# Constant-depth approximation of nowhere-differentiable functions

Bruce M. Kapron
Computer Science Department
University of Victoria
bmkapron@uvic.ca

## 1 Introduction

In his 1962 ICM lecture [Kol62], A.N. Kolmogorov proposed a program of approximating continuous functions by Boolean circuits and using circuit size to measure the complexity of approximation. In particular, circuits are viewed as functions on signed dyadic rationals. For a class $X$ of real-valued functions on $[-1,1]$ let $K(X,\epsilon)$ denote the smallest $n$ such that every $f \in X$ is $\epsilon$-approximated by a circuit of size $n$. Kolmogorov cites work of Ofman giving a tight characterization of $K(X,\epsilon)$ when $X$ is a certain class of smooth functions and upper and lower bounds on $K(X,\epsilon)$ when $X$ is a certain class of functions with an analytic continuation. Here we will cite improvements of these characterizations given by Asarin [Asa84]. Let $W_{p+\alpha,c}$, $(p \geq 0,\ 0 < \alpha < 1)$ denote the class of functions $f : [-1,1] \to [-1,1]$ that are $p$ times continuously differentiable and whose $p$-th derivative satisfies a Hölder condition with exponent $\alpha$ and coefficient $c$:

$$|f^{(p)}(x) - f^{(p)}(y)| \leq c|x-y|^{\alpha} \quad \text{for any } x, y \in [-1,1]$$

and let $A_{r,c}$ denote the functions $f : [-1,1] \to \mathbb{R}$ that can be continued analytically to the Bernstein ellipse with foci $-1$ and $1$ and sum of semi-axes $r$ with modulus bounded by $c$ on the ellipse. Then

$$K(W_{q,c},\epsilon) = \Theta\left(\left(\frac{1}{\epsilon}\right)^{\frac{1}{q}} \log\frac{1}{\epsilon}\right)$$

$$K(A_{r,c},\epsilon) = \Omega\left(\log^2\frac{1}{\epsilon} / \log\log\frac{1}{\epsilon}\right) \qquad K(A_{r,c},\epsilon) = \tilde{O}\left(\log^2\frac{1}{\epsilon}\right)$$

The upper bounds are obtained constructively while the lower bounds rely on entropy arguments (Asarin [Asa84] also gives a constructive exponential lower bound for $2^{-n}$-approximating a 1-Lipschitz function using an incompressibility-based argument.) Note that in the analytic case, the upper bound gives a $2^{-n}$-approximation by size $\tilde{O}(n^2)$ circuits. Also note that $K(W_1, 2^{-n}) = \Theta(n2^n)$, where $W_1 = \bigcup_c W_{1,c}$.

An important aspect of the classical theory of polynomial approximation is the existence of so-called "inverse theorems", roughly stating that being "well-approximated" by low-degree polynomials implies smoothness (or even analytic continuation when convergence is rapid enough) [Ber12]. Kolmogorov observes that corresponding results for small circuit size do not hold. In particular, he notes that a variant of van der Waerden's nowhere-differentiable function can be $2^{-n}$ approximated by circuits of size $O(n^2)$. Kolmogorov's observation is the starting point for the current note. In particular, we ask whether the lack of an inverse theorem also holds in the case of circuit *depth*, recalling that certain classes of analytic functions may be approximated by constant-depth threshold circuits ([RT92, MT99].) Unfortunately, in this setting we also obtain a negative result by showing that Takagi's nowhere-differentiable function can be $2^{-n}$ approximated in $\mathsf{TC}^0$ and by defining a variant of the function that can be $2^{-n}$ approximated in $\mathsf{AC}^0$.

## 2 Takagi's Function

In the sequel, we will work over $[0,1]$, which is more appropriate for the functions we consider. For $x \in \mathbb{R}$, let $\langle\langle x \rangle\rangle$ denote the distance from $x$ to the nearest integer. In 1901, T. Takagi [Tak01] introduced a continuous nowhere-differentiable function $\tau : [0,1] \to [0,1]$, which may be defined as follows (see [Lag11] for a complete discussion):

$$\tau(x) = \sum_{k=0}^{\infty} \frac{1}{2^k} \langle\!\langle 2^k x \rangle\!\rangle.$$

Suppose $x, \tilde{x} \in [0,1]$ where $\tilde{x} = \sum_{i=1}^{2n+2} b_i 2^{-i}$ and $|x - \tilde{x}| \le 2^{-(2n+2)}$. We have

$$\begin{aligned}
|\tau(x) - \tau(\tilde{x})| &= \left| \sum_{k=0}^{\infty} \frac{1}{2^k} \langle\!\langle 2^k x \rangle\!\rangle - \sum_{k=0}^{\infty} \frac{1}{2^k} \langle\!\langle 2^k \tilde{x} \rangle\!\rangle \right| \\
&= \left| \sum_{k=0}^{\infty} \frac{1}{2^k} \langle\!\langle 2^k x \rangle\!\rangle - \sum_{k=0}^{2n+1} \frac{1}{2^k} \langle\!\langle 2^k \tilde{x} \rangle\!\rangle \right| \\
&\le \sum_{k=0}^{2n+1} \frac{1}{2^k} \left| \langle\!\langle 2^k x \rangle\!\rangle - \langle\!\langle 2^k \tilde{x} \rangle\!\rangle \right| + \sum_{k=2n+2}^{\infty} \frac{1}{2^k} \langle\!\langle 2^k x \rangle\!\rangle \\
&\le \sum_{k=0}^{2n+1} \frac{1}{2^k} \cdot \frac{1}{2^{2n-k+2}} + \sum_{k=2n+2}^{\infty} \frac{1}{2^{k+1}} \\
&= \frac{2n+2}{2^{2n+2}} + \frac{1}{2^{2n+2}} \le \frac{1}{2^n}
\end{aligned}$$

So $\tau(\tilde{x})$ provides a $2^{-O(n)}$ approximation to $\tau(x)$. We note that $\tau(\tilde{x})$ is obtained by adding $2n+2$ numbers of $2n+2$ bits, where each number is obtained via a constant-depth transformation of $\tilde{x}$ involving only Boolean operations (see the discussion in the following paragraph.) Since such sums may be computed in $\mathsf{TC}^0$ (see, e.g., [Vol99]), we conclude that $\tau$ may be $2^{-O(n)}$-approximated in $\mathsf{TC}^0$.

Could it be possible that $\tau$ is approximated by an even simpler class of circuits? We will show that $\tau$ cannot be $2^{-n}$-approximated in $\mathsf{AC}^0$. To start, we recall [Lag11] that for $x \in [0,1]$

$$\langle\!\langle x \rangle\!\rangle = \begin{cases} x & \text{if } 0 \le x < \frac{1}{2} \quad \text{i.e., } b_1 = 0 \\ 1 - x & \text{if } \frac{1}{2} \le x \le 1 \quad \text{i.e., } b_1 = 1. \end{cases}$$

Furthermore, if $x = \sum_{i=1}^{\infty} b_i 2^{-i} = 0.b_1 b_2 b_3 \ldots$ and $i \ge 0$, we have

$$\langle\!\langle 2^i x \rangle\!\rangle = \begin{cases} 0.b_{i+1} b_{i+2} \ldots & \text{if } b_{i+1} = 0 \\ 0.\overline{b}_{i+1} \overline{b}_{i+2} \ldots & \text{if } b_{i+1} = 1 \end{cases}$$

where for $b \in \{0,1\}$, $\overline{b} = 1 - b = b$. Using the fact that $\overline{b} = b \oplus 1$, where $\oplus$ denotes exclusive-or, we can rewrite this as

$$\langle\!\langle 2^i x \rangle\!\rangle = 0.(b_{i+1} \oplus b_{i+1})(b_{i+1} \oplus b_{i+2}) \ldots$$

and so

$$\frac{1}{2^i} \langle\!\langle 2^i x \rangle\!\rangle = 0.\underbrace{00\ldots0}_{i}(b_{i+1} \oplus b_{i+1})(b_{i+1} \oplus b_{i+2}) \ldots$$

Suppose that we have a family $\{C_n\}$ of constant-depth poly-size Boolean circuits, $C_n : \{0,1\}^{\alpha(n)} \to \{0,1\}^{\beta(n)}$ such that for all $x, \tilde{x} \in [0,1]$ where $\tilde{x} = \sum_{i=1}^{\alpha(n)} b_i 2^{-i}$ and $|x - \tilde{x}| \le 2^{-\alpha(n)}$, we have $|\tau(x) - C_n(\tilde{x})| \le 2^{-n}$, where $C_n(\tilde{x})$ denotes $C_n(b_1 b_2 \ldots b_{\alpha(n)})$. In particular, $|\tau(\tilde{x}) - C_n(\tilde{x})| \le 2^{-n}$, which means that $\tau(\tilde{x})$ and $C_n(\tilde{x})$ must agree on their first $n$ bits. Assuming $\alpha(n) \ge n+1$, consider $\tilde{x}$ where $b_{n+1} = 1$ and $b_i = 0$ for $i > n+1$. From the discussion above we have, for $0 \le k < n$,

$$\langle\!\langle 2^k \tilde{x} \rangle\!\rangle = \begin{cases} 0.b_{k+1} \ldots b_{n-1} b_n 100 \ldots & \text{if } b_{k+1} = 0 \\ 0.\overline{b}_{k+1} \ldots \overline{b}_{n-1} \overline{b}_n 011 \cdots = 0.\overline{b}_{k+1} \ldots \overline{b}_{n-1} \overline{b}_n 100 \ldots & \text{if } b_{k+1} = 1 \end{cases}$$

From this we see that for $0 \le k < n$, the $n$th bit of $\frac{1}{2^k}\langle\!\langle 2^k\tilde{x}\rangle\!\rangle$ is $b_{k+1} \oplus b_n$, while the $(n+1)$st bit is always 1, so that

$$\begin{array}{rlllll}
\tau(\tilde{x}) = 0. & (b_1 \oplus b_1) & (b_1 \oplus b_2) & \ldots & (b_1 \oplus b_n) & 1 \\
+\ 0. & 0 & (b_2 \oplus b_2) & \ldots & (b_2 \oplus b_n) & 1 \\
+\ 0. & 0 & 0 & \ldots & (b_3 \oplus b_n) & 1 \\
 & & \cdots & & & \\
+\ 0. & 0 & 0 & \ldots & (b_n \oplus b_n) & 1 \\
+\ 0. & 0 & 0 & \ldots & 0 & 1
\end{array}$$

Letting $c$ denote the $n$th bit of $\tau(\tilde{x})$, we have

$$b_1 \oplus \cdots \oplus b_n = \begin{cases} c \oplus b_n \oplus 1 & \text{if } n+1 \text{ is even and } \left\lfloor \frac{n+1}{2} \right\rfloor \text{ is odd} \\ c \oplus b_n & \text{if } n+1 \text{ is even and } \left\lfloor \frac{n+1}{2} \right\rfloor \text{ is even} \\ c \oplus 1 & \text{if } n+1 \text{ is odd and } \left\lfloor \frac{n+1}{2} \right\rfloor \text{ is odd} \\ c & \text{if } n+1 \text{ is odd and } \left\lfloor \frac{n+1}{2} \right\rfloor \text{ is even,} \end{cases}$$

where $\oplus$ denotes exclusive-or. Here the parity of $n+1$ determines parity of the $b_n$'s in the $n$th position while the parity of $\left\lfloor \frac{n+1}{2} \right\rfloor$ determines whether there is a carry-out from the $(n+1)$st position to the $(n-1)st$ position. So, by the assumption on $\{C_n\}$, we may obtain a family $\{D_n\}$ of constant-depth poly-sized circuits such that $D_n(b_1 \ldots b_n) = b_1 \oplus \cdots \oplus b_n$, contrary to the fact that $\text{PARITY} \notin \mathsf{AC}^0$.

We have ruled out $\mathsf{AC}^0$-approximation of Takagi's function. Now consider the following function:

$$T(x) = \sum_{k=0}^{\infty} \frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}x \rangle\!\rangle.$$

Suppose $x, \tilde{x} \in [0,1]$ where $\tilde{x} = \sum_{i=1}^{2n+2} b_i 2^{-i}$ and $|x - \tilde{x}| \le 2^{-(2n+2)}$. Here we have

$$\begin{aligned}
|T(x) - T(\tilde{x})| &= \left| \sum_{k=0}^{\infty} \frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}x \rangle\!\rangle - \sum_{k=0}^{\infty} \frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}\tilde{x} \rangle\!\rangle \right| \\
&= \left| \sum_{k=0}^{\infty} \frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}x \rangle\!\rangle - \sum_{k=0}^{\lceil \log(n+1) \rceil + 1} \frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}\tilde{x} \rangle\!\rangle \right| \\
&\le \sum_{k=0}^{\lceil \log(n+1) \rceil + 1} \frac{1}{2^{2^k-1}} \left| \langle\!\langle 2^{2^k-1}x \rangle\!\rangle - \langle\!\langle 2^{2^k-1}\tilde{x} \rangle\!\rangle \right| + \sum_{k=\lceil \log(n+1) \rceil + 2}^{\infty} \frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}x \rangle\!\rangle \\
&\le \sum_{k=0}^{\lceil \log(n+1) \rceil + 1} \frac{1}{2^{2^k-1}} \cdot \frac{1}{2^{2n-(2^k-1)+2}} + \sum_{k=\lceil \log(n+1) \rceil + 2}^{\infty} \frac{1}{2^{2^k}} \\
&\le \frac{\lceil \log(n+1) \rceil + 2}{2^{2n+2}} + \frac{1}{2^{2n+2}} \le \frac{1}{2^n}
\end{aligned}$$

So $T(\tilde{x})$ provides a $2^{-O(n)}$ approximation to $T(x)$. $T(\tilde{x})$ is obtained by adding $\lceil \log(n+1) \rceil + 2$ of numbers of $2n+2$ bits where each number can be obtained in $\mathsf{AC}^0$. Since such sums may be performed in $\mathsf{AC}^0$ (see, e.g., [Vol99]). In conclusion, we can $2^{-O(n)}$-approximate $T$ in $\mathsf{AC}^0$. Here we can use an immediate adaptation to $T$ of similar proofs (e.g., [Bil82]) for $\tau$. The continuity of $T$ follows by the Weierstrass M-test (or from the fact that it can be approximated.) Finally, letting $\varphi_k(x)$ denote $\frac{1}{2^{2^k-1}} \langle\!\langle 2^{2^k-1}x \rangle\!\rangle$, if $T$ were differentiable at $x \in [0,1]$, we would have $T'(x) = \sum_{k=0}^{\infty} \varphi'_k(x)$. However, this sum diverges since for all $k$, $|\varphi'_k(x)| = 1$. This intuition can be formalized to show that $T$ is nowhere differentiable.

## References


[Asa84] Eugene A Asarin. On convergence of uniform approximations of continuous functions. *Russian Mathematical Surveys*, 39(3):179, 1984.

[Ber12] Sergei N Bernstein. On the best approximation of continuous functions by polynomials of a given degree. *Comm. Soc. Math. Kharkow, Ser*, 2(13):49–194, 1912.

[Bil82] Patrick Billingsley. Van der Waerden's continuous nowhere differentiable function. *The American Mathematical Monthly*, 89(9):691–691, 1982.

[Kol62] Andrei N Kolmogorov. Various approaches to the estimate of difficulty of the approximate definition and calculation of functions. *Proc. Internat. Congr. Mathematicians*, pages 351–356, 1962.

[Lag11] Jeffrey C Lagarias. The Takagi function and its properties. *arXiv preprint arXiv:1112.4205*, 2011.

[MT99] Alexis Maciel and Denis Thérien. Efficient threshold circuits for power series. *Inf. Comput.*, 152(1):62–73, 1999.

[RT92] John H. Reif and Stephen R. Tate. On threshold circuits and polynomial computation. *SIAM J. Comput.*, 21(5):896–908, 1992.

[Tak01] Teiji Takagi. A simple example of the continuous function without derivative. *Tokyo Sugaku-Butsurigakkwai Hokoku*, 1:F176–F177, 1901.

[Vol99] Heribert Vollmer. *Introduction to Circuit Complexity - A Uniform Approach.* Texts in Theoretical Computer Science. An EATCS Series. Springer, 1999.