SUBDIFFERENTIAL ANALYSIS OF THE VAN DER WAERDEN FUNCTION

P. GÓRA AND R. J. STERN

Abstract. A concise and direct proof is given that Hölder subdifferentials of the (continuous but nowhere differentiable) Van der Waerden function $H(\cdot)$ exhibits the same behaviour as the Weierstrass function: There exists a countable dense set $\Gamma \subset R$ (the dyadic rationals) such that each Hölder subdifferential $\partial_\alpha H(x)$ is all of $\mathbb{R}$ for every $x \in \Gamma$, while $\partial_\alpha H(x) = \emptyset$ for $x \notin \Gamma$.

1. Introduction

The classical *Weierstrass function* is given by

$$W(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x) \ , \quad 0 < x < 1 \ ,$$

where $0 < a < 1$, $b > 0$ is an odd integer, and $ab > 1 + \frac{3\pi}{2}$. The function $W(\cdot)$ is known to be continuous but nowhere differentiable.

In [8], Garg applied general results on the structure of nondifferentiable functions as well as specific properties of $W(\cdot)$ in order to describe its Dini-subderivates, and in [11], Wang noted that Garg's results imply that the proximal and Dini subdifferentials of $W(\cdot)$ are nonempty only on a countable dense set, and that on this set these subdifferentials equal $\mathbb{R}$. This is in a similar spirit to the main result of Borwein, Girgensohn and Wang [3], which involved the construction of Lipschitz functions whose Hölder subdifferentials are empty except on a countable dense set, where they take a constant interval value.

McCarthy [9] may have been the first to observe that a "fractal sawtooth" construction also yielded a continuous nondifferentiable function, but in a very simple and direct way. One such sawtooth function is given as follows:

Denote by $G(x)$ the distance from the real number x to the nearest integer. Let

$$G_n(x) = \frac{1}{2^n} G(2^n x) \ , \quad n = 0, 1, 2, \dots \ ,$$

and let $H : \mathbb{R} \to \mathbb{R}$ be defined as

$$H(x) = \sum_{n=0}^{\infty} G_n(x) \ .$$

$H(\cdot)$ is known as the *Van der Waerden function.* (See e.g. Billingsley [1], Cater [5] as well as Shidfar and Sabetfakhri [10] for discussions of further properties of this function.) Clarke, Ledyaev and Wolenski [7] employed a variant of $H(\cdot)$ in order to

---

2000 *Mathematics Subject Classification.* Primary 26A27, 49J52.

Key words and phrases. Van der Waerden function, Hölder subdifferentials, nonsmooth analysis.

Research supported by the Natural Sciences Engineering Research Council of Canada.

produce a C^1 function whose proximal subdifferential is nonempty only on a dense set of measure zero which does not contain a dense G_δ.

Our purpose in the present note is to show that just as the Van der Waerden function provides a much simpler example than the Weierstrass function when it comes to demonstrating continuity plus nondifferentiablilty, it also serves to exhibit, in a particularly direct way, the aforementioned conclusion of Wang in [11] (which is easily extended to all Hölder subdifferentials).

1.1. **Some useful properties of $H(\cdot)$.** Let

$$H_n(x) := \sum_{m=0}^{n-1} G_m(x),$$

and denote the dyadic rationals by

$$\Gamma := \left\{ \frac{k}{2^n} \,:\, k, n \text{ are integers}, \quad n \geq 0 \right\}.$$

Denote the slope of $H_n(\cdot)$ at a point x by $sl(H_n(x))$. The following facts are readily verified; the figure below is useful for this purpose. (Note that the functions in the figure ae defined on the entire real line, but for convenience are only plotted on the interval $[0, 2]$. Note also that the function $H(\cdot)$ has period 1.)

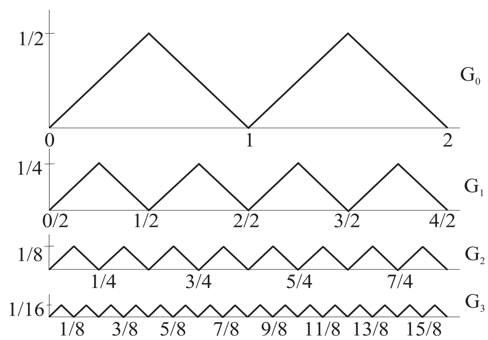

FIGURE 1. The family $\{G_n\}$

(a) $G_n(x) = 0$ for all $x = \frac{k}{2^n}$, k, $n \geq 0$ integers.

(b) G_n is piecewise linear with slopes ± 1 and G_n has constant slopes on intervals $(k/2^{n+1}, (k+1)/2^{n+1})$, and in general $(k/2^m, (k+1)/2^m)$ with $m > n$.

(c) If I is a maximal interval for H_n with constant slope S, then $sl(H_{n+1}(\cdot))$ has slope $S + 1$ on first half of I and $S - 1$ on the second.

(d) If $x = k/2^m$, then for $n \geq m$ the slope of G_n changes at x from -1 to 1.

1.2. **Hölder subgradients and statement of main result.** For $\alpha > 1$ and $x \in \mathbb{R}$, we say that $\zeta \in \mathbb{R}$ is an *$\alpha-$Hölder subgradient* of $H(\cdot)$ at x if there exist $\sigma > 0$ and an open interval $I = I(x, \sigma)$ containing x such that

$$H(t) - H(x) + \sigma |t - x|^\alpha \geq \zeta (t - x) \quad \forall\, t \in I \,. \tag{1}$$

Also, ζ is said to be a 1-Hölder subgradient of $H(\cdot)$ at x provided that for *any* given $\sigma > 0$ there is an open interval $I = I(x, \sigma)$ such that (1) holds. For $\alpha \geq 1$, the set of all α-Hölder subgradients at x is called the *α-Hölder subdifferential* of $H(\cdot)$ at x, denoted $\partial_\alpha H(x)$. (In common nonsmooth analysis parlance, $\partial_1 H(x)$ is called the *Dini subdifferential*, while $\partial_2 H(x)$ is known as the *proximal subdifferential*.)

We have that $\zeta \in \partial_1 H(x)$ if for each $\sigma > 0$ there exists $r = r(\sigma)$ such that

$$H(y) \geq H(x) + \zeta(y-x) - \sigma|y-x| \quad \forall\, y \in (x-r, x+r)\ . \tag{2}$$

This means that $H(\cdot)$ majorizes the wedge W_x^σ formed by the graph of

$$f_x(y) := H(x) + \zeta(y-x) - \sigma|y-x|,$$

for $y \in (x-r, x+r)$, while $f_x(x) = H(x)$. The graph of $f_x(y)$ for $y \in (x-r, x)$ will be called the left arm of the wedge and the graph of $f_x(y)$ for $y \in (x, x+r)$ will be called the right arm of the wedge.

It is readily noted that for every $x \in \mathbb{R}$ one has

$$1 \leq \alpha < \alpha' \implies \partial_{\alpha'} H(x) \subset \partial_\alpha H(x)\ , \tag{3}$$

which is vacuously true if $\partial_{\alpha'} H(x) = \emptyset$. Due to results in Borwein and Preiss [2], we know that for each $\alpha \geq 1$, the set $\{x \in \mathbb{R}\ :\ \partial_\alpha H(x) \neq \emptyset\}$ is dense in $\mathbb{R}$. (See e.g. Clarke, Ledyaev, Stern and Wolenski [6] for density proofs specialized to the proximal case.)

The main result of this note is the following, which asserts that the pathological subdifferential behavior of the Weierstrass functon is duplicated by the Van der Waerden function:

Theorem 1. *For each $\alpha \geq 1$ one has*

$$x \in \Gamma \implies \partial_\alpha H(x) = \mathbb{R}\ , \tag{4}$$

and

$$x \notin \Gamma \implies \partial_\alpha H(x) = \emptyset\ . \tag{5}$$

Theorem 1 asserts that while the set

$$\{x \in \mathbb{R}\ :\ \partial_\alpha H(x) \neq \emptyset\} = \Gamma$$

is as "small as possible" (i.e. countably dense), the $\alpha-$Hölder subdifferential is "as large as possible" (namely all of $\mathbb{R}$) on Γ, and that this is true for all $\alpha \geq 1$.

Remark 2. Let us denote the binary expansion by of $x \in [0,1]$ by

$$x = \sum_{j=0}^{\infty} \frac{b_j}{2^{j+1}}\ ,\quad b_j \in \{0,1\}\ .$$

Then the slopes of the functions $G_m(\cdot)$ at x are $(-1)^{b_m}$, $m = 0, 1, 2, \ldots$, and the slope of $H_n(\cdot)$ at x (denoted below by $sl(H_n(x))$) is $\sum_{m=0}^{n-1}(-1)^{b_m}$, $n = 1, 2, 3, \ldots$ A moment's reflection yields that the sequence $\{sl(H_n(x))\}_{n=1}^{\infty}$ can be viewed as a random walk, whence the law of the iterated logarithm (see e.g. Breiman [4]) implies that almost everywhere in $(0,1]$ one has

$$\limsup_{n\to\infty} \frac{sl(H_n(x))}{\sqrt{2n\log\log n}} = -\liminf_{n\to\infty} \frac{sl(H_n(x))}{\sqrt{2n\log\log n}} = 1\ . \tag{6}$$

It transpires that if $x \in (0,1]$ satisfies (6), then $\partial_\alpha H(x) = \emptyset$ for all $\alpha \geq 1$. In view of (3), it is enough to consider $\alpha = 1$. Suppose that $H(\cdot)$ majorizes a wedge at x as described above, and assume the left arm has slope S (which is necessarily nonnegative). Then (6) implies that for infinitely many values of n, the slopes

$sl(H_n(x)) > S$. Let us fix such an n and a k such that the interval $(\frac{k}{2^n}, \frac{k+1}{2^n})$ containing x. (H_n has constant slope on this interval.) We have $H(\frac{k}{2^n}) = H_n(\frac{k}{2^n})$ and $H(x) \geq H_n(x)$. Since $sl(H_n(x)) > S$, the point $(\frac{k}{2^n}, H(\frac{k}{2^n}))$ is in the interior of the wedge. We have a sequence of such points converging to $(x, H(x))$. The treatment for the right arm is similar and we draw the required conclusion.

Although the random walk idea has appeal, the assertion of Theorem 1 is of course stronger.

2. Proof of main result

Lemma 3. *The point x is a local minimum for $H(\cdot)$ if and only if $x \in \Gamma$. Furthermore, for any integer $\Lambda \geq 1$ there exists an open interval containing x such that $H(t) \geq H(x) + \Lambda|t - x|$ for all t in this interval.*

Proof: Let $x = k/2^n$ with k odd and n a nonnegative integer. (The case of nonzero even k reduces to the odd case upon division, and the case $k = 0$ is handled similarly to what follows.)

Let $h = H_n(x)$ and let the integer S be the slope of $H_n(\cdot)$ in the interval $(\frac{k-1}{2^n}, \frac{k}{2^n})$. (Note that it is constant). Then, the slope of $H_n(\cdot)$ in the interval $(\frac{k}{2^n}, \frac{k+1}{2^n})$ is constant and equals $S - 2$. Let

$$F_n = H_{n+|S|+\Lambda+2} = H_n + G_n + G_{n+1} + \cdots + G_{n+|S|+\Lambda+1} \ .$$

By (a), $F_n(x) = h$. Also, By (d) the slope of $F_n(\cdot)$ is $S - |S| - \Lambda - 2 < -\Lambda$ on the interval $(k/2^n - 1/2^{n+|S|+\Lambda+2}, k/2^n)$ and is $S - 2 + |S| + \Lambda + 2 \geq \Lambda$ on $(k/2^n, k/2^n + 1/2^{n+|S|+\Lambda+2})$. Let

$$K(\cdot) := H(\cdot) - F_n(\cdot) = \sum_{m=n+|S|+\Lambda+2}^{\infty} G_m(\cdot) \ .$$

We have $K(x) = 0$ and $K(t) \geq 0$ for all $t \in \mathbb{R}$. Thus, x is local minimum of $H(\cdot)$ and

$$H(t) \geq H(x) + \Lambda|t - x| \quad \forall\, t \in (x - 1/2^{n+|S|+\Lambda+2}, x + 1/2^{n+|S|+\Lambda+2}) \ .$$

Now assume that $x \notin \Gamma$ and that x is a local minimum of $H(\cdot)$. Then there is an interval J containing x and such that

$$H(t) \geq H(x) \quad \forall\, t \in J \ . \tag{7}$$

The interval J contains (strictly) an interval $I = (k/2^n, (k+1)/2^n)$ containing x. The function $H_n(\cdot)$ has constant slope on $I = (k/2^n, (k+1)/2^n)$. We can assume it is not 0. (Note that if it is, then we could consider halves of I, the intervals $I_1 = ((k/2^n, (2k+1)/2^{n+1})$ and $I_2 = ((2k+1)/2^{n+1}, (k+1)/2^n)$. One of them must contain x and the slope of $H_{n+1}(\cdot)$ is 1 on I_1 and is -1 on I_2.)

Then one of the values $v_1 = H_n(k/2^n)$ or $v_2 = H_n((k+1)/2^n)$, say v_1, satisfies $v_1 < H_n(x)$. We have

$$G_k((k/2^n)) = G_k((k+1)/2^n) = 0 \quad \forall\, k \geq n \ ,$$

and

$$G_k(t) \geq 0 \quad \forall\, t \in \mathbb{R}, \ k \geq n \ .$$

Thus, $H(k/2^n) = v_1 < H(x)$, which contradicts (7). $\square$

Lemma 4. *Let $x \in \mathbb{R}$. If $x \notin \Gamma$, then $\partial_1 H(x) = \emptyset$.*

Proof: If $x \notin \Gamma$, then the sequence $(sl(H_n(x)))_{n\geq 1}$ satisfies at least one of the conditions

(a) $\limsup_{n\to\infty} sl(H_n(x) = +\infty$;

(b) $\liminf_{n\to\infty} sl(H_n(x) = -\infty$;

(c) $(sl(H_n(x)))_{n\geq 1}$ is bounded and there exist at least two integers, say $a > b$ such that $(sl(H_n(x)))_{n\geq 1}$ contains infinitely many a's and infinitely many b's.

Let us fix $x \neq \frac{k}{2^n}$ and assume that $\zeta \in \partial_1(x)$. We can assume that $\zeta \geq 0$. The proof for $\zeta < 0$ is similar.

We will first deal with cases (a) and (b). For fixed $\sigma > 0$, the slope of left arm of W_x^σ is $\zeta + \sigma > 0$. The slope of the right arm is $\zeta - \sigma$. In case (a) we choose n_1 such that $2^{-n_1} < r$ and $sl(H_{n_1}(x)) > \zeta + \sigma$.

There exists $k = k_{n_1}$ such that the interval $(\frac{k}{2^{n_1}}, \frac{k+1}{2^{n_1}})$ contains x. (H_{n_1} has constant slope on this interval.) We have $H(\frac{k}{2^{n_1}}) = H_n(\frac{k}{2^{n_1}})$ and $H(x) \geq H_{n_1}(x)$. Since $sl(H_{n_1}(x)) > \zeta + \sigma$ the point $(\frac{k}{2^{n_1}}, H(\frac{k}{2^{n_1}}))$ is in the interior of the wedge W_x^σ. This contradicts (2).

In case (b) we choose n_2 such that $2^{-n_2} < r$ and $sl(H_{n_2}(x)) < \zeta - \sigma$. There exists $k = k_{n_2}$ such that interval $(\frac{k}{2^{n_2}}, \frac{k+1}{2^{n_2}})$ contains x. (H_{n_2} has constant slope on this interval.) We have $H(\frac{k+1}{2^{n_2}}) = H_n(\frac{k+1}{2^{n_2}})$ and $H(x) \geq H_{n_2}(x)$. Since $sl(H_{n_2}(x)) < \zeta - \sigma$ the point $(\frac{k+1}{2^{n_2}}, H(\frac{k+1}{2^{n_2}}))$ is in the interior of the wedge W_x^σ. This contradicts (2).

Case (c). Let α be the obtuse angle between a straight line of slope a and a straight line of slope b. We have $\alpha < \pi$ so we can find an $\sigma > 0$ such that $\alpha < \alpha_x^\sigma$, the angle between the left and right arms of the wedge. Let $r = r(\sigma)$ correspond to this σ. We choose $n_1 < n_2$ such that $2^{-n_1}, 2^{-n_2} < r$ and $sl(H_{n_1}(x)) = a$, $sl(H_{n_2}(x)) = b$. There exist k_1, k_2 such that intervals $I_1 = (\frac{k_1}{2^{n_1}}, \frac{k_1+1}{2^{n_1}})$, $I_2 = (\frac{k_2}{2^{n_2}}, \frac{k_2+1}{2^{n_2}})$ both contain x. (H_{n_1} has constant slope on I_1 and H_{n_2} has constant slope on I_2.) We have $H(\frac{k_1}{2^{n_1}}) = H_{n_1}(\frac{k_1}{2^{n_1}})$ and $H(x) \geq H_{n_1}(x)$. We have $H(\frac{k_2+1}{2^{n_2}}) = H_{n_2}(\frac{k_2+1}{2^{n_2}})$ and $H(x) \geq H_{n_2}(x)$. Thus, the angle $\angle(H(\frac{k_1}{2^{n_1}}), H(x), H(\frac{k_2+1}{2^{n_2}}))$ is smaller then angle α and therefore smaller than angle α_x^σ. Thus, at least one of the points $(\frac{k_1}{2^{n_1}}, H(\frac{k_1}{2^{n_1}}))$, $(\frac{k_2+1}{2^{n_2}}, H(\frac{k_2+1}{2^{n_2}}))$ is in the interior of the wedge W_x^σ. This contradicts (2). $\square$

Proof of Theorem 1: The implication (4) follows directly from Lemma 3, while Lemma 4 and (3) immediately yield (5). $\square$

For $\alpha \geq 1$, let us define the *limiting α-Hölder subdifferential* of $H(\cdot)$ at x as

$$\partial_\alpha^L H(x) := \{\lim_i \zeta_i \,:\, \zeta_i \in \partial_\alpha H(x_i)\,,\, x_i \to x\}\ .$$

Basic nonsmooth analysis (see [6]) asserts equality in the Dini and proximal cases; that is, $\partial_1^L H(x) = \partial_2^L H(x)$, which in view of (3) gives equality for all $\alpha \in [1, 2]$. However, Theorem 1 implies more than this due to the special nature of the Van der Waerden function:

Corollary 5. *For all $\alpha \geq 1$, we have $\partial_\alpha^L H(x) = \mathbb{R}$ for every $x \in \mathbb{R}$.*

Acknowledgement: The authors benefited from discussions with F. H. Clarke, and from the comments of the referee.

References

[1] P. Billingsley, *Van der Waerden's continuous nowhere differentiable function*, A.M.S. Monthly 89 (1982), p. 691.

[2] J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. Amer. Math. Soc. 303 (1987), pp. 517–527.

[3] J. M. Borwein, R. Girgensohn and X. Wang, *On the construction of Hölder and proximal subderivatives*, Canad. Math. Bull. 41 (1998), pp. 497–507.

[4] L. Brieman, *Probability*, Addison-Wesley 1968

[5] F. S. Cater, *On Van der Waerden's nowhere differentiable function*, A.M.S. Monthly 91 (1984), pp. 307–308.

[6] F. H. Clarke, Yu. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Graduate Texts in Mathematics, 178, Springer-Verlag, New York (1998).

[7] F. H. Clarke, Yu. S. Ledyaev and P. R. Wolenski, *Proximal analysis and minimization principles*, J. Math. Anal. Appl. 196 (1995), pp. 722–735.

[8] K. M. Garg, *On asymmetrical derivatives of nondifferentiable functions*, Canadian J. Math. 20 (1968), pp. 135–143.

[9] J. McCarthy, *An everywhere continuous nowhere differentiable function*, A.M.S. Monthly 15 (1953).

[10] A. Shidfar and K. Sabetfakhri, *On the continuity of Van der Waerden's function in the Hölder sense*, A.M.S. Monthly 93 (1986), pp. 375–376.

[11] S. X Wang, *Fine and Pathological Properties of Subdifferentials*, Ph.D. Dissertation in Mathematics, Simon Fraser University (1999).

Department of Mathematics and Statistics, Concordia University, 1400 De Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8, Canada
E-mail address: pgora@mathstat.concordia.ca

Department of Mathematics and Statistics, Concordia University, 1400 De Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8, Canada
E-mail address: stern@mathstat.concordia.ca