CALCULUS ON THE CANTOR SET

A.C.M. van Rooij and W.H. Schikhof

§1. Introduction to the subject.

1.1. In the classical theory of real functions of one variable the domain of definition of the functions is mostly a very nice subset of $\mathbb{R}$. Admittedly, in the theory of Lebesgue integration arbitrary measurable sets are used as domains, but when dealing with derivation and antiderivation one hardly ever encounters other sets than intervals or unions of intervals. Nevertheless, calculus for real functions defined on wilder subsets of $\mathbb{R}$ can be quite interesting, if only because it may give us new insight in the "ordinary" theory by revealing the roles of the special properties of intervals.

We are going to consider functions defined on the Cantor set, which we denote by the letter $\mathbb{D}$. Our purpose is to display some of the contrasts between calculus on $\mathbb{D}$ and calculus on $[0,1]$. The reader will find our main results in Theorems 5.9 and 6.2.

Let $f : \mathbb{D} \to \mathbb{R}$. We say that f is <u>differentiable</u> at $a \in \mathbb{D}$ if

$$\lim_{\substack{x \to a \\ x \in \mathbb{D}}} \frac{f(x)-f(a)}{x-a}$$

exists. Definitions of terms like "derivative" and "antiderivative" are confidently left to the reader.

The well-known rules for differentiation of sums, products, quotients and compositions of functions remain valid. A rational function, defined on all of $\mathbb{D}$, is everywhere differentiable. A differentiable function is continuous.

1.2. So far, nothing new. The above could have been done for any non-empty subset of $\mathbb{R}$ that has no isolated points. (Observe, however, that for some points of $\mathbb{D}$ no reasonable definition of one-sided derivatives is possible.)

But the Cantor set has very special properties. Some of them, like compactness, it shares with $[0,1]$. These and their immediate consequences ("Every continuous function is uniformly continuous and attains a largest value") are not going to interest us. The feature of $\mathbb{D}$ that really sets it apart from $[0,1]$ is its total disconnectedness. This is the property that will concern us in the following pages.

For later use it is convenient to recall some facts and fix a few notations.

The elements of the Cantor set are precisely the numbers $a \in \mathbb{R}$ for which there exist $\alpha_1, \alpha_2, \ldots \in \{0,1\}$ such that $a = 2\sum_{i=1}^{\infty} \alpha_i 3^{-i}$. (The α_i are uniquely determined by a.) In the sequel we will often use expressions like "let $a = 2\sum_{i=1}^{\infty} \alpha_i 3^{-i} \in \mathbb{D}$", tacitly assuming that the α_i all lie in $\{0,1\}$.

For $a = 2\sum_{i=1}^{\infty} \alpha_i 3^{-i} \in \mathbb{D}$ and $n \in \mathbb{N}$ we define

$$a_n := 2\sum_{i=1}^{n} \alpha_i 3^{-i}.$$

The functions $a \mapsto a_n$ are continuous and locally constant; so are the <u>coördinate functions</u> $a \mapsto \alpha_i$.

1.3. Let $n \in \mathbb{N}$, $x, a \in \mathbb{D}$. Then

(i) $|a-a_n| \le 3^{-n}$,

(ii) if $x_1 = a_1, \ldots, x_{n-1} = a_{n-1}$, $x_n \neq a_n$, then $3^{-n} \le |x-a| \le 3^{-n+1}$,

(iii) if $|x-a| < 3^{-n}$, then $x_1 = a_1, \ldots, x_n = a_n$,

(iv) $|x_n - a_n| \le 3|x-a|$.

The reader will have no trouble proving these statements.

§2. Derivation and antiderivation: a first analysis.

If $f : [0,1] \to \mathbb{R}$ is differentiable and $f' = 0$ everywhere, then f is constant. This follows from the Mean Value Theorem: if $x,y \in [0,1]$ and $x \neq y$, then $\frac{f(x)-f(y)}{x-y}$ is a value of f', whence $f(x) = f(y)$ if $f' = 0$. But the Mean Value Theorem depends heavily on the connectedness of $[0,1]$. This may be made clear by the fact that there are many non-constant functions on $\mathbb{D}$ whose derivatives vanish identically. Examples are the coördinate functions, that are locally constant. We even have:

2.1. THEOREM. <u>Every continuous function on $\mathbb{D}$ is a uniform limit of locally constant functions, hence of functions with vanishing derivatives.</u>

PROOF. Let $f : \mathbb{D} \to \mathbb{R}$ be continuous. For $n \in \mathbb{N}$, let $f_n(x) := f(x_n)$. Then each f_n is locally constant. Since $|x-x_n| \leq 3^{-n}$ for all x and n, by the uniform continuity of f we have $\lim_{n\to\infty} f_n = f$ uniformly.

Does $f' = 0$ imply local constantness of f? The answer is negative, as we can see from:

2.2. EXAMPLE. <u>The following function</u> $f : \mathbb{D} \to \mathbb{R}$ <u>is strictly increasing but</u> $f' = 0$ <u>everywhere</u>.

$$f\left(2 \sum_{i=1}^{\infty} \alpha_i 3^{-i}\right) := \sum_{i=1}^{\infty} \alpha_i 9^{-i}.$$

PROOF. Let $a = 2\sum_{i=1}^{\infty} \alpha_i 3^{-i} \in \mathbb{D}$, $b = 2\sum_{i=1}^{\infty} \beta_i 3^{-i} \in \mathbb{D}$ and $a < b$. Let $n := \min\{i \in \mathbb{N} : \alpha_i \neq \beta_i\}$. Then $0 < \frac{b-a}{2} = (\beta_n-\alpha_n)3^{-n} + \sum_{i>n} (\beta_i-\alpha_i)3^{-i} \leq$ $\leq (\beta_n-\alpha_n)3^{n} + \sum_{i>n} 3^{-i} = (\beta_n-\alpha_n+\frac{1}{2})3^{-n}$, so $\beta_n = 1$, $\alpha_n = 0$. Therefore,

$f(b)-f(a) = \sum_{i\geq n} (\beta_i-\alpha_i)9^{-i} = 9^{-n} + \sum_{i>n} (\beta_i-\alpha_i)9^{-i} \geq 9^{-n} - \frac{1}{8} 9^{-n} > 0.$

It follows that f is strictly increasing. The fact that $f' = 0$ is a consequence of the formula $|f(b)-f(a)| \leq \sum_{i\geq n} |\beta_i-\alpha_i| 9^{-i} \leq \frac{9}{8} 9^{-n} \leq \frac{9}{8}(b-a)^2$ (by 1.3).

From Theorem 2.1 we see that, if a function on $\mathbb{D}$ has an antiderivative, then it has many. What functions on $\mathbb{D}$ have antiderivatives? The following theorem will come as no surprise.

2.3. THEOREM. All continuous functions on $\mathbb{D}$ have antiderivatives.

2.4. Before we start giving the actual proof we introduce a standard method for extending functions on $\mathbb{D}$ to functions on $[0,1]$.

Let $f : \mathbb{D} \to \mathbb{R}$ (continuous or not). For $x \in [0,1]$ we define $\tilde{f}(x) \in \mathbb{R}$ by:

(1) If $x \in \mathbb{D}$, then $\tilde{f}(x) := f(x)$.

(2) Otherwise, x lies in a component interval (a,b) of the open set $[0,1]\setminus\mathbb{D}$: then $\tilde{f}(x) := f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$.

Thus, $\tilde{f}$ coincides with f on $\mathbb{D}$ and interpolates f linearly on the intervals contiguous to $\mathbb{D}$. The following facts are easily verified.

(i) The correspondence $f \mapsto \tilde{f}$ is linear.

(ii) If $f \geq g$, then $\tilde{f} \geq \tilde{g}$.

(iii) If f is increasing, then so is $\tilde{f}$.

(iv) If $s,t \in \mathbb{R}$ and $s \leq f \leq t$, then $s \leq \tilde{f} \leq t$.

(v) If f is continuous, then so is $\tilde{f}$.

(vi) If $\lim_{n\to\infty} f_n = f$, then $\lim_{n\to\infty} \tilde{f}_n = f$.

(vii) If f is semicontinuous, then so is $\tilde{f}$.

PROOF of Theorem 2.3. Let $f : \mathbb{D} \to \mathbb{R}$ be continuous. Define

$$F(x) := \int_0^x \tilde{f}(t)\,dt \qquad (x \in \mathbb{D}).$$

Then $F' = f$.

2.5. If g is any differentiable function on $[0,1]$, then, trivially, the restriction of g' to $\mathbb{D}$ has an antiderivative. It follows easily that some discontinuous functions on $\mathbb{D}$ have antiderivatives. There is nothing remarkable about that. More striking than this general observation may be the following concrete example.

2.6. EXAMPLE. *For $a \in \mathbb{D}$ let $e_a : \mathbb{D} \to \mathbb{R}$ be the characteristic function of the set $\{a\}$. Then e_a has an antiderivative, e.g., the function $f : \mathbb{D} \to \mathbb{R}$ defined by:*

$$f(x) := \begin{cases} a \text{ if } x = a, \\ x_{2n} \text{ if } x \in \mathbb{D},\ x \neq a,\ n = \min\{i : x_i \neq a_i\}. \end{cases}$$

Indeed, since f is locally constant on $\mathbb{D}\setminus\{a\}$, we clearly have $f'(x) = 0$ for all $x \in \mathbb{D}$, $x \neq a$. Now take $x \in \mathbb{D}$, $x \neq a$: setting $n = \min\{i : x_i \neq a_i\}$ and applying 1.3 we obtain

$$|f(x)-f(a)-(x-a)| = |x_{2n}-a-x+a| = |x_{2n}-x| \le 3^{-2n} \le |x-a|^2.$$

Consequently, $f'(a) = 1$.

Not *all* functions on $\mathbb{D}$ have antiderivatives:

2.7. EXAMPLE *of a (bounded) function on $\mathbb{D}$ without antiderivative.*

Let D_r be the set of all right end points of the components of $\mathbb{R}\setminus\mathbb{D}$. Define $f : \mathbb{D} \to \mathbb{R}$ by

$$f(x) := \begin{cases} 1 \text{ if } x \in D_r \\ -1 \text{ if } x \in \mathbb{D}^r\setminus D_r \end{cases}$$

Suppose we have an $F : \mathbb{D} \to \mathbb{R}$ with $F' = f$; we will obtain a contradiction. Being continuous, F assumes a largest value at some point a of $\mathbb{D}$. If $a \in D_r$, then a is an accumulation point of $\mathbb{D} \cap (a,1]$, so

$$f(a) = \lim_{\substack{x\downarrow a \\ x\in\mathbb{D}}} \frac{F(x)-F(a)}{x-a} \le 0$$

which is false. But if $a \notin D_r$, then a is an accumulation point of $[0,a) \cap \mathbb{D}$ and we find $f(a) \geq 0$, which is also false.

In §5 we will go deeper into the question what functions on $\mathbb{D}$ possess an antiderivative.

§3. Uniformly differentiable functions.

If $f : \mathbb{D} \to \mathbb{R}$ is differentiable and if f' is of constant sign, is there anything in general to be said about monotony of f? Example 2.2 is not encouraging and the following examples do not help.

3.1. EXAMPLE. Define $g : \mathbb{D} \to \mathbb{R}$ by

$$g(x) := \sum_{i=1}^{\infty} (-1)^i \xi_i 9^{-i} \text{ where } x = 2\sum_{i=1}^{\infty} \xi_i 3^{-i} \in \mathbb{D}.$$

Then g is differentiable, $g'(x) = 0$ for all $x \in \mathbb{D}$ but g is on no open interval of $\mathbb{D}$ monotone.

(A subset of $\mathbb{D}$ is called an open interval of $\mathbb{D}$ if it is the intersection of $\mathbb{D}$ with an open interval of $\mathbb{R}$ and is not empty. Then automatically it contains infinitely many points.)

PROOF. The identity $g' = 0$ is proved in the same way as $f' = 0$ in Example 2.2. To prove the lack of monotony, observe that for all $x \in \mathbb{D}$ and $n \in \mathbb{N}$, x_n and $x_n + 2\cdot 3^{-n-1}$ are elements of $\mathbb{D}$ and $g(x_n + 2\cdot 3^{-n-1}) - g(x_n) = (-1)^{n+1} 9^{-n-1}$.

3.2. EXAMPLE. A function $f : \mathbb{D} \to \mathbb{R}$ with $f' = 1$ everywhere that is not injective on any open interval of $\mathbb{D}$ containing 0.

Define $h : \mathbb{D} \to \mathbb{R}$ by

$$\begin{cases} h(0) := 0 \\ h(x) := 2\xi_{2n} 3^{-2n} \text{ if } x = 2\sum_{i=1}^{\infty} \xi_i 3^{-i} \in \mathbb{D}, x \neq 0,\ n = \min\{i : \xi_i \neq 0\}. \end{cases}$$

h is locally constant on $\mathbb{D}\setminus\{0\}$ and $|h(x)| \le x^2$ for all $x \in \mathbb{D}$, so $h' = 0$. Setting $f(x) := x-h(x)$ $(x \in \mathbb{D})$ we have $f' = 1$ but $f(2\cdot 3^{-k}+2\cdot 3^{-2k}) = f(2\cdot 3^{-k})$ for all $k \in \mathbb{N}$.

However, not all is lost:

3.3. THEOREM. Let $f : \mathbb{D} \to \mathbb{R}$ be differentiable, $f'(x) > 0$ for all $x \in \mathbb{D}$. Then on some open interval of $\mathbb{D}$, f is strictly increasing. Indeed, the open intervals of $\mathbb{D}$ on which f is strictly increasing form a dense set in $\mathbb{D}$.

PROOF. Let D_r be the set of all right end points of the components of $\mathbb{R}\setminus\mathbb{D}$.

Let J be an open interval of $\mathbb{D}$ and assume that J does not contain any open interval of $\mathbb{D}$ in which f is strictly increasing. If $\varepsilon > 0$, then surely J contains points a and b with $a < b$ and $f(b)-f(a) < \varepsilon(b-a)$. As D_r is dense in $\mathbb{D}$, for b we can choose an element of D_r. Then $J \cap (b,2b-a)$ is an open interval of $\mathbb{D}$ that is contained in J.

This observation enables one to construct $a_1,a_2,\ldots \in J$ and $b_1,b_2,\ldots \in J \cap D_r$ such that for all $i \in \mathbb{N}$

(1) $\qquad a_i < b_i,\ f(b_i)-f(a_i) < 2^{-i}(b_i-a_i)$

(2) $\qquad a_i$ and b_i lie in $J \cap \bigcap_{j<i} (b_j,2b_j-a_j)$.

The sequence $b_1,b_2,\ldots$ is increasing: let b be its limit. For all i, $a_i < b_i < a_{i+1}$, so $b = \lim_{i\to\infty} a_i$. It follows from (2) that, for all $j \in \mathbb{N}$, $b \le 2b_j-a_j$, whence $2(b-b_j) \le b-a_j$. Observing that $f'(b) > 0$, for large j we obtain $f(b)-f(b_j) > 0$, whence

$$\frac{f(b)-f(a_j)}{b-a_j} < \frac{f(b)-f(b_j)}{b-a_j} + 2^{-j}\,\frac{b_j-a_j}{b-a_j} < \frac{f(b)-f(b_j)}{2(b-b_j)} + 2^{-j}.$$

This leads to the inequality $f'(b) \le \frac{1}{2}f'(b)$, contradicting $f'(b) > 0$.

3.4. We get better results by restricting ourselves to tamer functions.

Let $f : \mathbb{D} \to \mathbb{R}$ be differentiable. We call f uniformly differentiable if

$$(1)\quad \begin{cases} \text{for every } \varepsilon > 0 \text{ there exists a } \delta > 0 \text{ such that if } a,x,y \in \mathbb{D}, \\ |x-a| < \delta,\ |y-a| < \delta, \text{ then } |f(x)-f(y)-f'(a)(x-y)| \le \varepsilon|x-y|. \end{cases}$$

It is not hard to see that (1) is equivalent to

$$(2)\quad \begin{cases} \text{the function } \Phi f : \mathbb{D}\times\mathbb{D} \to \mathbb{R}, \text{ defined by} \\ \qquad \Phi f(x,y) := \begin{cases} \dfrac{f(x)-f(y)}{x-y} & \text{if } x \neq y \\ f'(x) & \text{if } x = y \end{cases} \\ \text{is continuous (hence, uniformly continuous).} \end{cases}$$

In a similar way one defines uniform differentiability for functions on $[0,1]$. It is well known that the analogs of (1) and (2) for $[0,1]$ are equivalent and also that a function on $[0,1]$ is uniformly differentiable if and only if it has a continuous derivative ("is continuously differentiable"). For $\mathbb{D}$, however, the matter is different. It is clear that a uniformly differentiable function on $\mathbb{D}$ has a continuous derivative (use (2)), but the function f of Example 3.2, whose derivative is the constant function 1, is not uniformly differentiable. For all $k \in \mathbb{N}$ we have $\Phi f(2\cdot 3^{-k}+2\cdot 3^{-2k}, 2\cdot 3^{-k}) = 0$.)

3.5. For uniformly differentiable functions we can strengthen Theorem 3.3:

THEOREM. Let $f : \mathbb{D} \to \mathbb{R}$ be uniformly differentiable. If $a \in \mathbb{D}$ and $f'(a) > 0$, then f is strictly increasing on some open interval of $\mathbb{D}$ that contains a. If $f'(x) > 0$ for all $x \in \mathbb{D}$, then there exists a locally constant function h on $\mathbb{D}$ such that $f+h$ is strictly increasing.

(Caution. Do not expect too much from uniform differentiability. The functions mentioned in Examples 2.2 and 3.1 are uniformly differentiable.)

PROOF. Let $a \in \mathbb{D}$, $f'(a) > 0$. There is a $\delta > 0$ such that for all $x,y \in \mathbb{D} \cap (a-\delta,a+\delta)$, $|\Phi f(x,y)-f'(a)| < f'(a)$, whence $\Phi f(x,y) > 0$. It follows that f is strictly increasing on $\mathbb{D} \cap (a-\delta,a+\delta)$.

Now assume that $f' > 0$ everywhere on $\mathbb{D}$. Let $\varepsilon := \inf\{f'(x) : x \in \mathbb{D}\}$; then $\varepsilon > 0$. By the uniform continuity of Φf there exists an $n \in \mathbb{N}$ such that, if $x,y \in \mathbb{D}$ and $|x-y| \le 3^{-n}$, then $|\Phi f(x,y)-f'(x)| < \varepsilon \le f'(x)$, so that $\Phi f(x,y) > 0$. Thus, if $J \subset \mathbb{R}$ is a closed interval of length 3^{-n}, then f is strictly increasing on $\mathbb{D} \cap J$. Now $\mathbb{D}$ can be covered by 2^n pairwise disjoint closed intervals of length 3^{-n}. The construction of h is obvious.

§4. A brief look at Functional Analysis.

4.1. We consider three vector spaces of functions on $\mathbb{D}$:

$C(\mathbb{D})$: the space of all continuous functions,

$\mathcal{D}(\mathbb{D})$: the space of all differentiable functions,

$C^1(\mathbb{D})$: the space of all uniformly differentiable functions.

Trivially, $C^1(\mathbb{D}) \subset \mathcal{D}(\mathbb{D}) \subset C(\mathbb{D})$.

4.2. $C(\mathbb{D})$ is a Banach space under the sup-norm $\|\ \|_\infty$; according to Theorem 2.1 the locally constant functions are dense in $C(\mathbb{D})$.

If $f \in C^1(\mathbb{D})$, then Φf (see 3.4) is a continuous, hence bounded, function on $\mathbb{D}\times\mathbb{D}$. Therefore, f satisfies a Lipschitz condition and we can define a norm $\|\ \|_{Lip}$ on $C^1(\mathbb{D})$ by

$$\|f\|_{Lip} := \max\{\|f\|_\infty , \|\Phi f\|_\infty\}.$$

4.3. THEOREM. <u>Under $\|\ \|_{Lip}$, $C^1(\mathbb{D})$ is a Banach space. Differentiation is a continuous linear map of $C^1(\mathbb{D})$ onto $C(\mathbb{D})$. Its kernel, $\{f \in C^1(\mathbb{D}) : f' = 0\}$, is a closed linear subspace of $C^1(\mathbb{D})$ and is, in fact, the closure in $C^1(\mathbb{D})$ of the set of all locally constant functions.</u>

PROOF. Most of the theorem is quite easy to prove; we only show that every $f \in C^1(\mathbb{D})$ with $f' = 0$ is a $\|\ \|_{Lip}$-limit of locally constant functions. Let $f \in C^1(\mathbb{D})$, $f' = 0$.

For $n \in \mathbb{N}$, define the locally constant function f_n on $\mathbb{D}$ by setting $f_n(x) := f(x_n)$ $(x \in \mathbb{D})$. In the proof of Theorem 2.1 we have already seen that $\lim_{n\to\infty} f_n = f$ uniformly: it remains to prove that $\lim_{n\to\infty} \|\Phi f - \Phi f_n\|_\infty = 0$.

Take $\varepsilon > 0$. By the uniform continuity of Φf and the fact that $\Phi f(x,x) = 0$ for all $x \in \mathbb{D}$, there exists a $\delta > 0$ such that

$$\text{if } x,y \in \mathbb{D} \text{ and } |x-y| < \delta, \text{ then } |\Phi f(x,y)| \le \varepsilon/4.$$

Let $n \in \mathbb{N}$ be so large that $\|f-f_n\|_\infty \le \varepsilon\delta/6$: we show that $\|\Phi f - \Phi f_n\|_\infty \le \varepsilon$. Take $x,y \in \mathbb{D}$. If $|x-y| \ge \delta/3$, then

$$|\Phi f(x,y) - \Phi f_n(x,y)| \le \frac{2\|f-f_n\|_\infty}{|x-y|} \le \varepsilon.$$

On the other hand, if $|x-y| < \delta/3$, then $|x-y| < \delta$ while, by 1.3 (iv), $|x_n - y_n| \le 3|x-y| < \delta$, so

$$|\Phi f(x,y) - \Phi f_n(x,y)| = |\Phi f(x,y) - \frac{x_n - y_n}{x-y}\Phi f(x_n,y_n)| \le$$

$$\le |\Phi f(x,y)| + 3|\Phi f(x_n,y_n)| \le \varepsilon.$$

4.4. Theorem 3.2 and its proof yield a linear map $P : C(\mathbb{D}) \to C^1(\mathbb{D})$:

$$(Pf)(x) := \int_0^x \tilde{f}(t)\,dt \qquad (x \in \mathbb{D};\ f \in C(\mathbb{D})).$$

This P is a one-sided inverse of the differentiation in the sense that $(Pf)' = f$ for all $f \in C(\mathbb{D})$. Clearly, P is continuous (even isometric) relative to the norms we have imposed on $C(D)$ and $C^1(D)$.

Let $\chi \in C(\mathbb{D})$ be the function $x \mapsto x$ $(x \in \mathbb{D})$. It is easy to see that P produces χ and $\frac{1}{2}\chi^2$ as antiderivatives of the functions 1 and χ, respectively.

But Px^2 is not at all $\frac{1}{3}x^3$. Does there exist a "decent" antiderivation operation that sends each x^n into the corresponding $\frac{1}{n+1}x^{n+1}$? The answer is negative in the following sense.

4.5. THEOREM. <u>There is no linear map</u> $A : C(\mathbb{D}) \to C(\mathbb{D})$ <u>for which</u>

(i) A <u>is continuous relative to the sup-norm</u>,

(ii) $Ax^{n-1} = \frac{1}{n}x^n \quad (n \in \mathbb{N})$.

PROOF. Suppose we have an A satisfying (i) and (ii). Let P be the polynomial function $x \mapsto \frac{9}{2}x(1-x)$ $(x \in [0,1])$. As $\mathbb{D} \subset [0,\frac{1}{3}] \cup [\frac{2}{3},1]$, we have $0 \le P(x) \le$ $\le P(\frac{1}{3}) = 1$ for all $x \in \mathbb{D}$. For $n \in \mathbb{N}$, let $f_n(x) := P(x)^n$ $(x \in \mathbb{D})$. Then $\|f_n\|_\infty = 1$, so, by (i), the sequence $Af_1, Af_2, \ldots$ is bounded. However, as $P(x) \ge \frac{10}{9}$ for all $x \in (\frac{4}{9},\frac{5}{9})$, it follows from (ii) that for every n

$$(Af_n)(1)-(Af_n)(0) = \int_0^1 P(x)^n dx \ge \int_{4/9}^{5/9}(\tfrac{10}{9})^n dx = \frac{1}{9}\cdot(\tfrac{10}{9})^n$$

and we have a contradiction.

§5. Derivation and antiderivation. A second analysis.

5.1. What functions on $\mathbb{D}$ have antiderivatives? A necessary condition for a function to have an antiderivative can be found as follows.

For $n \in \mathbb{N}$, define $\sigma_n : \mathbb{D} \to \mathbb{D}$ by

$$\sigma_n(x) := \begin{cases} x - 2\cdot 3^{-n} & \text{if } \xi_n = 1 \\ x + 2\cdot 3^{-n} & \text{if } \xi_n = 0 \end{cases} \qquad (x = 2\sum_{i=1}^{\infty}\xi_i 3^{-i} \in \mathbb{D}).$$

Each σ_n is continuous; $\lim_{n\to\infty}\sigma_n(x) = x$ for all $x \in \mathbb{D}$; $\sigma_n(x) \ne x$ for all x and n. Hence, if $g : \mathbb{D} \to \mathbb{R}$ is differentiable, then

$$g'(x) = \lim_{n\to\infty}\frac{g(x)-g(\sigma_n(x))}{x-\sigma_n(x)} \qquad (x \in \mathbb{D})$$

and the moral of <u>that</u> is that every derivative is a limit of continuous functions.

Let $\mathcal{D}'(\mathbb{D})$ be the space of all derivatives of differentiable functions $\mathbb{D} \to \mathbb{R}$; let $\mathcal{B}^1(\mathbb{D})$ be the space of all functions $\mathbb{D} \to \mathbb{R}$ that are limits of sequences of continuous functions. ($\mathcal{B}^1(\mathbb{D})$ is the so-called <u>first class of Baire on</u> $\mathbb{D}$.) We have proved:

$$\mathcal{D}'(\mathbb{D}) \subset \mathcal{B}^1(\mathbb{D}).$$

This result immediately leads to a question:

5.2. PROBLEM. <u>Are</u> $\mathcal{D}'(\mathbb{D})$ <u>and</u> $\mathcal{B}^1(\mathbb{D})$ <u>equal</u>? In other words, is a function on $\mathbb{D}$ that is of the first class of Baire necessarily a derivative?

5.3. We have no complete answer to this question, but there is a large subset of $\mathcal{B}^1(D)$ that we will show to be contained in $\mathcal{D}'(\mathbb{D})$.

We define $\mathcal{S}^1(D)$, the <u>first class of Sierpiński on</u> $\mathbb{D}$, to be the linear hull of the set of all lower semicontinuous functions $\mathbb{D} \to \mathbb{R}$. It is well known that every lower semicontinuous function is the limit of an incrasing sequence of continuous functions. (Indeed, $\mathcal{B}^1(\mathbb{D})$ and $\mathcal{S}^1(\mathbb{D})$ can be described in a way that makes this inclusion very clear. If $f : \mathbb{D} \to \mathbb{R}$, then $f \in \mathcal{B}^1(\mathbb{D})$ if and only if there exist continuous functions $f_1, f_2, \ldots$ on $\mathbb{D}$ such that $f = \sum_{n=1}^{\infty} f_n$; while $f \in \mathcal{S}^1(\mathbb{D})$ if and only if there exist continuous functions $f_1, f_2, \ldots$ on $\mathbb{D}$ with $f = \sum_{n=1}^{\infty} f_n$ and such that the series Σf_n converges absolutely. See [2].) Consequently, $\mathcal{S}^1(\mathbb{D}) \subset \mathcal{B}^1(\mathbb{D})$. We are going to prove

$$\mathcal{S}^1(\mathbb{D}) \subset \mathcal{D}'(\mathbb{D}).$$

Unfortunately, certain elements of $\mathcal{B}^1(\mathbb{D})$ do not belong to $\mathcal{S}^1(D)$. They are, however, very complicated and we do not know whether they have antiderivatives.

Our problem can be stated differently. It is known ([2]) that $\mathcal{S}^1(\mathbb{D})$ is uniformly dense in $\mathcal{B}^1(\mathbb{D})$ and also that $\mathcal{B}^1(\mathbb{D})$ itself is uniformly closed. Thus, Problem 5.2 is equivalent to:

5.4. PROBLEM. Is the space $\mathcal{D}'(\mathbb{D})$ uniformly closed?

One can also define spaces $\mathcal{D}'[0,1]$, $\mathcal{B}^1[0,1]$ and $\mathcal{S}^1[0,1]$ of functions on $[0,1]$. It is easy to see that $\mathcal{D}'[0,1] \subset \mathcal{B}^1[0,1]$ but the inclusion $\mathcal{S}^1[0,1] \subset \mathcal{D}'[0,1]$ is evidently false. A good characterization of the elements of $\mathcal{D}'[0,1]$ does not yet exist although many people have been working at it. See [1] for partial results. (It is interesting to observe that $\mathcal{D}'[0,1]$ is uniformly closed.)

One encounters problems analogous to 5.2 and 5.4 when studying calculus over other fields than $\mathbb{R}$. As this is the origin of our proof of Theorem 5.9, it may be good to digress for a moment. Let K be a complete non-Archimedean valued field. (The reader who has lost interest may as well move to 5.5.) Let X be a nonempty subset of K without isolated points. One can define differentation of functions $X \to K$ and introduce the spaces $\mathcal{D}'(X)$, $\mathcal{B}^1(X)$. It is easy to see that $\mathcal{D}'(X) \subset \mathcal{B}^1(X)$. In this situation the converse inclusion can be proved to hold. This was done by Schikhof in [4]. The proof given there can be adapted to the "real" case and yields Theorem 5.9. (It is interesting to note, however, that there is no such space as $\mathcal{S}^1(X)$, owing to the fact that K is not ordered.)

5.5. The proof of the inclusion $\mathcal{S}^1(\mathbb{D}) \subset \mathcal{D}'(\mathbb{D})$ requires a bit of preparation. First a few notations.

For $f : [0,1] \to \mathbb{R}$ we denote by $f|\mathbb{D}$ the restriction of f to $\mathbb{D}$.

If $f : [0,1] \to \mathbb{R}$ is differentiable, we define the function Φf on the square $[0,1]\times[0,1]$ in analogy to what we did in 3.4:

$$\Phi f(x,y) := \begin{cases} \dfrac{f(x)-f(y)}{x-y} & \text{if } x \neq y \\ f'(x) & \text{if } x = y \end{cases} \qquad (x,y \in [0,1]).$$

If g is a function on a set X, we define functions g^+, g^- by:

$$g^+(x) := \max\{0, g(x)\},\ g^-(x) := \max\{0, -g(x)\} \quad (x \in X).$$

5.6. LEMMA. *Let* $f \in C(\mathbb{D})$, $f \geq 0$. *Take* $\varepsilon > 0$. *Then there exists a continuous function* h *on* $[0,1]$ *such that*

$$0 \leq h \leq \tilde{f}$$

$$h|\mathbb{D} = f$$

$$\int_0^1 h(t)\,dt \leq \varepsilon$$

if $0 \leq x \leq y \leq 1$, *then* $\int_x^y h(t)\,dt \leq (\tilde{f}(x)+\varepsilon)(y-x)$.

PROOF. For $x \in [0,1]$ set $d(x) := \inf\{|x-y| : y \in \mathbb{D}\}$; then d is a continuous function on $[0,1]$, $d(x) = 0$ for all $x \in \mathbb{D}$ and $d(x) > 0$ for all $x \in [0,1]\backslash\mathbb{D}$. Consequently, if we define functions $f_1, f_2, \ldots$ on $[0,1]$ by $f_n := (\tilde{f}-nd)^+$, then

$$f_n|\mathbb{D} = f \qquad (n \in \mathbb{N})$$

$$\tilde{f} \geq f_1 \geq f_2 \geq \ldots$$

$$\lim_{n\to\infty} f_n = 0 \text{ pointwise on } [0,1]\backslash\mathbb{D}.$$

Set

$$F_n(x) := \int_0^x f_n(t)\,dt \qquad (x \in [0,1];\ n \in \mathbb{N}).$$

For each n, $||F_n||_\infty = F_n(1) = \int_0^1 f_n(t)\,dt$, so $\lim_{n\to\infty} F_n = 0$ uniformly, according to Lebesgue's Theorem.

Each ΦF_n is a continuous function on the unit square $[0,1]\times[0,1]$. Furthermore, $\Phi F_1 \geq \Phi F_2 \geq \ldots$, and for all $x,y \in [0,1]$,

$$\lim_{n\to\infty} \Phi F_n(x,y) = \begin{cases} 0 & \text{if } x \neq y \\ 0 & \text{if } x = y \in [0,1]\backslash\mathbb{D} \\ f(x) & \text{if } x = y \in \mathbb{D}. \end{cases}$$

With $k(x,y) := \tilde{f}(x)$ $(x,y \in [0,1])$ we have $\lim_{n\to\infty} \Phi F_n \le k$ pointwise, so $\lim_{n\to\infty} (\Phi F_n - k)^+ = 0$ pointwise; by Dini's Theorem $\lim_{n\to\infty} (\Phi F_n - k)^+ = 0$ uniformly. Hence, there is an $N \in \mathbb{N}$ with

$$||F_N||_\infty < \varepsilon$$

$$\Phi F_N - k \le (\Phi F_N - k)^+ \le \varepsilon \text{ everywhere on } [0,1].$$

Choose $h := f_N$. It is perfectly easy to check that this h satisfies the requirements.

5.7. THEOREM. Let $f : \mathbb{D} \to \mathbb{R}$ be lower semicontinuous, $f \ge 0$. Then there is a differentiable $F : [0,1] \to \mathbb{R}$ such that F' is lower semicontinuous, $F' \ge 0$ and $F'|\mathbb{D} = f$.

PROOF. Choose $f_1, f_2, \ldots \in C(\mathbb{D})$, $f_n \ge 0$, with $\sum_{n=1}^{\infty} f_n = f$ pointwise. The preceding lemma enables us to choose, for each n, a continuous $h_n : [0,1] \to \mathbb{R}$ for which $0 \le h_n \le \tilde{f}_n$, $\int_0^1 h_n(t)\,dt \le 2^{-n}$, $h_n|\mathbb{D} = f_n$ and $|\int_x^y h_n(t)\,dt| \le (\tilde{f}_n(x) + 2^{-n})|x-y|$ $(x,y \in [0,1])$. Define

$$F(x) := \sum_{n=1}^{\infty} \int_0^x h_n(t)\,dt \qquad (x \in [0,1]).$$

(The sum converges since $\int_0^1 h_n(t)\,dt \le 2^{-n}$ for all n.) We have $\sum_{n=1}^{\infty} f_n = f$ pointwise on $\mathbb{D}$, hence $\sum_{n=1}^{\infty} \tilde{f}_n = \tilde{f}$ pointwise on $[0,1]$ and therefore $0 \le \sum_{n=1}^{\infty} h_n \le \tilde{f}$ pointwise on $[0,1]$. Set $h := \sum_{n=1}^{\infty} h_n$. Then $h|\mathbb{D} = f$ and we are done if we can prove that $F' = h$.

Take $a \in [0,1]$ and $\varepsilon > 0$: we show that, if $x \in [0,1]\setminus\{a\}$ is close enough to a, then $\left|\frac{F(x)-F(a)}{x-a} - h(a)\right| \le \varepsilon$.

For all $N \in \mathbb{N}$ and $x \in [0,1]$, $x \ne a$,

$$\left|\frac{F(x)-F(a)}{x-a} - h(a)\right| = \left|\sum_{n=1}^{\infty} \{(x-a)^{-1}\int_a^x h_n(t)\,dt - h_n(a)\}\right| \le$$

$$(*) \qquad \le \sum_{n\le N} \left|(x-a)^{-1}\int_a^x h_n(t)\,dt - h_n(a)\right| + \sum_{n>N} \{(\tilde{f}_n(a)+2^{-n}) + h_n(a)\}.$$

Choose an N so large that the second sum in (*) is smaller than $\frac{1}{2}\varepsilon$: this can be done because the series $\Sigma_n(\tilde{f}_n(a)+2^{-n})$ and $\Sigma_n h_n(a)$ both converge. For this N, take a positive δ such that the first sum in (*) is smaller than $\frac{1}{2}\varepsilon$ for all $x \in [0,1]$ with $0 < |x-a| < \delta$. Then for such x we have the desired inequality.

5.8. THEOREM. <u>Let</u> f <u>and</u> g <u>be lower semicontinuous functions on</u> $\mathbb{D}$, $g \geq f \geq 0$. <u>Then there exist differentiable functions</u> F <u>and</u> G <u>on</u> $[0,1]$ <u>with</u> $F'|\mathbb{D} = f$, $G'|\mathbb{D} = g$ <u>and</u> $G' \geq F'$.

PROOF. By Theorem 5.7 there exists a differentiable $G : [0,1] \to \mathbb{R}$ with a lower semicontinuous nonnegative derivative G' that is an extension of g. Observe that $\tilde{f}$ is lower semicontinuous on $[0,1]$. Then so is the function $\tilde{f} \wedge G' : x \mapsto \min\{\tilde{f}(x), G'(x)\}$ $(x \in [0,1])$. Consequently, $\tilde{f} \wedge G'$ is a limit of an increasing sequence of continuous functions on $[0,1]$. As $\tilde{f} \wedge G' \geq 0$, we can choose continuous nonnegative functions $k_1, k_2, \ldots$ on $[0,1]$ for which $\sum_{n=1}^{\infty} k_n = \tilde{f} \wedge G'$. With $f_n := k_n|\mathbb{D}$ $(n \in \mathbb{N})$ we have $f_n \in C(\mathbb{D})$, $f_n \geq 0$ and $\sum_{n=1}^{\infty} f_n = f$. Choose h_n as in the proof of Theorem 5.7, with, in addition, $h_n \leq k_n$. Make h and F as in the proof of Theorem 5.7. Then F is differentiable, $F'|\mathbb{D} = f$ and $F' = h = \sum_{n=1}^{\infty} h_n \leq \sum_{n=1}^{\infty} k_n = \tilde{f} \wedge G' \leq G'$.

As an immediate consequence of this theorem we obtain:

5.9. THEOREM. <u>Let</u> $f \in \mathcal{S}^1(\mathbb{D})$. <u>Then</u> f <u>has an antiderivative. In fact, there is a differentiable function</u> F <u>on</u> $[0,1]$ <u>such that</u> $F' \in \mathcal{S}^1[0,1]$ <u>and</u> $f = F'|\mathbb{D}$. <u>If</u> $f \geq 0$, <u>for</u> F <u>one can take an increasing function</u>.

5.10. In connection with the last part of the above theorem it may be worth noticing that there is no linear map $A : \mathcal{S}^1(\mathbb{D}) \to C(\mathbb{D})$ with

(1) for all $f \in \mathcal{S}^1(\mathbb{D})$, Af is an antiderivative of f,

(2) if $f \in \mathcal{S}^1(\mathbb{D})$ and $f \geq 0$, then Af is increasing.

Indeed, suppose we have such an A. Let $\omega := (Ae)(1)-(Ae)(0)$ where e is the constant function 1 on $\mathbb{D}$; for $a \in \mathbb{D}$, let $\omega(a) := (Ae_a)(1)-(Ae_a)(0)$ where $e_a : \mathbb{D} \to \mathbb{R}$ is the characteristic function of the set $\{a\}$. If $a_1,\ldots,a_n \in \mathbb{D}$ are pairwise distinct, then $e-(e_{a_1}+\ldots+e_{a_n})$ is a nonnegative element of $\mathcal{S}^1(\mathbb{D})$, so $A(e-(e_{a_1}+\ldots+e_{a_n}))$ is increasing, whence $\omega \geq \omega(a_1)+\ldots+\omega(a_n)$. It follows that there can exist only countably many elements a of $\mathbb{D}$ with $\omega(a) > 0$. On the other hand, for <u>every</u> $a \in \mathbb{D}$, $A(e_a)$ is an increasing and nonconstant function, so $\omega(a) > 0$. Contradiction.

§6. Functions of two variables.

If $F : [0,1]\times[0,1] \to \mathbb{R}$ is a function that has enough continuous partial derivatives, then the mixed partial derivatives of the second order are equal. It is not always realized how heavily the proofs of this classical theorem rely on the connectedness of the domain of definition. Theorem 6.2 illustrates quite dramatically the difference between functions on $[0,1]\times[0,1]$ and functions on $\mathbb{D}\times\mathbb{D}$. It may be formulated as "all continuous differential forms on $\mathbb{D}\times\mathbb{D}$ are exact".

This section, like the preceding one, is based upon work done in non-Archimedean analysis. Our proof of Theorem 6.2 is a direct translation of the reasoning followed by D. Treiber in [5].

6.1. In $\mathbb{R}^2$ we adopt the maximum norm, $\| \ \|_\infty$, and the canonical inner product, $\langle \ , \ \rangle$. The definition of "partial derivative" of a function on $\mathbb{D}^2$ is obvious.

Let $F : \mathbb{D}^2 \to \mathbb{R}$ and $f : \mathbb{D}^2 \to \mathbb{R}^2$. We say that F is <u>uniformly differentiable</u> with <u>gradient</u> f if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that, if $a,x,y \in \mathbb{D}^2$ and $||x-a||_\infty < \delta$, $||y-a||_\infty < \delta$, then

(*) $$|F(x)-F(y)-\langle f(a),x-y\rangle| \le \varepsilon||x-y||_\infty.$$

(Then, of course, the coordinates of f are the partial derivatives of F and they are continuous.)

6.2. THEOREM. <u>If</u> $f_1,f_2 : \mathbb{D}^2 \to \mathbb{R}$ <u>are continuous, then there exists a uniformly differentiable</u> $F : \mathbb{D}^2 \to \mathbb{R}$ <u>whose partial derivatives are</u> f_1 <u>and</u> f_2.

PROOF. For every $x \in \mathbb{D}^2$ there exist unique $\alpha_i,\beta_i \in \{0,1\}$ $(i \in \mathbb{N})$ such that

$$x = (2\sum_{i=1}^{\infty}\alpha_i 3^{-i}, 2\sum_{i=1}^{\infty}\beta_i 3^{-i}).$$

Using this notation, for $x \in \mathbb{D}^2$ and $n \in \mathbb{N}$ we define $x_n \in \mathbb{D}^2$ by

$$x_n := (2\sum_{i=1}^{n}\alpha_i 3^{-i}, 2\sum_{i=1}^{n}\beta_i 3^{-i}).$$

Further, we set $x_0 := (0,0) \in \mathbb{D}^2$ and $\Delta x_n := x_n - x_{n-1} = 2\cdot 3^{-n}(\alpha_n,\beta_n)$ $(n \in \mathbb{N})$. For all $x,y \in \mathbb{D}^2$ we have

(i) $||x-x_n||_\infty \le 3^{-n}$ $(n = 0,1,2,\ldots)$

(ii) if $n \in \{0,1,2,\ldots\}$ and $||x-y||_\infty < 3^{-n}$, then $x_n = y_n$

(iii) $||\Delta x_n||_\infty \le 2\cdot 3^{-n}$ $(n = 1,2,\ldots)$

(iv) $x = \sum_{n=1}^{\infty} \Delta x_n$.

Now let $f_1,f_2 : \mathbb{D}^2 \to \mathbb{R}$ be continuous; define $f : \mathbb{D}^2 \to \mathbb{R}^2$ by

$$f(x) := (f_1(x),f_2(x)) \qquad (x \in \mathbb{D}^2)$$

By (iii), we can introduce a continuous $F : \mathbb{D}^2 \to \mathbb{R}$ as follows:

$$F(x) := \sum_{n=1}^{\infty} \langle f(x_n),\Delta x_n\rangle \qquad (x \in \mathbb{D}^2).$$

Let $\varepsilon > 0$.

Choose $N \in \mathbb{N}$ such that $||f(x)-f(y)||_\infty \le \varepsilon/12$ for all $x,y \in \mathbb{D}^2$ with $||x-y||_\infty \le 2\cdot 3^{-N}$ and put $\delta := 3^{-N}$.

Take $a,x,y \in \mathbb{D}^2$ with $||x-a||_\infty < \delta$, $||y-a||_\infty < \delta$: we prove (*) Without restriction, let $x \neq y$. Set $m := \min\{n \in \mathbb{N} : x_n \neq y_n\}$. (Here we use (i).) If $n \in \mathbb{N}$ and $n < m$, then $x_n = y_n$ and $\Delta x_n = \Delta y_n$. Hence, by (iv) and by the definition of F, we obtain

$$\begin{aligned} &F(x)-F(y)-\langle f(a),x-y\rangle = \\ &\quad = \sum_{n=1}^{\infty} \{\langle f(x_n),\Delta x_n\rangle - \langle f(y_n),\Delta y_n\rangle - \langle f(a),\Delta x_n-\Delta y_n\rangle\} = \\ &\quad = \sum_{n=m}^{\infty} \{\langle f(x_n),\Delta x_n\rangle - \langle f(y_n),\Delta y_n\rangle - \langle f(a),\Delta x_n-\Delta y_n\rangle\} = \\ &\quad = \sum_{n=m}^{\infty} \langle f(x_n)-f(a),\Delta x_n\rangle - \sum_{n=m}^{\infty} \langle f(y_n)-f(a),\Delta y_n\rangle. \end{aligned}$$

As $||x-a||_\infty < \delta = 3^{-N}$, by (ii) we have $x_N = a_N$ and, similarly, $y_N = a_N$. Hence, $x_N = y_N$, so $N < m$. Then, in view of (i), for all $n \ge m$ we have $||x_n-a||_\infty \le 2\cdot 3^{-N}$ and, consequently, $||f(x_n)-f(a)||_\infty \le \varepsilon/12$. For the same reasons, $||f(y_n)-f(a)||_\infty \le \varepsilon/12$ if $n \ge m$. Thus, from (iii) we deduce that

$$|F(x)-F(y)-\langle f(a),x-y\rangle| \le 2\sum_{n=m}^{\infty} 2\cdot\frac{\varepsilon}{12}\cdot 2\cdot 3^{-n} = \varepsilon\cdot 3^{-m}.$$

Now $x_m \neq y_m$, so $||x-y||_\infty \ge 3^{-m}$ (by (ii)) and we are done.

6.3. It follows easily that there exists a functions F on $\mathbb{D}^2$ with continuous partial derivatives of all orders and such that

$$\frac{\partial^2 F}{\partial x \partial y} = 0,\ \frac{\partial^2 F}{\partial y \partial x} = 1 \text{ everywhere.}$$

But if one defines C^2-functions in the spirit of 3.4(2) (continuity of second order partial difference quotients) then one proves without trouble that for such C^2-functions $F : \mathbb{D}\times\mathbb{D} \to \mathbb{R}$ one has

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x}.$$

6.4. Theorem 6.2 works just as well for complex valued functions on $\mathbb{D}^2$ as for real valued ones. Identifying $\mathbb{R}^2$ with $\mathbb{C}$ (and thus viewing $\mathbb{D}^2$ as a subset of $\mathbb{C}$) one easily derives that for every continuous function $f : \mathbb{D}^2 \to \mathbb{C}$ there exists an $F : \mathbb{D}^2 \to \mathbb{C}$ that is complex differentiable with derivative f.

This result can be generalized: if X is any compact totally disconnected subset of $\mathbb{C}$ without isolated points, then every continuous function $X \to \mathbb{C}$ has an antiderivative. (The same conclusion holds if X is a continuously differentiable simple arc, but it is false for the set $X = \{z \in \mathbb{C} : \mathrm{Re}\, z \in [0,1];\ \mathrm{Im}\, z = 0 \text{ or } (\mathrm{Im}\, z)^{-1} \in \mathbb{N}\}$.)

References.

1. A.M. Bruckner,
Differentiation of Real Functions,
Lecture Notes on Mathematics, 659, Springer-Verlag, Berlin (1978).

2. S. Kempisty,
Sur les séries itérées des fonctions continues,
Fund. Math. 2 (1921), 64-73.

3. W.H. Schikhof,
The Set of Derivatives in a non-Archimedean Field,
Math. Ann. 216 (1975), 67-70.

4. W. Sierpiński,
Sur les fonctions développables en séries absolument convergentes de fonctions continues,
Fund. Math. 2 (1921) 15-27.

5. D. Treiber,
A non-Archimedean "curve integral" and its applications to the construction of potentials and solutions of differential equations,
Bull. Soc. Math. France, Mém. 39-40 (1974), 359-374.