction does not affect the number of local level sets contained in any level set, which is all we are concerned with in this survey.

Theorem 6.12 (Lagarias and Maddock, 2010)**.** *The expected number of local level sets contained in a level set $L(y)$ with y chosen at random from $[0, \frac{2}{3}]$ is $\frac{3}{2}$. More precisely,*

$$\mathrm{E}[N^{loc}(y)] := \frac{3}{2} \int_0^{2/3} N^{loc}(y)\, dy = \frac{3}{2}.$$

A simpler proof of this theorem is given by Allaart [5]. In that paper, local level sets are also examined from the category point of view. The result is in marked contrast with the conclusion of Theorem 6.12. Define the sets

$$S_\infty^{loc} := \{y : L(y) \text{ contains infinitely many different local level sets}\},$$

$$S_\infty^{loc,uc} := \{y : L(y) \text{ contains uncountably many different local level sets}\}.$$

Theorem 6.13 (Allaart 2011)**.** *(i) The set S_∞^{loc} is residual (co-meager) in $[0, \frac{2}{3}]$.*

(ii) The set $S_\infty^{loc,uc}$ is dense in $[0, \frac{2}{3}]$, and intersects any subinterval of $[0, \frac{2}{3}]$ in a continuum.

6.5   Open problems

A number of interesting questions about the level sets of the Takagi function remain open. We give a brief selection here, and refer to Lagarias [49] for additional problems.

Problem 6.1. Describe the set S^{co}_∞ of ordinates y with a countably infinite level set. Or less ambitiously, determine whether S^{co}_∞ contains any points which are not the image of a dyadic rational.

Problem 6.2. (Knuth [41, p. 32, Exer. 83]) Does there exist, for each rational ordinate $y \in [0, \frac{2}{3}]$, a rational abscissa $x \in [0, 1]$ such that $T(x) = y$? If x is irrational but $y = T(x)$ is rational, must $L(y)$ be uncountable?

Problem 6.3. Determine the probability distribution of $|L(y)|$. In other words, find $\lambda(S_{2n})$ for each $n \in \mathbb{N}$. This could be very difficult, but the weaker problem of showing that $\lambda(S_{2n}) > 0$ for every n (or finding a counterexample) may be solvable.

Problem 6.4. Determine the probability distribution of $N^{loc}(y)$, the number of local level sets contained in $L(y)$. This too may be difficult.

Problem 6.5. The set $S^{loc,uc}_\infty$ intersects each subinterval of $[0, \frac{2}{3}]$ in a continuum. Is it residual? Does it have Hausdorff dimension 1? Weaker than the last question, does S^{loc}_∞ have Hausdorff dimension 1?

Problem 6.6. (Lagarias [49]) Determine the dimension spectrum of T. That is, determine the function

$$f(\alpha) := \dim_H\{y : \dim_H L(y) \geq \alpha\}.$$

7   Generalizations.

7.1   The Takagi-van der Waerden functions

An immediate generalization of the Takagi function is the sequence of functions

$$f_r(x) := \sum_{n=0}^{\infty} \frac{1}{r^n} \phi(r^n x), \qquad r = 2, 3, \ldots.$$

Thus, f_2 is the Takagi function and f_{10} is van der Waerden's function. Billingsley's argument (with only trivial modifications) shows that each f_r is nowhere differentiable. Van der Waerden's original and elegant proof for the case $r = 10$ works for all even $r \geq 4$, but not for odd r or for $r = 2$.

Each f_r is also Hölder continuous of any order $\alpha < 1$. This follows from the more general result of Shidfar and Sabetfakhri [70], but is also easy to prove directly. In fact, by analogy with (2.4), we have

$$-1 \leq \liminf_{h \to 0} \frac{f_r(x+h) - f_r(x)}{h \log_r(1/|h|)} \leq \limsup_{h \to 0} \frac{f_r(x+h) - f_r(x)}{h \log_r(1/|h|)} \leq 1.$$

It seems likely that f_r has an infinite derivative at many points, but we have not found any detailed study of this question in the literature.

Generalizing Kahane's result [34], Baba [13] determines for each $r \geq 2$ the maximum value M_r of f_r and the set of points $E_r = \{x \in [0, 1] : f_r(x) = M_r\}$.

Theorem 7.1 (Baba 1984). *(i) If r is odd, then $E_r = \{1/2\}$ and $M_r = r/(2r - 2)$.*

(ii) If r is even, then E_r is a Cantor set of dimension $1/2$, and $M_r = r^2/(2r^2 - 2)$.

7.2 The Takagi class

Another direct generalization of the Takagi function is obtained by replacing the factor $1/2^n$ in (1.1) with a general (real) constant c_n. This gives functions of the form

$$f(x) = \sum_{n=0}^{\infty} c_n \phi(2^n x). \tag{7.1}$$

It is immediately clear that the series converges uniformly, and hence defines a continuous function f, when $\sum_{n=0}^{\infty} |c_n| < \infty$. Hata and Yamaguti [31] show that this condition is also necessary, and call the collection of functions of the form (7.1) the *Takagi class*. They note that each member of the Takagi class solves a discrete version of the Dirichlet boundary value problem involving the "discrete Laplacian"

$$\Delta_{i,n} f := f\left(\frac{i}{2^n}\right) + f\left(\frac{i+1}{2^n}\right) - 2f\left(\frac{2i+1}{2^{n+1}}\right),$$

in the sense that $\Delta_{i,n} f = -c_n$ for $n \geq 0$ and $i = 0, 1, \ldots, 2^n - 1$, with $f(0) = f(1) = 0$.

Beside the Takagi function itself, the Takagi class contains a number of interesting examples from the classical literature. For instance, Faber [26] introduced the highly lacunary series

$$F(x) = \sum_{k=1}^{\infty} \frac{1}{10^k} \phi(2^{k!} x),$$

and showed that F has no derivative, finite or infinite, at any point. Moreover, he proved that F does not satisfy a Lipschitz condition of any order. Kahane [34] likewise constructs lacunary series of the form $f(x) = \sum_{\nu=1}^{\infty} p_\nu 2^{-k_\nu} \phi(2^{k_\nu} x)$, and shows how p_ν and k_ν can be chosen so that the modulus of continuity of f is majorized (respectively minorized) by a given function satisfying appropriate conditions.

Many of the functions in the Takagi class are fractals, in the sense that the Hausdorff dimension of their graph is strictly greater than one. Besicovitch and Ursell [16] showed that, if f is any function satisfying a Lipschitz condition of order $\delta \in (0, 1]$, then

$$1 \leq \dim_H(f) \leq 2 - \delta, \tag{7.2}$$

where $\dim_H(f)$ denotes the Hausdorff dimension of the graph of f. They show that within these bounds any dimension is possible. The implications of their work for the Takagi class are collected in the following theorem.

Theorem 7.2 (Besicovitch and Ursell, 1937)**.** *Let* $f(x) = \sum_{n=0}^{\infty} 2^{-\delta a_n} \phi(2^{a_n} x)$, *where* $0 < \delta < 1$ *and* $a_{n+1} - a_n \geq A$ *for some constant* $A > 0$. *Then* f *is Lipschitz of order* δ *but of no smaller order, and:*
   (i) If $a_{n+1}/a_n \to \infty$, *then* $\dim_H(f) = 1$.
   (ii) If $a_n = \mu^n$ *where* $\mu > 1$, *then* $1 < \dim_H(f) < 2 - \delta$. *Moreover, for each* $d \in (1, 2 - \delta)$ *there exists* $\mu > 1$ *such that, if* $a_n = \mu^n$, *then* $\dim_H(f) = d$.
   (iii) If $a_{n+1}/a_n \to 1$ *but* $a_{n+1} - a_n \to \infty$, *then* $\dim_H(f) = 2 - \delta$.

Examples of (i), (ii) and (iii) are, respectively: $a_n = 2^{n^2}$, $a_n = 2^n$, and $a_n = n^2$. Note that in all three cases the series (7.1) is lacunary. Generally speaking, the more lacunary the series, the smaller the dimension of the graph is. When the series is extremely lacunary as in (i), the fine structure of the graph virtually disappears and the function f becomes "almost differentiable". But the above theorem does not say anything about the important (nonlacunary) case $a_n = n$. In other words, it does not give the dimension of the functions

$$g_\delta(x) = \sum_{n=0}^{\infty} 2^{-\delta n} \phi(2^n x), \qquad 0 < \delta < 1. \tag{7.3}$$

This boundary case was addressed more than half a century later by Ledrappier [52], using modern tools that were not yet available to Besicovitch and Ursell.

Here too g_δ is Lipschitz of order δ. Ledrappier showed that typically, g_δ attains the upper bound in (7.2). More precisely, he proved that $\dim_H(g_\delta) = 2 - \delta$ whenever $2^{\delta - 1}$ is an Erdős number. A number $\lambda \in (0, 1)$ is called an Erdős number if the probability distribution of $\sum_{n=0}^{\infty} \lambda^n \varepsilon_n$ (the so-called *Bernoulli convolution*) has Hausdorff dimension 1, where $\{\varepsilon_n\}$ are i.i.d. random variables taking the values 1 and -1 each with probability $1/2$. Later progress on Bernoulli convolutions due to Solomyak [71] implies that almost every $\lambda \in (1/2, 1)$ is Erdős, and hence, by Ledrappier's result, $\dim_H(g_\delta) = 2 - \delta$ for almost every $\delta \in (0, 1)$. Ledrappier's approach is dynamical, viewing the graph of g_δ as the repeller for some expanding self-map of $[0, 1] \times \mathbb{R}$. The deep and

difficult proof uses ideas from smooth ergodic theory and a Marstrand-type lemma concerning projections of measures.

Recently, Katzourakis [37] has reported that, with $a_n = 2\nu n$ for a parameter $\nu \in \mathbb{N}$, the function f in Theorem 7.2 can be used to construct a "pathological" solution to the nonlinear Aronsson partial differential equation and to the Infinity-Laplace PDE system.

For $\delta > 1$, the function g_δ defined by (7.3) is Lipschitz and hence almost everywhere differentiable. When $1 \leq \delta \leq 2$, g_δ turns out to be the extremal function in a certain approximate convexity problem; see Section 8.3 below.

After the publication of Hata and Yamaguti's paper, Kôno [44] investigated the Takagi class in greater generality. Perhaps the most striking result, concerning the differentiability of f, is the following:

Theorem 7.3 (Kôno 1987). *Let f be defined by (7.1), and put $a_n := 2^n c_n$.*

(i) *If $\{a_n\} \in \ell^2$, then f is absolutely continuous and hence differentiable almost everywhere.*

(ii) *If $\{a_n\} \notin \ell^2$ but $\lim_{n\to\infty} a_n = 0$, then f is nondifferentiable at almost every point of $[0,1]$, but f is differentiable on an uncountably large set, and the range of f' is $\mathbb{R}$.*

(iii) *If $\limsup_{n\to\infty} |a_n| > 0$, then f is nowhere differentiable.*

Kôno [44] also considers the oscillations of f (stating more general forms of Theorems 2.3 and 2.4), and proves furthermore that the Takagi class contains only one function which is smooth in Zygmund's sense. That is, if

$$f(x + h) + f(x - h) - 2f(x) = o(h) \qquad \text{as } h \downarrow 0 \tag{7.4}$$

for all $x \in (0,1)$, then $c_n = a/4^n$ for some constant a, and $f(x) = 2ax(1-x)$.

A special case of the Takagi class arises when one takes $c_n = \pm 1/2^n$ for all n in (7.1). Precisely, let $\mathbf{r} = (r_0, r_1, \dots)$ be a sequence with $r_n \in \{-1, 1\}$ for each n, and define

$$F_{\mathbf{r}}(x) = \sum_{n=0}^{\infty} \frac{r_n}{2^n} \phi(2^n x). \tag{7.5}$$

For example, the *alternating Takagi function*

$$\hat{T}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{\phi(2^n x)}{2^n}, \tag{7.6}$$

depicted in Figure 4, is of the above form. It was introduced by Kawamura [38], who showed that it has a direct connection with a family of self-similar
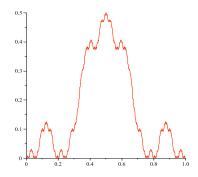
Figure 4: The alternating Takagi function

sets in the plane determined by two contractions which includes the Koch curve and Pólya's space-filling curve, among others. (Similarly, she showed that the Takagi function itself is related to another family of self-similar sets, which includes the Lévy dragon curve as an example.)

In general, (7.5) identifies an uncountably large class of functions which are "close" to the Takagi function T in the sense that their partial sums are all piecewise linear with integer slopes that change by ± 1 at each step. It should therefore be no surprise that these functions share a large number of properties with T. For instance, Allaart [5, Section 5] shows that many of the results from Section 6 concerning level sets hold for arbitrary $F_{\mathbf{r}}$: Almost all level sets of $F_{\mathbf{r}}$ are finite, but their average cardinality is infinite and the set of ordinates y with uncountably large level sets is residual in the range of $F_{\mathbf{r}}$.

An interesting random version of the Takagi function is obtained by taking the components of $\mathbf{r}$ to be independent random variables with $\mathrm{P}(r_n = 1) = p$ and $\mathrm{P}(r_n = -1) = 1 - p$, where $0 \le p \le 1$. The maximum value M of $F_{\mathbf{r}}$ is then a random variable, and the set $\mathcal{M} := \{x \in [0, 1) : F(x) = M\}$ is a random set. Allaart [2] determines the probability distributions of M and the size of $\mathcal{M}$. If $p < 1/2$, the distribution of M is purely atomic and $|\mathcal{M}|$ is almost surely finite, with range $\{2^l(2^m - 1) : l \in \mathbb{Z}_+, m \in \mathbb{N}\}$. (For instance, one can have exactly 24 maximum points with positive probability.) If $p \ge 1/2$, the distribution μ of M is singular continuous, and $\mathcal{M}$ is a Cantor set with almost-sure Hausdorff dimension $(2p - 1)/2p$. In the latter case, Allaart also determines the Hausdorff dimension and multifractal spectrum of μ.

7.3 The Zygmund spaces Λ_d^*, λ_d^* and $\Lambda_{d,1}^*$

In 1945, Zygmund [83] introduced the class λ^* of "smooth" functions of period 1, i.e. those functions f satisfying (7.4) uniformly in x, and the wider class Λ^* of "quasismooth" functions of period 1, which satisfy (7.4) with $O(h)$ replacing $o(h)$. Zygmund studied various properties of functions in these classes, and characterized them in terms of uniform approximation by polynomials.

In 1989, Anderson and Pitt [12] introduced the larger classes λ_d^* and Λ_d^* of periodic functions satisfying a weaker form of smoothness defined in terms of differences over dyadic intervals. These classes have simple characterizations in terms of the Schauder expansions of their members. Recall that every continuous function f of period 1 which vanishes at 0 has on $[0, 1)$ a unique Schauder expansion of the form (5.2). For easier comparison with the Takagi class, we will write the Schauder expansion in the form

$$f(x) = \sum_{n=0}^{\infty} \frac{r_n(x)}{2^n} \phi(2^n x), \tag{7.7}$$

where $r_n(x)$ depends only on the first n binary digits of x. More precisely, $r_n(x) = R_n(\varepsilon_1, \ldots, \varepsilon_n)$, where $x = \sum_{n=0}^{\infty} 2^{-n}\varepsilon_n$ and $\varepsilon_n \in \{0, 1\}$. The classes Λ_d^* and λ_d^* are defined as follows: $f \in \Lambda_d^*$ if and only if there exists a uniform bound M such that $|r_n(x)| < M$ for all n and all x; and $f \in \lambda_d^*$ if and only if $r_n(x) \to 0$ uniformly in x.

Anderson and Pitt [12] show that $\Lambda^* \subset \Lambda_d^*$ and $\lambda^* \subset \lambda_d^*$. The Takagi function is an example of a function which is in Λ_d^* but not in Λ^*. On the other hand, as shown by Abbott et al. [1], the alternating Takagi function $\hat{T}$ belongs to Λ^*. In general, functions in Λ^* can have corners but no cusps, whereas functions in Λ_d^* can have logarithmic cusps. It is shown in [12] that every member f of Λ_d^* is Lipschitz of order $h \log(1/h)$. That is, there is a constant C such that, for all $0 \le x < x + h \le 1$ with h sufficiently small,

$$|f(x+h) - f(x)| \le Ch \log(1/h).$$

Another result in [12] is that the graph of every $f \in \Lambda_d^*$ is of σ-finite linear measure (see Section 3). With regard to level sets $L(y) = \{x \in [0, 1) : f(x) = y\}$, Anderson and Pitt prove that (i) if $f \in \Lambda_d^*$, then $L(y)$ is countable for almost every y; and (ii) if $f \in \lambda_d^*$, then $L(y)$ is finite for almost every y. It does not appear to be known whether the condition $f \in \Lambda_d^*$ gives enough regularity to the graph of f in order that $L(y)$ be finite for almost all y.

The specific case of (7.7) where $|r_n(x)|$ is constant in x for each n was studied by Allaart [3]. We shall call the collection of such functions the *flexible Takagi class*. It is an immediate generalization of the Takagi class in which the
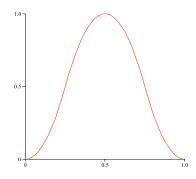
Figure 5: A smooth function in the flexible Takagi class, defined by (7.8).

individual "tents" at each level can point either upward or downward, but all tents within a given level have the same amplitude. This guarantees uniformity in the fine structure across the domain of f, while allowing for a wide variety of general shapes of the graph. Indeed, Allaart [3] manages to extend all of Kôno's results to this more general setting. Specifically, statements (i)-(iii) of Theorem 7.3 hold when $a_n = |r_n|$. Whereas the Takagi class contains (up to a multiplicative constant) only one function in the Zygmund space λ^*, the flexible Takagi class contains many. For example, it contains the "bell-shaped" curve

$$f(x) = \begin{cases} 8x^2, & x \leq 1/4 \\ 8x(1-x) - 1, & 1/4 \leq x \leq 3/4 \\ 8(1-x)^2, & x \geq 3/4, \end{cases} \tag{7.8}$$

depicted in Figure 5 and obtained by setting $r_0 = 1$, $r_1 = 0$, and $r_n = -2^{2-n}X_1X_2$ for $n \geq 2$, where $X_n(x) = (-1)^{\varepsilon_n(x)}$ is the nth Rademacher function. See [3, Section 4] for more examples. Whether all functions in the flexible Takagi class that belong to λ^* must be piecewise quadratic remains open.

A subcollection of the flexible Takagi class in which $|r_n| = 1$ for each n was studied by Abbott, Anderson and Pitt [1], who denote this subcollection by $\Lambda^*_{d,1}$. It contains the Takagi function, as well as several other interesting functions that have occurred in the literature. For instance, the *Gray Takagi function* of Kobayashi [42], which plays a role in the analysis of Gray code digital sums (see Section 8.1 below), belongs to $\Lambda^*_{d,1}$. It has $r_n = X_n$ for each n. Another example is the function T^3 of Kawamura [38], which has a connection with certain self-similar sets in the complex plane. It has $r_n = X_1 \cdots X_n$ for
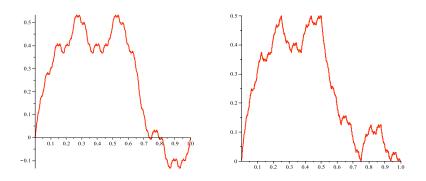
Figure 6: The Gray Takagi function (left) and Kawamura's T^3 (right)

each n. These functions are shown in Figure 6.

For a continuous function f on $[0, 1)$, define the dyadic difference quotients $D_n f$ as follows. If $x \in [0, 1)$, let k be the integer such that $k/2^n \leq x < (k+1)/2^n$, and put

$$D_n f(x) := \frac{f\left(\frac{k+1}{2^n}\right) - f\left(\frac{k}{2^n}\right)}{2^{-n}} = 2^n \left[f\left(\frac{k+1}{2^n}\right) - f\left(\frac{k}{2^n}\right) \right].$$

(For the Takagi function, $D_n f(x) = D_n(x)$ as in Section 2.1; see (2.1).) Abbott, Anderson and Pitt [1] consider the set S_K of *slow points* with constant K, that is, the set

$$S_K := S_K(f) := \{x \in [0, 1) : |D_n f(x)| \leq K \text{ for all } n\}.$$

They show that for $f \in \Lambda_{d,1}^*$, the Hausdorff dimension of S_K is given by

$$\dim_H S_K = 1 + \log_2 \left(\cos\left(\frac{\pi}{2(K+1)} \right) \right).$$

Moreover, their proof shows that S_K is an s-set, with s being the above dimension. In particular, the set $\bigcup_{K=1}^{\infty} S_K$ of all slow points of f has full Hausdorff dimension 1. On the other hand, for $f \in \Lambda_{d,1}^*$ the set of slow points is a null set, because $\{D_n f\}_n$ is a simple random walk and hence obeys the law of the iterated logarithm. The authors of [1] are particularly interested in the interplay between slow points and local Lipschitz properties of a function. Let $L = L(f)$ denote the set of points x at which f satisfies a local Lipschitz condition; that is, those points x for which there exists a constant M such that

$|f(x) - f(y)| \leq M|x - y|$ for all $y \neq x$. Abbott, Anderson and Pitt show that any $f \in \Lambda_{d,1}^*$ satisfies a local Lipschitz condition at most points of $S_K = S_K(f)$ in the sense that $\dim_H(S_K \cap L) = \dim_H S_K$. For the Takagi function and the alternating Takagi function the stronger statement $S_K \subset L$ holds, but this is exceptional: If a member $f \in \Lambda_{d,1}^*$ is chosen "at random", then $\mathcal{H}^\alpha(S_K \cap L) = 0$ with probability 1, where $\alpha = \dim_H S_K$. Since $\mathcal{H}^\alpha(S_K) > 0$, this can be interpreted as saying that the "typical" $f \in \Lambda_{d,1}^*$ satisfies a Lipschitz condition at very few of its slow points.

7.4   The functions of Sekiguchi and Shiota

Following the important paper of Hata and Yamaguti [31], Sekiguchi and Shiota [68] studied further generalizions of the Takagi function. Their first result concerns the system of difference equations

$$f\left(\frac{2j+1}{2^{n+1}}\right) - (1-r)f\left(\frac{j}{2^n}\right) - rf\left(\frac{j+1}{2^n}\right) = c_n, \tag{7.9}$$
$$j = 0, 1, \ldots, 2^n - 1, \quad n = 0, 1, 2, \ldots,$$

where $r \in (0,1)$ is a constant parameter. Let $L_r(x)$ be Lebesgue's singular function, and define the generalized Schauder function $S_r : \mathbb{R} \to [0,1]$ by

$$S_r(x) = \begin{cases} L_r(x)/r, & \text{if } 0 \leq x \leq 1/2 \\ \bigl(1 - L_r(x)\bigr)/(1-r), & \text{if } 1/2 \leq x \leq 1, \end{cases}$$

and $S_r(x+1) = S_r(x)$ for all $x \in \mathbb{R}$. Sekiguchi and Shiota show that the system (7.9) with boundary conditions $f(0) = a$ and $f(1) = b$ has a unique continuous solution f on $[0,1]$ if and only if $\sum_{n=0}^\infty |c_n| < \infty$, in which case

$$f(x) = a + (b-a)L_r(x) + \sum_{n=0}^\infty c_n S_r(2^n x).$$

The second result of [68] is that $L_r(x)$ is an analytic function of r, and a recursive construction is given in terms of the functions $S_r(x)$ and Rademacher functions of the (normalized) kth partial derivative

$$T_{r,k}(x) := \frac{1}{k!}\frac{\partial^k L_r(x)}{\partial r^k}. \tag{7.10}$$

Sekiguchi and Shiota use martingale theory (based on the connection of $L_r(x)$ with unfair coin tossing explained in Section 5) to prove their results. They

show moreover that the functions $T_{r,k}$ satisfy the system of difference equations

$$T_{r,k}\left(\frac{2j+1}{2^{n+1}}\right) - (1-r)T_{r,k}\left(\frac{j}{2^n}\right) - rT_{r,k}\left(\frac{j+1}{2^n}\right)$$
$$= T_{r,k-1}\left(\frac{j+1}{2^n}\right) - T_{r,k-1}\left(\frac{j}{2^n}\right)$$

for $k \geq 1$, $n \geq 0$ and $j = 0, 1, \ldots, 2^n - 1$, where we set $T_{r,0} := L_r$. Note that by (5.6), $T_{1/2,1} = 2T$.

The special case of the functions $T_{r,k}$ in which $r = 1/2$ was investigated further by Allaart and Kawamura [7]. In that paper we show that $T_{1/2,n}(1 - x) = T_{1/2,n}(x)$ when n is odd, and $T_{1/2,n}(1 - x) = -T_{1/2,n}(x)$ when n is even. We derive from (5.7) and (7.10) the representation

$$T_{1/2,n}(x) = \sum_{k=1}^{\infty} \varepsilon_k \left(\frac{1}{2}\right)^{k-n} \sum_{i=0}^{n} (-1)^i \binom{I_k - 1}{i}\binom{k - I_k + 1}{n - i}, \qquad 0 \leq x \leq 1.$$

This is then used to prove that for each n, $T_{1/2,n}$ is nowhere differentiable and Hölder continuous of order $h\big(\log(1/h)\big)^n$. That is, there is a constant C_n such that, for $0 \leq x < x + h \leq 1$ and h sufficiently small,

$$|T_{1/2,n}(x + h) - T_{1/2,n}(x)| \leq C_n h\big(\log(1/h)\big)^n,$$

and this bound is the best possible. As a result, the graph of $T_{1/2,n}$ has Hausdorff dimension 1. We also determine the global and local extrema of $T_{1/2,2}$ and $T_{1/2,3}$. The sets of points where these functions attain there absolute maximum are shown to be Cantor sets of Hausdorff dimension zero, and their members have binary expansions that follow a remarkable pattern. By contrast, we conjecture that $T_{1/2,n}$ has only finitely many absolute maximum points when $n \geq 4$. Regarding the growth rate of $M_n := \max_{0 \leq x \leq 1} T_{1/2,n}$, we conjecture that

$$M_n \sim \frac{2^n}{\sqrt{\pi n}}, \qquad \text{as } n \to \infty. \tag{7.11}$$

(Proposition 6.25 in [7] shows that M_n grows at least this fast.)

For general $r \in (0,1)$, the functions $T_{r,n}$ have not yet been thoroughly investigated. However, a new paper by de Amo et al. [10] announces the surprising result that if $r \neq 1/2$, then $T_{r,n}$ is differentiable (with vanishing derivative) almost everywhere for each n. This is in sharp contrast with the fact that $T_{r,n}$ is nowhere differentiable if $r = 1/2$. One might expect a further study of these functions to reveal many more interesting properties.

7.5 Other generalizations and variations

Many other generalizations and variations of the Takagi function have been studied. We briefly mention a few, but the list below is by necessity far from complete.

Frankl et al. [28] define a generalized Takagi function using arbitrary base $1 + c$ instead of base 2, where c is any positive real number. These functions are used in connection with the Kruskal-Katona theorem in combinatorics; see Section 8.2 below. The resulting graphs look like the graph of T with the wind blowing in from the side. Another way to create skewed versions of the Takagi function is of course to take the composition $T \circ h$ of T with an arbitrary homeomorphism $h : [0, 1] \to [0, 1]$; Kawamura [39] considers the case where $h = L_r^{-1}$ and looks at the set of points where $T \circ h$ has a vanishing derivative. Tsujii [79] constructs a Takagi-like function of two variables. Sumi [73] observes that Lebesgue's singular function $L_r(x)$ can be interpreted as the probability of "tending to $+\infty$" in a certain kind of random dynamics in the real line. He then extends this notion to random dynamics in the complex plane, obtaining a complex version of Lebesgue's singular function and (by extending the relationship (5.6)) a complex version of the Takagi function.

For the more general setting where the tent map ϕ is replaced with an arbitrary periodic Lipschitz function, see Mauldin and Williams [59] and the many papers citing this work.

7.6 Open problems

There are many natural questions regarding the functions which have occurred in this section. Some of the problems listed below may have a known answer, some have not yet been seriously investigated, and others may be hard.

Problem 7.1. Characterize exactly which functions in the flexible Takagi class belong to Zygmund's space λ^*. (A partial characterization is given in [3, Theorem 4.1].) Must these functions necessarily be piecewise quadratic?

Problem 7.2. The classes Λ^* and $\Lambda_{d,1}^*$ are both contained in Λ_d^*, but they are not disjoint. (For instance, the alternating Takagi function $\hat{T}$ belongs to both.) Can one characterize the intersection $\Lambda^* \cap \Lambda_{d,1}^*$?

Problem 7.3. Is it true that for every $f \in \Lambda_{d,1}^*$, the level set $L_f(y) := \{x \in [0, 1] : f(x) = y\}$ is finite for almost every y?

Problem 7.4. A *Besicovitch function* is one that does not possess a one-sided (finite or infinite) derivative at any point. Does the flexible Takagi class contain any Besicovitch functions?

Problem 7.5. At which points do the Takagi-van der Waerden functions f_r possess an infinite derivative? (See [8] for the case $r = 2$.)

Problem 7.6. Determine the level set structure of f_r for $r \geq 3$. In particular, is the level set of f_r at almost every level y finite?

Problem 7.7. Prove or disprove that $T_{1/2,n}$ has a unique global maximum point and a unique global minimum point when $n \geq 4$.

Problem 7.8. Prove or disprove (7.11).

8   Applications.

In this final section, we discuss a number of applications of the Takagi function and its generalizations. We begin with an application to the digital sum problem in number theory.

8.1   Applications in Number theory

Let n be a positive integer with binary expansion $n = \sum_{i=0}^{\infty} \alpha_i(n)2^i$, where $\alpha_i(n) \in \{0, 1\}$. We define the following arithmetical sums:

$$s(n) = \sum_{i=0}^{\infty} \alpha_i(n), \qquad \text{(the binary digital sum)}, \qquad (8.1)$$

$$S_k(N) = \sum_{n=0}^{N-1} s(n)^k, \qquad \text{(the power sum)}, \qquad (8.2)$$

$$F(\xi, N) = \sum_{n=0}^{N-1} e^{\xi s(n)}, \qquad \text{(the exponential sum)}, \qquad (8.3)$$

where k, N are positive integers, and ξ is a real number. Note that $s(n)$ is the number of ones in the binary expansion of n. As a special case, $F(\log 2, N)$ is the number of odd numbers in the first N rows of Pascal's triangle.

The power and exponential sums of digital sums were first studied in connection with divisibility problems involving factorials and binomial coefficients, but they occur in other areas of mathematics as well: algebraic number theory, topology, combinatorics and computational algorithms. For more information, see Stolarsky [72].

If N is a power of 2, it is clear that

$$S_1(N) = \frac{N \log_2 N}{2}, \qquad F(\xi, N) = N^{\log_2(1+e^{\xi})}.$$

However, it is difficult to get an explicit formula for arbitrary $N \in \mathbb{N}$.

The first author to give an exact expression for $S_1(N)$ was Trollope [78] in 1968. Delange [25] gave a simpler proof and extended the result to digits in arbitrary bases. Trollope's expression is

$$S_1(N) = \frac{N \log_2 N}{2} + E(N), \tag{8.4}$$

where the error term $E(N)$ involves the Takagi function: if N is written as $N = 2^m(1 + x)$ with $m \in \mathbb{Z}_+$ and $0 \leq x < 1$, then

$$E(N) = 2^{m-1} \{2x - T(x) - (1 + x) \log_2(1 + x)\}. \tag{8.5}$$

We can easily derive this from (4.7) and (5.1). First, note that (4.7) can be written as

$$T\left(\frac{k}{2^m}\right) = \frac{mk}{2^m} - \frac{1}{2^{m-1}} S(k).$$

So if N is expressed as above, then $N \leq 2^{m+1}$ and

$$\begin{aligned} S(N) &= \frac{(m + 1)N}{2} - 2^m T\left(\frac{1 + x}{2}\right) \\ &= 2^{m-1}\big(m(1 + x) + 2x - T(x)\big), \end{aligned}$$

where the last step follows since T satisfies (5.1). Since

$$\frac{N \log_2 N}{2} = 2^{m-1}(1 + x)\big(m + \log_2(1 + x)\big),$$

(8.5) follows.

It is natual to ask if we can generalize the expressions (8.4) and (8.5) to arbitrary $k \in \mathbb{N}$. For $k = 2$, Coquet [23] obtained an explicit formula and proved that $S_2(N)$ also has a close relationship with a nowhere differentiable function. Unfortunately, the general formula for $S_k(N)$ given by Coquet included a function specified only by a complicated recursion.

It seemed quite difficult to find a direct formula for $S_k(N)$ with $k \geq 3$, but finally, in 1995, Okada et al. [61] gave the complete answer not only for the power sum but also for the exponential sum. The key was to connect $F(\xi, N)$ with Lebesgue's singular function $L_r(x)$. Set $r = (1 + e^\xi)^{-1}$ and $t = \log_2 N$. Denote the integer part of t by $[t]$, and the fractional part by $\{t\}$. Then

$$F(\xi, N) = \frac{1}{r^{[t]+1}} L_r\left(\frac{1}{2^{1-\{t\}}}\right), \tag{8.6}$$

for $\xi \in \mathbb{R}$ and $N \in \mathbb{N}$. Using Salem's expression for $L_r(x)$ (See [65]), an explicit expression for $F(\xi, N)$ is obtained.

For $S_k(N)$, observe that the following equality holds for $k \in \mathbb{N}$:

$$S_k(N) = \frac{1}{2} \left. \frac{\partial^k}{\partial \xi^k} F(\xi, N) \right|_{\xi=0}. \tag{8.7}$$

An essential part of this derivative is $(\partial^k/\partial r^k) L_r(x)|_{r=1/2}$, and hence, $S_k(N)$ can be expressed explicitly in terms of the nowhere differentiable function $T_{1/2,k}$ of Section 7.4.

Later, several analogous problems were studied. For instance, instead of the binary expansion of natural numbers, Kobayashi [42] considered the *Gray code*, which is an encoding of natural numbers as sequences of 0's and 1's with the property that the representations of adjacent integers differ in exactly one position. Though the Gray code was introduced initially as a solution to a communications problem involving digitization of analogue data, it has since been used in a wide variety of other applications, including databases, experimental design and even puzzle solving; see Savage [66].

Kobayashi defined a probability measure $\tilde{\mu}_r$ on $[0, 1]$ analogous to the binomial measure, but whereby each *Gray code* digit (rather than binary digit) of $x \in [0, 1]$ is 0 with probability r and 1 with probability $1 - r$, independently of the other digits. Adapting the methods of [68] and [61], Kobayashi gave explicit expressions for the Gray digital sum, the Gray power sum and the Gray exponential sum, which are defined just as in (8.1)-(8.3), but with Gray code digits replacing binary digits. The expressions are in terms of the distribution function $\tilde{L}_r$ of $\tilde{\mu}_r$ and a nowhere-differentiable continuous function $\tilde{T}$ which Kobayashi calls the *Gray Takagi function*; see Section 7.3.

Krüppel [46] modified the Trollope-Delange formula in a different direction, considering instead of $s(n)$ the alternating binary sum $\hat{s}(n) = \sum_{i=0}^{\infty} (-1)^i \alpha_i(n)$. He derived an expression for $\hat{S}_1(N) = \sum_{n=0}^{N-1} \hat{s}(n)$ in terms of the alternating Takagi function $\hat{T}$ defined in (7.6).

A comprehensive review of the role of nowhere-differentiable functions and singular measures in the study of digital sum problems is found in Kobayashi et al. [43].

8.2   Applications in Combinatorics

A few interesting connections between the Takagi function and the discrete isoperimetric problems in combinatorics have been discovered.

The first surprising result was given by Frankl et al. [28] in 1995. Let $\binom{\mathbb{N}}{k}$ denote the collection of subsets of $\mathbb{N}$ having k elements. For any family $\mathcal{F} \subset \binom{\mathbb{N}}{k}$ and positive integer $l < k$, let the lth *shadow* of $\mathcal{F}$ be

$$\Delta_l(\mathcal{F}) := \{G \in \binom{\mathbb{N}}{l} : \exists F \in \mathcal{F} \text{ s.t. } G \subset F\}.$$

It is clear that $|\Delta_l(\mathcal{F})|$, the size of the shadow, depends on $\mathcal{F}$. Thus, it is natural to ask for a fixed $m \in \mathbb{N}$, which family $\mathcal{F}$ such that $|\mathcal{F}| = m$ attains the minimum size of the shadow. This problem can be viewed as a special case of the *vertex discrete isoperimetric problem* by considering a graph whose vertices correspond to finite subsets of $\binom{\mathbb{N}}{k}$, with an edge connecting a set of cardinality k to each of its subsets of cardinality $k - 1$.

The shadow minimization problem was solved by Kruskal and Katona independently. For $A, B \in \binom{\mathbb{N}}{k}$, define *the colex order* by

$$A <_{colex} B \qquad \Longleftrightarrow \qquad \max\{a \in A \backslash B\} < \max\{b \in B \backslash A\}.$$

Denote by $Colex(k, m)$ the family of the first m elements in $\binom{\mathbb{N}}{k}$ with the colex order. The Kruskal-Katona theorem says that for all $\mathcal{F} \subset \binom{\mathbb{N}}{k}$ such that $|\mathcal{F}| = m$, and $l < k$,

$$\min |\Delta_l(\mathcal{F})| = |\Delta_l(Colex(k, m))|.$$

Finally, define the *Kruskal-Katona function* by

$$K_l^k(m) = -m + |\Delta_l(Colex(k, m))|.$$

Frankl et al. show that, properly normalized, $K_l^k(m)$ converges uniformly to $T(x)$ if $l = k-1$. More precisely, define the shadow function S_k by normalizing K_l^k, where $l = k - 1$:

$$S_k(x) := k\binom{2k - 1}{k}^{-1} K_{k-1}^k\left(\left\lfloor \binom{2k - 1}{k} x \right\rfloor\right), \qquad \text{for} \qquad 0 \leq x \leq 1.$$

Then $S_k \to T$ uniformly on $[0, 1]$.

In general, an estimation of the Kruskal-Katona function is hard to calculate. The result above clearly shows the reason. Frankl et al. exploit the above relationship and various known properties of the Takagi function to give new estimates of the Kruskal-Katona function. They also define a generalized Takagi function (using the base $1 + c$ expansion of $x \in [0, 1]$ for an arbitrary real $c > 0$) and show that it has a similar connection to the minimum shadow size for $l = \lfloor ck \rfloor$.

Another connection of the Takagi function with combinatorics was given by Guu [30]. Assume a graph (V, E) with a set of vertices V and a set of edges E is given. For $S \subset V$, let $\theta(S)$ denote the number of edges in E which connect a vertex in S to a vertex in $V \backslash S$. For given $0 \le k \le |V|$, the *edge discrete isoperimetric problem* is to minimize $\theta(S)$ over all S having k elements. Guu considers as a special case the n-cube $Q_n = (V_n, E_n)$, and defines the function

$$\theta(n, k) = \min\{\theta(S) : S \subset V_n, |S| = k\}.$$

He shows that

$$\lim_{n \to \infty} \frac{\theta(n, \lfloor 2^n x \rfloor)}{2^n} = T(x).$$

8.3   Applications in Real Analysis

In this subsection we sketch two applications of the Takagi function in real analysis. One concerns the zero sets of continuous nowhere-differentiable functions, while the other relates to the study of approximate convexity of real functions.

8.3.1   Zero sets of continuous nowhere-differentiable functions

If a continuous function does not have a finite derivative anywhere, what can one say about its set of zeros? This question was answered in 1966 by Lipinski [53], who gave the following characterization.

Theorem 8.1. *Let $C \subset [0, 1]$. Then C is the zero set of some nonnegative continuous nowhere-differentiable function $f : [0, 1] \to \mathbb{R}$ if and only if C is closed and nowhere dense.*

It is easy to see that the condition on C is necessary. To prove its sufficiency, Lipinski used the Takagi function to construct an example of a function f with the required properties. To begin, write $[0, 1] \backslash C = \bigcup_{n=1}^{\infty} (a_n, b_n)$ with the union disjoint. (If $[0, 1] \backslash C$ is a finite union of open intervals, the construction is easy.) Define functions

$$T_{a,b}(x) = (b - a) T\left(\frac{x - a}{b - a}\right), \qquad a \le x \le b,$$

where $0 \le a < b \le 1$. Thus, the graph of $T_{a,b}$ is a smaller copy of the graph of T confined to the interval $[a, b]$. It is tempting to construct f by putting

$$f(x) = \begin{cases} 0, & x \in C \\ T_{a_n, b_n}(x), & x \in (a_n, b_n), n \in \mathbb{N}. \end{cases}$$

Indeed, Schubert [67] had mistakenly believed that this creates a nowhere differentiable function. But, as Lipinski points out, f defined this way can be differentiable (with $f'(x) = 0$) at many points of C. In fact, if $\lambda(C) > 0$, then f is differentiable almost everywhere on C. To correct this problem, Lipinski slightly modified the above construction as follows. Enumerate the dyadic open subintervals of $[0,1]$ (i.e. those of the form $(j/2^n, (j+1)/2^n)$) in some arbitrary order as $\{A_n : n \in \mathbb{N}\}$. For each n, let F_n be any interval (a_{i_n}, b_{i_n}) contained in A_n if such an interval exists, and $F_n = \emptyset$ otherwise, with the restriction that no interval (a_i, b_i) is chosen more than once. Now define a function $\tilde{f}$ by

$$\tilde{f}(x) = \begin{cases} 0, & x \in C \\ T_{a_i, b_i}(x), & x \in (a_i, b_i), (a_i, b_i) \notin \{F_n\} \\ \lambda(A_n) T_{a_{i_n}, b_{i_n}}(x)/(b_{i_n} - a_{i_n}), & x \in F_n. \end{cases}$$

Lipinski shows that $\tilde{f}$ is continuous and nowhere differentiable, and hence has the desired properties. Of course, in this construction we can replace T with any continuous nowhere-differentiable function which is strictly positive in $(0,1)$ and vanishes at 0 and 1.

8.3.2   Approximate convexity

Our next application is to approximate convexity of real functions. Let V be a convex subset of a normed space and let $\varepsilon \geq 0$, $p > 0$ be given constants. Say a function $f : V \to \mathbb{R}$ is (ε, p)-*midconvex* if

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2} + \varepsilon\|x - y\|^p, \qquad x, y \in V.$$

It was shown by Házy and Páles [32] that if $f : V \to \mathbb{R}$ is continuous and $(\varepsilon, 1)$-midconvex, then for all $x, y \in V$ and $t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \varepsilon\omega(t)\|x - y\|, \qquad (8.8)$$

where $\omega = 2T$. A natural question is whether ω in this last inequality can be replaced by a smaller function. The negative answer came a few years later, when Boros [18] proved that ω is itself $(1,1)$-convex. In terms of T, this comes down to the inequality

$$T\left(\frac{x+y}{2}\right) \leq \frac{T(x) + T(y)}{2} + \left|\frac{x-y}{2}\right|.$$

(To see that this implies the minimality of ω in (8.8), take $f = \varepsilon\omega$, $x = 1$ and $y = 0$.) Boros' proof was somewhat laborious, with no fewer than eight separate cases in the induction step. But the important thing was that it confirmed Páles' conjecture of the minimality of ω.

The next question, then, is what function takes the role of ω when $p \neq 1$. More precisely, what is the smallest function ω_p such that, whenever $f : V \to \mathbb{R}$ is (ε, p)-midconvex, we have

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \varepsilon\omega_p(t)\|x - y\|^p \qquad (8.9)$$

for all $x, y \in V$ and $t \in [0, 1]$? For $p \in [1, 2]$, Tabor and Tabor [74, 75] showed that the answer is the function

$$\omega_p(x) = 2 \sum_{n=0}^{\infty} \frac{1}{2^{np}} \phi(2^n x), \qquad (8.10)$$

which belongs to the Takagi class. In [74] they show that any (ε, p)-midconvex function satisfies (8.9), and in [75] they establish that ω_p is itself $(1, p)$-midconvex; that is,

$$\omega_p\left(\frac{x + y}{2}\right) \leq \frac{\omega_p(x) + \omega_p(y)}{2} + |x - y|^p. \qquad (8.11)$$

Note that $\omega_2(x) = 4x(1 - x)$ for $x \in [0, 1]$, and $\omega_1(x) = 2T(x)$. Tabor and Tabor prove (8.11) first for $p = 2$, and then deduce it for all $p \in [1, 2]$ by expressing ω_p as an infinite series in terms of ω_2. A different proof of their result (which includes that of Boros) is given by Allaart [4], who derives the formula

$$\omega_p\left(\frac{m}{2^n}\right) = \sum_{k=0}^{m-1} \sum_{i=0}^{n-1} \frac{(-1)^{\varepsilon_i(k)}}{2^{(n-i-1)p+i}}, \qquad (8.12)$$

where $\varepsilon_i(k) \in \{0, 1\}$ is determined by $\sum_{i=0}^{n-1} 2^i \varepsilon_i(k) = k$. Using this expression, (8.11) can be reduced to a simple inequality for weighted sums of binary digits, which has an easy induction proof; see [4].

Shortly after Tabor and Tabor's work appeared, Máko and Páles [56] proved a much more general result, which has the following remarkable consequence: For $p \in (0, 1]$, the minimal function ω_p in (8.9) is given by

$$\omega_p(x) = 2 \sum_{n=0}^{\infty} \frac{1}{2^n} (\phi(2^n x))^p. \qquad (8.13)$$

Comparing (8.10) and (8.13), it is fascinating to see how at the boundary case $p = 1$, the exponent p "jumps" to the other factor in the series' summand.

For more references to papers concerning approximate convexity, we refer to [56].

8.4   Connection with the Riemann Hypothesis

To end this section, we briefly mention a recent result connecting the Takagi function with the Riemann Hypothesis. For a more detailed account, we refer to Lagarias [49].

Let F_n be the nth Farey sequence (also called Farey series), that is, the set of irreducible fractions in $(0, 1]$ with denominator less than or equal to n. Then $|F_n| = \Phi(n)$, where $\Phi(n) = \sum_{k \le n} \varphi(k)$, φ being the Euler totient function. The connection between Farey series and the Riemann Hypothesis is well documented and goes back to a 1924 paper by Franel [27]. Recently, Balasubramanian, Kanemitsu and Yoshimoto [14] showed, as an interesting example of their more general theory, that the Riemann Hypothesis is equivalent to the statement

$$\sum_{r \in F_n} T(r) = \frac{1}{2} \Phi(n) + O(n^{\frac{1}{2} + \varepsilon}) \qquad \text{for every } \varepsilon > 0. \qquad (8.14)$$

Note that the left hand side involves only function values of rational numbers, and can hence be computed by the method of Section 5.1. While it may seem unlikely that the Riemann Hypothesis will some day be solved by directly proving (8.14), the connection is nonetheless surprisingly elegant and beautiful.

References

[1] S. Abbott, J. M. Anderson and L. D. Pitt, Slow points for functions in the Zygmund class Λ_d^*, *Real Anal. Exchange* **32** (2006/07), no. 1, 145–170.

[2] P. C. Allaart, Distribution of the extrema of random Takagi functions, *Acta Math. Hungar.* **121** (2008), no. 3, 243–275.

[3] P. C. Allaart, On a flexible class of continuous functions with uniform local structure, *J. Math. Soc. Japan* **61** (2009), no. 1, 237–262.

[4] P. C. Allaart, An inequality for sums of binary digits, with application to Takagi functions, *J. Math. Anal. Appl.* **381** (2011), no. 2, 689–694.

[5] P. C. Allaart, How large are the level sets of the Takagi function?   *Preprint*, arXiv:1102.1616 (2011).

[6] P. C. Allaart, On the distribution of the cardinalities of level sets of the Takagi function, *preprint*, arXiv:1107.0712 (2011).

[7] P. C. Allaart and K. Kawamura, Extreme values of some continuous, nowhere differentiable functions, *Math. Proc. Camb. Phil. Soc.* **140** (2006), no. 2, 269–295.

[8] P. C. ALLAART and K. KAWAMURA, The improper infinite derivatives of Takagi's nowhere-differentiable function, *J. Math. Anal. Appl.* **372** (2010), no. 2, 656–665.

[9] E. DE AMO, I. BHOURI, M. DÍAZ CARRILLO, and J. FERNÁNDEZ-SÁNCHEZ, The Hausdorff dimension of the level sets of Takagi's function, *Nonlinear Anal.* **74** (2011), no. 15, 5081–5087.

[10] E. DE AMO, M. DÍAZ CARRILLO, and J. FERNÁNDEZ-SÁNCHEZ, Singular functions with applications to fractals and generalized Takagi functions, *preprint* (2011).

[11] E. DE AMO and J. FERNÁNDEZ-SÁNCHEZ, Takagi's function revisited from an arithmetical point of view, *Int. J. Pure Appl. Math.* **54** (2009), no. 3, 407–427.

[12] J. M. ANDERSON and L. D. PITT, Probabilistic behavior of functions in the Zygmund spaces Λ^* and λ^*, *Proc. London Math. Soc.* **59** (1989), no. 3, 558-592.

[13] Y. BABA, On maxima of Takagi-van der Waerden functions, *Proc. Amer. Math. Soc.* **91** (1984), no. 3, 373–376.

[14] R. BALASUBRAMANIAN, S. KANEMITSU and M. YOSHIMOTO, Euler products, Farey series, and the Riemann hypothesis. II, *Publ. Math. Debrecen* **69** (2006), no. 1-2, 1–16.

[15] E. G. BEGLE and W. L. AYRES, On Hildebrandt's example of a function without a finite derivative, *Amer. Math. Monthly* **43** (1936), no. 5, 294–296.

[16] A. S. BESICOVITCH and H. D. URSELL, Sets of fractional dimensions (V). On dimensional numbers of some continuous curves, *J. London Math. Soc.* **12** (1937), 18–25.

[17] P. BILLINGSLEY, Van Der Waerden's Continuous Nowhere Differentiable Function, *Amer. Math. Monthly* **89** (1982), no. 9, 691.

[18] Z. BOROS, An inequality for the Takagi function. *Math. Inequal. Appl.* **11** (2008), no. 4, 757–765.

[19] J. B. BROWN and G. KOZLOWSKI, Smooth interpolation, Hölder continuity, and the Takagi-van der Waerden function. *Amer. Math. Monthly* **110** (2003), no. 2, 142–147.

[20] Z. BUCZOLICH, Micro tangent sets of continuous functions, *Math. Bohem.* **128** (2003), 147–167.

[21] Z. BUCZOLICH, Irregular 1-sets on the graphs of continuous functions. *Acta Math. Hungar.* **121** (2008), no. 4, 371–393.

[22] F. S. CATER, On van der Waerden's nowhere differentiable function. *Amer. Math. Monthly* **91** (1984), no. 5, 307–308.

[23] J. COQUET, Power sums of digital sums *J. Number Theory* **22** (1986), 161–176.

[24] W. F. DARSOW, M. J. FRANK and H.-H. KAIRIES, Errata: "Functional equations for a function of van der Waerden type", *Rad. Mat.* **5** (1989), no. 1, 179–180.

[25] H. DELANGE, Sur la fonction sommatoire de la fonction "somme des chiffres", *Enseignement Math.* **21** (1975), 31–47.

[26] G. FABER, Einfaches Beispiel einer stetigen nirgends differenzierbaren Funktion, *Jahresber. Deutschen Math.-Verein* **16** (1907), 538–540.

[27] J. FRANEL, Les suites de Farey et les problémes des nombres premiers, *Nachr. Ges. Wiss. Göttingen*, Math.-Phys. Kl (1924), 198–201.

[28] P. FRANKL, M. MATSUMOTO, I. Z. RUZSA and N. TOKUSHIGE, Minimum shadows in uniform hypergraphs and a generalization of the Takagi function. *J. Combin. Theory Ser. A* **69** (1995), no. 1, 125–148.

[29] N. G. Gamkrelidze, On a probabilistic properties of Takagi's function (sic), *J. Math. Kyoto Univ.* **30** (1990), no. 2, 227–229.

[30] C. J. Guu, The McFunction [The Takagi function]. Selected topics in discrete mathematics (Warsaw, 1996). *Discrete Math.* **213** (2000), no. 1-3, 163–167.

[31] M. Hata and M. Yamaguti, Takagi function and its generalization, *Japan J. Appl. Math.* **1** (1984), 183–199.

[32] A. Házy and Zs. Páles, On approximately midconvex functions, *Bull. London Math. Soc.* **36** (2004), 339–350.

[33] T. H. Hildebrandt, A simple continuous function with a finite derivative at no point, *Amer. Math. Monthly* **40** (1933), no. 9, 547–548.

[34] J.-P. Kahane, Sur l'exemple, donné par M. de Rham, d'une fonction continue sans dérivée, *Enseignement Math.* **5** (1959), 53–57.

[35] H.-H. Kairies, Takagi's function and its functional equations, *Rocznik Nauk.-Dydakt. Prace Mat.* **15** (1998), 73–83.

[36] H.-H. Kairies, W. F. Darsow and M. J. Frank, Functional equations for a function of van der Waerden type. *Rad. Mat.* **4** (1988), no. 2, 361–374.

[37] N. I. Katzourakis, A Hölder continuous nowhere improvable function with derivative singular distribution, *preprint*, http://arxiv.org/abs/1011.6071v2 (2011).

[38] K. Kawamura, On the classification of self-similar sets determined by two contractions on the plane, *J. Math. Kyoto Univ.* **42** (2002), 255–286.

[39] K. Kawamura, On the set of points where Lebesgue's singular function has the derivative zero, *Proc. Japan Acad. Ser. A* **87** (2011), 162–166.

[40] K. Knopp, Ein einfaches Verfahren zur Bildung stetiger nirgends differenzierbarer Funktionen. *Math. Zeitschr.* **2** (1918), 1–26.

[41] D. E. Knuth, *The art of computer programming, Vol. 4, Fasc. 3*, Addison-Wesley: Upper Saddle River, NJ, 2005.

[42] Z. Kobayashi, Digital sum problems for the Gray code representation of natural numbers, *Interdiscip. Inform. Sci.* **8** (2002), 167–175.

[43] Z. Kobayashi, T. Okada, T. Sekiguchi and Y. Shiota, Applications of measure theory to digital sum problems, *Sem. Math. Sci., Keio Univ.* **35** (2006), 95–109.

[44] N. Kôno, On generalized Takagi functions, *Acta Math. Hungar.* **49** (1987), 315–324.

[45] M. Krüppel, On the extrema and the improper derivatives of Takagi's continuous nowhere differentiable function, *Rostock. Math. Kolloq.* **62** (2007), 41–59.

[46] M. Krüppel, Takagi's continuous nowhere differentiable function and binary digital sums, *Rostock. Math. Kolloq.* **63** (2008), 37–54.

[47] M. Krüppel, On the improper derivatives of Takagi's continuous nowhere differentiable function, *Rostock. Math. Kolloq.* **65** (2010), 3–13

[48] S. T. Kuroda, Private communication.

[49] J. C. Lagarias, The Takagi function and its properties, preprint.

[50] J. C. Lagarias and Z. Maddock, Level sets of the Takagi function: local level sets, arXiv:1009.0855 (2010).

[51] J. C. Lagarias and Z. Maddock, Level sets of the Takagi function: generic level sets, arXiv:1011.3183 (2010).

[52] F. LEDRAPPIER, On the dimension of some graphs, *Contemp. Math.* **135** (1992), 285–293.

[53] J. S. LIPINSKI, On zeros of a continuous nowhere differentiable function, *Amer. Math. Monthly* **73** (1966), no. 2, 166–168.

[54] Z. LOMNICKI and S. ULAM, Sur la théorie de la mesure dans les espaces combinatoires et son application au calcul des probabilités I. Variables indépendantes. *Fund. Math.* **23** (1934), 237–278.

[55] Z. MADDOCK, Level sets of the Takagi function: Hausdorff dimension, *Monatsh. Math.* **160** (2010), no. 2, 167–186.

[56] J. MAKÓ and ZS. PÁLES, Approximate convexity of Takagi type functions, *J. Math. Anal. Appl.* **369** (2010), 545–554.

[57] B. MARTYNOV, On maxima of the van der Waerden function, *Kvant* (1982) June, 8–14 (in Russian).

[58] B. MARTYNOV, Van der Waerden's pathological function: examining a "miserable sore". *Quantum* **8** (1998), no. 6, 12–19.

[59] R. MAULDIN and S. WILLIAMS, On the Hausdorff dimension of some graphs, *Trans. Amer. Math. Soc.* **298** (1986), 793–803.

[60] K. ODANI, On Takagi's nowhere differentiable function, *Sugaku* **47** (1995), 422-423 (in Japanese)

[61] T. OKADA, T. SEKIGUCHI and Y. SHIOTA, An explicit formula of the exponential sums of digital sums, *Japan J. Indust. Appl. Math.* **12** (1995), 425–438.

[62] H. OKAMOTO, *The evolution of the function concept* (in Japanese), Kindai Kagakusha, in press.

[63] G. DE RHAM, Sur un exemple de fonction continue sans dérivée. *Enseignement Math.* **3** (1957), 71–72.

[64] G. DE RHAM, Sur quelques courbes definies par des equations fonctionnelles, *Rend. Sem. Mat. Torino* **16** (1957), 101–113.

[65] R. SALEM, On some singular monotonic functions which are strictly increaing, *Trans. Amer. Math. Soc.* **53** (1943), 427–439.

[66] C. D. SAVAGE, A survey of combinatorial Gray codes, *SIAM Rev.* **39** (1997), 605–629.

[67] S. R. SCHUBERT, On A Function of Van Der Waerden. *Amer. Math. Monthly* **70** (1963), no. 4, 402.

[68] T. SEKIGUCHI and Y. SHIOTA, A generalization of Hata-Yamaguti's results on the Takagi function. *Japan J. Appl. Math.* **8** (1991), 203–219.

[69] A. SHIDFAR and K. SABETFAKHRI, On the Continuity of Van Der Waerden's Function in the Hölder Sense, *Amer. Math. Monthly* **93** (1986), no. 5, 375–376.

[70] A. SHIDFAR and K. SABETFAKHRI, On the Hölder continuity of certain functions, *Exposition. Math.* **8** (1990), 365–369.

[71] B. SOLOMYAK, On the random series $\sum \pm \lambda^n$ (an Erdős problem), *Ann. Math. (2)* **142** (1995), no. 3, 611–625.

[72] K. B. STOLARSKY, Power and exponential sums of digital sums related to binomial coefficient parity, *SIAM J. Appl. Math.* **32** (1977), no. 4, 713–730.

[73] H. Sumi, Cooperation principle, stability and bifurcation in random complex dynamics, *preprint*, http://arxiv.org/abs/1008.3995 (2010).

[74] J. Tabor and J. Tabor, Generalized approximate midconvexity, *Control Cybernet.* **38** (2009), no. 3, 655–669.

[75] J. Tabor and J. Tabor, Takagi functions and approximate midconvexity, *J. Math. Anal. Appl.* **356** (2009), no. 2, 729–737.

[76] T. Takagi, A simple example of the continuous function without derivative, *Proc. Phys.-Math. Soc. Japan* **1** (1903), 176-177. *The Collected Papers of Teiji Takagi*, S. Kuroda, Ed., Iwanami (1973), 5–6.

[77] R. Tambs-Lyche, Une fonction continue sans dérivée, *Enseignement Math.* **38** (1939/40), 208–211.

[78] J. R. Trollope, An explicit expression for binary digital sums, *Math. Mag.* **41** (1968), 21–25.

[79] Y. Tsujii, On modified Takagi functions of two variables, *J. Math. Kyoto Univ.* **25** (1985), no. 3, 577–581.

[80] B. W. van der Waerden, Ein einfaches Beispiel einer nicht-differenzierbaren stetigen Funktion, *Math. Z.* **32** (1930), 474–475.

[81] H. Watanabe, On the scaling exponents of Takagi, Levy and Weierstrass functions (English summary), *Hokkaido Math. J.* **30** (2001), no. 3, 589–604.

[82] Y. Yamaguchi, K. Tanikawa and N. Mishima, Fractal basin boundary in dynamical systems and the Weierstrass-Takagi functions, *Phys. Lett. A* **128** (1988), no. 9, 470–478.

[83] A. Zygmund, Smooth functions, *Duke Math. J.* **12** (1945), 47–76.