Simple proofs of nowhere-differentiability for Weierstrass' function and cases of slow growth

Johnsen, Jon

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Johnsen, J. (2008). *Simple proofs of nowhere-differentiability for Weierstrass' function and cases of slow growth*. Department of Mathematical Sciences, Aalborg University. Research Report Series, No. R-2008-02

AALBORG UNIVERSITY

Simple proofs of nowhere-differentiability for Weierstrass' function and cases of slow growth

by

Jon Johnsen

R-2008-02 February 2008

DEPARTMENT OF MATHEMATICAL SCIENCES
AALBORG UNIVERSITY
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 96 35 80 80 ▪ Telefax: +45 98 15 81 29
URL: http://www.math.aau.dk

ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

SIMPLE PROOFS OF NOWHERE-DIFFERENTIABILITY FOR WEIERSTRASS' FUNCTION AND CASES OF SLOW GROWTH

JON JOHNSEN

ABSTRACT. Using a few basics from integration theory, a short proof of nowhere-differentiability of Weierstrass functions is given. Restated in terms of the Fourier transformation, the method consists of a second microlocalisation, which is used to derive two general results on existence of nowhere differentiable functions. Examples are given in which the frequencies are of polynomial growth and of almost quadratic growth as a limiting case.

1. INTRODUCTION

In 1872, K. Weierstrass presented his famous example of a nowhere differentiable function W on the real line. With two real parameters $b \geq a > 1$, this may be written as

$$W(t) = \sum_{j=0}^{\infty} a^{-j} \cos(b^j t), \quad t \in \mathbb{R}. \tag{1.1}$$

Weierstrass proved that W is continuous at every $t_0 \in \mathbb{R}$, but not differentiable at any $t_0 \in \mathbb{R}$, at least if

$$\frac{b}{a} > 1 + \frac{3\pi}{2}, \quad b \text{ is an odd integer.} \tag{1.2}$$

Subsequently several mathematicians attempted to relax condition (1.2), but with limited luck. Much later G. H. Hardy [Har16] was able to remove it:

Theorem 1.1 (Hardy 1916)**.** *For every real number $b \geq a > 1$ the functions*

$$W(t) = \sum_{j=0}^{\infty} a^{-j} \cos(b^j t), \qquad S(t) = \sum_{j=0}^{\infty} a^{-j} \sin(b^j t), \tag{1.3}$$

are bounded and continuous on $\mathbb{R}$, but have no points of differentiability.

The assumption $b \geq a$ here is optimal for every $a > 1$, for W is in $C^1(\mathbb{R})$ whenever $\frac{b}{a} < 1$, due to uniform convergence of the derivatives. (Strangely this was not observed in [Har16].) Hardy also proved that $S'(0) = +\infty$ for

$$1 < a \leq b < 2a - 1, \tag{1.4}$$

2000 *Mathematics Subject Classification.* 26A27.

Key words and phrases. nowhere-differentiability, Weierstrass function, lacunary Fourier series, second microlocalisation.

so that in such cases the graph of $S(t)$ is not entirely rough at $t = 0$ (similarly $W'(\pi/2) = +\infty$ if in addition $b \in 4\mathbb{N}+1$). However, Hardy's treatment is not entirely elementary and yet it fills around 15 pages.

It is perhaps partly for this reason that various attempts have been made over the years to find other examples. These have often involved a replacement of the sine and cosine above by a function with a zig-zag graph, albeit at the price that already the partial sums are not C^1. Indeed, for such functions every $x \in \mathbb{R}$ is a limit $x = \lim_N r_N$ where each $r_N \in \mathbb{Q}$ is a point at which the N^{th} partial sum has no derivatives; whence nowhere-differentiability of the sum function is less startling. Nevertheless, a fine example of this sort was given in just 13 lines by J. McCarthy [McC53].

Somewhat surprisingly, there is an equally short proof of nowhere-differentiability for W and S, using a few basics of integration theory. This is explained here in the rest of the introduction, with a concise proof of Theorem 1.1.

It is a major purpose of this paper to show that the simple method has an easy extension to large classes of nowhere differentiable functions. Thus the main part of the paper contains two general theorems, of which at least the last should be a novelty. It is also shown that the method covers several old and new examples, including some that are rather different from W because of a slow increase of the frequencies.

To present the ideas in a clearer way, one may consider the following function f_θ which (in this paper) serves as a typical nowhere differentiable function,

$$f_\theta(t) = \sum_{j=0}^{\infty} 2^{-j\theta} e^{\mathrm{i}2^j t}, \qquad 0 < \theta \leq 1. \tag{1.5}$$

It is convenient to choose an auxiliary function $\chi\colon \mathbb{R} \to \mathbb{C}$ thus: the Fourier transformed function $\mathscr{F}\chi(\tau) = \hat{\chi}(\tau) = \int_{\mathbb{R}} e^{-\mathrm{i}t\tau}\chi(t)\,dt$ is chosen as a C^∞-function fulfilling

$$\hat{\chi}(\tau) = 0 \text{ for } \tau \notin]\tfrac{1}{2}, 2[, \qquad \hat{\chi}(1) = 1; \tag{1.6}$$

for example by setting

$$\hat{\chi}(\tau) = \exp\Big(2 - \frac{1}{(2-\tau)(\tau-1/2)}\Big) \quad \text{for} \quad \tau \in]\tfrac{1}{2}, 2[. \tag{1.7}$$

Using (1.6) it is easy to show that $\chi(t) = \mathscr{F}^{-1}\hat{\chi}(t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{\mathrm{i}t\tau}\hat{\chi}(\tau)\,d\tau$ is continuous and that $t^k\chi(t)$ is bounded for all $k \in \mathbb{N}_0$. Therefore χ is integrable, ie $\chi \in L_1(\mathbb{R})$, and clearly $\int \chi\,dt = \hat{\chi}(0) = 0$.

With this preparation, the function f_θ is particularly simple to treat, using only ordinary exercises in integration theory: First one may introduce the convolution $\chi * f_\theta$, or rather

$$2^k\chi(2^k\cdot) * f_\theta(t_0) = \int_{\mathbb{R}} 2^k\chi(2^k t) f_\theta(t_0 - t)\,dt, \tag{1.8}$$

which is in $L_1(\mathbb{R})$ since $f_\theta \in L_\infty(\mathbb{R})$ and $\chi \in L_1(\mathbb{R})$. Secondly this will be analysed in two different ways in the proof of

Proposition 1.2. *For $0 < \theta \le 1$ the function $f_\theta(t) = \sum_{j=0}^\infty 2^{-j\theta} e^{\mathrm{i}2^j t}$ is a continuous 2π-periodic, hence bounded function $f_\theta\colon \mathbb{R} \to \mathbb{C}$ without points of differentiability.*

Proof. By uniform convergence f_θ is for $\theta > 0$ a continuous 2π-periodic and bounded function; this follows from Weierstrass' majorant criterion as $\sum 2^{-j\theta} < \infty$.

After insertion of the series defining f_θ into (1.8), Lebesgue's theorem on majorised convergence allows the sum and integral to be interchanged (eg with $\frac{2^k}{1-2^{-\theta}}|\chi(2^k t)|$ as a majorant), which gives

$$\begin{aligned} 2^k\chi(2^k\cdot) * f_\theta(t_0) &= \lim_{N\to\infty} \sum_{j=0}^N 2^{-j\theta} \int_{\mathbb{R}} 2^k \chi(2^k t) e^{\mathrm{i}2^j(t_0-t)}\,dt \\ &= \sum_{j=0}^\infty 2^{-j\theta} e^{\mathrm{i}2^j t_0} \int_{\mathbb{R}} e^{-\mathrm{i}z2^{j-k}} \chi(z)\,dz = 2^{-k\theta} e^{\mathrm{i}2^k t_0} \hat{\chi}(1) = 2^{-k\theta} e^{\mathrm{i}2^k t_0}. \end{aligned} \tag{1.9}$$

Here it was also used that $\hat{\chi}(2^{j-k}) = 1$ for $j = k$ and equals 0 for $j \ne k$.

Moreover, since $f_\theta(t_0)\int_{\mathbb{R}} \chi\,dz = 0$ (cf the note prior to the proposition) this gives

$$2^{-k\theta} e^{\mathrm{i}2^k t_0} = 2^k\chi(2^k\cdot) * f_\theta(t_0) = \int_{\mathbb{R}} \chi(z)(f_\theta(t_0 - 2^{-k}z) - f_\theta(t_0))\,dz. \tag{1.10}$$

So if f_θ were differentiable at t_0, $F(h) := \frac{1}{h}(f_\theta(t_0+h) - f_\theta(t_0))$ would define a function in $C(\mathbb{R}) \cap L_\infty(\mathbb{R})$ for which $F(0) = f'(t_0)$, and Lebesgue's theorem, applied with $|z\chi(z)|\sup_{\mathbb{R}}|F|$ as the majorant, would imply that

$$-2^{(1-\theta)k} e^{\mathrm{i}2^k t_0} = \int F(-2^{-k}z) z\chi(z)\,dz \xrightarrow[k\to\infty]{} f'(t_0)\int_{\mathbb{R}} z\chi(z)\,dz = f'(t_0)\,\mathrm{i}\frac{d\hat{\chi}}{d\tau}(0) = 0; \tag{1.11}$$

hence that $1 - \theta < 0$. This would contradict the assumption that $\theta \le 1$. □

By now this argument is of course of a classical nature. But it is seemingly not well established in the literature. Eg, recently R. Shakarchi and E. M. Stein treated the nowhere-differentiability of f_θ in Thm. 3.1 of Chap. 1 in their treatise [SS03] with a method they described thus: "The proof of the theorem is really the story of three methods of summing a Fourier series…partial sums…Cesaro summability…delayed means." However, they cover $0 < \theta < 1$ in a few pages, while the case $\theta = 1$ relies on refinements sketched in the page-long Problem 5.8 based on the Poisson summation formula.

The present proofs are rather concise, covering all cases at once. They are also not confined to periodic functions (cf the next section), for the theory of lacunary Fourier series is here simply replaced by the Fourier transformation $\mathscr{F}$ and its basic properties.

Moreover, also Hardy's theorem can be derived in this way. The main point is to keep the factor $e^{\mathrm{i}2^k t_0}$ instead of introducing $\cos(2^k t_0)$ and $\sin(2^k t_0)$, that do not necessarily stay away from 0 as $k \to \infty$ (one of the difficulties dealt with in [Har16]). Indeed, with a few modifications one obtains a

Proof of Theorem 1.1. As $a > 1$, clearly $W \in C(\mathbb{R}) \cap L_\infty$ with $|W(t)| \le \sum_{j=0}^{\infty} a^{-j} = \frac{a}{a-1}$.

Since $b > 1$ it may in this proof be arranged that $\overset{\wedge}{\chi}(1) = 1$ and $\overset{\wedge}{\chi}(\tau) \ne 0$ only for $\frac{1}{b} < \tau < b$. In the same way as for f_θ this gives, by Euler's formula,

$$b^k \chi(b^k \cdot) * W(t_0) = \sum_{j=0}^{\infty} a^{-j} \int_{\mathbb{R}} b^k \chi(b^k t) \tfrac{1}{2}(e^{\mathrm{i} b^j (t_0 - t)} + e^{\mathrm{i} b^j (t - t_0)})\, dt. \tag{1.12}$$

The term $e^{\mathrm{i} b^j (t-t_0)}$ is redundant here, for $z := t b^k$ yields $\int e^{\mathrm{i} b^j t} \chi(b^k t) b^k \, dt = \overset{\wedge}{\chi}(-b^{j-k}) = 0$, as $\overset{\wedge}{\chi}$ vanishes on $]-\infty, 0]$. So as in (1.9), one has $b^k \chi(b^k \cdot) * W(t_0) = \frac{e^{\mathrm{i} b^k t_0}}{2a^k}$.

Hence existence of $W'(t_0)$ would imply that $\lim_k (\frac{b}{a})^k e^{\mathrm{i} b^k t_0} = 0$; cf (1.10)–(1.11). This would contradict that $b \ge a$, so W is nowhere differentiable. Similarly $S(t)$ is so. $\square$

It is known that Hölder regularity and differentiability of W can be analysed with the theory of wavelets; cf [Hol95]. But the above proofs have the advantage of relying only on "first principles".

In Section 2 the proofs are reinforced using the Fourier transformation consistently, leading to a general result on nowhere differentiable functions. Refining a dilation argument, a further extension is found in Section 3, where functions with polynomial growth of the frequencies are covered. Examples with much slower growth are given in Section 4.

Remark 1.3. By a well-known reasoning, nowhere-differentiability of W is obtained because the j^{th} term cannot cancel the oscillations of the previous ones: it is out of phase with previous terms since $b > 1$, and moreover the amplitudes decay exponentially since $\frac{1}{a} < 1$; as $b \ge a > 1$ the combined effect is large enough (vindicated by the optimality of $b \ge a$ noted after Theorem 1.1). However, it will be clear from Section 4 below that frequencies growing almost quadratically suffice for nowhere-differentiability; cf Remark 4.4.

Remark 1.4. Inspection of the proof of Theorem 1.1 shows that Lebesgue's theorem on majorised convergence is the most advanced part. As this result appeared in 1908, cf [Leb08, p. 12], it seems that the argument could have been written down a century ago.

2. Proof by microlocalisation

Although the proofs above are short, it is useful to rephrase them with a few facts from the distribution theory of L. Schwartz (cf [Sch66]). As shown here, this gives a better insight and leads directly to a general result.

Recall that the Fourier transformation $\mathscr{F}f(\tau)=\hat{f}(\tau)=\int_{\mathbb{R}} e^{-\mathrm{i}t\tau}f(t)\,dt$ has a well-known extension to the space $\mathscr{S}'(\mathbb{R})$ of so-called tempered distributions. In particular it applies to exponential functions $e^{\mathrm{i}bt}$, and as a basic exercise this yields 2π times Dirac's delta measure δ_b, ie the point measure at $\tau=b$,

$$\mathscr{F}(e^{\mathrm{i}b\cdot})(\tau)=2\pi\delta(\tau-b)=2\pi\delta_b(\tau). \tag{2.1}$$

For $f_\theta(t)=\sum_{j=0}^{\infty}2^{-j\theta}e^{\mathrm{i}2^jt}$ the Fourier transformation gives, cf (2.1),

$$\mathscr{F}f_\theta=\sum_{j=0}^{\infty}2^{-j\theta}\mathscr{F}(e^{\mathrm{i}2^j\cdot})=2\pi\sum_{j=0}^{\infty}2^{-j\theta}\delta_{2^j}. \tag{2.2}$$

($\mathscr{F}$ applies termwisely because it is continuous on the space $\mathscr{S}'(\mathbb{R})$, and the right-hand side converges there for every $\theta\in\mathbb{R}$.) The formula (2.2) just expresses the intuitively obvious fact that f_θ is synthesized precisely from the frequencies $1,2,\dots,2^j,\dots$.

Each of the frequencies may be picked out in a well-known way: In the space $\mathscr{S}(\mathbb{R})$ of rapidly decreasing C^∞-functions one can choose χ such that $\mathscr{F}\chi(1)=1$ whereas $\mathscr{F}\chi(\tau)=0$ for $\tau\notin]\frac{2}{3},\frac{3}{2}[$. Clearly $\hat{\chi}(2^{-j}\cdot)\neq0$ only on $[\frac{2}{3}2^j,\frac{3}{2}2^j]$, so

$$\hat{\chi}(2^{-k}\cdot)\delta_{2^j}=\begin{cases}0 & \text{for } j\neq k\\ \delta_{2^k} & \text{for } j=k.\end{cases} \tag{2.3}$$

The general rule $\mathscr{F}(\chi*f)=\hat{\chi}\cdot\hat{f}$ applies to $\chi\in\mathscr{S}(\mathbb{R})$ and $f_\theta\in L_\infty\subset\mathscr{S}'(\mathbb{R})$, whence

$$\mathscr{F}(2^k\chi(2^k\cdot)*f_\theta)=\hat{\chi}(2^{-k}\cdot)\cdot\mathscr{F}f_\theta=2\pi\sum_{j=0}^{\infty}2^{-j\theta}\hat{\chi}(2^{-k}\cdot)\delta_{2^k}=2\pi2^{-k\theta}\delta_{2^k}. \tag{2.4}$$

So by use of $\mathscr{F}^{-1}$ and (2.1),

$$2^k\chi(2^k\cdot)*f_\theta(t)=2^{-k\theta}e^{\mathrm{i}2^kt}. \tag{2.5}$$

This formula coincides with (1.9), but from the above procedure it is clear that the convolution precisely gives the part of f_θ having frequency 2^k.

Remark 2.1. The process in (2.4)–(2.5) has of course been known for ages, but with the distribution theory it is fully justified although $\mathscr{F} f_\theta$ consists of measures. In principle, it is a so-called second microlocalisation of f_θ, since $\hat{\chi}(2^{-k}\tau)\mathscr{F} f_\theta(\tau)$ is localised to frequencies τ restricted in both size and direction; ie $|\tau| \approx 2^k$ and $\tau > 0$.

The second microlocalisation is more visible in a separate treatment of

$$\operatorname{Re} f_\theta(t) = \sum_{j=0}^{\infty} 2^{-j\theta} \cos(2^j t), \qquad \operatorname{Im} f_\theta(t) = \sum_{j=0}^{\infty} 2^{-j\theta} \sin(2^j t). \tag{2.6}$$

They are continuous and 2π-periodic like f_θ, and also both without derivatives anywhere. Indeed, by Euler's formula and (2.1),

$$\mathscr{F} \cos(2^j\cdot) = \tfrac{2\pi}{2}(\delta_{2^j} + \delta_{-2^j}), \qquad \mathscr{F} \sin(2^j\cdot) = \tfrac{2\pi}{2\mathrm{i}}(\delta_{2^j} - \delta_{-2^j}). \tag{2.7}$$

Here multiplication by $\hat{\chi}(2^{-j}\cdot)$ *removes* the contribution from δ_{-2^j}. Therefore one can replace f_θ in (1.10) by $\operatorname{Re} f_\theta$ and $\operatorname{Im} f_\theta$ if only $2^{-k\theta}$ is replaced by $\frac{1}{2}2^{-k\theta}$ and $\frac{1}{2\mathrm{i}}2^{-k\theta}$, respectively. The rest of the proof is similar.

It is not difficult to carry this observation over to the context of Theorem 1.1, which will explain why the proof of this was saved by the redundancy of the term $e^{\mathrm{i}b^j(t-t_0)}$.

However, one can just as well pass to the following result on possibly non-periodic functions of the form $f(t) = \sum a_j \exp(\mathrm{i} b_j t)$, with more general sequences (a_j) of amplitudes and frequencies (b_j). The latter are assumed to be monotone increasing, ie $b_1 < b_2 < \dots < b_j < \dots$ Throughout it is assumed that $b_j > 0$ for all j, for a finite number of negative frequencies would only contribute a C^∞-term. For short this is written $0 < b_j \nearrow b$, where $b = \lim_j b_j$ belongs to $\mathbb{R} \cup \{\infty\}$.

Recall that $f\colon \mathbb{R} \to \mathbb{C}$ is said to be Lipschitz continuous at t_0 if there exist two constants $L > 0$, $\eta > 0$ such that $|f(t) - f(t_0)| \le L|t - t_0|$ for every $t \in]t_0 - \eta, t_0 + \eta[$.

Theorem 2.2. *Let $f\colon \mathbb{R} \to \mathbb{C}$ be given as*

$$f(t) = \sum_{j=0}^{\infty} a_j \exp(\mathrm{i} b_j t) \tag{2.8}$$

by means of a complex sequence $(a_j)_{j\in\mathbb{N}_0}$ and a real sequence $(b_j)_{j\in\mathbb{N}_0}$ fulfilling

$$\sum_{j=0}^{\infty} |a_j| < \infty, \qquad 0 < b_j \nearrow \infty, \tag{2.9}$$

$$\liminf_{j\to\infty} \frac{b_{j+1}}{b_j} > 1, \qquad a_j b_j \not\to 0 \quad \text{for} \quad j \to \infty. \tag{2.10}$$

Then f is bounded and continuous on $\mathbb{R}$, but has no points of differentiability. If in addition $\sup_k |a_k| b_k = \infty$, then f is moreover not Lipschitz continuous at any point. The same conclusions are valid for $\operatorname{Re} f$ *and* $\operatorname{Im} f$.

Clearly $\limsup |a_j| b_j > 0$ is equivalent to $a_j b_j \not\to 0$; cf (2.10). The latter is a natural condition because termwise differentiation yields $\sum a_j b_j e^{\mathrm{i} b_j t}$, which cannot converge unless $a_j b_j \to 0$.

Proof. Using (2.9) to invoke the majorant criterion, f is easily seen to be in $C(\mathbb{R}) \cap L_\infty(\mathbb{R})$. Hence $f \in \mathscr{S}'(\mathbb{R})$, ie f is a tempered distribution on $\mathbb{R}$, and

$$\mathscr{F} f(\tau) = \sum_{j=0}^{\infty} a_j \mathscr{F}(e^{\mathrm{i} b_j \cdot}) = 2\pi \sum_{j=0}^{\infty} a_j \delta_{b_j}(\tau). \tag{2.11}$$

In view of (2.10) one can fix $\lambda \in]1, \liminf \frac{b_{j+1}}{b_j}[$ so that $b_{k+1} > \lambda b_k$ for all $k \geq K$, if K is chosen appropriately. Taking $\overset{\wedge}{\chi} \in C^\infty(\mathbb{R})$ such that $\overset{\wedge}{\chi}(1) = 1$ and $\overset{\wedge}{\chi}(\tau) \neq 0$ only holds for $\lambda^{-1} < \tau < \lambda$, then $\overset{\wedge}{\chi}(\tau / b_k) \neq 0$ only for $\tau \in]\frac{b_k}{\lambda}; \lambda b_k[$.

Because $[\frac{b_k}{\lambda}; \lambda b_k] \subset]b_{k-1}; b_{k+1}[$ and the sequence (b_k) is monotone increasing, clearly it holds that $\overset{\wedge}{\chi}(\tau / b_k) \delta_{b_j}(\tau) = 0$ for all $j \neq k$. Thence

$$\mathscr{F}(b_k \chi(b_k \cdot) * f)(\tau) = \overset{\wedge}{\chi}(\tau / b_k) \hat{f}(\tau) = 2\pi a_k \delta_{b_k}(\tau). \tag{2.12}$$

By inverse Fourier transforming this,

$$a_k e^{\mathrm{i} b_k t_0} = b_k \chi(b_k \cdot) * f(t_0) = \int_{\mathbb{R}} \chi(z) f(t_0 - z/b_k)\, dz. \tag{2.13}$$

If f were differentiable at t_0, then $F(t) = (f(t_0 + t) - f(t_0))/t$ would be in L_∞ (like f), so since $\int \chi(t)\, dt = 0$, majorised convergence would imply

$$-a_k b_k e^{\mathrm{i} b_k t_0} = \int_{\mathbb{R}} z \chi(z) \frac{f(t_0 - z/b_k) - f(t_0)}{-z/b_k}\, dz \xrightarrow[k \to \infty]{} f'(t_0)\, \mathrm{i} \frac{d\overset{\wedge}{\chi}}{d\tau}(0) = 0. \tag{2.14}$$

This would entail $|a_k| b_k \to 0$ for $k \to \infty$, in contradiction of (2.10).

In addition, if f were Lipschitz continuous at t_0, then again F would be bounded since $f \in L_\infty$, so the integral in (2.14) would be uniformly bounded with respect to k, in which case $\sup_k |a_k| b_k < \infty$.

Finally, using (2.7) it is easy to see that one can replace f in (2.12) by $\operatorname{Re} f$ or $\operatorname{Im} f$ if only a_k is replaced by $a_k/2$ and $a_k/(2\mathrm{i})$, respectively. Eg

$$\overset{\wedge}{\chi}(\cdot / b_k) \mathscr{F} \operatorname{Re} f = 2\pi \sum_{j=0}^{\infty} \overset{\wedge}{\chi}(\cdot / b_k) (\frac{\operatorname{Re} a_j}{2}(\delta_{b_j} + \delta_{-b_j}) - \frac{\operatorname{Im} a_j}{2\mathrm{i}}(\delta_{b_j} - \delta_{-b_j})) = 2\pi \frac{a_k}{2} \delta_{b_k}. \tag{2.15}$$

Proceeding as for f itself, it follows that neither $\operatorname{Re} f$ nor $\operatorname{Im} f$ can be differentiable at some $t_0 \in \mathbb{R}$, respectively Lipschitz continuous if $\sup_j |a_j| b_j = \infty$. □

Remark 2.3. A necessary condition for Hölder continuity of order $\alpha \in]0, 1[$ follows at once from a modification of the above argument: replacing $a_k b_k$ on the left-hand side of (2.14) by $a_k b_k^\alpha$, the resulting integral will be uniformly bounded with respect to k since $\int |z|^\alpha |\chi(z)|\,dz < \infty$. Hence

$$\sup_k |a_k| b_k^\alpha < \infty \tag{2.16}$$

whenever $f(t)$ in Theorem 2.2 is Hölder continuous of order α at a single point t_0.

Example 2.4. Sequences of power type like $a_j = a^{-j}$ and $b_j = b^j$ for parameters $b \geq |a| > 1$ give $f(t) = \sum_{j=0}^\infty a^{-j} e^{\mathrm{i} b^j t}$, which is covered by Theorem 2.2 as $|a_j| b_j = |\frac{b}{a}|^j \geq 1$ and $\frac{b_{j+1}}{b_j} = b > 1$. Therefore Theorem 2.2 contains Theorem 1.1 and extends it to complex amplitudes.

For $W(t)$ Remark 2.3 reduces to $\frac{b^\alpha}{a} \leq 1$, hence to $\alpha \leq \frac{\log a}{\log b}$. In case $b > a > 1$ it is known from [Har16, p. 311] that W is globally Hölder continuous of order $\alpha = \log a / \log b$; whereas for $b = a > 1$ it was only obtained there that $W(t+h) - W(t) = O(|h| \log 1/|h|)$. So Remark 2.3 at once gives a sharp upper bound for the Hölder exponent of W (mentioned as a difficult task in [Jaf97]). The reader may also consult Theorem 4.9 in Ch. II of Zygmund's book [Zyg59] for the Hölder continuity of W.

Example 2.5. In the same way, Theorem 2.2 also covers Darboux' function

$$f(t) = \sum_{j=0}^\infty \frac{\sin((j+1)!t)}{j!}, \tag{2.17}$$

for $a_j = 1/j!$ and $b_j = (j+1)!$ fulfil in particular $\frac{b_{j+1}}{b_j} = j+2 \nearrow \infty$ and $a_j b_j = j+1 \nearrow \infty$.

Example 2.6. Setting $a_j = a^{-j}$ for some $a > 1$ and defining (b_j) by $b_{2m} = a^{2m}$ and $b_{2m+1} = (1+a^{-p})a^{2m}$, it is seen directly that when the power p is so large that $1 + a^{-p} < a^2$, then the sequences (a_j) and (b_j) fulfil the conditions of Theorem 2.2. Eg (2.10) holds as $\frac{b_{j+1}}{b_j} \in \{1 + a^{-p}, a^2(1+a^{-p})^{-1}\} \subset]1,\infty[$ and $a_j b_j \in \{1, (1+a^{-p})/a\}$. Thus $f(t)$ is nowhere differentiable in this case. If further p is so large that $1 + a^{-2} < a^p(1-a^{-2})$ it is easily verified that $b_{2m+1} - b_{2m} < b_{2m} - b_{2m-1}$ so that $(b_{j+1} - b_j)$ is not monotone increasing. Eg if $a = 5$, both requirements are met by $p = 1$ and the values of (b_j) are

$$1, \tfrac{6}{5}, 25, 30, 625, 750, 15625, 18750, \dots$$

Clearly these frequencies have a distribution with lacunas of rather uneven size.

To shed light on the assumptions in Theorem 2.2, note that for every sequence (b_j) of positive reals,

$$\liminf \frac{b_{j+1}}{b_j} > 1 \iff \exists \varepsilon > 0\colon\ \liminf \frac{b_{j+1} - b_j}{b_j} = \varepsilon \tag{2.18}$$

$$\iff \exists \varepsilon > 0\ \exists J \in \mathbb{N}\ \forall j > J\colon\ \varepsilon b_j < b_{j+1} - b_j < \tfrac{2+\varepsilon}{1+\varepsilon} b_{j+1}. \tag{2.19}$$

Clearly (2.19) gives a control over the spectral gaps $b_{j+1} - b_j$; by the equivalence the requirement $\liminf \frac{b_{j+1}}{b_j} > 1$ in Theorem 2.2 may therefore be seen as a *spectral gap* condition. Since $b_j \nearrow \infty$, the last line shows that $b_{j+1} - b_j \to \infty$ in all cases covered by Theorem 2.2; but the spectral gaps need not be monotone increasing, cf Example 2.6.

3. Dilation by differences

The spectral gaps were used in Theorem 2.2 to ensure that the cut-off function $\hat{\chi}(\tau/b_k)$ would only yield the single frequency b_k. This rules out cases in which b_k is of a much larger order of magnitude than $b_{k+1} - b_k$. This may also be seen from (2.18).

This drawback is felt in case of polynomial growth such as $a_j = j^{-p}$ and $b_j = j^q$, for since $\lim \frac{b_{j+1}}{b_j} = \lim(1 + \frac{1}{j})^q = 1$, Theorem 2.2 does not apply.

But it should suffice to dilate a cut-off function just by the gap $b_{j+1} - b_j$, or by $b_j - b_{j-1}$ if this is the smaller gap at b_j, at least if $\lim(b_{j+1} - b_j) = \infty$, which could be assumed (replacing the spectral gap condition by one of its consequences, cf (2.19)). For a subsequence, this is done implicitly in condition (3.3) below:

Theorem 3.1. *Let $f\colon \mathbb{R} \to \mathbb{C}$ be given as*

$$f(t) = \sum_{j=0}^{\infty} a_j \exp(\mathrm{i} b_j t) \tag{3.1}$$

for a complex sequence (a_j) for which $\sum_{j=0}^{\infty} |a_j| < \infty$ and a real sequence $0 < b_j \nearrow \infty$. If

$$\Delta b_j = \min(b_j - b_{j-1}, b_{j+1} - b_j) \tag{3.2}$$

has the property that

$$a_j \Delta b_j \not\to 0 \quad \text{for } j \to \infty, \tag{3.3}$$

then f is bounded and continuous on $\mathbb{R}$, but f is nowhere differentiable. Moreover, f is not Lipschitz continuous at any $t_0 \in \mathbb{R}$ when $\sup_j |a_j| \Delta b_j = \infty$. The conclusions are also valid for $\operatorname{Re} f$ *and* $\operatorname{Im} f$.

Proof. $f \in C(\mathbb{R}) \cap L_\infty(\mathbb{R})$ is shown as in Theorem 2.2. Let now $\mathscr{F}\psi \in C^\infty(\mathbb{R})$ fulfil $\mathscr{F}\psi(0) = 1$ and $\mathscr{F}\psi(\tau) \neq 0$ only for $|\tau| < 1/2$, and take the spectral cut-off function as

$$\hat{\psi}_k(\tau) = \hat{\psi}\Big(\frac{\tau - b_k}{\Delta b_k}\Big). \tag{3.4}$$

Then the definition of Δb_k as a minimum entails

$$\hat{\psi}_k(\tau) \neq 0 \implies b_k - \tfrac{1}{2}(b_k - b_{k-1}) < \tau < b_k + \tfrac{1}{2}(b_{k+1} - b_k). \tag{3.5}$$

Since (b_j) is increasing, the τ-interval specified here only contains b_j for $j=k$, whence

$$\overset{\wedge}{\psi}_k(\tau)\hat{f}(\tau) = 2\pi \sum_{j=0}^{\infty} a_j \overset{\wedge}{\psi}_k(\tau)\delta_{b_j}(\tau) = 2\pi a_k \delta_{b_k}(\tau). \tag{3.6}$$

Note that by a change of variables,

$$\mathscr{F}^{-1}\overset{\wedge}{\psi}_k(t) = \tfrac{1}{2\pi}\int_{\mathbb{R}} e^{\mathrm{i}t(b_k+\sigma\Delta b_k)}\overset{\wedge}{\psi}(\sigma)\Delta b_k\,d\sigma = (\Delta b_k)e^{\mathrm{i}tb_k}\psi(t\Delta b_k). \tag{3.7}$$

Since the integral of the right-hand side is 0 by (3.5), application of $\mathscr{F}^{-1}$ to (3.6) gives,

$$\begin{aligned} a_k(\Delta b_k)e^{\mathrm{i}b_k t_0} &= (\Delta b_k) f * \psi_k(t_0) \\ &= \int_{\mathbb{R}} (f(t_0-t)-f(t_0))(\Delta b_k)^2 e^{\mathrm{i}b_k t}\psi(t\Delta b_k)\,dt. \\ &= \int_{\mathbb{R}} \frac{f(t_0-z/\Delta b_k)-f(t_0)}{z/\Delta b_k} z\psi(z)e^{\mathrm{i}z\frac{b_k}{\Delta b_k}}\,dz. \end{aligned} \tag{3.8}$$

In case f is Lipschitz continuous at t_0 one has again $F\in L_\infty$, so this yields that for some suitably large $L\in\mathbb{R}$,

$$\sup_k |a_k|\Delta b_k \leq \sup_k \int_{\mathbb{R}} \left|\frac{f(t_0-z/\Delta b_k)-f(t_0)}{z/\Delta b_k}\right| |z\psi(z)|\,dz \leq L\int_{\mathbb{R}} |z\psi(z)|\,dz < \infty. \tag{3.9}$$

Moreover, because $b_k/\Delta b_k \geq b_k/(b_k-b_{k-1}) > 1$,

$$\int_{\mathbb{R}} z\psi(z)e^{\mathrm{i}z\frac{b_k}{\Delta b_k}}\,dz = \mathrm{i}\frac{d\overset{\wedge}{\psi}}{d\tau}(-\frac{b_k}{\Delta b_k}) = 0. \tag{3.10}$$

So were f differentiable at t_0, it would follow from (3.8) by majorised convergence that

$$a_k(\Delta b_k)e^{\mathrm{i}t_0 b_k} = -\int_{\mathbb{R}} \Big(\frac{f(t_0-z/\Delta b_k)-f(t_0)}{-z/\Delta b_k} - f'(t_0)\Big) z\psi(z)e^{\mathrm{i}z\frac{b_k}{\Delta b_k}}\,dz \xrightarrow[k\to\infty]{} 0, \tag{3.11}$$

in contradiction of (3.3). Finally the same arguments apply to $\operatorname{Re} f$, $\operatorname{Im} f$ by dividing a_k by 2 and $2\mathrm{i}$, respectively, as in the previous theorem. □

Remark 3.2. It is immediate from (3.8) that when $f(t)$ in Theorem 3.1 is Hölder continuous of order $\alpha\in]0,1[$ at some t_0, then

$$\sup_j |a_j|(\Delta b_j)^\alpha < \infty. \tag{3.12}$$

When applied to W, this condition gives the same result as Remark 2.3, for $\Delta b_j = cb^j$ with $c = 1-1/b > 0$ when $b>1$. Consequently the gap growth condition (3.12) cannot be sharpened in general.

Seemingly, nowhere differentiability has not been obtained under the weak assumptions of Theorem 3.1 before. In the first application of this result it is seen, like for f_θ and W, that the regularity of the sum function improves when the growth of the frequencies is taken smaller, here by reducing q:

Example 3.3 (Polynomial growth)**.** For $p>1$ one has uniformly continuous functions

$$f_{p,q}(t)=\sum_{j=1}^{\infty}\frac{\exp(\mathrm{i}tj^q)}{j^p},\qquad \operatorname{Re}f_{p,q}(t),\qquad \operatorname{Im}f_{p,q}(t),\tag{3.13}$$

that moreover are C^1 and bounded with bounded derivatives on $\mathbb{R}$ in case $q<p-1$. However, for $q\geq p+1$ they are nowhere differentiable according to Theorem 3.1: (3.3) follows since by the mean value theorem the spectral gaps increase, and

$$\limsup j^{-p}(j^q-(j-1)^q)\geq \limsup qj^{q-p-1}(1-1/j)^{q-1}=\begin{cases} q & \text{for } q=p+1,\\ \infty & \text{for } q>p+1.\end{cases}\tag{3.14}$$

Moreover, for $q>p+1$ there is not Lipschitz continuity at any point.

But the functions in (3.13) are globally Hölder continuous of order $\alpha=(p-1)/q$ if only $q\geq p-1$. This results from the splitting

$$|f_{p,q}(t+h)-f_{p,q}(t)|\leq \sum_{j\leq N} j^{q-p}|h|+\sum_{j>N}2j^{-p}\leq N^{q-p+1}|h|+\frac{2}{p-1}N^{1-p}.\tag{3.15}$$

For $0<|h|\leq\frac12$ this is exploited for the unique N such that

$$N\leq |h|^{-1/q}<N+1.\tag{3.16}$$

For the Hölder exponents, this is optimal among the powers $|h|^{-\theta}$, for clearly $\theta=1/q$ is the value that maximises

$$\min(\theta(p-1),1-\theta(q-p+1)).\tag{3.17}$$

Using (3.16) in (3.15) gives a $C<\infty$ so that for $|h|\leq 1/2$,

$$|f_{p,q}(t+h)-f_{p,q}(t)|\leq C|h|^{\alpha},\qquad \alpha=\frac{p-1}{q}.\tag{3.18}$$

As $f_{p,q}\in L_\infty$, this holds for all h,t if C is sufficiently large.

The inequality $\Delta b_j<qj^{q-1}$ shows that the necessary condition in Remark 3.2 is fulfilled for $\alpha(q-1)-p\leq 0$, which only gives the upper bound $\alpha\leq\frac{p}{q-1}$. So in view of (3.18) there remains a gap for these functions.

The progression of the frequencies in Example 3.3 is clearly slower than in Example 2.6, and in this sense Theorem 3.1 improves Theorem 2.2 a good deal. The condition $|a_j|\Delta b_j\not\to 0$ in (3.3) cannot be sharpened in general, for already for W it amounts to $b\geq a$, that is equivalent to nowhere-differentiability.

However, (3.3) does not give optimal results for $f_{p,q}$. Eg the case with $p = q = 2$ has been completely clarified and shown to have a delicate nature, as it is known from several investigations that the so-called Riemann function

$$R(t) = \sum_{j=1}^{\infty} \frac{\sin(\pi j^2 t)}{j^2} \tag{3.19}$$

is differentiable with $R'(t) = -1/2$ exactly at $t = r/s$ for odd integers r, s. For properties of this function the reader is referred to the paper of J. Duistermaat [Dui91].

As $f_{p,q}$ is in $C^1(\mathbb{R})$ for every $q < p-1$ when $p > 1$, transition to nowhere-differentiability occurs (perhaps gradually) as q runs through the interval $[p-1, p+1[$. Nowhere-differentiability for $q \geq p+1$ was also mentioned for $\operatorname{Im} f_{p,q}$ by W. Luther [Lut86] as an outcome of a very general Tauberian theorem. (In addition $\operatorname{Im} f_{p,2}$ was covered with nowhere-differentiability for $p \leq 3/2$ providing cases in $[p+\frac{1}{2}, p+1[$; for t irrational Luther's result relied on Hardy's investigation [Har16], that covered $\operatorname{Im} f_{p,2}$ for $p < 5/2$ thus giving cases of almost nowhere differentiability in $]p-\frac{1}{2}, p+\frac{1}{2}]$ for $q = 2$.)

From (3.18) it is seen that $R(t)$ is globally Hölder continuous of order $\alpha = 1/2$, which is well known; cf [Dui91]. At the points of differentiability, this is of course not optimal, but it is also known that the local Hölder regularity of R attains every value $\alpha \in [\frac{1}{2}, \frac{3}{4}]$ in a non-empty set; cf the paper of S. Jaffard [Jaf97].

In view of this, it is envisaged that the global Hölder exponent $\alpha = \frac{p-1}{q}$ is optimal in Example 3.3 whenever $q \geq p-1$, $p > 1$.

Remark 3.4. $f_{p,q}$ was recently studied by F. Chamizo and A. Ubis for $q \in \mathbb{N}$, $p > 1$. In [CU07, Prop. 3.3] they treated nowhere-differentiability by convolving $f_{p,q}$ with the Fejér kernel; this method was proposed as an alternative to those of [Lut86], and it is similar in spirit to the above proof of Theorem 3.1. However, other statements are flawed. Eg, in [CU07, Thm. 3.1], $f_{p,q}$ is claimed differentiable at an irreducible fraction $r/s \in \mathbb{Q}$, $s > 0$, if and only if both $q < p+1/2$ and q divides $\gamma - 1$ but is relatively prime with $\sigma - 1$ for some maximal prime power σ^γ in the factorisation of s. But as noted above $f_{p,q}$ is C^1 for every $q < p-1$, and for the cases with $q \in \mathbb{N} \cap [2, p-1]$ the condition on q and the prime factors of s is violated even in a dense subset of $\mathbb{R}$ (eg where $2 \leq \gamma \leq q$ for each prime factor of s), so the claim is not correct for such q.

4. Slowly growing frequencies

As a result of Theorem 3.1, nowhere-differentiability follows in several new cases where the ratio b_j/j^2 can have arbitrarily slow growth.

This will be clear from the examples of this section. They all relate to the limiting case $p = 1$, $q = 2$ in Example 3.3.

Example 4.1. Setting $\log^a t = (\log t)^a$ for $a \in \mathbb{R}$ and $t > 1$, the functions

$$F_1(t) = \sum_{j=2}^{\infty} \frac{\exp(\mathrm{i}tj^2 \log^b j)}{j \log^a j}, \qquad \operatorname{Re} F_1(t), \qquad \operatorname{Im} F_1(t) \tag{4.1}$$

are for $b \geq a > 1$ continuous, bounded and nowhere differentiable on $\mathbb{R}$; for $b > a > 1$ not even Lipschitz continuous at any point.

Indeed, the auxiliary function $\varphi(t) = t^2 \log^b t$ has derivative $\varphi'(t) = t \log^{b-1} t(b + 2\log t)$; since this is increasing, so is $\varphi(t) - \varphi(t-1)$. (Actually φ is convex as $\varphi'' > 0$.) This gives

$$\begin{aligned} a_j \Delta b_j &= a_j(b_j - b_{j-1}) \\ &= (j \log^a j)^{-1}(j^2 \log^b j - j^2 \log^b (j-1) + 2j \log^b (j-1) - \log^b (j-1)) \\ &= j \log^{b-a} j(1 - [\frac{\log j(1-1/j)}{\log j}]^b) + 2\log^{b-a} j(1 + \frac{\log(1-1/j)}{\log j})^b - \frac{\log^b (j-1)}{j \log^a j}. \end{aligned} \tag{4.2}$$

For the square bracket the binomial series gives

$$\begin{aligned} [1 + \frac{\log(1-1/j)}{\log j}]^b &= [1 + \frac{-j^{-1} + O(j^{-2})}{\log j}]^b \\ &= 1 + b\frac{-j^{-1} + O(j^{-2})}{\log j} + O(\frac{-j^{-1} + O(j^{-2})}{\log j})^2 = 1 - \frac{b}{j \log j} + O(j^{-2}). \end{aligned} \tag{4.3}$$

Hence

$$a_j \Delta b_j = b \log^{b-a-1} j + 2(1 + o(1)) \log^{b-a} j + o(1) \xrightarrow[j \to \infty]{} \begin{cases} 2 & \text{for } \ b = a, \\ \infty & \text{for } \ b > a. \end{cases} \tag{4.4}$$

Therefore Theorem 3.1 yields the stated properties of F_1. Taking eg $a = b = 2$ the frequencies are (with decimals rounded off) 2, 11, 31, 65, 116, 186, 277, 391, ...

To avoid repetition of lengthy details as above, one can adopt the next result, that also shows that many examples arise simply by taking the frequencies as $b_j = j/|a_j|$:

Corollary 4.2. *When $\sum_{j>J} |a_j| < \infty$ for a sequence with $|a_j| \geq |a_{j+1}|$ for all $j > J$ and there is a convex function $\varphi\colon\ ]J,\infty[\to \mathbb{R}$ such that $\varphi(j) = j/|a_j|$ for every $j \in \mathbb{N} \cap]J,\infty[$, then*

$$f(t) = \sum_{j>J} a_j \exp(\mathrm{i}tj/|a_j|) \tag{4.5}$$

is a bounded continuous function on $\mathbb{R}$, for which f, $\operatorname{Re} f$, $\operatorname{Im} f$ are nowhere differentiable.

Proof. To analyse Δb_j, note that

$$\tfrac{j}{|a_j|} - \tfrac{j-1}{|a_{j-1}|} \leq \tfrac{j+1}{|a_{j+1}|} - \tfrac{j}{|a_j|} \iff \tfrac{j}{|a_j|} \leq \tfrac{1}{2}(\tfrac{j+1}{|a_{j+1}|} + \tfrac{j-1}{|a_{j-1}|}); \tag{4.6}$$

the last inequality holds true by the assumptions on φ. Therefore

$$|a_j| \Delta b_j = |a_j|(\tfrac{j}{|a_j|} - \tfrac{j-1}{|a_{j-1}|}) = \tfrac{|a_j|}{|a_{j-1}|} + j(1 - \tfrac{|a_j|}{|a_{j-1}|}). \tag{4.7}$$

If $\lambda := \limsup \frac{|a_j|}{|a_{j-1}|}$ fulfils $0 < \lambda \le 1$, one has $\limsup |a_j|\Delta b_j \ge \lambda > 0$ because $|a_j \Delta b_j| \ge \frac{|a_j|}{|a_{j-1}|}$. Otherwise $\lambda = 0$, but since $|a_j|\Delta b_j \ge j(1-1/2)$ eventually, $\limsup |a_j|\Delta b_j = \infty$. In both cases Theorem 3.1 yields the claim, for the $b_j = j/|a_j|$ are strictly increasing. □

The corollary applies to the example preceeding it for $b = a > 1$, for the function $\varphi(t) = t^2 \log^a t$ has for all $t > 1$ the derivative

$$\varphi'(t) = t(a + 2\log t)\log^{a-1} t, \tag{4.8}$$

which is increasing for $a > 1$, so that φ is convex. Clearly $\varphi(j) = j/(j\log^a j)^{-1}$ for $j > 1$.

A slower growth is obtained as follows:

Example 4.3. Setting $\log_{\circ 2}{}^a t = (\log\log t)^a$ for $a \in \mathbb{R}$, $t > e$, there is a bounded, continuous but nowhere differentiable function given by

$$F_2(t) = \sum_{j=3}^{\infty} \frac{\exp(\mathrm{i} t j^2 \log j \log_{\circ 2}{}^a j)}{j \log j \log_{\circ 2}{}^a j}, \qquad a > 1. \tag{4.9}$$

Indeed, $\sum (j \log j \log_{\circ 2}{}^a j)^{-1} < \infty$ by the integral criterion since the 'primitive' function $(1-a)^{-1} \log_{\circ 2}{}^{1-a} t \to 0$ for $t \to \infty$. Here $\varphi(t) = t^2 \log t \log_{\circ 2}{}^a t$, and since

$$\varphi'(t) = t \log_{\circ 2}{}^{a-1} t (a + (1 + 2\log t)\log_{\circ 2} t), \qquad t > e \tag{4.10}$$

is monotone increasing, φ is convex so the corollary applies. For $a = 2$ the (rounded) frequencies for $j \ge 2$ are 0.37, 0.09, 2, 9, 22, 42, 71, 110, 160,...

To elucidate Example 4.3, the imaginary part $\operatorname{Im} F_2$ is skecthed in Figure 1. All figures are made using Maple11 with a plot of a partial sum with 1000 terms and 1000 partition points in the interval, and possibly a further subdivision into at most 10 bits as decided by the program. However, the plots are by no means optimal or even faithful. In fact, neither the details are rendered correctly (absent or blurred) and nor are the macroscopic properties, as the graphs appear to be C^1 in certain subintervals. This is unavoidable due to the line thickness; anyhow each figure should provide an overview of the function's behaviour and an impression of its highly oscillatory nature.

The quasi-periodic behaviour visible in Figure 1 comes about because the first term of the series is dominating in the sense that it is much larger in magnitude than the remainder.

Remark 4.4. In Examples 4.1, 4.3 and 4.5 below, the series $\sum a_j$ converges very slowly, so most properties of $t \mapsto \sum_{j \le N} a_j \exp(\mathrm{i} t j/|a_j|)$ are likely to be destroyed by addition of the terms with $j > N$. Especially because the frequencies in $b_j = j/|a_j|$ do not change very much. Therefore the inference in Remark 1.3 is unlikely to apply in general.

One can give more pronounced examples of slow growth:

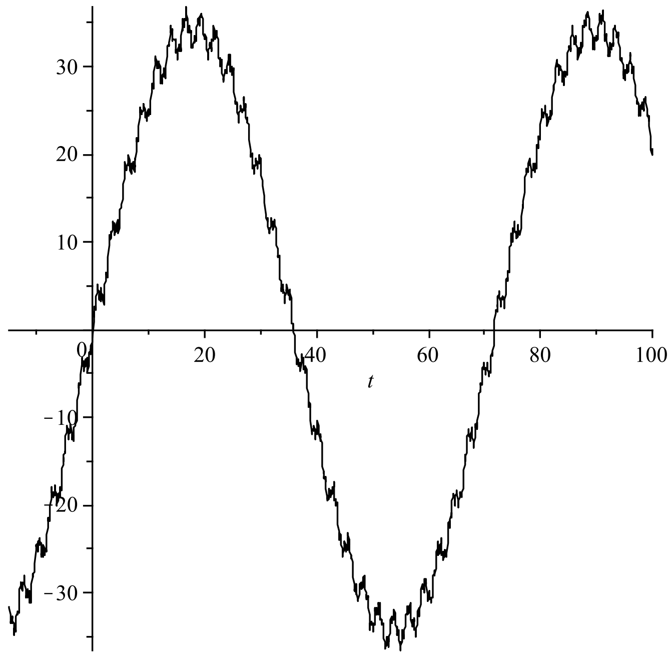

FIGURE 1. $\operatorname{Im} F_2(t)$ for the function in Example 4.3

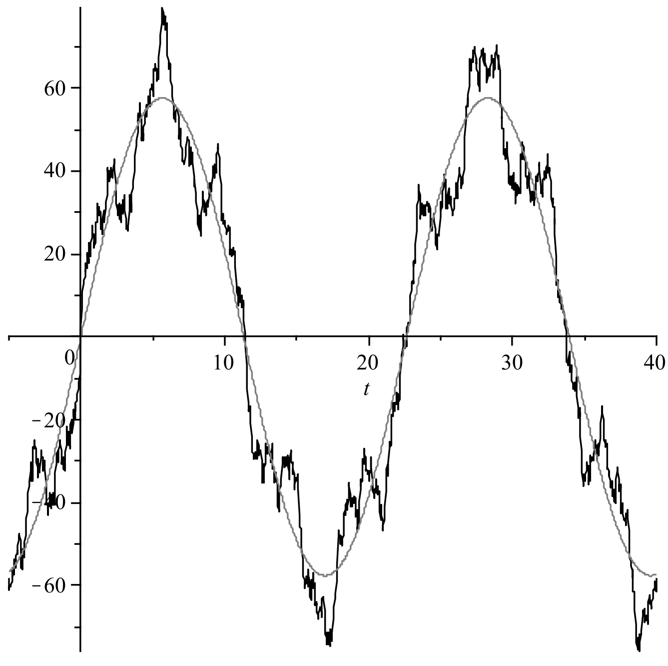

FIGURE 2. $\operatorname{Im} F_3(t)$ for the function in Example 4.5, $n = 3$, $a = 2$

Example 4.5. Denoting the n-fold logarithm by $\log_{\circ n} t := \log \dots \log t$, defined for $t > E_{n-2} := \exp \dots \exp 1$ (ie exp applied $n-2$ times), and setting $\log^a_{\circ n} t = (\log_{\circ n} t)^a$ for $a \in \mathbb{R}$ and $t > E_{n-1}$ (so that $s \mapsto s^a$ is defined at $s = \log_{\circ n} t$), there is a continuous nowhere differentiable function given by

$$F_n(t) = \sum_{j > E_{n-1}} \frac{\exp(\mathrm{i} t j^2 \log j \dots \log_{\circ(n-1)} j \cdot \log^a_{\circ n} j)}{j \log j \dots \log_{\circ(n-1)} j \cdot \log^a_{\circ n} j}. \tag{4.11}$$

Indeed, $\sum |a_j| < \infty$ for $a_j = 1/(j \log j \dots \log_{\circ(n-1)} j \log^a_{\circ n} j)$; this may be seen from the integral criterion since a_j is the value at $t = j$ of the derivative of

$$g(t) = \tfrac{1}{1-a} \log^{1-a}_{\circ n} t = \tfrac{1}{1-a} (\log \dots \log t)^{1-a} \xrightarrow[t \to \infty]{} 0 \quad \text{for} \quad a > 1. \tag{4.12}$$

It is clear that $a_j \geq a_{j+1}$, since all iterated logarithms are monotone increasing and positive for $j > E_{n-1}$. To apply the corollary one needs convexity of

$$\varphi_{a,n}(t) = t^2 \log t \dots \log_{\circ(n-1)} t \cdot \log^a_{\circ n} t, \qquad t > E_{n-1}. \tag{4.13}$$

Writing $S_n(x) = (x+1) \log_{\circ n} t$ for $n \in \mathbb{N}$, the derivative equals

$$\varphi'_{a,n}(t) = t \log^{a-1}_{\circ n} t (a + S_n \circ \dots \circ S_2(1 + 2\log t)), \qquad a > 0. \tag{4.14}$$

In fact, in a proof by induction the basis is provided by the previous example; and if this formula has been shown for all $a > 0$ for some integer $n \geq 2$, Leibniz' rule gives

$$\begin{aligned} \varphi'_{a,n+1}(t) &= \varphi'_{1,n}(t) \log^a_{\circ(n+1)} t + \varphi_{1,n}(t) (\log^a_{\circ(n+1)} t)' \\ &= t \log^{a-1}_{\circ(n+1)} t ((1 + S_n \circ \dots \circ S_2(1 + 2\log t)) \log_{\circ(n+1)} t + a) \\ &= t \log^{a-1}_{\circ(n+1)} (a + S_{n+1} \circ S_n \circ \dots \circ S_2(1 + 2\log t)). \end{aligned} \tag{4.15}$$

The formula for $\varphi'_{a,n}$ shows that, for $a > 1$, this function is monotone increasing on $]E_{n-1}, \infty[$, since $\log^{a-1}_{\circ n}$ is so and all iterated logarithms are positive on this interval. Hence $\varphi_{a,n}$ is convex on $]E_{n-1}, \infty[$ as required.

For $a = 2$ and $n = 3$ one has $E_{n-1} = e^e \approx 15.15$, and for $j \geq 16$ the (rounded) frequencies are

$$0.28,\ 1.4,\ 3.5,\ 6.8,\ 11,\ 17,\ 25,\ 34,\ 45,\ 57, \dots \tag{4.16}$$

(As $a = 2$ it is also possible to sum over $j \in \{3, 4, \dots, 15\}$, but these indices are best left out since the frequencies decrease from 11 for $j = 6$ to 0.009 for $j = 15$.) The function f_2 is sketched in Figure 2, where also the first term is drawn.

Example 4.6. As a last comparison, for $a = 2$ and $n = 4$, summation begins in (4.11) after $E_3 = e^{e^e} = 3814279.1\dots$ Cf Figure 3. A few of the resulting frequencies b_j are given in Table 1.

This clearly shows that, despite the larger number of j-dependent factors, one gets slower growth of the frequencies by using iterated logarithms of higher order. This may also be seen

term no.	1	..	10	..	20	..	100	..	1000
index j	3814280		3814290		3814300		3814380		3815280
frequency	0.02		2.4		10.6		247		24326

TABLE 1. Selected frequencies in Example 4.5 for $a = 2$, $n = 4$

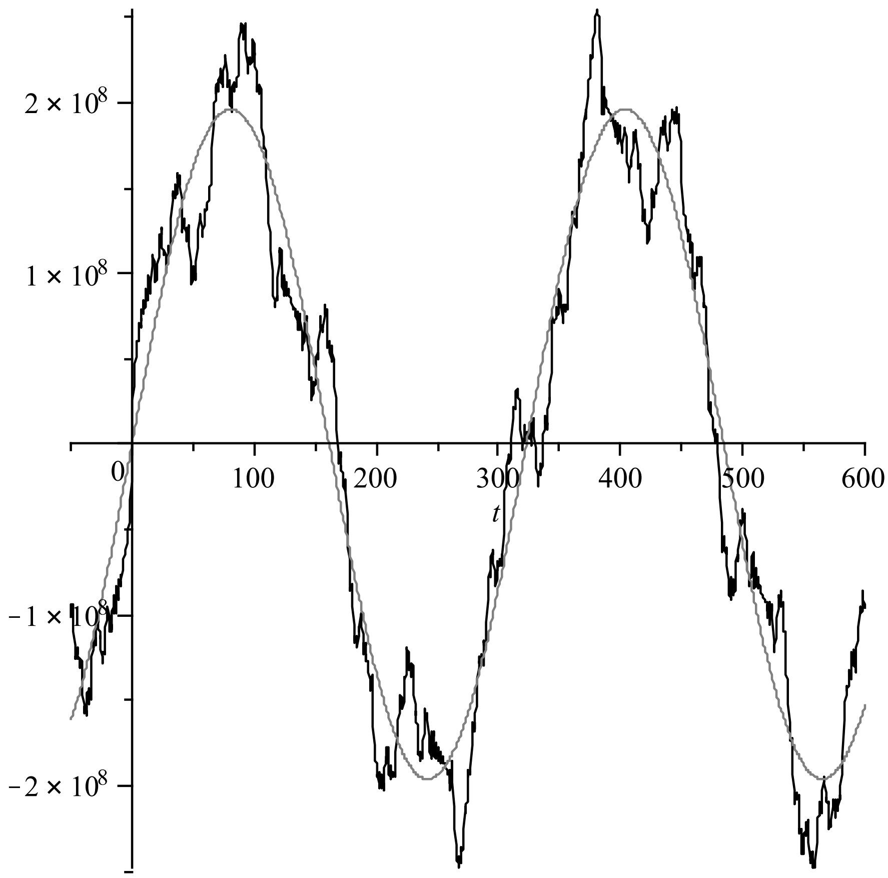

FIGURE 3. $\operatorname{Im} F_4(t)$ for the function in Example 4.3, $n = 4$, $a = 2$

analytically since a substitution in $\lim_{t\to\infty} t^{-\alpha} \log^{\beta} t = 0$, yields that for $a > 1$

$$\log_{\circ n} t \log^{a}_{\circ(n+1)} t = o(\log^{a}_{\circ n} t) \quad \text{for} \quad t \to \infty. \tag{4.17}$$

It appears from Figures 1, 2 and 3 that as the frequency growth is reduced, one gets increasingly larger deviations from a sinusoidal curve. Ie the first term becomes less dominating, cf Remark 4.4.

To illustrate that, Figure 4 shows the deviation from the first term, ie the sum over $j \geq 3814281$. Notice that the sinusoidal structure is almost completely lost, ie the first term is not dominating here. So this case seems to corroborate Remark 4.4 more clearly.

In addition to the vertical tangent at the origin in Figure 4, there are many approximate self-similarities, like those for $R(t)$ analysed by J. Duistermaat [Dui91]. Eg the behaviour for ca. $40 < t < 75$ seems similar to that found for $25 < t < 40$ and so on for $t \to 0_+$.

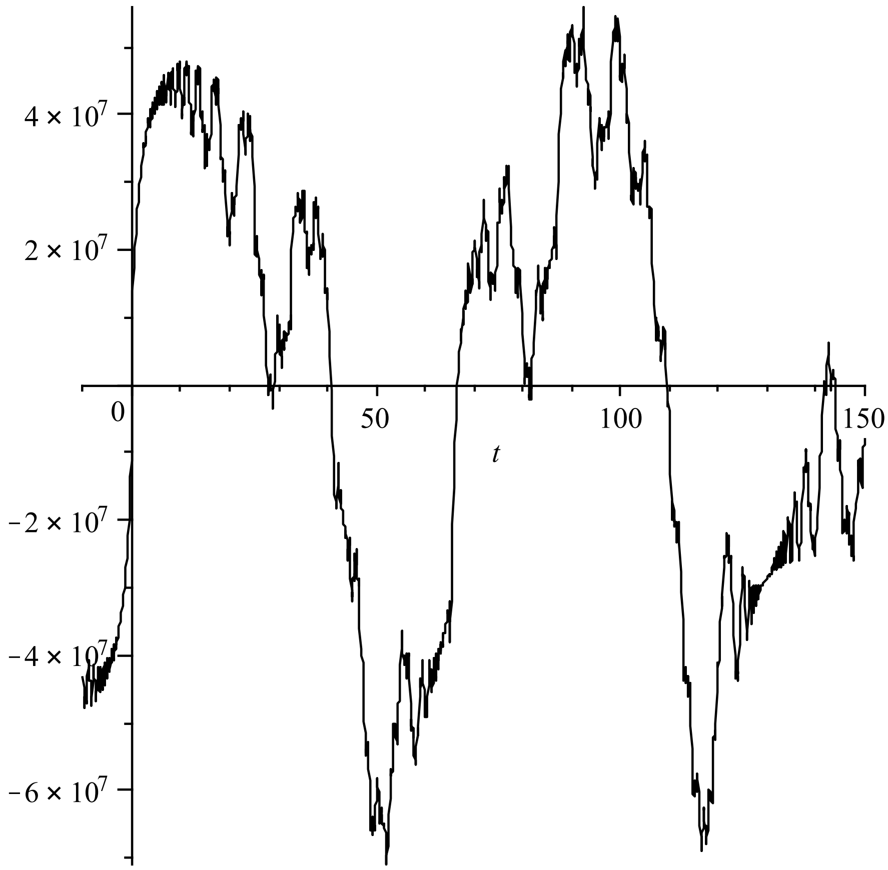

FIGURE 4. Deviation from the first term of $\operatorname{Im} F_4(t)$ in Example 4.5, $n = 4$, $a = 2$

For larger n I have not found it worthwhile to attempt any plots, for already for $n = 5$ the summation runs over $j > E_4 = e^{3814279.1\cdots}$. This tremendous number would make even a sketch of the partial sums a tricky task in scientific computing.

5. FINAL REMARKS

The first example of a nowhere differentiable function is due to R. Bolzano (ca. 1830, discovered 1920). T. Takagi [Tak03] introduced $t \mapsto \sum_{j=0}^{\infty} 2^{-j} \operatorname{dist}(2^j t, \mathbb{Z})$ where the terms are only piecewise C^1. Recently examples of nowhere-differentiability were given by means of infinite products; cf [Wen02]. For a review of the historical development of the subject the reader could consult the illustrated thesis of J. Thim [Thi03].

My interest in the subject developed during a recent study of pseudo-differential operators of order $d \in \mathbb{R}$ and type $1,1$; cf [Joh]. The basic pathologies of this operator class are deduced by means of functions on $\mathbb{R}^n$ of the form

$$u_{\xi_0}(x) = v(x) \sum_{j=1}^{\infty} 2^{-jd} \exp(\mathrm{i}\, 2^j \langle x, \xi_0 \rangle). \tag{5.1}$$

Eg when v is in $\mathscr{S}(\mathbb{R}^n)$ and has a small compact spectrum, it may be seen that the above u_{ξ_0} has $\mathbb{R}^n \times (\mathbb{R}_+ \xi_0)$ as its wavefront set. This enters the proof that not all operators A of type $1,1$ preserve wavefront sets: A can be taken such that $\mathbb{R}^n \times (\mathbb{R}_+(-\xi_0))$ is the wavefront set of Au_{ξ_0}; cf [Joh]. However, an interesting aspect of this is that the sum defines a nowhere differentiable function for $0 < d \leq 1$ (nowhere C^∞ for general $d \in \mathbb{R}$). Multiplication by the analytic function v gives u_{ξ_0} the same lack of differentiability almost everywhere, which explains why u_{ξ_0} has its singular support spread over the entire $\mathbb{R}^n$.

When the manuscript was almost complete, it was discovered that a few elements of the arguments exist sporadically in the literature; cf Remark 3.4 for comments on [CU07]. Moreover, the real and imaginary parts of f_θ, $\theta = 1$ have been analysed by Y. Meyer [Mey93, Ch. 9.2] with a method partly based on wavelets and partly similar to the proof of Theorem 2.2. The method was attributed to G. Freud but without any references.

Subsequently an inspection of [Fre62, Satz VI] revealed that G. Freud showed that an integrable periodic function f with Fourier series $\sum \rho_k \sin(n_k t + \varphi_k)$, $\inf n_{k+1}/n_k > 1$ is differentiable at a point only if $\lim \rho_k n_k = 0$, similarly to Theorem 2.2. His proof was based on estimates of the differentiated Cesaro means using inequalities for the corresponding Fejer kernel (as done also in [SS03]), so it is applicable only for periodic functions.

There is much more overlap with the work of Y. Meyer, especially in the use of the Fourier transformation $\mathscr{F}$ on $\mathbb{R}$. Whereas the purpose in [Mey93, Ch. 9.2] was to explain that the lack of differentiability of $\operatorname{Re} f_1$, $\operatorname{Im} f_1$ can be derived with wavelet theory, the present paper goes much beyond this. Eg nowhere-differentiability of f_θ, or W, is shown to follow directly from basic facts in integration theory; cf the introduction. And using only $\mathscr{F}$, differentiability was in Theorem 2.2 linked to the growth of the frequencies b_j. Moreover, the removal of the spectral gap condition $\liminf b_{j+1}/b_j > 1$ in Theorem 3.1 seems to be a novelty, which shows that the growth of the frequency increments Δb_j is equally important.

In the context of distribution theory, the present proofs are rather straightforward. But even so they may be valuable, eg because of the easy access to the examples in Section 4, that have a rather different nature than the previously known ones.

Acknowledgement. I am grateful to Prof. L. Rodino and Prof. H. Cornean for their interest in this work and for their call to publish it.

REFERENCES

[CU07] F. Chamizo and A. Ubis, *Some Fourier series with gaps*, J. Anal. Math. **101** (2007), 179–197.

[Dui91] J. J. Duistermaat, *Self-similarity of "Riemann's nondifferentiable function"*, Nieuw Arch. Wisk. (4) **9** (1991), no. 3, 303–337.

[Fre62] G. Freud, *Über trigonometrische Approximation und Fouriersche Reihen*, Math. Z. **78** (1962), 252–262.

[Har16] G. H. Hardy, *Weierstrass's non-differentiable function*, Trans. Amer. Math. Soc. **17** (1916), no. 3, 301–325.

[Hol95] M. Holschneider, *Wavelets*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995, An analysis tool, Oxford Science Publications.

[Jaf97] S. Jaffard, *Old friends revisited: the multifractal nature of some classical functions*, J. Fourier Anal. Appl. **3** (1997), 1–22.

[Joh] J. Johnsen, *Type* $1,1$*-operators defined by vanishing frequency modulation*, Preprint R-2008-1, Aalborg University, 2008.

[Leb08] H. Lebesgue, *Sur la méthode de M. Goursat pour la résolution de l'équation de Fredholm*, Bull. Soc. Math. France **36** (1908), 3–19.

[Lut86] W. Luther, *The differentiability of Fourier gap series and "Riemann's example" of a continuous, nondifferentiable function*, J. Approx. Theory **48** (1986), 303–321.

[McC53] J. McCarthy, *An everywhere continuous nowhere differentiable function*, Amer. Math. Monthly **60** (1953), 709.

[Mey93] Y. Meyer, *Wavelets*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1993, Algorithms & applications.

[Sch66] L. Schwartz, *Théorie des distributions*, revised and enlarged ed., Hermann, Paris, 1966.

[SS03] E. M. Stein and R. Shakarchi, *Fourier analysis*, Princeton Lectures in Analysis, vol. 1, Princeton University Press, Princeton, NJ, 2003.

[Tak03] T. Takagi, *A simple example of the continous function without erivative*, Proc. Rhys. Math. Soc. Japan **1** (1903), 176–177.

[Thi03] J. Thim, *Continuous nowhere differentiable functions*, Master's thesis, Luleå University of Technology, Sweden, 2003.

[Wen02] Liu Wen, *A nowhere differentiable continuous function constructed by infinite products*, Amer. Math. Monthly **109** (2002), 378–380.

[Zyg59] A. Zygmund, *Trigonometric series. 2nd ed. Vols. I, II*, Cambridge University Press, New York, 1959.

DEPARTMENT OF MATHEMATICAL SCIENCES, AALBORG UNIVERSITY, FREDRIK BAJERS VEJ 7G, DK-9220 AALBORG ØST, DENMARK

E-mail address: `jjohnsen@math.aau.dk`