Solution to the gradient problem of C. E. Weil

Zoltán Buczolich

Abstract

In this paper we give a complete answer to the famous gradient problem of C. E. Weil. On an open set $G \subset \mathbb{R}^2$ we construct a differentiable function $f : G \to \mathbb{R}$ for which there exists an open set $\Omega_1 \subset \mathbb{R}^2$ such that $\nabla f(\mathbf{p}) \in \Omega_1$ for a $\mathbf{p} \in G$ but $\nabla f(\mathbf{q}) \notin \Omega_1$ for almost every $\mathbf{q} \in G$. This shows that the Denjoy-Clarkson property does not hold in higher dimensions.

1. Introduction

The gradient problem of C. E. Weil is the following question: Assume $n \geq 2$ and $G \subset \mathbb{R}^n$ is open, $f : G \to \mathbb{R}$ is a differentiable function. Then ∇f maps G into $\mathbb{R}^n$. Assume $\Omega \subset \mathbb{R}^n$ is open. Is it true that $(\nabla f)^{-1}(\Omega) = \{\mathbf{p} \in G : \nabla f(\mathbf{p}) \in \Omega\}$ is either empty, or of positive n-dimensional Lebesgue measure?

When $n = 1$ then the answer is positive and is the so called Denjoy-Clarkson property of the derivative functions [8], [7]: If $f : (a, b) \to \mathbb{R}$ is a differentiable function and (α, β) is an open interval then $(f')^{-1}(\alpha, \beta)$ is either empty, or of positive (one dimensional) Lebesgue measure. In [12] it was shown that the kth Peano derivatives and the approximate derivatives also have the Denjoy-Clarkson property.

The gradient problem was one of the well-known and famous unsolved problems in Real Analysis. It was around since the paper [12] appeared in the 1960s. I have learned it in 1987 and have worked on it since then. In 1990 at the Fourteenth Summer Symposium on Real Analysis it was advertised and appeared in print in [13].

In this paper we answer this question by giving a two dimensional counterexample. We show that there exists a nonempty open set $G \subset \mathbb{R}^2$, a differentiable function $f : G \to \mathbb{R}$ and an open set $\Omega_1 \subset \mathbb{R}^2$ for which there exists a $\mathbf{p} \in G$ such that $\nabla f(\mathbf{p}) \in \Omega_1$ but for almost every (in the sense of two dimensional Lebesgue measure, λ_2,) $\mathbf{q} \in G$ the gradient $\nabla f(\mathbf{q})$ is not in Ω_1.

The result in [2] which was reproved by different methods in [10] shows that in the above example, although $\lambda_2((\nabla f)^{-1}(\Omega_1)) = 0$, if $\mathcal{H}^1$ denotes the one dimensional Hausdorff measure then $\mathcal{H}^1((\nabla f)^{-1}(\Omega_1)) > 0$. The results in [10] imply more than this. It turns out that any projection of $(\nabla f)^{-1}(\Omega_1)$ onto a line is of positive λ_1 measure, $(\nabla f)^{-1}(\Omega_1)$ is non-σ-porous and is porous at none of its points. The results in [5] also show that if $\mathbf{v}$ is an interior point of

$$\mathcal{R}_1 \overset{\text{def}}{=} \Omega_1 \cap \{\nabla f(\mathbf{p}) : \mathbf{p} \in G\}$$

then $\mathcal{H}^1((\nabla f)^{-1}(\mathbf{v})) > 0$, which implies that $(\nabla f)^{-1}(\Omega_1)$ is of non-σ-finite $\mathcal{H}^1$ measure. Furthermore, the range $\mathcal{R}_1$ satisfies a certain interesting convexity/concavity property. Actually, this convexity/concavity property is behind the construction of the counterexample given in this paper. These results from [5] show that our counterexample function has some strange properties.

On the other hand, [3] shows that if $\nabla f \neq (0,0)$ on G (this will be true in our counterexample) then the level set structure of f is relatively simple, $f^{-1}(\{y\})$ has components which are homeomorphic to differentiable arcs and there are no bifurcation points on these components. Related to our work on the gradient problem in [6] a pathological C^1-function, with "too many tangent planes" was constructed. This result pointed in the direction that maybe the counterexample function of this paper exists.

In [11] G. Petruska verified that if $F : \mathbb{R} \to \mathbb{R}$ is a differentiable function of one variable and $f = F'$ then f takes each of its values on A_f, where A_f denotes the approximate continuity points of f. In [4] we showed that if $F : \mathbb{R}^n \to \mathbb{R}$ is differentiable and $i \in \{1, \ldots, n\}$ then $f = \partial_i F$ takes each of its values on A_f. On the other hand, there exists a continuous function $F : \mathbb{R}^2 \to \mathbb{R}$ such that $f = \partial F / \partial x_1$ exists everywhere and f does not take each of its values on A_f. In Theorem 3 of [4] a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ is constructed such that ∇f does not take each of its values on $A_{\nabla f}$.

Assume $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function. We call a point $\mathbf{y} \in \mathbb{R}^n$ a regular value of ∇f if there is an $\mathbf{x} \in A_{\nabla f}$ such that $\nabla f(\mathbf{x}) = \mathbf{y}$. Denote the set of regular values by $REG(\nabla f)$. Our question in [4] about the density of $REG(\nabla f)$ in the range of ∇f receives a negative answer by the negative answer to the gradient problem. On the other hand our question about the characterization of $REG(\nabla f)$ remains open.

There is a new question related to the gradient problem: Assume that $G \subset \mathbb{R}^n$ is an open set, $f : G \to \mathbb{R}$ is a differentiable function and $\Omega \subset \mathbb{R}^n$ is such an open set that $(\nabla f)^{-1}(\Omega) \neq \emptyset$. Then what can we say about the Hausdorff-dimension of $(\nabla f)^{-1}(\Omega)$. The results discussed above imply that $(\nabla f)^{-1}(\Omega)$ is of Hausdorff dimension at least one, and our counterexample in this paper shows that it is possible that $\lambda_2((\nabla f)^{-1}(\Omega)) = 0$.

2. Main Result

This section is devoted to the proof of our main result, Theorem 1.

We put $G = (-1, 1) \times (-1, 1)$ and $\Omega_0 = [-\frac{1}{2}, \frac{1}{2}] \times [0, 2]$, $\Omega_1 = (-0.49, 0.49) \times (0, 1.99)$, and $\Omega_2 = [-0.51, 0.51] \times [0, 2.01]$.

Theorem 1. *There exists a differentiable function $f : G \to \mathbb{R}$ such that $\nabla f(0,0) = (0,1)$ and $\nabla f(\mathbf{p}) \notin \Omega_1$ for almost every $\mathbf{p} \in G$.*

The construction of f is quite complicated and is done in Subsection 2.2. Here is an informal plan of what we are doing. We start with a function $h_{-1}(x, y) = y$. Then $\nabla h_{-1} = (0, 1)$ everywhere on G. Now our aim is to choose a sequence of functions h_n so that

$$f(x,y) \overset{\text{def}}{=} \sum_{n=-1}^{\infty} h_n(x,y)$$

satisfies Theorem 1. Each function h_{n+1} can be regarded as a perturbation of the previous partial sum $\sum_{k=-1}^{n} h_k$. Our aim is to push $\nabla f(\mathbf{p})$ outside of Ω_1 for almost every $\mathbf{p} \in G$. In the construction we will have a nested sequence of open sets G_n such that for almost every $\mathbf{p}$ outside of G_n we will have $\nabla f(\mathbf{p}) \notin \Omega_1$. We use a "stopping time argument" by not perturbing the function at almost every point, $\mathbf{p}$ any more, once $\nabla f(\mathbf{p}) \notin \Omega_0$. For these points $\sum_{n=-1}^{\infty} h_n(x,y) = \sum_{n=-1}^{n_0} h_n(x,y)$ for a suitable n_0. We will show that $\lambda_2(G_n) \to 0$. The main difficulty is to show that f is differentiable at those points $\mathbf{p} \in \cap_{n=0}^{\infty} G_n$ which are subject to infinitely many perturbations. To handle this difficult case an argument which originates from dynamical systems is used, I learned the heuristic behind this argument when I worked on the proof of [1]. This important argument is discussed first in Subsection 2.1. Reading this subsection one should think that $\mathbf{p}$ is a given point in G and the trajectory x_n is the first coordinate of $\nabla(\sum_{k=-1}^{n-1} h_k(\mathbf{p}))$ (apart from some vanishing error term) and y_n is the second coordinate (again with some vanishing error term). A sequence $\{g_n\}$ of concave down auxiliary functions is also defined. These functions will determine the way we move in the gradient space and the motivation for the introduction of these functions originates from the convexity/concavity result in [5].

2.1. Trajectories in the gradient space, the dynamics behind our construction

Set $x_0 = 0$, $y_0 = 1$, $a(0, 1) = 0$, $k_0 = 1$, $\mathcal{A}_0 = \{0\}$, $g_0(x) = 1 - \frac{1}{4}x^4$.

Suppose $n \geq 0$ and the trajectory $x_k \in [-1, 1]$, for $k = 1, \ldots, n$ is given together with another sequence $y_k \in \mathbb{R}$. This second sequence is computed based on the trajectory $\{x_k\}$.

We also have a set $\mathcal{A}_n = \{a(n, j) : j = 1, \ldots, k_n\}$ which gives the indices of the so called *active points* of the trajectory: $x_{a(n,1)}, \ldots, x_{a_{a(n,k_n)}}$. We assume that $a(n, j) < a(n, j+1)$, $j = 1, \ldots, k_n - 1$ and $a_{(n,k_n)} = n$. The point $x_0 = 0$ will always be regarded to be active. Hence $a(n, 1) = 0$ for all n. We also have a collection of functions g_k, $k = 1, \ldots, n$ which are concave down and $g_k(x_k) = y_k$. We also assume, $x_{j+1} - x_j \neq 0$ for $j = 0, \ldots, n-1$. Suppose that the next term $x_{n+1} \neq x_n$ of the trajectory is given. Denote by $\tau(x)$ the tangent of g_n at the point $(x_n, g_n(x_n))$. Set $y_{n+1} = \tau(x_{n+1})$. Since g_n is concave we clearly have $y_{n+1} \geq g_n(x_{n+1})$. See Figure 1.
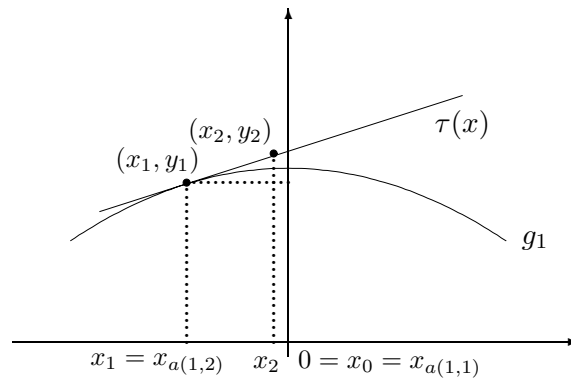

FIGURE 1: The functions g_n.

If the trajectory at step $n + 1$ is not changing direction, that is,

$$(x_n - x_{n-1})(x_{n+1} - x_n) > 0$$

then x_n becomes inactive and this means that it will not belong to $\mathcal{A}_{n+1}$. (At the first step, by definition, we do not change direction.)

If the trajectory at step $n + 1$ is changing direction, that is,

$$(x_n - x_{n-1})(x_{n+1} - x_n) < 0$$

then x_n stays active, that is, $x_n \in \mathcal{A}_{n+1}$, provided it is not decativated by the "passing over" rule given below.

The last point of the trajectory x_{n+1} is always active, so

$$n + 1 = a(n + 1, k_{n+1}) \in \mathcal{A}_{n+1}.$$

Inactive points cannot become active but some active points might get deactivated when x_n is "passing over" them, or over their predecessors. This means that if $x_{a(n,j)} \in (x_n, x_{n+1}]$ for a $j \in \{1, \ldots, k_n\}$ then all $x_{a(n,j')}$ with $j \leq j' \leq k_n$, $j' \neq 1$ become inactive and these $a(n,j')$'s will not belong to $\mathcal{A}_{n+1}$. If $0 = x_0 \in (x_n, x_{n+1}]$ we say that we are resetting the active trajectory, and $\mathcal{A}_{n+1} = \{0, n+1\}$ in this case.

The elements of $\mathcal{A}_n$ which are not deactivated by the above process will belong to $\mathcal{A}_{n+1}$.

The only way x_n can survive step $n+1$ as an active point is by being a turning point of the trajectory. Hence, all active points are (some earlier) turning points, that is,

$$(x_{a(n,j)} - x_{a(n,j-1)})(x_{a(n,j+1)} - x_{a(n,j)}) < 0.$$

Assuming $x_{a(n,2)} < x_{a(n,1)} = 0$ the "cancellation by passing over" rule implies that we have $x_{a(n,2(j-1))} \leq x_{a(n,2j)}$, $x_{a(n,2j-1)} \geq x_{a(n,2j+1)}$ and $x_{a(n,2j)} \leq x_n \leq x_{a(n,2j'-1)}$ for all j, j' for which the arguments of $a(n, .)$ are defined. See Figure 2. In case $x_{a(n,2)} > x_{a(n,1)} = 0$ we have a mirror image and all inequalities are reversed.
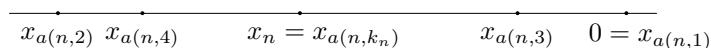

$$\underline{\quad \overset{\bullet}{x_{a(n,2)}} \; \overset{\bullet}{x_{a(n,4)}} \qquad \overset{\bullet}{x_n = x_{a(n,k_n)}} \qquad \overset{\bullet}{x_{a(n,3)}} \quad \overset{\bullet}{0 = x_{a(n,1)}} \quad}$$

FIGURE 2: Active points at step n.

If $a, b \in \mathbb{R}$ we denote the closed interval with endpoints a and b by $[a, b]$ irregardless whether $a \leq b$, or $b \leq a$, we adopt similar convention for other type of intervals as well.

Next we show how g_n is defined. We define its derivative g_n' on $\mathbb{R}$ and assume that it also satisfies the initial condition $g_n(x_n) = y_n$.

We still assume $x_{a(n,2)} < x_{a(n,1)} = 0$, the case $x_{a(n,2)} > x_{a(n,1)} = 0$ is similar, the case $x_{a(n,2)} = x_{a(n,1)} = 0$ is trivial, we set $g_n'(x) = -x^3$. On $[x_{a(n,1)}, +\infty)$ and on $(-\infty, x_{a(n,2)}]$ we put $g_n'(x) = -(x - x_{a(n,1)})^3 = -x^3$. This also defines $g_n'(x_{a(n,2)}) = -(x_{a(n,2)} - x_{a(n,1)})^3$. Assume

$$g_n'(x_{a(n,j)}) = -(x_{a(n,2)} - x_{a(n,1)})^3 - (x_{a(n,3)} - x_{a(n,2)})^3 - \cdots - (x_{a(n,j)} - x_{a(n,j-1)})^3.$$

For a $j \in \{2, \ldots, k_n - 1\}$ on $[x_{a(n,j+1)}, x_{a(n,j-1)})$ we put $g_n'(x) = g_n'(x_{a(n,j)}) - (x - x_{a(n,j)})^3$. Which implies

$$\begin{aligned} g_n'(x_{a(n,j+1)}) = &-(x_{a(n,2)} - x_{a(n,1)})^3 - (x_{a(n,3)} - x_{a(n,2)})^3 \\ &- \cdots - (x_{a(n,j)} - x_{a(n,j-1)})^3 - (x_{a(n,j+1)} - x_{a(n,j)})^3. \end{aligned}$$

Finally, on the interval $(x_{a(n,k_n-1)}, x_{a(n,k_n)})$ put $g_n'(x) = g_n'(x_{a(n,k_n)}) - (x - x_{a(n,k_n)})^3$. Since $g_n'(x_{a(n,k_n)}) = g_n'(x_{a(n,k_n-1)}) - (x_{a(n,k_n)} - x_{a(n,k_n-1)})^3$ the function g_n' is continuous at $x_{a(n,k_n-1)}$.

We set $g_n(x) = y_n + \int_{x_n}^x g'_n(x)dx$.

Observe that g'_n is continuous at the points $x_{a(n,j)}$, $j = 1, \ldots, k_n$ and is monotone decreasing on the intervals determined by these active points.

Hence g_n is concave (down) on $\mathbb{R}$.

We will only be interested in g_n when the trajectory, $\{x_n\}$, is in $[-1, 1]$ and with this assumption we also have $|g'_n(x)| \leq 1$ for $x \in [-1, 1]$.

Lemma 2. *For all $x \in \mathbb{R}$ and $n = 0, 1, \ldots$ we have $g_{n+1}(x) \geq g_n(x)$.*

Proof. We separate three cases.

Case I: There is no cancellation of active points and at step $n + 1$ the trajectory is not turning its direction. This implies $k_{n+1} = k_n$.

Case II: There is no cancellation and at step $n+1$ the trajectory is changing its direction. Then x_n stays active and x_{n+1} becomes active. Thus $k_{n+1} = k_n + 1 > k_n$.

Case III: There is some cancellation of active points. Then $k_{n+1} \leq k_n$ and $k_{n+1} = k_n$ is possible only when $k_{n+1} = k_n = 2$ and at step $n + 1$ we reset the trajectory.

During the whole proof we assume $x_{a(n,2)} < x_{a(n,1)} = 0$, the case $x_{a(n,2)} > x_{a(n,1)} = 0$ is similar, the case $x_{a(n,2)} = x_{a(n,1)} = 0$ is trivial.

Case I: Without limiting generality we can also suppose that $u \overset{\text{def}}{=} x_{a(n,k_n-1)}$ $= x_{a(n+1,k_{n+1}-1)} < x_n < x_{n+1} < w = x_{a(n,k_n-2)} = x_{a(n+1,k_{n+1}-2)} \leq x_{a(n,1)}$, see Figure 3. By using the definition of the functions g_n and g'_{n+1} on $(-\infty, u]$ and on $[x_{n+1}, +\infty)$ we have $g'_n = g'_{n+1}$, this might not completely clear on (x_{n+1}, w) but on this interval $g'_n(x) = g'_{n+1}(x) = g'_n(u) - (x - u)^3$. Thus it is enough to show $g_{n+1} \geq g_n$ on $[u, x_{n+1}]$.
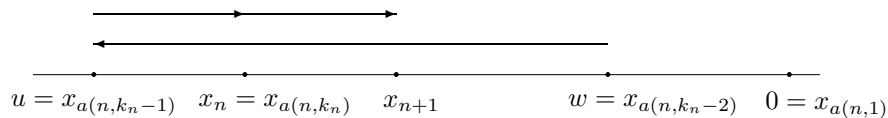

FIGURE 3: Case I.

On $[x_n, x_{n+1}]$ we have $g'_n(x) = g'_n(u) - (x - u)^3$ and $g'_{n+1}(x) = g'_{n+1}(x_{n+1}) - (x - x_{n+1})^3 = g'_n(u) - (x_{n+1} - u)^3 - (x - x_{n+1})^3$. So

$$g'_{n+1}(x) - g'_n(x) = -(x_{n+1} - u)^3 - (x - x_{n+1})^3 + (x - u)^3$$
$$= -((x_{n+1} - x) + (x - u))^3 + (x_{n+1} - x)^3 + (x - u)^3 < 0,$$

therefore, using $x < x_{n+1}$ we infer

$$g_{n+1}(x) - g_{n+1}(x_{n+1}) - (g_n(x) - g_n(x_{n+1})) = \int_{x_{n+1}}^{x} (g'_{n+1}(t) - g'_n(t))dt > 0,$$

which based on $y_{n+1} = g_{n+1}(x_{n+1}) \geq g_n(x_{n+1})$, yields $g_{n+1}(x) \geq g_n(x)$ for all $x \in [x_n, x_{n+1}]$.

On $[u, x_n]$ we have

$$g'_n(x) = g'_n(x_n) - (x - x_n)^3 = g'_n(u) - (x_n - u)^3 - (x - x_n)^3$$

and

$$g'_{n+1}(x) = g'_{n+1}(x_{n+1}) - (x - x_{n+1})^3 = g'_{n+1}(u) - (x_{n+1} - u)^3 - (x - x_{n+1})^3$$
$$= g'_n(u) - (x_{n+1} - u)^3 - (x - x_{n+1})^3.$$

Now,

$$g'_{n+1}(x) - g'_n(x) = -(x_{n+1} - u)^3 - (x - x_{n+1})^3 + (x_n - u)^3 + (x - x_n)^3,$$
$$g'_{n+1}(u) - g'_n(u) = 0, \qquad \text{and}$$
$$g''_{n+1}(x) - g''_n(x) = -3(x - x_{n+1})^2 + 3(x - x_n)^2 < 0.$$

Thus $g'_{n+1} - g'_n \leq 0$ on $[u, x_n]$. Hence, again

$$g_{n+1}(x) - g_{n+1}(x_n) - (g_n(x) - g_n(x_n)) = \int_{x_n}^{x} (g'_{n+1}(t) - g'_n(t))dt \geq 0.$$

Since we have already verified that $g_{n+1}(x_n) \geq g_n(x_n)$ this implies $g_{n+1}(x) \geq g_n(x)$ on $[u, x_n]$.

Case II. Now $k_{n+1} > k_n$.

Again, without limiting generality, we assume $x_{n+1} < x_n < x_{a(n,1)} = x_{a(n+1,1)} = 0$ and setting

$$u \overset{\text{def}}{=} x_{a(n,k_n-1)} < x_{n+1} \qquad \text{and} \qquad w \overset{\text{def}}{=} x_{a(n,k_n-2)}$$

we have $u < x_{n+1} < x_n < w \leq x_{a(n,1)}$, see Figure 4. Now, by our assumptions $x_n = x_{a(n+1,k_n)}$ and $x_{n+1} = x_{a(n+1,k_n+1)}$.
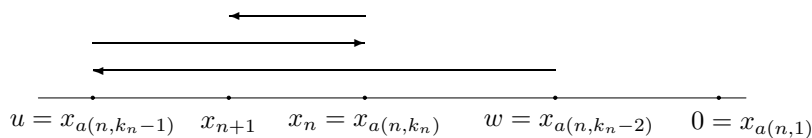

$$u = x_{a(n,k_n-1)} \quad x_{n+1} \quad x_n = x_{a(n,k_n)} \quad w = x_{a(n,k_n-2)} \quad 0 = x_{a(n,1)}$$

FIGURE 4: Case II.

On $(-\infty, u]$ and on $[w, +\infty)$ by our definitions we have $g_n' = g_{n+1}'$.

On $[x_n, w]$ we have

$$g_n'(x) = g_n'(u) - (x - u)^3 = g_{n+1}'(u) - (x - u)^3 = g_{n+1}'(x).$$

On $[u, x_{n+1}]$ we also have $g_n'(x) = g_n'(x_n) - (x - x_n)^3$ and

$$g_{n+1}'(x) = g_{n+1}'(x_n) - (x - x_n)^3 = g_n'(x).$$

On $[x_{n+1}, x_n]$,

$$g_{n+1}'(x) = g_{n+1}'(x_{n+1}) - (x - x_{n+1})^3 = g_n'(x_n) - (x_{n+1} - x_n)^3 - (x - x_{n+1})^3,$$

and $g_n'(x) = g_n'(x_n) - (x - x_n)^3$.

Hence

$$\begin{aligned} g_{n+1}'(x) - g_n'(x) &= -(x_{n+1} - x_n)^3 - (x - x_{n+1})^3 + (x - x_n)^3 \\ &= ((x_n - x) + (x - x_{n+1}))^3 - (x - x_{n+1})^3 - (x_n - x)^3 > 0 \end{aligned}$$

and

$$g_{n+1}(x) - g_{n+1}(x_{n+1}) - (g_n(x) - g_n(x_{n+1})) = \int_{x_{n+1}}^x (g_{n+1}'(t) - g_n'(t))dt > 0.$$

This, together with $y_{n+1} = g_{n+1}(x_{n+1}) \geq g_n(x_{n+1})$, implies again that $g_{n+1}(x) \geq g_n(x)$.

Case III. Now $k_{n+1} \leq k_n$ and there are some deactivated points of the trajectory.

First we assume that the trajectory at step $n+1$ is not changing its direction, this is called **Case IIIa**. Without limiting generality we can assume $x_n < x_{a(n,1)}$, $x_n < x_{n+1}$ and, due to the cancellation, $u_j = x_{a(n,k_n-2j)} \in (x_n, x_{n+1}]$ for $j = 1, \ldots, j_0$ with some $j_0 \geq 1$. Set $w_j = x_{a(n,k_n-2j+1)}$, $j = 1, \ldots, j_0$. Then $w_{j_0} < \cdots < w_2 < w_1 < x_n < u_1 < \cdots < u_{j_0} \leq x_{n+1}$, see Figure 5.
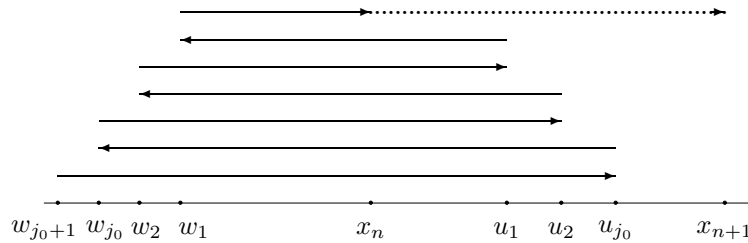

FIGURE 5: Case IIIa.

For $t \in [x_n, x_{n+1}]$ we define some auxiliary functions $g_{n,t}$.

For $t \in [x_n, u_1)$ set $g_{n,t}(t) = g_n(t)$ and $g'_{n,t}(x) = g'_n(x)$ for $x \leq w_1$ or $x \geq u_1$. For $x \in [t, u_1)$ set

$$g'_{n,t}(x) = g'_{n,t}(w_1) - (x - w_1)^3 = g'_n(w_1) - (x - w_1)^3 = g'_n(x).$$

See Figure 6. This implies $g_{n,t}(x) = g_n(x)$ for all $x \geq t$. For $x \in (w_1, t]$ set

$$g'_{n,t}(x) = g'_{n,t}(t) - (x - t)^3 = g'_{n,t}(w_1) - (t - w_1)^3 - (x - t)^3.$$
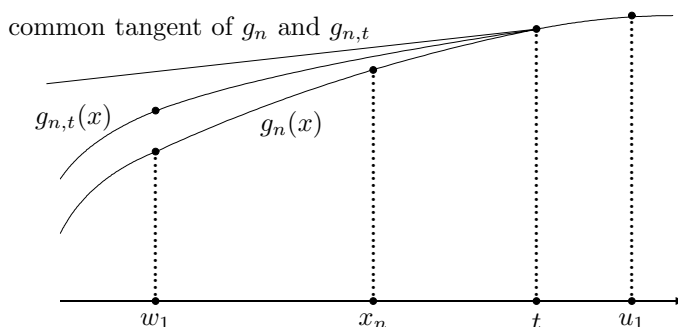

FIGURE 6: The functions g_n and $g_{n,t}$.

Observe that the definition of $g_{n,t}$ is exactly the same what we would have obtained in Case I if we had $x_{n+1} = t$ and $y_{n+1} = g_n(t)$. So, as in Case I, one could see that $g_{n,t}(x) \geq g_n(x)$ for all $x \in \mathbb{R}$. (See Figure 6.)

Furthermore, taking $t, t' \in (x_n, u_1)$, $t < t'$ and thinking of t as being the next point of the trajectory, x_{n+1}, and t' as x_{n+2}, $y_{n+1} = g_n(t)$, $y_{n+2} = g_n(t')$ one could see like in Case I that $g_{n,t'}(x) \geq g_{n,t}(x)$.

Set $g_{n,u_1^-}(x) = \lim_{t \to u_1^-} g_{n,t}(x)$. Then $g'_{n,u_1^-}(x) = \lim_{t \to u_1^-} g'_{n,t}(x)$. Set $n_1 = a(n, k_n - 2)$, then $u_1 = x_{n_1}$ and next we show that $g'_{n,u_1^-}(x) = g'_{n_1}(x)$ which shows that after the cancellation at u_1 we are "dropping down" to a function which is a vertically shifted copy of g_{n_1}. Indeed, for $x \notin (w_1, u_1)$ we have $g'_{n,u_1^-}(x) = g'_n(x) = g'_{n_1}(x)$. For $x \in (w_1, u_1)$ we have

$$g'_{n,u_1^-}(x) = g'_n(u_1) - (x - u_1)^3 = g'_{n_1}(u_1) - (x - u_1)^3.$$

It is also clear that $g_{n,u_1^-}(x) \geq g_n(x)$.

Again we want to reduce our argument to Case I. Here is the idea of what we are planning to do: Due to cancellation we can forget the trajectory x_n for $n > n_1$ and when studying $g_{n,t}$ for $t \in [u_1, u_2)$ we can use the argument from Case I with $y_{n_1} = g_{n,u_1^-}(x_{n_1})$, $x_{n_1+1} = t$ and $y_{n_1+1} = g_{n,u_1^-}(t)$.

Furthermore, like in the earlier case, we can also obtain the monotonicity property

$$(2.1) \qquad g_{n,t'}(x) \geq g_{n,t}(x) \quad \text{for all } t, t' \in [u_1, u_2),\ t < t' \text{ and } x \in \mathbb{R}.$$

So assume $t \in [u_1, u_2)$ and set $g_{n,t}(t) = g_{n,u_1^-}(t) = g_n(t)$.

For $x \geq t$, or for $x \leq w_2$ set

$$g'_{n,t}(x) = g'_{n,u_1^-}(x) = g'_n(x) = g'_{n_1}(x).$$

This again implies $g_{n,t}(x) = g_n(x)$ for $x \geq t$. For $w_2 < x < t$ set

$$g'_{n,t}(x) = g'_{n,t}(t) - (x-t)^3 = g'_{n_1}(t) - (x-t)^3.$$

Now, as earlier, by using an argument similar to Case I (now applied for $g_{n_1}(t)$) one could see that $g_{n,t}(x) - g_{n,t}(t) - (g_{n_1}(x) - g_{n_1}(t)) \geq 0$ for $w_2 < x < t$ and hence $g_{n,t}(x) \geq g_{n,u_1^-}(x)$ for all $x \in \mathbb{R}$. The monotonicity property (2.1) can be established as well. Then one can define $g_{n,u_2^-}(x)$ and continue the argument on $[u_2, u_3)$. By induction we can prove (2.1) for all intervals $[u_{l-1}, u_l)$ for $l = 2, \ldots, j_0$ and setting $n_l = a(n, k_n - 2l)$ we have $g_{n,u_l^-}(x) = g'_{n_l}(x)$ on $\mathbb{R}$.

Hence we can define $g_{n,u_{j_0}^-}(x)$ and

$$g'_{n,u_{j_0}^-}(x) = g'_n(u_{j_0}) - (x - u_{j_0})^3 = g'_{n_{j_0}}(u_{j_0}) - (x - u_{j_0})^3$$

on (w_{j_0}, u_{j_0}) while $g'_{n,u_{j_0}^-}(x) = g'_{n_{j_0}}(x) = g'_n(x)$ for $x \notin (w_{j_0}, u_{j_0})$.

Set $w_{j_0+1} = x_{a(n,k_n-2j_0-1)}$ if $k_n - 2j_0 - 1 \geq 1$.

If $k_n - 2j_0 - 1 < 1$ then $x_{a(n,1)} = x_0 = 0 \in (x_n, x_{n+1}]$ and at step $n + 1$ we reset the trajectory. In this case $u_{j_0} = 0$ and we put $w_{j_0+1} = 0$. We also have $g'_{n,u_{j_0}^-}(x) = g'_0(x) = -x^3$ and $g_{n,u_{j_0}^-}(0) = g_{n,u_{j_0}}(u_{j_0}) = g_n(0)$.

Now we return to the proof of this lemma and no longer assume $u_{j_0} = 0$. For $t \in [u_{j_0}, x_{n+1}]$ we define $g_{n,t}(t) = g_n(t)$ and $g'_{n,t}(x) = g'_{n_{j_0}}(x) = g'_n(x)$ when $x \notin (w_{j_0+1}, t)$. When $x \in (w_{j_0+1}, t)$ set $g'_{n,t}(x) = g'_{n,t}(t) - (x-t)^3$. The monotonicity property (2.1) can be established as before. Hence

$$g_{n,x_{n+1}}(x) \geq g_{n,t}(x) \geq g_{n,u_{j_0}^-}(x) \geq \cdots \geq g_n(x) \quad \text{for all } x \in \mathbb{R}.$$

Now, $g_{n+1}(x_{n+1}) > g_n(x_{n+1}) = g_{n,x_{n+1}}(x_{n+1})$ and $g'_{n+1} = g'_{n,x_{n+1}}$ implies $g_{n+1} \geq g_n$.

In the proof of Lemma 4 we need the following estimate for the case when $u_{j_0} = 0$, that is, when we reset the trajectory:

$$(2.2) \qquad g_{n+1}(x) \geq g_{n,u_{j_0}^-}(x) = g_{n,u_{j_0}^-}(0) + \int_0^x -t^3 dt = g_n(0) - \frac{x^4}{4}$$

(we remark that this inequality holds in Case IIIb as well, provided that $u_{j_0} = 0$).

Case IIIb. Now we assume that at step $n+1$ the trajectory is changing direction. The other cases being similar we assume $x_n < x_{a(n,1)}$, $x_n < x_{n+1}$, $x_n < x_{n-1}$. Set $u_1 = x_{a(n,k_n-1)}$. Now $x_{n-1} \in (x_n, u_1]$ and due to cancellation $u_j = x_{a(n,k_n-2j+1)} \in (x_n, x_{n+1}]$ for $j = 1, \ldots, j_0$ with some $j_0 \geq 1$, see Figure 7. Set $w_j = x_{a(n,k_n-2j+2)}$, $j = 1, \ldots, j_0+1$. When $k_n - 2j_0 < 1$ which corresponds to the case when $u_{j_0} = 0$, that is, we reset the trajectory, we put $w_{j_0+1} = x_0 = 0$. Observe that $w_1 = x_n$.
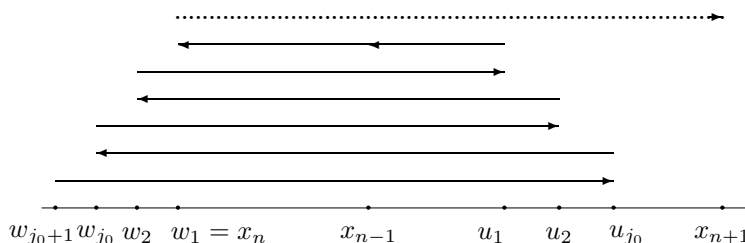

FIGURE 7: Case IIIb.

For $t \in [x_n, u_1)$ we define $g_{n,t}$ exactly as it was defined in Case IIIa. To verify $g_{n,t} \geq g_n$ now we need to use an argument similar to Case II instead of Case I. To show the monotonicity property (2.1) for $t, t' \in (x_n, u_1)$ for the first point $t < t'$ an argument similar to Case II, and then for t' the argument of Case IIIa, referring to Case I can be used. Now we have $g'_{n,u_1^-}(x) = g'_{n_1}(x)$ with $n_1 = a(n, k_n - 1)$ and we "drop down" to level n_1. On the intervals $[u_{l-1}, u_l)$, $l = 2, \ldots, j_0$ and on $[u_{j_0}, x_{n+1})$ we can argue exactly like in Case IIIa, that is, Case I can be used. ∎

Lemma 3. *If $x_n \in [-1, 1]$ and $x_n \to x^*$ then there exists $y^* \in \mathbb{R} \cup \{+\infty\}$ such that $y_n \to y^*$.*

Proof. Set $g_n(x^*) = y_n^*$. By Lemma 2, y_n^* is monotone increasing and hence there exists $y^* \in \mathbb{R} \cup +\infty$ such that $y_n^* \to y^*$. Since $|g'_n| \leq 1$ we also have $|y_n^* - y_n| = |g_n(x^*) - g_n(x_n)| \leq |x^* - x_n|$. This implies $|y^* - y_n| \leq |y^* - y_n^*| + |x^* - x_n| \to 0$ as $n \to \infty$. ∎

Lemma 4. *If $x_n \in [-1, 1]$, $\liminf x_n = x_* < \limsup x_n = x^*$ and $c = (x_* + x^*)/2$ then $g_n(c) \to \infty$. This, by the uniform Lipschitz property of g_n, implies $y_n \to \infty$ as well.*

Proof. First we assume $x_* < x^* < 0$. By definition of x_* and x^* we have infinitely many "zig-zags" between x_* and x^*. Using the cancellation law of active points we can choose infinitely many $n_0 < n_1 < n_2$ such that setting $u = x_{n_0}$, $v = x_{n_1}$, and $w = x_{n_2}$ we have $u, v, w < 0$, $v = x_{n_1} \leq x_n < u = x_{n_0}$ and $x_n < w = x_{n_2}$ for all $n_0 < n < n_2$, furthermore, $|u - x^*| < 0.01|x^* - x_*|^4$, $|v - x_*| < 0.01|x^* - x_*|^4$, $|w - x^*| < 0.01|x^* - x_*|^4$, and $n_0 \in \mathcal{A}_{n_1}$.

Then, as x_n moves between u and v all points $x_{n'}$ with $n_0 < n' < n_1$ are deactivated by the time g_{n_1} is defined. Hence $a(n_1, k_{n_1}) = n_1$ and $a(n_1, k_{n_1} - 1) = n_0$. Now

$$g'_{n_1}(x) = g'_{n_1}(v) - (x - v)^3$$

for all $x \in (u, v]$, therefore

$$g_{n_1}(c) = g_{n_1}(v) + \int_v^c g'_{n_1}(t)dt = g_{n_1}(v) + g'_{n_1}(v)(c - v) - \frac{(c - v)^4}{4},$$

similarly

$$g_{n_1}(u) = g_{n_1}(v) + g'_{n_1}(v)(u - v) - \frac{(u - v)^4}{4},$$

by the Lipschitz property of g_{n_1}, we also have

$$g_{n_1}(w) \geq g_{n_1}(u) - |u - w| \geq g_{n_1}(u) - 0.01|x^* - x_*|^4.$$

As x_n moves between v and w all points $x_{n'}$ with $n_1 < n' < n_2$ are deactivated by the time g_{n_2} is defined. Hence

$$a(n_2, k_{n_2}) = n_2 \qquad \text{and} \qquad a(n_2, k_{n_2} - 1) \leq n_1.$$

First we suppose that the active point $u = x_{n_0}$ is not deactivated at step n_2, that is, $a(n_2, k_{n_2} - 1) = n_1$ and $w = x_{n_2} < u = x_{n_0}$. Therefore,

$$g'_{n_2}(w) = g'_{n_2}(v) - (w - v)^3 = g'_{n_1}(v) - (w - v)^3$$

and

$$g'_{n_2}(x) = g'_{n_2}(w) - (x - w)^3$$

for all $x \in (v, w]$, hence

$$
\begin{aligned}
g_{n_2}(c) = {} & g_{n_2}(w) + \int_w^c g'_{n_2}(t)dt = g_{n_2}(w) + g'_{n_2}(w)(c - w) - \frac{(c - w)^4}{4} \\
\geq {} & g_{n_1}(w) + g'_{n_2}(w)(c - w) - \frac{(c - w)^4}{4} \\
\geq {} & g_{n_1}(u) - 0.01|x^* - x_*|^4 + g'_{n_2}(w)(c - w) - \frac{(c - w)^4}{4} \\
= {} & g_{n_1}(v) + g'_{n_1}(v)(u - v) - \frac{(u - v)^4}{4} - 0.01|x^* - x_*|^4 \\
& + (g'_{n_1}(v) - (w - v)^3)(c - w) - \frac{(c - w)^4}{4}.
\end{aligned}
$$

Now,

$$g_{n_2}(c) - g_{n_1}(c) \geq$$

$$\geq g'_{n_1}(v)(u-v) - \frac{(u-v)^4}{4} - 0.01|x^* - x_*|^4 +$$

$$+ (g'_{n_1}(v) - (w-v)^3)(c-w) - \frac{(c-w)^4}{4} - g'_{n_1}(v)(c-v) + \frac{(c-v)^4}{4}$$

$$= g'_{n_1}(v)(u-w) - \frac{(u-v)^4}{4} - 0.01|x^* - x_*|^4$$

$$- (w-v)^3(c-w) - \frac{(c-w)^4}{4} + \frac{(c-v)^4}{4}$$

$$= (w-c)(w-v)^3 + g'_{n_1}(v)(u-w) - 0.01|x^* - x_*|^4 - \frac{(u-v)^4}{4}$$

$$- \frac{(c-w)^4}{4} + \frac{(c-v)^4}{4} \geq 0.49|x^* - x_*| \cdot 0.98|x^* - x_*|^3$$

$$- 0.02|x^* - x_*|^4 - 0.01|x^* - x_*|^4 - \frac{1.02^4|x^* - x_*|^4}{4} - \frac{0.51^4|x^* - x_*|^4}{4}$$

$$> 0.1|x^* - x_*|^4.$$

Using that $g_n(c)$ is monotone increasing as $n \to \infty$ and it increases infinitely often by $0.1|x^* - x_*|^4$ we obtain $g_n(c) \to \infty$.

Suppose now that the active point $u = x_{n_0}$ is deactivated at step n_2. This implies $a(n_2, k_{n_2} - 1) < n_0 < n_1$, that is, $u = x_{n_0} \leq w = x_{n_2}$. From $x_{n_2-1} < u < 0$ it follows that $u = x_{n_0} \in \mathcal{A}_{n_2-1}$. Now, using the notation of Case III in Lemma 2 one can show that

$$g'_{n_2-1,u}(u) = g'_{n_2-1,u}(v) - (u-v)^3 = g'_{n_1}(v) - (u-v)^3$$

and $g'_{n_2-1,u}(x) = g'_{n_2-1,u}(u) - (x-u)^3$ for all $x \in (v, u)$. The above argument used now with g_{n_2} replaced by $g_{n_2-1,u}$ and w replaced by u shows that $g_{n_2-1,u}(c) > g_{n_1}(c) + 0.1|x^* - x_*|^4$. Since

$$g_{n_2}(c) \geq g_{n_2-1,u}(c) > g_{n_1}(c) + 0.1|x^* - x_*|^4$$

this completes the proof of the case $x_* < x^* < 0$.

The case $0 < x_* < x^*$ is similar to the previous one.

So assume $x_* \leq 0 \leq x^*$.

Without limiting generality we can suppose $x_* < 0$. If there are only finitely many times when the trajectory is being reset then x^* should equal 0 and $x_n < 0$ for all $n \geq N$ for some N. In this case the argument used for the case $x_* < x^* < 0$ works. So finally we assume that the trajectory is being reset infinitely many times. Then we can choose infinitely many $n_0 < n_1 < n_2$ such that $x_{n_0}, x_{n_2} > 0$, $x_{n_1} < 0$, $|x_{n_1} - x_*| < 0.001|x_*|^4$, $x_{n_1} \leq x_n < 0$ for $n_0 < n < n_2$.

This means that the trajectory is being reset at steps n_0+1 and n_2 but is not being reset for steps $n_0 + 1 < n < n_2$. From $x_{n_1} \leq x_n$ for $n = n_0, \ldots, n_2$ it also follows that $x_{n_1} \in \mathcal{A}_{n_2-1}$ and due to cancellation $\mathcal{A}_{n_1} = \{0, n_1\}$. Now (2.2) implies

$$g_{n_1}(x_{n_1}) \geq g_{n_0+1}(x_{n_1}) \geq g_{n_0}(x_{n_1}) = g_{n_0}(0) - \frac{x_{n_1}^4}{4}.$$

By our construction we also have $g'_{n_1}(x_{n_1}) = -x_{n_1}^3$ and

$$g_{n_1}(0) = g_{n_1}(x_{n_1}) + \int_{x_{n_1}}^0 g'_{n_1}(x)dx = g_{n_1}(x_{n_1}) + \int_{x_{n_1}}^0 (g'_{n_1}(x_{n_1}) - (x - x_{n_1})^3)dx$$

$$= g_{n_1}(x_{n_1}) + (-x_{n_1})(-x_{n_1}^3) - \frac{x_{n_1}^4}{4}$$

$$\geq g_{n_0}(0) + \frac{x_{n_1}^4}{2} \geq g_{n_0}(0) + 0.4|x_*|^4.$$

Hence

$$g_{n_2}(0) \geq g_{n_1}(0) \geq g_{n_0}(0) + 0.4|x_*|^4.$$

Since $g_n(0)$ is monotone increasing as $n \to \infty$ and increases infinitely often by $0.4|x_*|^4$ we obtain $g_n(0) \to \infty$. By the Lipschitz property of g_n we have $g_n(c) \to \infty$ as well. ∎

2.2. Construction of the example

We start with a function $h_{-1}(x, y) = y$. At step n, $(n = 0, 1, \ldots)$ we are choosing some disjoint open squares $B_{n,k}$, called perturbation blocks. For each perturbation block we choose a perturbation function $\phi_{B_{n,k}}$ such that this function is zero outside $B_{n,k}$. This way the sum

$$h_n(x, y) = \sum_k \phi_{B_{n,k}}(x, y)$$

converges everywhere.

Our function f will be defined as

$$f(x, y) \overset{\text{def}}{=} \sum_{n=-1}^{\infty} h_n(x, y).$$

For the partial sums we use the notation

$$f_n(x, y) = \sum_{k=-1}^{n} h_k(x, y).$$

Assume $n = 0, 1, \ldots,$ is fixed. We will fix a constant $c_n > 0$ so that $c_n \to 0$ as $n \to \infty$.

Next we define our perturbation blocks, B and functions ϕ_B. Assume that the open square B is centered at $\mathbf{o}_B$ and its sides are parallel to the perpendicular unit vectors $\mathbf{v}_B$ and $\mathbf{w}_B$, see Figure 8.
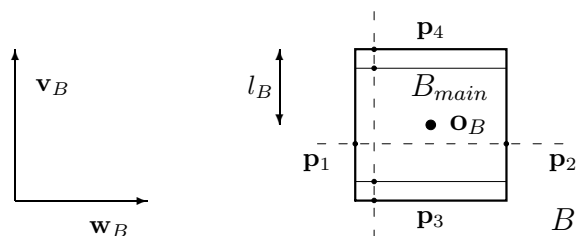

FIGURE 8: The perturbation block B.

The first vector $\mathbf{v}_B$ is called the direction vector of the block and the angle between this vector and and $(0,1)$ will be in $[-\pi/4, \pi/4]$ and we choose $\mathbf{w}_B$ so that its first component is positive. We assume that $B = \{\mathbf{o}_B + \alpha\mathbf{v}_B + \beta\mathbf{w}_B$ for $|\alpha|, |\beta| < l_B\}$.

The differentiable perturbation function ϕ_B used at level n will have the following properties:

It will be continuously differentiable on B. For $\mathbf{p} \in B$ we have

$$(2.3) \qquad |\phi_B(\mathbf{p})| \leq \frac{(dist(\mathbf{p}, \partial B))^2}{2^{n+1}},$$

and $\phi_B = 0$ outside B (in the displayed equation, ∂B denotes the boundary of B). We assume $|\nabla\phi_B| \leq 2c_n$ everywhere. The main region of B is defined as

$$B_{main} = \{\mathbf{o}_B + \alpha\mathbf{v}_B + \beta\mathbf{w}_B \ \text{ for } |\alpha| < (1 - 2^{-n-2})l_B, \ |\beta| < l_B\}$$

and $B_{trans} = B \setminus B_{main}$ is called the transitional region. If $\mathbf{p} \in B_{main}$ we choose ϕ_B so that $\nabla\phi_B(\mathbf{p})$ is parallel to $\mathbf{w}_B$ (this gradient can be zero). In fact, the cross section of ϕ_B in the direction of $\mathbf{w}_B$ will consist of piecewise linear parts with some sanded corners, see Figure 9.
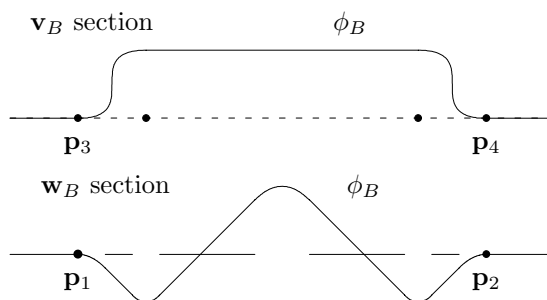

FIGURE 9: The $\mathbf{w}_B$ and $\mathbf{v}_B$ sections of ϕ_B.

All these $\mathbf{w}_B$ cross sections are the same in B_{main} and by choosing a sufficiently large number of zig-zags one can ensure that (2.3) holds. We choose ϕ_B so that the linear parts dominate. This means that there exists a set $B_{good} \subset B_{main}$ such that

$$\lambda_2(B_{good}) \geq (1 - 2^{-n})\lambda_2(B) \quad \text{and} \quad |\partial_{\mathbf{w}_B}\phi_B(\mathbf{p})| = ||\nabla\phi_B(\mathbf{p})|| = c_n$$

if $\mathbf{p} \in B_{good}$, that is, B_{good} consists of those points where the cross section of ϕ in the direction of $\mathbf{w}_B$ is linear and of slope $\pm c_n$, these points are marked by a solid line on the segment $[\mathbf{p}_1, \mathbf{p}_2]$ of Figure 9. Denote by B_{zero} the set of those $\mathbf{p} \in B$ for which $\nabla\phi_B(\mathbf{p}) = (0,0)$. We also assume that $\lambda_2(B_{zero}) = 0$, in fact, one can choose ϕ_B so that B_{zero} is the union of finitely many line segments pointing in the direction $\mathbf{v}_B$.

If $\mathbf{p} \in B_{trans}$ we assume that

$$(2.4) \qquad\qquad\qquad |\partial_{\mathbf{v}_B}\phi_B(\mathbf{p})| < \frac{0.001}{2^{n+1}}.$$

What we want to use is the $\mathbf{w}_B$ component of ∇f, the small transitional component $\partial_{\mathbf{v}_B}\phi_B(\mathbf{p})$ estimated by (2.4) just gives a small error in the end. Restriction (2.4) is again a restriction on the heights of the amplitudes of the $\mathbf{w}_B$ sections, but this just means that we might need more zig-zags.

It is not difficult to see that given a perturbation block the function ϕ_B satisfying the above assumptions can be chosen.

Since the perturbation blocks are nonoverlapping at level n any point $\mathbf{p} \in G$ can belong to at most one perturbation block, which will be denoted by $B_n(\mathbf{p})$ and the corresponding perturbation function will be $\phi_{B_n(\mathbf{p})}$. In case $\mathbf{p}$ is not belonging to a perturbation block at level n then we set $B_n(\mathbf{p}) = \emptyset$ and $\phi_{B_n(\mathbf{p})}$ will be the identically zero function. This will also imply by the rules we follow in our construction that for any $k > n$ the point $\mathbf{p}$ will not belong to any (nonempty) $B_k(\mathbf{p})$ and we will have $\phi_{B_k(\mathbf{p})} = 0$ everywhere. Now

$$h_n(\mathbf{p}) \overset{\text{def}}{=} \phi_{B_n(\mathbf{p})}(\mathbf{p})$$

and we will show that

$$(2.5) \qquad \nabla f(\mathbf{p}) = (0,1) + \sum_{n=0}^{\infty} \nabla\phi_{B_n(\mathbf{p})}(\mathbf{p}) = \sum_{n=-1}^{\infty} \nabla h_n(\mathbf{p}).$$

It is clear that

$$\nabla f_n(\mathbf{p}) = (0,1) + \sum_{k=0}^{n} \nabla\phi_{B_k(\mathbf{p})}(\mathbf{p}).$$

In order to do our estimates later we will have to introduce "reduced sums". This means that we are not considering the "error term" coming from the transitional components, that is, we consider

$$\nabla_r f_n(\mathbf{p}) = (0, 1) + \sum_{k=0}^{n} \partial_{\mathbf{w}_{B_k(\mathbf{p})}} \phi_{B_k(\mathbf{p})}(\mathbf{p}) \mathbf{w}_{B_k(\mathbf{p})},$$

and

$$\nabla_r f(\mathbf{p}) = (0, 1) + \sum_{n=0}^{\infty} \partial_{\mathbf{w}_{B_n(\mathbf{p})}} \phi_{B_n(\mathbf{p})}(\mathbf{p}) \mathbf{w}_{B_n(\mathbf{p})}.$$

In Section 2.1 we were talking about trajectories in the gradient space. To generate these trajectories assume that $B_n(\mathbf{p}) \neq \emptyset$. This implies that $B_k(\mathbf{p}) \neq \emptyset$ for $k < n$ and hence the center $\mathbf{o}_{B_k}(\mathbf{p})$ is given. We define the "reduced centered sum" at level n as:

$$\nabla_{rc,n} f(\mathbf{p}) = (0, 1) + \sum_{k=0}^{n-1} \partial_{\mathbf{w}_{B_k(\mathbf{p})}} \phi_{B_k(\mathbf{p})}(\mathbf{o}_{B_{k+1}(\mathbf{p})}) \mathbf{w}_{B_k(\mathbf{p})}.$$

Now set $x_0(\mathbf{p}) = 0$ and $x_n(\mathbf{p}) = \pi_x(\nabla_{rc,n} f(\mathbf{p}))$, (by π_x we denote the projection onto the x-axis). Our construction will imply that $y_0(\mathbf{p}) = 1$ and $y_n(\mathbf{p}) = \pi_y(\nabla_{rc,n} f(\mathbf{p}))$. In case $B_n(\mathbf{p}) \neq \emptyset$ for all $n \in \mathbb{N}$ then we can define the infinite reduced centered sum

$$\nabla_{rc} f(\mathbf{p}) = (0, 1) + \sum_{n=0}^{\infty} \partial_{\mathbf{w}_{B_n(\mathbf{p})}} \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})}) \mathbf{w}_{B_n(\mathbf{p})}.$$

Finally, during the definition of the perturbation blocks by induction we will need a "modified reduced centered sum" at level $n + 1$ which is

$$\nabla_{mrc,n+1} f(\mathbf{p}) = \nabla_{rc,n} f(\mathbf{p}) + \partial_{\mathbf{w}_{B_n(\mathbf{p})}} \phi_{B_n(\mathbf{p})}(\mathbf{p}) \mathbf{w}_{B_n(\mathbf{p})}.$$

Observe that if $\mathbf{p} = \mathbf{o}_{B_{n+1}(\mathbf{p})}$ then $\nabla_{mrc,n+1} f(\mathbf{p}) = \nabla_{rc,n+1} f(\mathbf{p})$, actually the modified reduced centered sums are used to define the centers of the perturbation blocks at level $n + 1$.

Next we show by mathematical induction how the perturbation blocks at level n are defined.

We set $c_0 = 0.004$. We will use four perturbation blocks at level 0, these are: $B_{0,1} = (0, 1) \times (0, 1)$, $B_{0,2} = (-1, 0) \times (0, 1)$, $B_{0,3} = (-1, 0) \times (-1, 0)$, and $B_{0,4} = (0, 1) \times (-1, 0)$. We have $\mathbf{v}_{B_{0,k}} = (0, 1)$ and $\mathbf{w}_{B_{0,k}} = (1, 0)$ for $k = 1, \ldots, 4$. We choose the functions $\phi_{B_{0,k}}$, $k = 1, \ldots, 4$. By our assumptions about the functions $\phi_{B_{0,k}}$ the set $Z_0 = \cup_k B_{zero,0,k}$ is of measure zero and is relatively closed in $\cup_k B_{0,k}$ so setting $G_0 = \cup_k B_{0,k} \setminus Z_0$ we define an open set and $G \setminus G_0$ is of measure zero. For $\mathbf{p} \in G_0$ set $x_0(\mathbf{p}) = 0 = \pi_x(\nabla_{rc,0} f(\mathbf{p}))$, and $g_{0,\mathbf{p}}(x) = 1 - \frac{x^4}{4}$. Clearly, ∇f_0 is continuous on G_0.

Assume $n \geq 0$, the constant c_n and the open set G_n are given and for all $\mathbf{p} \in G_n$ the trajectory $\{x_0(\mathbf{p}), \ldots, x_n(\mathbf{p})\} \subset [-1, 1]$, the perturbation blocks $B_0(\mathbf{p}), \ldots, B_n(\mathbf{p})$ and the functions $g_{n,\mathbf{p}}$ are defined, furthermore $|g'_{n,\mathbf{p}}(x)| \leq 1$ for all $x \in [-1, 1]$ and ∇f_n is continuous on G_n.

Set

$$G_n^* = \{\mathbf{p} \in G_n : \nabla f_n(\mathbf{p}) \in \Omega_0\}.$$

We use Vitali's covering theorem to select the centers of the perturbation blocks. (The version of Vitali's covering theorem we need is Theorem 2.8.17 in [9]. This theorem is applicable if we take fine coverings of sets by squares which are of sides not necessarily parallel to the coordinate axes and the measure we use is λ_2.) So we need to define a Vitali cover first.

For $\mathbf{p} \in G_n^*$ we define $\mathbf{v}_{\mathbf{p},n+1}$ as the "upward" normal vector of $g_{n,\mathbf{p}}$ at the point with x coordinate $x_{n+1}^*(\mathbf{p}) = \pi_x(\nabla_{mrc,n+1} f(\mathbf{p}))$. The vector $\mathbf{w}_{\mathbf{p},n+1}$ will be the unit vector which has positive x-coordinate and which is perpendicular to $\mathbf{v}_{\mathbf{p},n+1}$. The assumption that $|g'_{n,\mathbf{p}}| \leq 1$ on $[-1, 1]$ implies that the angle between $(0, 1)$ and $\mathbf{v}_{\mathbf{p},n+1}$ is in $[-\pi/4, \pi/4]$. Similarly the angle between $(1, 0)$ and $\mathbf{w}_{\mathbf{p},n+1}$ is in $[-\pi/4, \pi/4]$. We choose $\delta_{0,n+1,\mathbf{p}} > 0$ such that for any $\delta \in (0, \delta_{0,n+1,\mathbf{p}})$ the square

$$Q_{\delta,n+1}(\mathbf{p}) = \{\mathbf{p} + \alpha \mathbf{v}_{\mathbf{p},n+1} + \beta \mathbf{w}_{\mathbf{p},n+1} : |\alpha|, |\beta| < \delta\}$$

is a subset of $G_n \cap B_n(\mathbf{p})$, furthermore for any $\mathbf{q} \in Q_{\delta,n+1}(\mathbf{p})$ we have

$$(2.6) \qquad ||\nabla \phi_{B_n(\mathbf{p})}(\mathbf{p}) - \nabla \phi_{B_n(\mathbf{p})}(\mathbf{q})|| < \frac{0.001}{2^{n+1}},$$

and

$$(2.7) \qquad ||\nabla f_n(\mathbf{p}) - \nabla f_n(\mathbf{q})|| < \frac{0.001}{2^{n+1}}.$$

The squares $Q_{\delta,n+1}(\mathbf{p})$ form a Vitali covering of G_n^* and hence by Vitali's covering theorem we can choose a disjoint system of these squares denoted by $B_{n+1,k} = Q_{\delta_k,n+1}(\mathbf{p}_{k,n+1})$ such that setting $G_{n+0.5} = \cup_k B_{n+1,k}$ almost every point of G_n^* belongs to $G_{n+0.5}$. Clearly, $\mathbf{o}_{B_{n+1,k}} = \mathbf{p}_{k,n+1}$.

We choose the perturbation functions $\phi_{B_{n+1,k}}$ on the perturbation blocks $B_{n+1,k}$. Now $Z_{n+1} = \cup_k B_{zero,n+1,k}$ is of measure zero and is relatively closed in $G_{n+0.5}$. Hence if we put $G_{n+1} = G_{n+0.5} \setminus Z_{n+1}$ we obtain an open set.

Furthermore $G_n \setminus G_{n+1} = F_{n+1} \cup E_{n+1}$, where F_{n+1} is of measure zero and $\nabla f_n(\mathbf{p}) \notin \Omega_0$ for $\mathbf{p} \in E_{n+1}$.

It is also clear that ∇f_{n+1} is continuous on G_{n+1} and from (2.7) it follows that $\nabla f_{n+1}(\mathbf{p}) \in \Omega_2$ for all $\mathbf{p} \in G_{n+1}$.

Set $c_{n+1} = 0.001\lambda_2(G_{n+1})$. Now for $\mathbf{p} \in G_{n+1}$ we denote by $B_{n+1}(\mathbf{p})$ the unique block of the form $B_{n+1,k}$ which contains $\mathbf{p}$. We set

$$x_{n+1}(\mathbf{p}) = \pi_x(\nabla_{rc,n+1}f(\mathbf{p}))$$

and define $g_{n+1,\mathbf{p}}$ by using the trajectory $\{x_0(\mathbf{p}), \ldots, x_{n+1}(\mathbf{p})\}$ and the functions $g_{0,\mathbf{p}}, \ldots, g_{n,\mathbf{p}}$ as it was explained in Section 2.1. Observe that by our definition of $\nabla_{rc,n+1}f$ and by the definition of the functions $g_{k,\mathbf{p}}$ we have

$$y_{n+1}(\mathbf{p}) = g_{n+1,\mathbf{p}}(x_{n+1}(\mathbf{p})) = \pi_y(\nabla_{rc,n+1}f(\mathbf{p})).$$

Now Lemmas 2-4 can be used.

Next we show that $\lambda_2(G_n) \to 0$ as $n \to \infty$.

Proceeding towards a contradiction assume $\lambda_2(G_n) > \gamma > 0$ for all $n \in \mathbb{N}$. Then $G_{n,good} = \cup_k B_{good,n,k}$ satisfies

$$\lambda_2(G_{n,good}) \geq (1 - 2^{-n})\lambda_2(G_n).$$

Hence $\lambda(\cap_{n=1}^\infty G_{n,good}) > \gamma/2$ and there exists $\mathbf{p} \in \cap_{n=1}^\infty G_{n,good}$. This implies $\nabla f_n(\mathbf{p}) \in \Omega_2$ and $\pi_x(\nabla f_n(\mathbf{p})) \in [-0.51, 0.51]$. Since $c_n \not\to 0$ one can see that the x coordinates of the vectors

$$\pm c_n \mathbf{w}_{\mathbf{p},n} = \partial_{\mathbf{w}_{B_n(\mathbf{p})}}\phi_{B_n(\mathbf{p})}(\mathbf{p})\mathbf{w}_{B_n(\mathbf{p})} = \nabla h_n(\mathbf{p})$$

are bounded away from zero and hence by (2.6), $x_n(\mathbf{p})$ diverges. This implies by Lemma 4 that $y_n(\mathbf{p}) \to \infty$. By using the estimate (2.6) with $\mathbf{p} := \mathbf{o}_{B_{n+1}(\mathbf{p})}$ and $\mathbf{q} := \mathbf{p}$ this would imply $\pi_y(\nabla f_n(\mathbf{p})) \to \infty$ which contradicts that $\nabla f_n(\mathbf{p}) \in \Omega_2$ for all n.

Now we show that f is differentiable on G.

If $\mathbf{p} \notin \cap_{n=0}^\infty G_n$ then let $n_0(\mathbf{p}) = \min\{k : \mathbf{p} \notin G_k\} - 1$. This means $f(\mathbf{p}) = f_{n_0(\mathbf{p})}(\mathbf{p})$. Since $\mathbf{p} \notin G_n$ for all $n > n_0$, for any $\mathbf{q}$, by using (2.3) we obtain $|h_n(\mathbf{q}) - h_n(\mathbf{p})| \leq ||\mathbf{q} - \mathbf{p}||^2/2^n$, which implies $\nabla f(\mathbf{p}) = \nabla f_{n_0(\mathbf{p})}(\mathbf{p})$. Especially, we have $\nabla f((0,0)) = (0,1)$. Since $\lambda_2(\cap_{n=0}^\infty G_n) = 0$ almost every $\mathbf{p} \in G$ belongs to $\cup_{n=0}^\infty E_n$ and for these $\mathbf{p}$ we have

$$\nabla f(\mathbf{p}) = \nabla f_{n_0(\mathbf{p})}(\mathbf{p}) \notin \Omega_0 \supset \Omega_1.$$

Now assume $\mathbf{p} \in \cap_{n=0}^\infty G_n$. This implies that $\nabla f_n(\mathbf{p})$ and hence $x_n(\mathbf{p})$, $y_n(\mathbf{p})$ are bounded. Then there exists x^* such that $x_n(\mathbf{p}) \to x^*$. Otherwise by Lemma 4, $y_n(\mathbf{p}) \to \infty$ which is impossible. By Lemma 3 there exists y^* such that $y_n(\mathbf{p}) \to y^*$ and

$$\nabla_{rc,n}f(\mathbf{p}) \to (x^*, y^*) = \nabla_{rc}f(\mathbf{p}).$$

If $\mathbf{o}_{B_{n+1}(\mathbf{p})} \in B_{main,n}(\mathbf{p})$ then

$$\partial \mathbf{w}_{B_n(\mathbf{p})} \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})}) \mathbf{w}_{B_n(\mathbf{p})} = \nabla \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})}).$$

If $\mathbf{o}_{B_{n+1}(\mathbf{p})} \in B_{trans,n}(\mathbf{p})$ then $\nabla \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})})$ has a small component pointing in the direction of $\mathbf{v}_{B_n(\mathbf{p})}$ and we can estimate this by using (2.4).

So,

$$(2.8) \qquad ||\partial \mathbf{w}_{B_n(\mathbf{p})} \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})}) \mathbf{w}_{B_n(\mathbf{p})} - \nabla \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})})|| \le \frac{0.001}{2^{n+1}}.$$

Next by the choice of $B_{n+1}(\mathbf{p})$ and (2.6) we have

$$(2.9) \qquad ||\nabla \phi_{B_n(\mathbf{p})}(\mathbf{p}) - \nabla \phi_{B_n(\mathbf{p})}(\mathbf{o}_{B_{n+1}(\mathbf{p})})|| < \frac{0.001}{2^{n+1}}.$$

From (2.8) and (2.9) it follows that

$$(2.10) \qquad ||\nabla f_n(\mathbf{p}) - \nabla_{rc,n+1} f(\mathbf{p})|| < \sum_{k=0}^{n} \frac{0.002}{2^{k+1}} < 0.002,$$

furthermore, one can also see that $\nabla f_n(\mathbf{p})$ forms a Cauchy sequence and hence converges to $\nabla f(\mathbf{p})$ defined in (2.5). Next we verify that $\nabla f(\mathbf{p})$ is indeed the gradient of f at $\mathbf{p}$. From (2.8) and (2.9) we can also deduce that for all $\mathbf{p} \in G_n$

$$(2.11) \qquad ||(\nabla f(\mathbf{p}) - \nabla f_n(\mathbf{p})) - (\nabla_{rc} f(\mathbf{p}) - \nabla_{rc,n+1} f(\mathbf{p}))|| < \frac{0.001}{2^n}.$$

Given $\epsilon > 0$ choose N such that for all $n \ge N$ we have

$$||\nabla_{rc} f(\mathbf{p}) - \nabla_{rc,n+1} f(\mathbf{p})|| < \frac{\epsilon}{3} \qquad \text{and} \qquad \frac{0.001}{2^n} < \frac{\epsilon}{3}.$$

Then

$$(2.12) \qquad ||\nabla f(\mathbf{p}) - \nabla f_n(\mathbf{p})|| < \frac{2\epsilon}{3} \text{ for all } n \ge N.$$

To show that f is differentiable at $\mathbf{p}$ we need to estimate

$$f(\mathbf{q}) - f(\mathbf{p}) - \nabla f(\mathbf{p}) \cdot (\mathbf{q} - \mathbf{p}).$$

Assume $n \ge N$, $\mathbf{p} \in G_n$, $\mathbf{p}, \mathbf{q} \in B_k(\mathbf{p})$ for $k \le n$, and $\mathbf{q} \notin B_{n+1}(\mathbf{p})$. Then by (2.3) used for all $k \ge n+1$ and added together

$$||f(\mathbf{q}) - f(\mathbf{p}) - (f_n(\mathbf{q}) - f_n(\mathbf{p}))|| \le 2\frac{0.001}{2^{n+1}}||\mathbf{q} - \mathbf{p}||^2.$$

By (2.12),

$$||\nabla f(\mathbf{p}) - \nabla f_n(\mathbf{p})|| \cdot ||\mathbf{q} - \mathbf{p}|| < \frac{2\epsilon}{3}||\mathbf{q} - \mathbf{p}||.$$

Using $||\nabla h_n|| \leq 2c_n$ and the Mean Value Theorem

$$||h_n(\mathbf{q}) - h_n(\mathbf{p}) - \nabla h_n(\mathbf{p}) \cdot (\mathbf{q} - \mathbf{p})|| \leq 4c_n||\mathbf{q} - \mathbf{p}||.$$

By (2.7) used with $n - 1$ and by the Mean Value Theorem

$$||f_{n-1}(\mathbf{q}) - f_{n-1}(\mathbf{p}) - \nabla f_{n-1}(\mathbf{p}) \cdot (\mathbf{q} - \mathbf{p})|| \leq \frac{0.001}{2^n}||\mathbf{q} - \mathbf{p}||.$$

Hence,

$$||f_n(\mathbf{q}) - f_n(\mathbf{p}) - \nabla f_n(\mathbf{p}) \cdot (\mathbf{q} - \mathbf{p})|| \leq (4c_n + \frac{0.001}{2^n})||\mathbf{q} - \mathbf{p}||$$

and

$$||f(\mathbf{q}) - f(\mathbf{p}) - \nabla f(\mathbf{p}) \cdot (\mathbf{q} - \mathbf{p})|| < (4c_n + \frac{0.001}{2^n} + \frac{2\epsilon}{3})||\mathbf{q} - \mathbf{p}|| + 2\frac{0.001}{2^{n+1}}||\mathbf{q} - \mathbf{p}||^2.$$

Since $c_n \to 0$ as $n \to \infty$ this implies that f is differentiable at $\mathbf{p}$. This concludes the proof of Theorem 1.

References

[1] BRUCKS, K. M. AND BUCZOLICH, Z.: Trajectory of the turning point is dense for a co-σ-porous set of tent maps. *Fund. Math.* **165** (2000), no. 2, 95–123.

[2] BUCZOLICH, Z.: The n-dimensional gradient has the 1-dimensional Denjoy-Clarkson property. *Real Anal. Exchange*, **18** (1992/93), no. 1, 221–224.

[3] BUCZOLICH, Z.: Level sets of functions $f(x, y)$ with nonvanishing gradient. *J. Math. Anal. Appl.*, **185** (1994), no. 1, 27–35.

[4] BUCZOLICH, Z.: Approximate continuity points of derivatives of functions of several variables. *Acta Math. Hungar.* **67** (1995), no. 3, 229–235.

[5] BUCZOLICH, Z.: Another note on the gradient problem of C. E. Weil. *Real Anal. Exchange*, **22** (1996/97), no. 2, 775–784.

[6] BUCZOLICH, Z.: Functions of two variables with large tangent plane sets. *J. Math. Anal. Appl.*, **220** (1998), no. 2, 562–570.

[7] CLARKSON, J. A.: A property of derivatives. *Bull. Amer. Math. Soc.* **53** (1947), 124–125.

[8] DENJOY, A.: Sur une proprieté des fonctions dérivées. *Enseignement Math.* **18** (1916), 320–328.

[9] FEDERER, H.: *Geometric measure theory*. Die Grundlehren der matematischen Wissenschaften **153**. Springer-Verlag, New York, 1969.

[10] HOLICKÝ, P., MALÝ, J., ZAJÍČEK, L. AND WEIL, C. E.: A note on the gradient problem. *Real Anal. Exchange*, **22** (1996/97), no. 1, 225–235.

[11] PETRUSKA, G.: Derivatives take every value on the set of approximate continuity points. *Acta Math. Hungar.* **42** (1983) no. 3-4, 355–360.

[12] WEIL, C. E.: On properties of derivatives. *Trans. Amer. Math. Soc.* **114** (1965), 363–376.

[13] Query 1. *Real Anal. Exchange*, **16** (1990/91), no. 1, 373.

Zoltán Buczolich
Department of Analysis
Eötvös Loránd University
Pázmány Péter Sétány 1/c
1117 Budapest, Hungary
buczo@cs.elte.hu