On maximum functions with a dense set of points of non-differentiability

R. Henrion

Introduction

There are various reasons for investigating differentiability properties of maximum functions of the type $g^{max}(z) := \max_{x\in K} g(x,z)$, where $K \subseteq \mathbb{R}^m$ is a a compact subset and $g \in C^1(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R})$. For instance, by g^{max} one can understand an optimal value function of some n- parametric optimization problem (with z as parameter). On the other hand there is a close relation to smoothness of boundaries of constraint sets in finite or semi-infinite optimization, depending on whether the index set K is finite or not: if the Mangasarian-Fromovitz constraint qualification is satisfied in its conventional or extended (see [1]) version respectively, then these boundaries may be locally described by the graphs of such maximum functions (see[2]). For a measure-theoretic study of nonsmoothness of maximum functions we refer to [3].

Since g^{max} is locally lipschitzian (see [4]) it has to be differentiable on a subset of $\mathbb{R}^n$, the complement of which has Lebesgue-measure zero by Rademacher's theorem. In contrast to this one can easily construct locally lipschitzian functions being non-differentiable on a dense subset of $\mathbb{R}^n$ (see [5]). For the special case of maximum functions one has to distinguish between finite and infinite sets K. In the first case it is immediately verified that there exists an open and dense subset of $\mathbb{R}^n$ on which g^{max} is differentiable, hence the set of non-differentiability cannot be dense. As a consequence, the Mangasarian-Fromovitz constraint qualification implies some nice structure of constraint sets since their boundary contain a C^1-manifold as an open and dense subset (see [2]). If, however, K is infinite then again counter-examples can be constructed to show that g^{max} may be non-differentiable on a dense subset of $\mathbb{R}^n$. Accordingly, in semi-infinite optimization constraint sets have - in spite of some constraint qualification - a much more complicated structure. Concerning the generality of such counter-examples it turns out that, depending on the geometric structure of K, they are even typical in a certain sense. More precisely, as the main result we shall prove that there is a dense set of functions g (in a topology to be specified below) for which the corresponding maximum functions g^{max} have a dense set of non-differentiability, provided K is a compact differentiable manifold (e.g. a sphere or a torus). The following example makes evident the impact of geometric structure of K on this result:

Example 1 *Let $K = [0,1]$ and $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x,z) := xz$. Then there is an open neighborhood $U(1)$, such that the maximum of $g(\cdot,z)$ over K is attained at $x = 1$ for all $z \in U$. Consequently, for $z \in U$ one has $g^{max}(z) = z$, hence g^{max} is differentiable on an open set and the set of non-differentiability cannot be dense. One easily checks, that this situation is stable for small C^1- perturbartions of g. Therefore it is not possible by an arbitrarily small C^1- perturbation of g to generate a pathological maximum function with a dense set of non-differentiability. Note that in this example - in contrast to the result stated above - K is not a differentiable manifold at the boundary point $x = 1$.* □

In this way a nicer structure of K implies a larger set of pathological maximum functions. The very technical reason for this is that the gradient of $g(\cdot,z)$ needs not vanish at the global maxima of this function unless K is a differentiable manifold.

Results

First some notation shall be introduced. Let A, B, S be arbitrary sets and assume $S \subseteq A, b \in B, h : A \times B \to \mathbb{R}$. Then denote by $E(h,S,b) := \{s \in S \mid h(s,b) \geq h(s',b) \; \forall s' \in S\}$ the (possibly empty) set of global maximizers of $h(\cdot,b)$. For $y \in \mathbb{R}^p$ and $r > 0$ let $B(y;r)$ be the closed p- dimensional euclidean ball of radius r around y. Finally, for a differentiable real valued function h define:

$$\|h\|_{B(y;r)} := \max_{y' \in B(y;r)} | h(y') | \quad \text{and} \quad \|\nabla h\|_{B(y;r)} := \max_{y' \in B(y;r)} \|\nabla h(y')\|,$$

where $\|\nabla h(y')\| := \max_{i=1,\dots,p} | \frac{\partial h}{\partial y_i}(y') |$. Finally the symbol C^1 refers as usually to differentiable functions having continous partial derivatives.

The following lemma is an essential technical tool for proving the desired result. It indicates how to generate a point of non-differentiability for g^{max}

at a pre-defined position z^* by arbitrarily small perturbartion of g. The considered functions are assumed to be defined only locally which will be necessary when proving the theorem below.

Lemma 1 *Let U_x and U_z be open neighborhoods of $x^* \in \mathbb{R}^m$ and of $z^* \in \mathbb{R}^n$. Given a function $h \in C^1(U_x \times U_z, \mathbb{R})$ with $x^* \in E(h, U_x, z^*)$ and given arbitrary $\epsilon_1, \epsilon_2 > 0$ such that $B((x^*, z^*); \epsilon_1) \subseteq U_x \times U_z$, there exists a function $\bar{h} \in C^1(U_x \times U_z, \mathbb{R})$ satisfying the following properties:*

$$\bar{h}(x,z) = h(x,z) \quad \forall (x,z) \in (U_x \times U_z) \setminus B((x^*, z^*); \epsilon_1) \tag{1}$$

$$\exists\, x^1, x^2 \in U_x : E(\bar{h}, U_x, z^*) = \{x^1, x^2\}, \tag{2}$$

$$\bar{h}(x^1, z^*) = \bar{h}(x^2, z^*) > h(x^*, z^*), \quad \frac{\partial \bar{h}}{\partial z_1}(x^1, z^*) \neq \frac{\partial \bar{h}}{\partial z_1}(x^2, z^*)$$

$$\max\left\{ \left\| h - \bar{h} \right\|_{B((x^*,z^*);\epsilon_1)}, \left\| \nabla h - \nabla \bar{h} \right\|_{B((x^*,z^*);\epsilon_1)} \right\} < \epsilon_2, \tag{3}$$

Proof:
For the purpose of abbreviation in this lemma $B(r)$ shall denote the closed ball of radius r around (x^*, z^*). First we define a parameter dependent function $h^*_\alpha \in C^1(U_x \times U_z, \mathbb{R})$ by

$$\begin{aligned} h^*_\alpha(x,z) \quad := \quad & -(x_1 - x_1^*)^4 + \alpha(x_1 - x_1^*)^2 - \sum_{i=2}^{m}(x_i - x_i^*)^2 + \\ & (z_1 - z_1^*)(x_1 - x_1^*) + \nabla_z h(x^*, z^*)(z - z^*) + h(x^*, z^*) \end{aligned}$$

The x_1- part of this function is illustrated in figure 1. For $0 < 2\sqrt{\alpha} < \epsilon_1$ one gets by the assumption of the lemma:

$$B(2\sqrt{\alpha}) \subseteq B(\epsilon_1) \subseteq U_x \times U_z. \tag{4}$$

It is easily verified that (for any α) h^*_α has first order contact at (x^*, z^*) with h, i.e. at this point functional values and partial derivatives (w.r.t x and z) coincide for both functions (recall that, by assumption, x^* is a global maximizer of $h(\cdot, z^*)$, hence $\nabla_x h(x^*, z^*) = 0$). Accordingly one may assume $\alpha > 0$ sufficiently small to meet the condition

$$\max\left\{ \|h^*_\alpha - h\|_{B(2\sqrt{\alpha})}, \|\nabla h^*_\alpha - \nabla h\|_{B(2\sqrt{\alpha})} \right\} < \frac{\epsilon_2}{2}. \tag{5}$$

Furthermore the set of global maximizers of $h^*_\alpha(\cdot, z^*)$ consists of the two points $x^1/x^2 := (x_1^* \pm \sqrt{\alpha/2}, x_2^*, \ldots, x_m^*)$ (note that due to (4) both points belong to U_x), which yield the functional value

$$h^*_\alpha(x^1, z^*) = h^*_\alpha(x^2, z^*) = h(x^*, z^*) + \alpha^2/4. \tag{6}$$

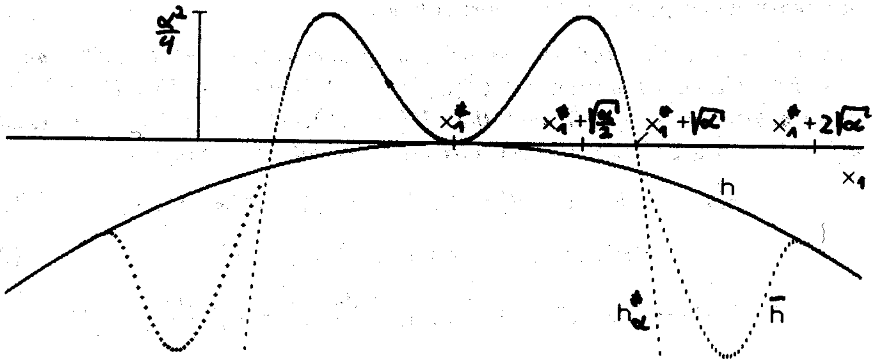

Figure 1: Illustration of the x_1- part of the functions $h, h^*_\alpha, \bar{h}$ at the point (x^*, z^*).

Now, the function $\bar{h}$, which is looked for in the lemma, is constructed as follows: $\bar{h} := \phi \cdot h^*_\alpha + (1-\phi) \cdot h$, where $\phi \in C^\infty(\mathbb{R}^{m+n}, \mathbb{R})$ is a function with the properties $0 \le \phi \le 1$, $\phi\, |_{B(\sqrt{\alpha})} \equiv 1$, $\phi\, |_{\mathbb{R}^{m+n} \setminus B(2\sqrt{\alpha})} \equiv 0$ and $\|\nabla\phi(y)\| \le \frac{2}{\sqrt{\alpha}}\ \forall y \in \mathbb{R}^{m+n}$, where $\|\cdot\|$ refers to the maximum norm (for details compare [6], p.42). The last property of ϕ results from the fact that $\sqrt{\alpha}$ is the width of the annulus where ϕ varies between 0 and 1. Then ϕ may be constructed in such a way that the modulus of all its partial derivatives is bounded by a number which is greater than but arbitrarily close to $1/\sqrt{\alpha}$. By definition $\bar{h}$ coincides with h outside $B(2\sqrt{\alpha})$, hence (1) is proved by (4). Inside $B(\sqrt{\alpha})$ $\bar{h}$ is identical with h^*_α (compare figure 1), thus in this ball $\bar{h}(\cdot, z^*)$ attains a maximum value which is identical to (6). Outside $B(\sqrt{\alpha})$ $\bar{h}$ is a convex combination of h and h^*_α. In this region all values of $h(\cdot, z^*), h^*_\alpha(\cdot, z^*)$ are strictly smaller than the maximum value in (6). This is shown by recalling that $h(x^*, z^*)$ is a global upper bound of $h(\cdot, z^*)$ and that the points $(x^1, z^*), (x^2, z^*)$ (where $\{x_1, x_2\}$ is the set of global maximizers of $h^*_\alpha(\cdot, z^*)$) are located within the interior of $B(\sqrt{\alpha})$. Consequently the first two relations of (2) are proved and the last one is derived from

$$\frac{\partial \bar{h}}{\partial z_1}(x^1, z^*) \quad = \quad \frac{\partial h^*_\alpha}{\partial z_1}(x^1, z^*) = \frac{\partial h}{\partial z_1}(x^*, z^*) + \sqrt{\alpha/2} \ \neq$$

$$\frac{\partial h}{\partial z_1}(x^*, z^*) - \sqrt{\alpha/2} = \frac{\partial h_\alpha^*}{\partial z_1}(x^2, z^*) = \frac{\partial \bar{h}}{\partial z_1}(x^2, z^*)$$

Concerning (3) it is obviously sufficient (see (4)), to consider the corresponding norm differences over $B(2\sqrt{\alpha})$. The definitions of $\bar{h}$ and ϕ imply:

$$| h(x,z) - \bar{h}(x,z) |=| \phi(x,z) \, || \, h_\alpha^*(x,z) - h(x,z) |< \frac{\epsilon_2}{2} \quad \forall (x,z) \in B(2\sqrt{\alpha})$$

which results from (5). It remains to estimate the difference of gradients within the same closed ball (omitting the argument (x, z) in all expressions and using $\|\cdot\|$ as maximum norm):

$$\|\nabla\bar{h} - \nabla h\| = \|\phi \cdot (\nabla h_\alpha^* - \nabla h) + (h_\alpha^* - h) \cdot \nabla\phi\| \leq \frac{\epsilon_2}{2} + 2\frac{| h_\alpha^* - h |}{\sqrt{\alpha}} \quad (7)$$

Here again (5) as well as the properties of ϕ (stated above) were exploited. This last inequality holds for all $(x, z) \in B(2\sqrt{\alpha})$, and one has to estimate the second term of the right-hand side within this ball. To this aim, define $\hat{h}(x,z,\alpha) := h_\alpha^*(x,z) - h(x,z)$ with the properties $\hat{h}(x^*, z^*, 0) = 0$ and $\nabla\hat{h}(x^*, z^*, 0) = 0$. As a consequence, there is a function $\epsilon(x, z, \alpha)$, satisfying

$$\lim_{(x,z,\alpha)\to(x^*,z^*,0)} \epsilon(x,z,\alpha) = 0$$

$$\text{and} \quad | \hat{h}(x,z,\alpha) | = \epsilon(x,z,\alpha) \left\|(x - x^*, z - z^*, \alpha)^T\right\|,$$

where again the maximum norm is used for convenience. Putting $\hat{\epsilon}(\alpha) := \max_{(x,z)\in B(2\sqrt{\alpha})} \epsilon(x,z,\alpha)$ one gets for small enough $\alpha > 0$:
$\max_{(x,z)\in B(2\sqrt{\alpha})} | \hat{h}(x,z,\alpha) |\leq \hat{\epsilon}(\alpha) \cdot 2\sqrt{\alpha}$. Furthermore, $\lim_{\alpha\to 0} \hat{\epsilon}(\alpha) = 0$ by the corresponding property of ϵ. Now, the remaining term of (7) may be estimated by $4\hat{\epsilon}(\alpha)$ which becomes smaller than $\epsilon_2/2$ for small enough $\alpha > 0$.
□

Suppose now, that the set $K \subseteq \mathbb{R}^p$ is a compact m- dimensional C^1-manifold. This means: Each point $y \in K$ posseses an open neighborhood $V \subseteq \mathbb{R}^p$ and a mapping $f \in C^1(V, \mathbb{R}^m)$ such that $U := f(K \cap V)$ is open in $\mathbb{R}^m$, the restriction $f |_{K\cap V}: K \cap V \to U$ is bijective and the inverse mapping $f^{-1} : U \to \mathbb{R}^p$ is of class The pair $(K \cap V, f |_{K\cap V})$ is called a chart around y. Since K is compact, one can select a finite number of charts $(K \cap V_i, f_i |_{K\cap V_i}) (i = 1, \ldots, l)$.

Given a fixed compact m-dimensional C^1- manifold K as well as some open set $U_z \subseteq \mathbb{R}^n$, one obtains a $m+n$- dimensional C^1- manifold $K \times U_z$

(note that by $\left([K\cap V_i]\times U_z, F_i\mid_{[K\cap V_i]\times U_z}\right)$, with $F_i(y,z) := (f_i(y), z)$ a finite system of charts for this manifold is defined). By $\mathcal{C}^1(K\times U_z, \mathbb{R})$ we denote the set of all functions $g : K\times U_z \to \mathbb{R}$, such that $g\circ F_i^{-1}\in \mathcal{C}^1\left(F_i\left([K\cap V_i]\times U_z,\right)\mathbb{R}\right)$ for all $i = 1,\ldots,l$ (recall that the domains of these functions are open subsets of $\mathbb{R}^{m+n}$ by definition of F_i). For instance, one has $g\in\mathcal{C}^1(K\times U_z,\mathbb{R})$, if there exists $\tilde{g}\in\mathcal{C}^1(\mathbb{R}^p\times U_z,\mathbb{R})$ with $\tilde{g}\mid_{K\times U_z}\equiv g$. The space $\mathcal{C}^1(K\times U_z,\mathbb{R})$ can be endowed with the so-called weak topology ([6], p.34) as follows:

If $g\in\mathcal{C}^1(K\times U_z,\mathbb{R})$, $(W_i, F_i\mid_{W_i})$ is some chart of $K\times U_z$, $L\subseteq W_i$ is some compact subset and $\epsilon>0$, then a subbasic neighborhood of g will be defined to consist of all $\bar{g}\in\mathcal{C}^1(K\times U_z,\mathbb{R})$ satisfying:

$$\left\|g\circ F_i^{-1}-\bar{g}\circ F_i^{-1}\right\|_L<\epsilon,\ \left\|\nabla\left(g\circ F_i^{-1}\right)-\nabla\left(\bar{g}\circ F_i^{-1}\right)\right\|_L<\epsilon. \qquad (8)$$

The set of all such subbasic neighborhoods (by variation of $g,(W_i,F_i\mid_{W_i}),L$ and ϵ) defines a subbasis of the weak topology. The resulting topological space shall be denoted by $C^1_w(K\times U_z,\mathbb{R})$. This space is metrizable by some complete metric δ.

Finally observe that for all $g\in\mathcal{C}^1(K\times U_z,\mathbb{R})$ it holds:

$$\nabla_z g(y,z) = \nabla_z[g\circ F_i^{-1}](F_i(y,z))\ \forall (y,z)\in[K\cap V_i]\times U_z\ \forall i. \qquad (9)$$

To see this, recall that $g\mid_{[K\cap V_i]\times U_z}\equiv\left(g\circ F_i^{-1}\circ F_i\right)\mid_{[K\cap V_i]\times U_z}$, hence for $(y,z)\in[K\cap V_i]\times U_z$ one gets

$$\nabla_z g(y,z) = \nabla_z\left(g\circ F_i^{-1}\circ F_i\right)(y,z) = \nabla_z[g\circ F_i^{-1}](F_i(y,z))$$

since $\nabla_z F_i = \left(0_{(m,n)}\mid I_n\right)^T$ by definition of the F_i.

Now, to each $g\in\mathcal{C}^1(K\times U_z,\mathbb{R})$ the corresponding maximum function $g^{max}: U_z\to\mathbb{R}$ is assigned by $g^{max}(z) := \max_{y\in K} g(y,z)$. We introduce the following set of 'pathological' functions:

$$\mathcal{G} := \{g\in\mathcal{C}^1_w(K\times U_z,\mathbb{R})\mid \exists D\subseteq U_z, D\,\text{dense}, g^{max}\ \text{non-differentiable on D}\}$$

Concerning the generality of $\mathcal{G}$ we can formulate the following result:

Theorem 1 $\mathcal{G}\subseteq\mathcal{C}^1_w(K\times U_z,\mathbb{R})$ *is a dense subset.*

Proof:

We start by defining a dense subset of U_z. To this aim let for $k\in\mathbb{N}$

$$T^k := \{t\in\mathbb{R}\mid t=\pm i\cdot 2^{-k}\quad (i=0,\ldots,k\cdot 2^k)\}$$

$$\mathcal{Z}^k := \left(\underbrace{T^k \times \cdots \times T^k}_{n}\right) \bigcap U_z \text{ and } \mathcal{Z}^\infty := \bigcup_{k=0}^{\infty} \mathcal{Z}^k.$$

Obviously, each $\mathcal{Z}^k$ is finite and $\mathcal{Z}^\infty \subseteq U_z$ is a dense subset. In order to proove the theorem one has to show that, given an arbitrary number $\gamma > 0$ and an arbitrary function $g \in \mathcal{C}^1_w(K \times U_z, I\!R)$, there is a function $g^* \in \mathcal{G}$ with $\delta(g, g^*) < \gamma$, where δ refers to the complete metric of $\mathcal{C}^1_w(K \times U_z, I\!R)$. The desired function g^* will be constructed as the limit of a sequence of functions $\{g_k\}$ which meet the following conditions:

$$g_k \in \mathcal{C}^1_w(K \times U_z, I\!R) \tag{10}$$

$$g_k(y,z) = g_{k-1}(y,z),\ \nabla_z g_k(y,z) = \nabla_z g_{k-1}(y,z) \quad \forall (y,z) \in K \times \mathcal{Z}^{k-1} \tag{11}$$

$$\forall\, z \in \mathcal{Z}^k\ \exists y^1, y^2 \in K: \quad \begin{array}{l} E(g_k, K, z) = \{y^1, y^2\} \\ \dfrac{\partial g_k}{\partial z_1}(y^1, z) \neq \dfrac{\partial g_k}{\partial z_1}(y^2, z) \end{array} \tag{12}$$

$$\delta(g_{k-1}, g_k) < \gamma \cdot 2^{-(k+1)}, \tag{13}$$

Assuming, for the moment, that (10-13) hold true, one observes that the $\{g_k\}$ form a Cauchy sequence (by (13)) with continously differentiable limit function $g^* \in \mathcal{C}^1_w(K \times U_z, I\!R)$. Furthermore, putting formally $g_{-1} := g$, one has $\delta(g, g^*) < \gamma$, as desired. In order to verify $g^* \in G$ one first remarks that $\forall z \in \mathcal{Z}^\infty\ \exists k_0 \in I\!N\ \forall y \in K$:

$$g^*(y,z) = g_{k_0}(y,z) \quad \text{and} \quad \nabla_z g^*(y,z) = \nabla_z g_{k_0}(y,z) \tag{14}$$

To see this, recall that $z \in \mathcal{Z}^\infty$ implies $z \in \mathcal{Z}^{k_0}$ for some $k_0 \in I\!N$. Since $\mathcal{Z}^{k_0} \subseteq \mathcal{Z}^k\ \forall k \geq k_0$, (11) yields for all $(y,z) \in K \times \mathcal{Z}^{k_0}$:

$$g_{k_0}(y,z) = g_{k_0+1}(y,z) = \cdots \quad \text{and} \quad \nabla_z g_{k_0}(y,z) = \nabla_z g_{k_0+1}(y,z) = \cdots$$

Then (14) follows from the convergence $g_k \longrightarrow g^*$ in the $\mathcal{C}^1_w$- topology. From the first relations in (12) and (14) one derives: $\forall z \in \mathcal{Z}^\infty\ \exists k_0 \in I\!N$: $E(g^*, K, z) = E(g_{k_0}, K, z) = \{y^1, y^2\}$, where y^1, y^2 depend on z. Now the second relations in (12) and (14) imply

$$\frac{\partial g^*}{\partial z_1}(y^1, z) \neq \frac{\partial g^*}{\partial z_1}(y^2, z).$$

These partial derivatives, however, determine the left- and right-sided derivatives, respectively, of the maximum function g^{*max} at z restricted to the

z_1- axis. Consequently g^{*max} is not differentiable at all points $z \in \mathcal{Z}^\infty$, hence $g^* \in \mathcal{G}$ and the theorem is proved.

Now, the sequence $\{g_k\}$ fulfilling (10-13) will be defined inductively starting with the construction of an initial function g_0 $(k = 0)$. To give sense to (11) we put formally $g_{-1} := g$ and $\mathcal{Z}^{-1} := \emptyset$. Since K is compact, there exists a point $y^* \in E(g, K, 0)$. Let $([K \cap V_i] \times U_z, F_i \mid_{[K \cap V_i] \times U_z})$ be a chart around $(y^*, 0)$ and denote $(x^*, 0) := F_i(y^*, 0)$ as well as $U_x \times U_z := F_i([K \cap V_i] \times U_z)$ (recall that F_i equals the identity over the z- part). Next choose $\epsilon_1 > 0$ to meet $B((x^*, 0); \epsilon_1) \subseteq U_x \times U_z$. Given any $\epsilon_2 > 0$ one can apply Lemma 1 to the function $h := g \circ F_i^{-1} \in \mathcal{C}^1(U_x \times U_z, \mathbb{R})$. Then existence of $\bar{h} \in \mathcal{C}^1(U_x \times U_z, \mathbb{R})$ satisfying conditions (1-3) is provided. We define

$$g_0(y, z) := \begin{cases} \bar{h}(F_i(y, z)) & y \in K \cap V_i \\ g(y, z) & y \in K \setminus V_i \end{cases} \tag{15}$$

Exploiting the facts that by (1) g_0 and g coincide outside the compact subset $F_i^{-1}(B((x^*, 0); \epsilon_1)) \subseteq [K \cap V_i] \times U_z$ of a chart and that by assumption $g \in \mathcal{C}^1_w(K \times U_z, \mathbb{R})$ one easily verifies the $\mathcal{C}^1$- property of all composite functions $g_0 \circ F_j^{-1}$, proving (10). By definition of $\mathcal{Z}^{-1}$ (11) is trivially fulfilled. From Lemma 1 and (15) it follows that

$$E(g_0, K \cap V_i, 0) = \{y^1, y^2\}, \quad \text{where} \quad (y^j, 0) = F_i^{-1}(x^j, 0). \tag{16}$$

Exploiting (2) and (15) one arrives at

$$g_0(y^j, 0) = \bar{h}(x^j, 0) > h(x^*, 0) = g(y^*, 0) \geq g(y, 0) = g_0(y, 0) \quad \forall y \in K \setminus V_i$$

Together with (16) this leads to $E(g_0, K, 0) = \{y^1, y^2\}$, which is the first relation of (12) (recall that $\mathcal{Z}^0 = \{0_n\}$). The second relation of (12) is obtained from (2) and (9). Finally (3) and (15) allow, by using arbitrarily small $\epsilon_2 > 0$, to keep g_0 arbitrarily close to g in the $\mathcal{C}^1_w$- topology (i.e. on all compact subsets of all charts), hence (13) may be fulfilled.

Now suppose, that some function g_k satisfying (10-13) has already been constructed for some $k \in \mathbb{N}$. Then we are going to find g_{k+1} having the corresponding properties. Many arguments run similar to the start of induction. Assume $\mathcal{Z}^{k+1} \setminus \mathcal{Z}^k = \{z^{(1)}, \ldots, z^{(m)}\}$ for some m. For each j there exists (due to compactness of K) an $y^{(j)} \in E(g^k, K, z^{(j)})$. Let $\left([K \cap V_{(j)}] \times U_z, F_{(j)} \mid_{[K \cap V_{(j)}] \times U_z}\right)$ be charts around $(y^{(j)}, z^{(j)})$. Give $U_x^{(j)}$ and $x^{(j)}$ similar meanings as U_x and x^* above. Choose $\epsilon_1 > 0$ to meet $B((x^{(j)}, z^{(j)}); \epsilon_1) \subseteq U_x^{(j)} \times U_z$. Given any $\epsilon_2 > 0$ apply lemma 1 to $h(j) := g_k \circ F_{(j)}^{-1}$ yielding $\bar{h}^{(j)}$ with the corresponding properties. Additionally take $\epsilon_1 > 0$ small enough to provide disjoint n- dimensional balls

$$B(z^a;\epsilon_1) \cap B(z^b;\epsilon_1) = \emptyset \quad \forall z^a, z^b \in \mathcal{Z}^{k+1}, z^a \neq z^b. \tag{17}$$

Then the following definition makes sense:

$$g_{k+1}(y,z) := \begin{cases} \bar{h}^{(j)}\left(F_{(j)}(y,z)\right) & \text{if} \quad (y,z) \in [K \cap V_{(j)}] \times B(z^{(j)};\epsilon_1) \\ g_k(y,z) & \quad\quad\quad\quad\quad\quad \textit{else} \end{cases}$$

Now, similar to (15), (10) follows for g_{k+1} from the corresponding property of g_k by the induction assumption. From (17) and the construction of g_{k+1} one derives that for all $z \in \mathcal{Z}^k$ there exists an open neighborhood $U(z)$ such that $g_{k+1}(y,z) = g_k(y,z) \quad \forall (y,z) \in K \times U(z)$. This implies (11) for $k+1$. As a consequence of this fact one obtains, exploiting the induction assumption (12): $\forall z \in \mathcal{Z}^k \; \exists y^1, y^2 \in K$:

$$E(g_{k+1}, K, z) = \{y^1, y^2\}, \; \frac{\partial g_{k+1}}{\partial z_1}(y^1, z) \neq \frac{\partial g_{k+1}}{\partial z_1}(y^2, z)$$

But the same relation holds also true for all $z \in \mathcal{Z}^{k+1} \setminus \mathcal{Z}^k$, by the above given construction of g_{k+1} (again using Lemma 1 similar to the start of induction). Consequently (12) is proved. Finally (13) may be satisfied by reducing $\epsilon_2 > 0$ similar to the start of induction. Since g_{k+1} and g_k differ on a finite number of disjoint balls on which this difference may be kept arbitrarily small (see Lemma 1, (3)) one arrives at this last relation to be verified. This completes the proof. □

References

[1] Jongen, H.T., Twilt, F. Weber, G.W., *Semi-infinite optimization: Structure and stability of the feasible set*, Journal of Optimization Theory and Applications, Vol. 72,

[2] Henrion, R., *On Constraint Qualifications*, Journal of Optimization Theory and Applications, Vol. 72, pp. 187-197, 1992.

[3] Yomdin, Y., *On Functions Representable as a Supremum of a Family of Smooth Functions*, SIAM Journal of Mathematical Analysis, Vol. 14, pp. 239-246, 1983.

[4] Clarke, F.H., *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[5] Kummer, B., *Some pathological Lipschitz functions*, Humboldt-University of Berlin, unpublished manuscript.

[6] Hirsch, M.W., *Differential Topology*, Springer, New York, 1988.