SURVEY

Real Analysis Exchange
Vol. 41(1), 2016, pp. 19–54

Krzysztof Chris Ciesielski, Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310 and Department of Radiology, MIPG, University of Pennsylvania, Blockley Hall – 4th Floor, 423 Guardian Drive, Philadelphia, PA 19104-6021. email: KCies@math.wvu.edu

David Miller, Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310.

A CONTINUOUS TALE ON CONTINUOUS AND SEPARATELY CONTINUOUS FUNCTIONS

This paper is dedicated to the memory of prof. Zbigniew Piotrowski, who inspired our work on the subject of separate continuity

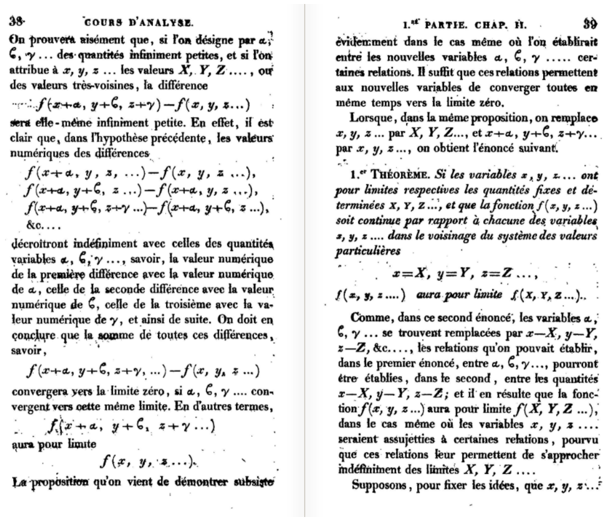
Abstract

Assume that you have developed a good set of tools allowing you to decide which real functions of one real variable, $f\colon \mathbb{R} \to \mathbb{R}$, are continuous. (Q): *How can such a tool-box be utilized to decide on the continuity of the functions* $g\colon \mathbb{R}^n \to \mathbb{R}$ *of* n *real variables?* This is one of the questions which must be faced by any student taking multivariable calculus. Of course, such a student is following the footsteps of many generations of mathematicians, which were, and still are, struggling with the same general question. The aim of this article is to present the history and the current research related to this subject in a real analysis perspective, rather than in a more general, topological perspective.

In addition to surveying the results published so far, this exposition includes also several original results (Theorems 1, 12 and 13), as well as some new simplified versions of the (sketches of the) proofs of older results. We also recall several intriguing open problems.

Mathematical Reviews subject classification: Primary: 26B05; Secondary: 26-02, 01-02, 54C05, 54C08
Key words: separate continuity, linear continuity
Received by the editors June 16, 2015
Communicated by: P. D. Humke

Contents



38 COURS D'ANALYSE.

On prouvera aisément que, si l'on désigne par α, β, γ ... des quantités infiniment petites, et si l'on attribue à x, y, z ... les valeurs X, Y, Z ...., ou des valeurs très-voisines, la différence

$$f(x+\alpha,\ y+\beta,\ z+\gamma)-f(x,\ y,\ z\ldots)$$

sera elle-même infiniment petite. En effet, il est clair que, dans l'hypothèse précédente, les valeurs numériques des différences

$$f(x+\alpha,\ y,\ z,\ \ldots)-f(x,\ y,\ z\ \ldots),$$
$$f(x+\alpha,\ y+\beta,\ z\ \ldots)-f(x+\alpha,\ y,\ z\ \ldots),$$
$$f(x+\alpha,\ y+\beta,\ z+\gamma\ \ldots)-f(x+\alpha,\ y+\beta,\ z\ \ldots),$$
&c....

décroîtront indéfiniment avec celles des quantités variables α, β, γ ..., savoir, la valeur numérique de la première différence avec la valeur numérique de α, celle de la seconde différence avec la valeur numérique de β, celle de la troisième avec la valeur numérique de γ, et ainsi de suite. On doit en conclure que la somme de toutes ces différences, savoir,

$$f(x+\alpha,\ y+\beta,\ z+\gamma,\ \ldots)-f(x,\ y,\ z\ \ldots)$$

convergera vers la limite zéro, si α, β, γ .... convergent vers cette même limite. En d'autres termes,

$$f(x+\alpha,\ y+\beta,\ z+\gamma\ \ldots)$$

aura pour limite

$$f(x,\ y,\ z\ \ldots).$$

La proposition qu'on vient de démontrer subsiste

1.^{re} PARTIE. CHAP. II. 39

évidemment dans le cas même où l'on établirait entre les nouvelles variables α, β, γ ..... certaines relations. Il suffit que ces relations permettent aux nouvelles variables de converger toutes en même temps vers la limite zéro.

Lorsque, dans la même proposition, on remplace x, y, z ... par X, Y, Z..., et $x+\alpha$, $y+\beta$, $z+\gamma$... par x, y, z..., on obtient l'énoncé suivant.

1.^{er} Théorème. *Si les variables x, y, z ... ont pour limites respectives les quantités fixes et déterminées X, Y, Z ..., et que la fonction $f(x, y, z\ldots)$ soit continue par rapport à chacune des variables x, y, z .... dans le voisinage du système des valeurs particulières*

$$x=X,\ y=Y,\ z=Z\ \ldots,$$

$f(x, y, z\ldots.)$ *aura pour limite* $f(X, Y, Z\ldots)$.

Comme, dans ce second énoncé, les variables α, β, γ ... se trouvent remplacées par $x-X$, $y-Y$, $z-Z$, &c...., les relations qu'on pouvait établir, dans le premier énoncé, entre α, β, γ..., pourront être établies, dans le second, entre les quantités $x-X$, $y-Y$, $z-Z$; et il en résulte que la fonction $f(x, y, z\ldots)$ aura pour limite $f(X, Y, Z\ldots)$, dans le cas même où les variables x, y, z .... seraient assujetties à certaines relations, pourvu que ces relations leur permettent de s'approcher indéfiniment des limites X, Y, Z ....

Supposons, pour fixer les idées, que x, y, z...

Figure 1: Theorem X in *Cours d'analyse*

1 Prehistory: Separate vs Joint Continuity

1.1 A Tale of Cauchy: Was his theorem false?

A function f of $n > 1$ variables $x_1, \ldots, x_n$ is said to be *separately continuous* provided f is continuous with respect to each variable separately. That is, for every fixed n-tuple $\langle x_1, \ldots, x_n \rangle$ from the domain of f and for every $i \in \{1, \ldots, n\}$, the mapping $t \mapsto f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$ is continuous. In addition, we say that f (of more than one variable) is *jointly continuous* if it is continuous in the usual (topological) sense.

Augustin-Louis Cauchy (1789-1857)[1] in his 1821 mathematical analysis textbook *Cours d'analyse* included the following result [9, pp. 38-39], see Figure 1. (Compare also [6, pp. 29].)

Theorem X: *A separately continuous function of real variables is continuous.*

This is a very appealing result. It has, however, one major flaw: the following simple function

$$g(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{for } \langle x,y \rangle \neq \langle 0,0 \rangle \\ 0 & \text{for } \langle x,y \rangle = \langle 0,0 \rangle \end{cases} \tag{1}$$

[1] We include the pictures and birth-to-death years of the main contributors to this story who are already deceased.

included in most modern texts of multivariable calculus, constitutes a counterexample for Theorem X! Thus, the conclusion seems inevitable:

Theorem X is false! Cauchy made a mistake!

Indeed, the claims of the falsehood of Theorem X can be found in [29, p. 952], [42, p. 115], and in the commented translation [6, pp. 29] of *Cours d'analyse*, while the papers [27] and [11, 12] talk openly on Cauchy's mistake/error. (Yes, the first author of this article is the same as for [11, 12].) However, as unbelievable as it may sound,

both Theorem X and its "counterexample" (1) are correct!

This discrepancy is caused by the fact, that Cauchy's text is written for the notion of (the ordered field of) real numbers different from the one we commonly use today:

Cauchy's set of real numbers, denoted here by $\mathscr{R}$, is non-Archimedean.

That is, it contains the infinitesimal numbers (i.e., positive but less than $1/n$ for all $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$), while we take for granted that the (standard) set $\mathbb{R}$ of real numbers is Archimedean, or, in other words, contains no infinitesimals. (See e.g. [5] and [19] for more on the subject.) The early calculus, including the work of Newton and Leibnitz, was all done with the infinitesimal real numbers (i.e., some version on $\mathscr{R}$). In particular, Cauchy defines a function $f(x)$ to be continuous at x provided

(C) *an infinitely small increment in the variable always produces an infinitely small increment in the function itself* (i.e., $f(x+dx) - f(x)$ is infinitesimal, provided so is dx).

Then, his proof of Theorem X (for $n = 2$) is by noticing that, if dx and dy are infinitesimal, then so is the quantity

$$\begin{aligned} |f(x+dx, y+dy) - f(x, y)| &\leq |f(x+dx, y+dy) - f(x+dx, y)| \\ &+ |f(x+dx, y) - f(x, y)|, \end{aligned}$$

as a sum of two infinitesimals is infinitesimal — a perfectly legitimate argument.

How does this argument relate to the example (1)? Basically, g given by (1), when considered on $\mathscr{R}$, is well defined, but *not* separately continuous. The easiest way to see this, is to take as both infinitesimals dx and

A. Cauchy

A. Robinson

dy a number $[1/i]_{i\in\mathbb{N}}$ given by a sequence $\langle 1/i\rangle_{i\in\mathbb{N}}$, in the sense of the nonstandard analysis [43] of Abraham Robinson (1918-1974), who incorporated infinitesimal and infinite numbers in a rigorous way into the mathematics based on $\mathbb{R}$.[2] The number $[1/i]_{i\in\mathbb{N}} \in \mathscr{R}$ is infinitesimal, since any standard number $r \in \mathbb{R}$ is, in Robinson's nonstandard analysis, identified with a constant sequence $\bar{r} = [r]_{i\in\mathbb{N}}$ and, for every $r > 0$, $\bar{0} < [1/i]_{i\in\mathbb{N}} < \bar{r}$ since $0 < 1/i < r$ for all but finitely many i. Now, for $dx = [1/i]_{i\in\mathbb{N}}$, the mapping $t \mapsto g(dx, t)$ is not continuous (in Cauchy's sense (C)) at $t = \bar{0}$, since $g(dx, dy) - g(dx, \bar{0}) = \left[\frac{(1/i)(1/i)}{(1/i)^2+(1/i)^2}\right]_{i\in\mathbb{N}} = [1/2]_{i\in\mathbb{N}}$ is not infinitesimal.

Cleave [16, p. 276], when discussing the discrepancy between Theorem X and example (1), writes that "Cauchy assumed that the neighborhood of a point contained all points infinitesimally close to it" and so, there is nothing wrong with Theorem X (in Cauchy's formalism). However, Cleave does not address an issue of whether Theorem X can be interpreted in our standard analysis of the Archimedean set $\mathbb{R}$ of real numbers.

Felscher [21, p. 857], after Robinson [43], writes that Cauchy's developments can be translated to today's analysis formalism by *reading it with today's standard knowledge of distinctions such as continuity versus uniform continuity, etc.* It is easy to see that the notion of uniform continuity itself does not help here, since the restriction of g from (1) to the square $[-1, 1]^2$ is still discontinuous, but it is separately uniformly continuous. However, the notion of equicontinuous functions does the job, as we have the following theorem. Recall that a family $\mathcal{F}$ of functions from $\mathbb{R}$ to $\mathbb{R}$ is *equicontinuous* provided for every $\varepsilon > 0$ and $x_0 \in \mathbb{R}$ there is a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for

[2]This means, for $\mathscr{R} = \mathbb{R}^{\mathbb{N}}/\mathcal{U}$, where $\mathcal{U}$ is an ultrafilter (i.e., a maximal family of subsets of $\mathbb{N}$, having finite intersection property) containing all co-finite sets and the sequences $[x_i]_{i\in\mathbb{N}}, [y_i]_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ are identified, when $\{i \in \mathbb{N} \colon x_i = y_i\} \in \mathcal{U}$. Similarly, for any relation $\ll$ among $\leq$, $<$, $\geq$, and $>$, we have $[x_i]_{i\in\mathbb{N}} \ll [y_i]_{i\in\mathbb{N}}$ if, and only if, $\{i \in \mathbb{N} \colon x_i \ll y_i\} \in \mathcal{U}$.

every $f \in \mathcal{F}$ and $x \in (x_0 - \delta, x_0 + \delta)$. To our best knowledge, the following theorem appears in print for the first time.

Theorem 1. *If $\bar{g}(x, y)$ is a separately continuous function in the sense of Theorem X (i.e, from $\mathscr{R}^2$ into $\mathscr{R}$), then its restriction g to $\mathbb{R}^2$ has the property*

(EC) *the families $\{g(x, \cdot)\colon x \in [-M, M]\}$ and $\{g(\cdot, y)\colon y \in [-M, M]\}$ are equicontinuous for every $M > 0$.*[3]

Moreover, if $g\colon \mathbb{R}^2 \to \mathbb{R}$ satisfies (EC), then it is continuous.

PROOF. To see the first part assume, by way of contradiction, that (EC) is false for some $M > 0$. So, either $\{g(x, \cdot)\colon x \in [-M, M]\}$ or $\{g(\cdot, y)\colon y \in [-M, M]\}$ is not equicontinuous. By symmetry, we can assume that this is the case for $\{g(\cdot, y)\colon y \in [-M, M]\}$. Therefore, there exist $\varepsilon > 0$ and $x \in \mathbb{R}$ such that for every $\delta = 1/n$ there exist $y_n \in [-M, M]$ and $x_n \in (x - 1/n, x + 1/n)$ with $|g(x_n, y_n) - g(x, y_n)| \geq \varepsilon$. Choosing a subsequence, if necessary, we can assume that $[y_n]_n$ converges to a $y \in [-M, M]$. Then $dx = [x_n - x]_{n\in\mathbb{N}} \in \mathscr{R}$ and $dy = [y_n - y]_{n\in\mathbb{N}} \in \mathscr{R}$ are infinitesimals and $\bar{g}(\cdot, y + dy)$ is discontinuous at $\bar{x}$, as $|g(\bar{x} + dx, \bar{y} + dy) - g(\bar{x}, \bar{y} + dy)| = [|g(x_n, y_n) - g(x, y_n)|]_{n\in\mathbb{N}} \geq \bar{\varepsilon}$ is not infinitesimal. This contradicts the assumption that $\bar{g}(x, y)$ is a separately continuous.

The additional part follows from the first part and Theorem X. The proof, without using infinitesimals, is as follows. Fix $\langle x_0, y_0\rangle \in \mathbb{R}^2$ and $\varepsilon > 0$. Choose an $M > 0$ such that $|x_0| < M - 1$ and $\delta \in (0, 1)$ for which:

- $|g(x, y_0) - g(x_0, y_0)| < \varepsilon/2$ whenever $|x - x_0| < \delta$ (ensured by the continuity of $g(\cdot, y_0)$ at x_0),
- $|g(x, y) - g(x, y_0)| < \varepsilon/2$ whenever $x \in [-M, M]$ and $|y - y_0| < \delta$ (ensured by the equicontinuity of $\{g(x, \cdot)\colon x \in [-M, M]\}$ at y_0).

Then, $|g(x, y) - g(x_0, y_0)| \leq |g(x, y) - g(x, y_0)| + |g(x, y_0) - g(x_0, y_0)| < \varepsilon$ whenever $|x - x_0| < \delta$ and $|y - y_0| < \delta$, insuring continuity of g at $\langle x_0, y_0\rangle$. $\square$

The equicontinuity condition (EC) is also related to the notion of *separate continuity in the strong sense* introduced in 1998-99 by Omar P. Dzagnidze [18]: a function $f\colon \mathbb{R}^n \to \mathbb{R}$ has this property provided

$$\lim_{x \to x^0} [f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i^0, x_{i+1}, \ldots, x_n)] = 0$$

[3] Actually, the intervals $[-M, M]$ in (EC) could be replaced with $\mathbb{R}$. But then, the number $y + dy$ in the proof would need to be replaced with $[y_n]_{n\in\mathbb{N}} \in \mathscr{R}$, which could be infinitely large (i.e., $> n$ for all $n \in \mathbb{N}$).

for every $x^0 = \langle x^0_1, \ldots, x^0_n\rangle$ and $i \in \{1, \ldots, n\}$. It is easy to see that separate continuity of f in the strong sense follows from the fact that the family $\{f(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_n)\}\colon x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in [-M.M]\}$ is equicontinuous for every $i \in \{1, \ldots, n\}$ and $M > 0$. In particular, the second part of Theorem 1 follows also from

Theorem 2. (Dzagnidze 1998) *If $f\colon \mathbb{R}^n \to \mathbb{R}$ is separately continuous function in the strong sense, then it is continuous.*

— 173 —

Ora si ha, considerando le $y_1 \ldots y_n$ funzioni di $x_1 \ldots x_n$ funzioni di $x_1 y_2 \ldots y_n$,

$$\frac{D(y_1 y_2 \ldots y_n)}{D(x_1 y_2 \ldots y_n)} = \frac{D(y_1 y_2 \ldots y_n)}{D(x_1 x_2 \ldots x_n)} \cdot \frac{D(x_1 x_2 \ldots x_n)}{D(x_1 y_2 \ldots y_n)};$$

ma

$$\frac{D(y_1 y_2 \ldots y_n)}{D(x_1 y_2 \ldots y_n)} = \frac{dy_1}{dx_1}; \quad \frac{D(y_1 \ldots y_n)}{D(x_1 \ldots x_n)} = 0 \quad \text{per ipotesi},$$

$$\frac{D(x_1 x_2 \ldots x_n)}{D(x_1 y_2 \ldots y_n)} = \frac{D(x_2 \ldots x_n)}{D(y_2 \ldots y_n)} = \frac{1}{\frac{D(y_2 \ldots y_n)}{D(x_2 \ldots x_n)}} = \text{quantità finita},$$

onde $\frac{dy_1}{dx_1} = 0$; ossia sostituendo in

$$y_1 = \varphi_1(x_1 \ldots\ldots x_n)$$

ad $x_2 \ldots x_n$ le loro funzioni di $x_1 y_2 \ldots y_n$, y_1 diventa indipendente da x_1, e solo funzione di $y_2 \ldots y_n$, c. v. d.

Esercizii.

123. — 1° La funzione

$$\frac{x-y}{x+y}$$

col tendere di x ed y a zero non tende ad alcun limite, ma assume in ogni intorno dei valori (0, 0) ogni valore.

2° — La funzione

$$f(x, y) = \frac{2xy}{x^2+y^2}$$

col tendere di x ed y a zero non tende ad alcun limite, ma assume nelle vicinanze dei valori (0, 0) delle variabili tutti i valori compresi fra -1 e $+1$. In essa

$$\lim_{x=0} f(x, 0) = \lim_{y=0} f(0, y) = 0$$

e se si attribuisce alla funzione il valore 0 quando si faccia $x=0$ ed $y=0$ essa è per tutte le coppie di valori di x ed y funzione

— 174 —

continua di x e funzione continua di y senz'essere funzione continua di x ed y considerate insieme.

3° — La funzione

$$f(x, y) = \frac{2xy^2}{x^2+y^4}$$

è tale che posto

$$x = ht, \quad y = kt,$$

e facendo tendere t a zero il limite di $f(x, y)$ è sempre zero qualunque siano h e k, ossia se x ed y sono le coordinate cartesiane d'un punto del piano, in qualunque direzione si faccia accostare il punto (x, y) al punto (0, 0) il limite della funzione è sempre zero; tuttavia $f(x, y)$ col tendere di x ed y a zero non tende verso alcun limite, ma in ogni intorno dei valori (0, 0) essa assume tutti i valori compresi fra -1 e $+1$.

4° — La derivata parziale d'un determinante rispetto ad un suo elemento è il suddeterminante complementare di quell'elemento.

5° — La funzione

$$f(x, y) = xy\frac{x^2-y^2}{x^2+y^2}$$

ove si faccia $f(0, 0) = 0$, è funzione continua delle variabili x, y, ed ammette le derivate prime

$$f'_x(x, y) = y\frac{x^2-y^2}{x^2+y^2} + 4y\frac{x^2y^2}{(x^2+y^2)^2},$$

$$f'_y(x, y) = x\frac{x^2-y^2}{x^2+y^2} - 4x\frac{x^2y^2}{(x^2+y^2)^2},$$

e

$$f'_x(0, 0) = f'_y(0, 0) = 0,$$

che sono pure continue; ma

$$f''_{xy}(0, 0) = -1, \quad f''_{yx}(0, 0) = 1,$$

onde non è lecito scambiare l'ordine delle derivazioni. In questo caso le derivate seconde sono discontinue.

Figure 2: Examples (1) and (2) in the Genocchi-Peano text

1.2 A Tale of Heine and Peano: Counterexamples

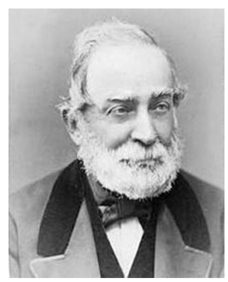
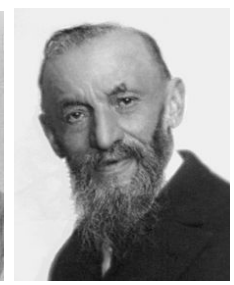
In the mid 19th century mathematicians began abandoning the use of infinitesimals in analysis. This, eventually, lead to realization of the existence of discrepancies described above. In particular, the first counterexample to Theorem X for the functions of standard real variables $\mathbb{R}$ appeared in the 1870 calculus text of J. Thomae [49, pp. 13–16]. It was due to E. Heine (see [42]) and was defined as $f(y, z) = \sin\left(4 \arctan \frac{y}{z}\right)$ for $z \neq 0$ and $f(y, 0) = 0$.

The example (1) as well as the following,

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{for } \langle x,y\rangle \neq \langle 0,0\rangle \\ 0 & \text{for } \langle x,y\rangle = \langle 0,0\rangle, \end{cases} \tag{2}$$

came from the 1884 treatise on calculus by Genocchi and Peano [23] (see Figure 2) and were given by Peano, see [42]. Function f from (2), while discontinuous, is not only separately continuous, but its restriction to any straight line is continuous. Functions with such property are often referred to as *linearly continuous*. Thus, example (2) shows that linearly continuous function need not be jointly continuous.



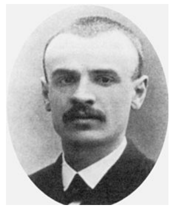



E. Heine J. Thomae A. Genocchi G. Peano

1.3 A Tale of Baire and Lebesgue: Structure of Separately Continuous Functions

As we pointed above, the class $SC(\mathbb{R}^n)$ of separate continuous functions on $\mathbb{R}^n$ does not coincide with the class $C(\mathbb{R}^n)$ of functions on $\mathbb{R}^n$ continuous in the usual sense. Although, the class $SC(\mathbb{R}^n)$ certainly is not as important as $C(\mathbb{R}^n)$, two great French analysts from the turn of 19th century, René-Louis Baire (1874–1932) and Henri Lebesgue (1875–1941), decided that the structure of separate continuous functions was worth studying.

Thus, independently Lebesgue, in his very first paper [33] of 1898, and Baire, in his Ph.D. thesis [1, Chapter 2] defended in 1899, proved that every separately continuous function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is a pointwise limit of a sequence of continuous functions or, in the contemporary terminology, it is of *Baire class 1*.



R. Baire

H. Lebesgue

Theorem 3. (Lebesgue 1898, Baire 1899) *A separately continuous function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is a pointwise limit of a sequence $f_k\colon \mathbb{R}^2 \to \mathbb{R}$ of continuous functions.*

To appreciate this result, we must have some knowledge on what kind of functions are of Baire class 1. Recall that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is of Baire class 1 if, and only if, $f^{-1}(U)$ is a G_δ-set (i.e., a countable intersection of open sets) for every open $U \subset \mathbb{R}$. It is also known (see Baire's thesis [1, pp. 66-67]), that every Baire class 1 function is continuous on a large set, a dense G_δ-set. However, a Baire class 1 can have a dense set of points of discontinuity, as shown by the function

$$f(x) = \begin{cases} \frac{1}{q} & \text{for } x = \pm\frac{p}{q}, \text{ where } p, q \in \mathbb{N} \text{ are relatively prime,} \\ 0 & \text{otherwise.} \end{cases}$$

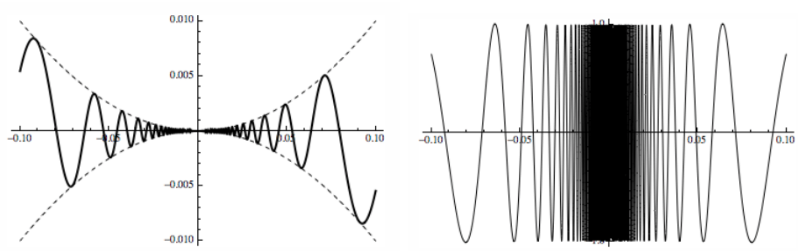


Figure 3: Graph of function h given by (3) and its derivative h'

A derivative of a differentiable function $h\colon \mathbb{R} \to \mathbb{R}$ is of Baire class 1, as h' is a pointwise limit of the continuous functions $h_n(x) = \frac{h(x+2^{-n})-h(x)}{2^{-n}}$. However, a derivative of a differentiable function need not be continuous. For example, the function

$$h(x)=\begin{cases} x^2\sin\left(x^{-1}\right) & \text{for } x\neq 0\\ 0 & \text{for } x=0\end{cases} \tag{3}$$

(see Figure 3) has a discontinuous derivative

$$h'(x)=\begin{cases} 2x\sin\left(x^{-1}\right)-\cos\left(x^{-1}\right) & \text{for } x\neq 0\\ 0 & \text{for } x=0.\end{cases}$$

It is perhaps worth to mention here the following old, but still open problem:

Problem 1. Find a non-trivial characterization of the derivatives, that is, the functions $h\colon \mathbb{R}\to\mathbb{R}$ such that $h=f'$ for some $f\colon\mathbb{R}\to\mathbb{R}$.

For more of this problem, see the 1947 paper [54] of Zygmunt Zahorski (1914-1998) or the monograph [7] of Andrew M. Bruckner (1932-). Note that (see e.g. [7]) the class of the derivatives contains all bounded *approximately continuous functions* $h\colon\mathbb{R}\to\mathbb{R}$; that is, bounded functions such that for every open $U\subset\mathbb{R}$, every point $x\in h^{-1}(U)$ is a Lebesgue density point of $h^{-1}(U)$ (i.e., $\lim_{h\to 0^+}\frac{\lambda(h^{-1}(U)\cap[x-h,x+h])}{2h}=1$, where λ is the Lebesgue measure). However, bounded derivative need not be approximatively continuous, as shown by the derivative of the function given by (3), see Figure 3. Moreover, there exist unbounded approximately continuous functions $h\colon\mathbb{R}\to\mathbb{R}$ which are not derivatives. (See [7].) We should also point out that, according to 1998 paper [22] of C. Freiling there is no "satisfactory" characterization of the derivatives, since "…the only way to characterize derivatives is by using some object or procedure which is at least as complicated as an integral."

Theorem 3 was generalized to functions of n variables, for any $n\geq 2$, in 1905 by Lebesgue [34, pp. 201-205]. To state it, we need to recall the definition of functions of Baire class α, for $\alpha<\omega_1$. The definition is recursive. Functions of Baire class 0 are simply the continuous functions. Functions of Baire class $\alpha+1$ are the pointwise limits of functions of Baire class α. Finally, if $\lambda<\omega_1$ is a limit ordinal (i.e., not of a successor form $\alpha+1$), then functions of Baire class λ consists of all functions of Baire class α for $\alpha<\lambda$. It is a standard fact that for every $\alpha<\omega_1$ there exists a function $f\colon\mathbb{R}\to\mathbb{R}$ of Baire class $\alpha+1$ which is not of Baire class α. For example, the characteristic function $\chi_{\mathbb{Q}}$ of the set $\mathbb{Q}$ of rational numbers is Baire class 2 but not Baire class 1. (It is of Baire class 2, since $\chi_{\mathbb{Q}}=\lim_{n\to\infty}\chi_{\{q_i\colon i\leq n\}}$, where $\{q_i\colon i\in\mathbb{N}\}=\mathbb{Q}$. It is not of Baire class 1 since it is everywhere discontinuous.)

The following theorem fully established the structure, in terms of Baire classes, of separately continuous functions $f\colon\mathbb{R}^n\to\mathbb{R}$.

Theorem 4. (Lebesgue 1905) *For every $n \geq 2$ the following holds.*

(i) A separately continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ is of Baire class $n-1$.

(ii) For every function $g\colon \mathbb{R} \to \mathbb{R}$ of Baire class $n-1$ there exists a separately continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $f(t,\ldots,t) = g(t)$ for every $t \in \mathbb{R}$. In particular, a separately continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ need not be of the Baire class $n-2$.

For example, by part (ii), there exists a separately continuous $f\colon \mathbb{R}^3 \to \mathbb{R}$ such that $f(t,t,t) = \chi_{\mathbb{Q}}(t)$. By part (i), it is of Baire class 2; however, it is not of Baire class 1, since then so would be its restriction to $\Delta_3 = \{\langle t,t,t\rangle \colon t \in \mathbb{R}\}$, which is not the case (as mentioned above).

On a proof of Theorem 4. A proof of (i) is relatively simple. It goes by noticing that a separately continuous $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ is a pointwise limit of maps $f_k\colon \mathbb{R}^{n+1} \to \mathbb{R}$, where each f_k is a linear interpolation of f restricted to $G_k = \mathbb{R}^n \times (2^{-k} \cdot \mathbb{Z})$ (i.e., $f_k \restriction G_k = f \restriction G_k$ and f_k is linear on each segment between any points $\langle x, m2^{-k}\rangle$ and $\langle x, (m+1)2^{-k}\rangle$, with $x \in \mathbb{R}^n$ and $m \in \mathbb{Z}$). Since, by inductive argument, $f \restriction G_k$ is of Baire class $n-1$, and the linear interpolation preserves this property, the result follows.

To see an idea behind the proof of (ii), let $g_k\colon \mathbb{R} \to \mathbb{R}$ be a sequence of Baire class $n-2$ functions converging pointwise to g. The proof is by induction on $n = 2, 3, \ldots$, with the case of $n = 2$ being quite different from the case of $n > 2$. Let $\Delta_n = \{\langle t,\ldots,t\rangle \in \mathbb{R}^n \colon t \in \mathbb{R}\}$.

For $n = 2$ the functions g_k are continuous. In this case, we define f as

$$f(x,y) = \begin{cases} g(x) & \text{for } \langle x,y\rangle \in \Delta_2 \\ \sum_{k=1}^{\infty} g_k(x)\varphi_k(x,y) & \text{for } \langle x,y\rangle \notin \Delta_2, \end{cases}$$

where the functions $\varphi_k\colon \mathbb{R}^2 \setminus \Delta_2 \to [0,1]$ form a particular partition of unity of $\mathbb{R}^2 \setminus \Delta_2$. That is, for every $\langle x,y\rangle \in \mathbb{R}^2 \setminus \Delta_2$ we have $\sum_{k=1}^{\infty} \varphi_k(x,y) = 1$ and there exists an open set U containing $\langle x,y\rangle$ such that $\varphi_k[U] = \{0\}$ for all but finitely many k. Notice, that this form insures the continuity of f on $\mathbb{R}^2 \setminus \Delta_2$, as well as the continuity of $f(x,\cdot)$ at x for every $x \in \mathbb{R}$. To insure the continuity of $f(\cdot,y)$ at y, the functions φ_k must be carefully chosen, using the fact that each g_k is uniformly continuous on $[-k,k]$.

Of course, for $n > 2$ a special choice of the partition of unity cannot rely on the continuity of functions g_k, since they need not be continuous. Thus, for $n > 2$ the above argument is modified as follows. Let functions $\varphi_k\colon \mathbb{R}^n \setminus \Delta_n \to [0,1]$ form an arbitrary partition of unity $\mathbb{R}^n \setminus \Delta_n$. By inductive assumption for every $k \in \mathbb{N}$ choose a separately continuous function $G_k : \mathbb{R}^{n-1} \to \mathbb{R}$ with $G_k(t, \ldots, t) = g_k(t)$ for every $t \in \mathbb{R}$. Define

$$f_n(x_1, \ldots, x_n) = \begin{cases} g(x_1) & \text{for } \langle x_1, \ldots, x_n \rangle \in \Delta_n \\ \sum_{k=1}^{\infty} G_k(x_1, \ldots, x_{n-1}) \varphi_k(x_1, \ldots, x_n) & \text{for } \langle x_1, \ldots, x_n \rangle \notin \Delta_n \end{cases}$$

and notice that f_n is continuous with any line parallel to the nth coordinate. It is also jointly continuous on $\mathbb{R}^n \setminus \Delta_n$ and we have $f_n(t, \ldots, t) = g(t)$ for every $t \in \mathbb{R}$. Repeating the construction, for every $i = 1, \ldots, n-1$, we can find a similar function f_i, which is continuous on any line parallel to the ith coordinate.

For every $i = 1, \ldots, n$, let D_i be the union of all lines in $\mathbb{R}^n$ parallel to the ith coordinate and intersect Δ_n. Then the sets $D_i \setminus \Delta_n$ are pairwise disjoint (since $n > 2$) and closed in $\mathbb{R}^n \setminus \Delta_n$. Since $\mathbb{R}^n \setminus \Delta_n$ is completely regular, for every i there exists a continuous $\psi_i : \mathbb{R}^n \setminus \Delta_n \to [0, 1]$ such that $\psi_i[D_i \setminus \Delta_n] = \{1\}$ and $\psi_i[D_j \setminus \Delta_n] = \{0\}$ for any $j \neq i$. Then, the function

$$f(x_1, \ldots, x_n) = \begin{cases} g(x_1) & \text{for } \langle x_1, \ldots, x_n \rangle \in \Delta_n \\ \sum_{i=1}^{n} \psi_i(x_1, \ldots, x_n) f_i(x_1, \ldots, x_n) & \text{for } \langle x_1, \ldots, x_n \rangle \notin \Delta_n \end{cases}$$

is as required.

Problem 2. Let $n > 2$. What is the smallest number m_n such that every linearly continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is of Baire class m_n?

Notice, that $m_n \geq 1$, since (as we will discuss below) there exist discontinuous linearly continuous functions $f : \mathbb{R}^n \to \mathbb{R}$. Also, by part (i) of Theorem 4, $m_n \leq n - 1$. This insures that $m_2 = 1$. So, the problem is for $n \geq 3$. It is worth to note that no easy modification of the proof of Theorem 4(ii) can show that $m_3 > 1$, since for linearly continuous functions $f : \mathbb{R}^3 \to \mathbb{R}$ there is no analogue of disjoint sets D_1, D_2, and D_3.

1.4 Other related results

Functions on $\mathbb{R}^{\mathbb{N}}$. It is worth mentioning that no analog of Lebesgue's result, Theorem 4, holds for separately continuous $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, as such function does not need to be of any Baire class α, for $\alpha < \omega_1$ (i.e., it does not need to be Borel-measurable). Indeed, this can be easily deduced from the following result of Edward Marczewski (1907-1976) and Czesław Ryll-Nardzewski (1926-), see [36].

Theorem 5. (Marczewski & Ryll-Nardzewski 1952) *Given an arbitrary function $g : \mathbb{R} \to \mathbb{R}$, there exists a separately continuous function $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ such that $f(x, x, \ldots) = g(x)$.*

The function f from Theorem 5 can be defined simply as

$$f(x_1, x_2, \ldots) = \begin{cases} g(a) & \text{if } x_n = a \text{ for all but finitely many } n \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

It is separately continuous, since f is constant on any line parallel to one of the axes.

The situation does not improve much, if one replaces $\mathbb{R}^{\mathbb{N}}$ with the space ℓ^2: a function $f\colon \ell^2 \to \mathbb{R}$ can be everywhere discontinuous, even if it is separately continuous in the strong sense (compareTheorem 2), as shown in 2004 by J. Činčura, T. Šalát, and T. Visnyai [15]. (See also [51].)

More on Baire class for functions on $\mathbb{R}^2$. A proof of the part (i) of Theorem 4 actually gives a stronger result for every Polish space X (e.g., for $X = \mathbb{R}^k$):

> For every $0 < \alpha < \omega_1$, a function $f\colon X \times \mathbb{R} \to \mathbb{R}$ is of Baire class $\alpha + 1$, provided every map $f(\cdot, y)$ is continuous and every map $f(x, \cdot)$ is of Baire class α.

For $0 < \alpha < \omega_1$, consider the following generalization of this statement (for $X = \mathbb{R}$):

Φ_α: If $f\colon \mathbb{R}^2 \to \mathbb{R}$ is such that $f(\cdot, y)$ is approximately continuous for every y and $f(x, \cdot)$ is of Baire class α for every x, then f is of Baire class $\alpha + 1$.

In 1995, M. Laczkovich and A. Miller [31] showed that Φ_1 is true. They also proved that, under the strong set theoretical assumption that there exists a real-valued measurable cardinal (which is a large cardinal), Φ_α is also true for any $2 \le \alpha < \omega_1$. However, there is no chance to prove Φ_2 in the standard set theory ZFC, since under the Continuum Hypothesis, CH, there are counter examples for Φ_2. This is noticed in [31], where the authors point out that the examples given independently in two different papers—a 1973 paper [17] by R. Davies and J. Dravecký and a 1974 paper [25] by Z. Grande—are in fact the counter examples for Φ_2.

Problem 3. Can consistency of Φ_2 be proved without the large cardinal assumption?

It should be mentioned that $\Phi_{\alpha+1}$ is a consequence of the statement:

Ψ_α: If $f\colon \mathbb{R} \times [0,1] \to \mathbb{R}$ is bounded, $f(x, \cdot)$ is Lebesgue measurable for every $x \in \mathbb{R}$, and $f(\cdot, y)$ is of Baire class α for every $y \in [0,1]$, then the function $x \mapsto \int_0^1 f(x,y)\,dy$ is of Baire class α.

The statement Ψ_1 is true, as proved in 1978 by J. Bourgain, D.H. Fremlin, and M. Talagrand [8]. Also, existence of a real-valued measurable cardinal implies that Ψ_α is true for any $\alpha < \omega_1$, see [31]. Now, a positive answer for Problem 3 would follow if the consistency of Ψ_2 could be proved without the large cardinal assumption.

2 Joint Continuity of $f(x, y)$ in Terms of Single Variable

In this section we discuss the question (Q), stated in the abstract, for the functions of two variables. More specifically, let $\mathcal{H}$ be a class of functions $h \colon \mathbb{R} \to \mathbb{R}$, where each h is identified with its graph: $h = \{\langle x, h(x)\rangle \colon x \in \mathbb{R}\}$. We say that a function $f \colon \mathbb{R}^2 \to \mathbb{R}$ is *$\mathcal{H}$-continuous* provided $f \restriction h$ is continuous for every $h \in \mathcal{H}$. It is *$\mathcal{H}^*$-continuous* provided for every $h \in \mathcal{H}$ both $f \restriction h$ and $f \restriction h^{-1}$ are continuous, where $h^{-1} = \{\langle h(x), x\rangle \colon x \in \mathbb{R}\}$. In other words, in $\mathcal{H}^*$-continuity, we examine the restrictions of f to the functions $h \in \mathcal{H}$ treated as functions from x to y and as functions from y to x. We examine for what classes $\mathcal{H}$

either $\mathcal{H}^*$-continuity or $\mathcal{H}$-continuity of $f(x, y)$ implies its joint continuity.

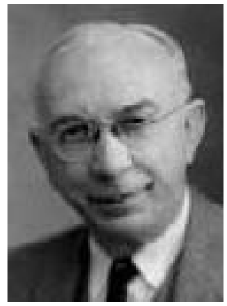
This is a version of a problem considered in 1905 by Henri Lebesgue [34, pp. 199–200].

Of course, the notion of separate continuity coincides with that of $\mathcal{H}^*$-continuity for $\mathcal{H}$ being the class CONST of constant functions, while linearly continuity is the $\mathcal{L}^*$-continuity, where $\mathcal{L}$ is the class of linear functions $h(x) = ax + b$. In particular, equation (2) shows that $\mathcal{L}^*$-continuity does not imply continuity and equation (1) shows that CONST*-continuity does not imply continuity.

2.1 $\mathcal{H}^*$-continuity: A Tale of Scheeffer, Lebesgue, Luzin, and A. Rosenthal

Scheeffer, Lebesgue, and Luzin: Let $\mathcal{A}$ be the class of real analytic functions, that is, representable by a Taylor series. The fact that $\mathcal{A}^*$-continuity at a point does not imply continuity at this point was noted in 1890 by Ludwig Scheeffer (1859-85) (see [45] or [44]) and in 1905 by Henri Lebesgue [34, pp. 199–200]. Lebesgue constructs such an example by induction. An easier example, due to Rosenthal, is given below by (4). Moreover, such an example can be also found in the 2005 paper [38] of Ollie Nanyes. Also, in 1948 text [35, pp. 173–176] of Nikolai Luzin (1883-1950) proves that $(C^0)^*$-continuity implies continuity, where $C^0 = C(\mathbb{R})$.






W. Sierpiński N. Luzin A. Rosenthal E. Marczewski

A. Rosenthal: In 1955 paper [44], Arthur Rosenthal (1887-1959) gave the ultimate solution of the $\mathcal{H}^*$ version of the problem in terms of the classes C^n of functions $h\colon \mathbb{R} \to \mathbb{R}$ having continuous n-th derivatives. Recall, that if D^n stands for the class of n-times differentiable functions, then we have $C^n \subsetneq D^n \subsetneq C^{n-1}$ for every $n \in \mathbb{N}$.[4]

Theorem 6. (A. Rosenthal 1955) *If $f\colon \mathbb{R}^2 \to \mathbb{R}$ is $(C^1)^*$-continuous, then it is continuous. However, there exist discontinuous $(D^2)^*$-continuous[5] functions $g\colon \mathbb{R}^2 \to \mathbb{R}$.*

The first part of the theorem follows easily from a fact [44] that:

(S) For every convergent sequence in $\mathbb{R}^2$ there exists a C^1 function h such that either h or h^{-1} contains infinitely many terms of that sequence. Moreover, h can be chosen to be either constant, or both strictly monotone and having derivative equal 0 at most in one point.[6]

For more details, see also an argument for Theorem 12.

The counterexample part is based on the function

$$g_0(x,y) = \begin{cases} 1 & \text{when } x \neq 0 \text{ and } y = x^{3/2}, \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

[4] The fact that the inclusions $D^1 \subset C^0$ and $C^1 \subset D^1$ are strict is justified, respectively, by $h(x) = |x|$ and h given by (3), see Figure 3. The examples for $n > 1$ are found by taking antiderivative of these examples $(n-1)$-times.

[5] Actually Rosenthal notices only that $(C^2)^*$-continuity does not imply continuity. But his example shows, in fact, the stronger statement.

[6] Actually, the non-constant function h constructed in [44] is either strictly convex or strictly concave. It must be modified, on one side of a limit of the sequence, to actually insure the property we claim here, as noted in [14].

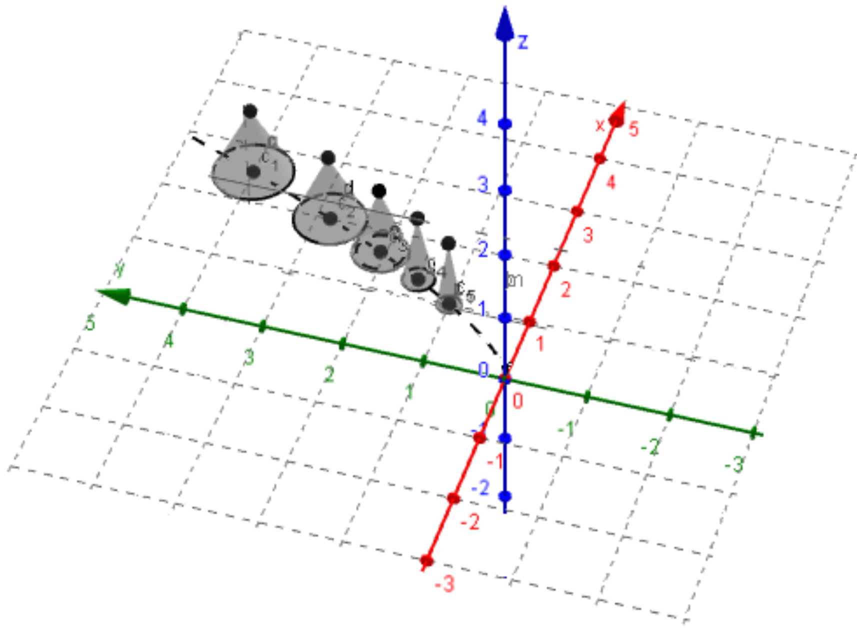


Figure 4: The graph of the function given by (5). The dashed curve is given by $\{\langle x, x^{3/2}, 0\rangle : x > 0\}$. The cones, each with base centered at c_n, are of height 1.

which shows that $(D^2)^*$-continuity at a point, the origin $\langle 0, 0\rangle$, does not imply continuity at this point. This can be transformed to the function g from Theorem 6 as follows. Choose a sequence $\langle c_n\rangle_{n=1}^{\infty}$ in $\{\langle x, x^{3/2}\rangle : x > 0\}$ converging to $\langle 0, 0\rangle$, the numbers $r_n > 0$, and let $g_n : \mathbb{R}^2 \to [0, 1]$ be a continuous "cone," with $g_n(c_n) = 1$ and $g(x) = 0$ whenever $\|x - c_n\| \geq r_n$. Then, for numbers r_n small enough the function (see Figure 4)

$$g = \sum_{n=1}^{\infty} g_n \tag{5}$$

is as desired.

Another proof of Theorem 6 can be found in the 2014 paper [39] of Ollie Nanyes.

Theorem 6 in terms of smooth curves. By a *curve* in the plane we understand any continuous map $h = \langle h_1, h_2 \rangle$ from an interval J into $\mathbb{R}^2$. A curve $h = \langle h_1, h_2 \rangle$ is *smooth*, if the coordinate functions h_1 and h_2 are continuously differentiable (i.e., are $\mathcal{C}^1$) and $\langle h_1'(t), h_2'(t) \rangle \neq \langle 0, 0 \rangle$ for every $t \in J$. It is *twice differentiable* ($\mathcal{C}^\infty$, respectively) provided it is smooth and $h_1, h_2 \in D^2$ ($h_1, h_2 \in \mathcal{C}^\infty$, respectively).

Since any $\mathcal{C}^1$ curve is locally equal to a $\mathcal{C}^1$ function either from x to y or from y to x, Theorem 6 can be expressed as follows.

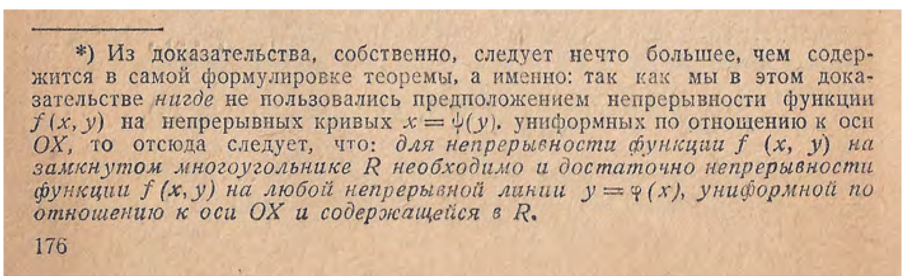
($*$) For any function $f: \mathbb{R}^2 \to \mathbb{R}$, if $f \circ h$ is continuous for every smooth curve h, then f is continuous. However, there exists a discontinuous function $g: \mathbb{R}^2 \to \mathbb{R}$ with $g \circ h$ being continuous for every smooth twice differentiable curve h.

In fact, it can easily be arranged that a function g from ($*$), given by (5), has a stronger property that $g \circ h$ is $\mathcal{C}^\infty$ for every smooth twice differentiable curve h—simply by taking functions g_n to be $\mathcal{C}^\infty$. In particular, at a first glance, such a g seems to contradict (for $n = 2$ and $k = 1$) the following 1967 theorem of Jan Boman [4] (see also [27]):

Theorem 7. (Boman 1967) *For every $n \geq 2$, $k \geq 1$, and $f: \mathbb{R}^n \to \mathbb{R}$, if $f \circ h$ is $\mathcal{C}^k$ for every $\mathcal{C}^\infty$ curve h, then $f \in \mathcal{C}^{k-1}$.*

However, there is no contradiction here, since in Theorem 7 the testing functions h include also non-smooth curves.

The proof of Theorem 6 as well as a weaker version of Theorem 7 can be also found in the 2011 paper [37] of Michael McAsey and Libin Mou.



*) Из доказательства, собственно, следует нечто большее, чем содержится в самой формулировке теоремы, а именно: так как мы в этом доказательстве *нигде* не пользовались предположением непрерывности функции $f(x, y)$ на непрерывных кривых $x = \psi(y)$, униформных по отношению к оси OX, то отсюда следует, что: *для непрерывности функции $f(x, y)$ на замкнутом многоугольнике R необходимо и достаточно непрерывности функции $f(x, y)$ на любой непрерывной линии $y = \varphi(x)$, униформной по отношению к оси OX и содержащейся в R.*

176

Figure 5: Theorem 8 in Luzin's 1948 text

2.2 $\mathcal{H}$-continuity: From N. Luzin's to contemporary results

N.N. Luzin: The first contribution to the $\mathcal{H}$ version of the problem appeared in 1948 text [35, p. 176] Luzin, as a footnote to his proof that $(C^0)^*$-continuity implies continuity, see Figure 5. It can be stated as follows.

Theorem 8. *For every function $f\colon \mathbb{R}^2 \to \mathbb{R}$, if $f(x, h(x))$ is continuous for every continuous $h\colon \mathbb{R} \to \mathbb{R}$, then f is continuous. In other words, C^0-continuity of f implies its continuity.*

The two statements in the theorem are equivalent, since: (1) the continuity of $f_h(x) = f(x, h(x))$ implies the continuity of $f \restriction h$ (as $f \restriction h = f_h \circ \pi$, where $\pi(x, y) = x$) and (2) the continuity of h and $f \restriction h$ implies the continuity of f_h (as $f_h = (f \restriction h) \circ \langle id, h\rangle$, where $id(x) = x$).

The theorem follows easily from (S) and the fact that if there exists a sequence $\langle x_n, y_n\rangle_{n=1}^{\infty}$ contradicting continuity of $f\colon \mathbb{R}^2 \to \mathbb{R}$, then there exists such a sequence with $\langle y_n\rangle_{n=1}^{\infty}$ being strictly monotone. See also the proof of Theorem 12.

K.C. Ciesielski & J. Rosenblatt: Let "C^1" be the family of all functions in C^0 having continuous derivatives, allowing infinite values. (For example, function $h(x) = \sqrt[3]{x}$ is "C^1" but not D^1.) Then, the property (S) and Theorem 6 imply immediately the first part of the following result, proved by the first author and J. Rosenblatt [14], which generalizes Theorem 8.

Theorem 9. (Ciesielski & Rosenblatt 2014) *If a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is "C^1"-continuous, then it is continuous. However, there exists a discontinuous D^1-continuous function $g\colon \mathbb{R}^2 \to \mathbb{R}$.*

A function g from Theorem 9 can be defined via (5), where disc centers c_n belong to a fixed vertical line.

Theorem 9 shows that there exists a reasonably small family $\mathcal{H}$ of test functions, namely $\mathcal{H} =$"C^1" or $\mathcal{H} = C^0$, for which $\mathcal{H}$-continuity implies continuity. How far can the family $\mathcal{H}$ be decreased while still preserving this implication? Can $\mathcal{H}$ consist of just a single test function?

Certainly a family $\mathcal{H}$ consisting of a single function $h\colon \mathbb{R} \to \mathbb{R}$ cannot have such property, since the testing functions must cover $\mathbb{R}^2$ for the implication to have a chance to hold true. Nevertheless, it seems that the $\mathcal{H}$-continuity may imply continuity if $\mathcal{H}$ is built just from a single template function h. That is, when $\mathcal{H}$ equals to the family $T(h)$ of all translates of h.

To convince the reader that such a thing might indeed be possible, notice that $T(|x|)$-continuity is, essentially, the separate continuity. More precisely, if r is a 45° rotation of $\mathbb{R}^2$ (e.g. with respect to the origin), then, as can be easily seen,

> *$f\colon \mathbb{R}^2 \to \mathbb{R}$ is $T(|x|)$-continuous if, and only if, $f \circ r$ is separately continuous.*

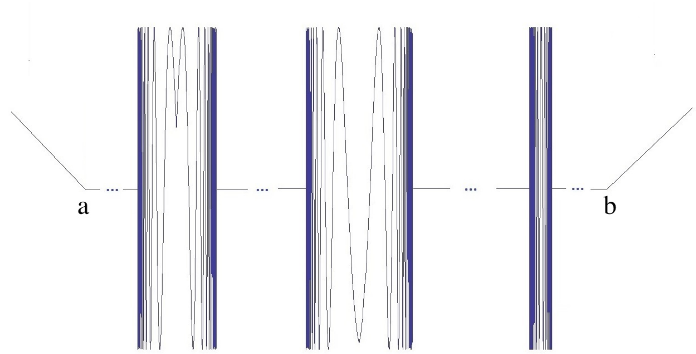


Figure 6: Graph of the function h given by (6)

Unfortunately, the next theorem, proved by the first author and J. Rosenblatt [14], shows that a single continuous template function cannot insure continuity of a functions of two variables. However, there exists a template function h that does the job, which is a pointwise limit of continuous functions.

Theorem 10. (Ciesielski & Rosenblatt 2014) *There exists a Baire class 1 function $h\colon \mathbb{R} \to \mathbb{R}$ such that $T(h)$-continuity of $f\colon \mathbb{R}^2 \to \mathbb{R}$ implies its continuity. However, for every continuous $h\colon \mathbb{R} \to \mathbb{R}$ there exists a discontinuous $T(h)$-continuous function $g\colon \mathbb{R}^2 \to \mathbb{R}$.*

The idea of constructing discontinuous $T(h)$-continuous function $g\colon \mathbb{R}^2 \to \mathbb{R}$ for a continuous h is relatively simple: g is of the form of a function given by (5), where disc centers c_n are chosen so that any $h_0 \in C^0$ containing infinitely many of them has worst uniform continuity parameter δ than $h \restriction [-n, n]$ for $\varepsilon = 4^{-n}$. Nevertheless, the detailed definition of c_n from h is a bit technical.

By the second part of the theorem, the function h from the first part cannot be continuous. However, it can still be of relative simple form, a variation of a $\sin(1/x)$ function, and can have a small set of points of discontinuity: nowhere dense and of measure 0. More specifically it can be defined via the formula

$$h(x) = \begin{cases} 0 & \text{when } x \in P, \\ \sin\left(\frac{1}{\operatorname{dist}(x,P)}\right) & \text{when } x \in [a,b] \setminus P, \\ \operatorname{dist}(x,P) & \text{when } x \in \mathbb{R} \setminus [a,b], \end{cases} \tag{6}$$

where $P \subset [a, b]$ is a "fat" closed nowhere dense subset of $\mathbb{R}$ with $a, b \in P$. (See Figure 6.) Here, "fat" means that for every converging sequence in $\mathbb{R}$, some translation of P contains infinitely many terms of the sequence. In particular, any set P of positive measure is "fat," as is the following set (see [14]):

$$P = \left\{ \sum_{k=2}^{\infty} \frac{a_k}{k!} \colon a_k \in \{0, \ldots, k-1\} \text{ for every } k \right\}.$$

Notice, that the function h from the first part of Theorem 10 (e.g., given by (6)) nicely relates to both parts of Theorem 8, since the property *"$f(x, h_0(x))$ is continuous for every $h_0 \in T(h)$"* implies *"$f \restriction h_0$ is continuous for every $h_0 \in T(h)$,"* which, in turn, implies the continuity of f.

Other related results. Another result, on insuring the continuity of $f(x, y)$ in terms of $f(x, \cdot)$ and $f(\cdot, y)$, is the following 1910 theorem of W.H. Young [52]. (See also [40].)

Theorem 11. (W.H. Young 1910) *Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be separately continuous. If $f(\cdot, y)$ is also monotone for every y, then f is continuous.*

This theorem was generalized to functions of n variables in the 1969 article [30] by R.L. Kruse and J.J. Deely, who apparently were not aware of W.H. Young earlier result from [52].

3 Continuity of $f(x_1, x_2, x_3, \ldots, x_n)$ in Terms of Single Variable

The results presented in this section have not been previously published.

For $n > 2$ let $\mathcal{H}_n$ stand for the class of functions from $\mathbb{R}$ to $\mathbb{R}^{n-1}$, identified with their graphs, where the variable goes to the first coordinate of $\mathbb{R}^n$ and the values to the remaining coordinates. Also, let $\mathcal{H}_n^*$ stand for the graphs of all functions in $\mathcal{H}_n$, where the coordinates order is randomly permuted. In particular, C_n^1 is the family of all maps from $\mathbb{R}$ to $\mathbb{R}^{n-1}$ (i.e., sets $\{\langle x, g_1(x), \ldots, g_{n-1}(x)\rangle \colon x \in \mathbb{R}\}$) for which each of the $(n-1)$-coordinate functions (i.e., g_is for $i = 1, \ldots, n-1$) is C^1.

Theorem 12. (Generalization of Theorems 6 and 9) *For every $n \geq 2$ a function $f \colon \mathbb{R}^n \to \mathbb{R}$ is continuous, provided it is either $(C_n^1)^*$-continuous or "C_n^1"-continuous.*

Proof. To see this, let f be discontinuous. It must be shown that f is neither $(C_n^1)^*$-continuous nor "C_n^1"-continuous. Clearly, there exists a sequence $\bar{s} = \langle s^k \rangle_{k=1}^{\infty}$ in $\mathbb{R}^n$ convergent to an $s^0 \in \mathbb{R}^n$ for which the limit $\lim_k f(s^k)$ (finite or infinite) exists, but is not equal $f(s^0)$. Let $s^k = \langle s_1^k, \ldots, s_n^k \rangle$ and notice that we can assume that

(•) the sequence $\langle s_i^k\rangle_k$ is strictly monotone for every $1\le i\le n$.

Indeed, if $f\restriction \ell$ is discontinuous on some line ℓ parallel to the diagonal $\Delta=\{\langle x,\ldots,x\rangle\colon x\in\mathbb{R}\}$, then we can choose $\bar{s}$ on such a line. So, assume that $f\restriction\ell$ is continuous for every such line ℓ. Choosing a subsequence, if necessary, we can ensure that for every $1\le i\le n$ the sequence $\langle s_i^k\rangle_k$ is either strictly monotone or constant. (This can be done either by induction on $i=1,\ldots,n$, or by using Ramsey's partition theorem.) For every k let ℓ_k be the line parallel to Δ containing s^k. Then, by induction on k and using continuity of $f\restriction\ell_k$, we can replace each s^k with an $\hat{s}^k\in\ell$ so that $|f(\hat{s}^k)-f(s^k)|\le 2^{-k}$, $\|\hat{s}^k-s^k\|\le 2^{-k}$, and each sequence $\langle \hat{s}_i^k\rangle_k$ is strictly monotone, insuring (•).

Next, by induction and choosing a subsequence of $\bar{s}$, if necessary, we can ensure that for every $1\le i<j\le n$, there is a "C^1" function $h_{i,j}$ as in (S) containing all points $\langle s_i^k,s_j^k\rangle$. Then $h_1=\{\langle x_1,h_{1,2}(x_1),\ldots,h_{1,n}(x_1)\colon x_1\in\mathbb{R}\rangle\}$ is "C_n^1" and contains $\bar{s}$, so f is not "C_n^1"-continuous.

To see that f is not $(C_n^1)^*$-continuous, for every $1\le i<j\le n$ put $h_{j,i}=h_{i,j}^{-1}\in$ "C_n^1". Also, for every i, let $h_{i,i}$ be the identity function and define $h_i=\{\langle h_{i,1}(x_i),\ldots,h_{i,n}(x_i)\rangle\colon x_i\in\mathbb{R}\}$. Then every h_i is ("C_n^1")* and it contains all terms of the sequence $\langle s^k\rangle_k$. Thus, it is enough to show that $h_i\in(C_n^1)^*$ for some i. To see this, choose an i for which the set $S_i=\{j\colon h_{i,j}(s_i^0)\notin\mathbb{R}\}$ has the smallest size. The result follows, if we show that S_i is empty.

So, by way of contradiction, assume that there is a j with $h_{i,j}(s_i^0)\notin\mathbb{R}$. Then, S_j has the smaller size than S_i, since $h_{j,i}(s_j^0)=0\in\mathbb{R}$, while for every l with $h'_{i,l}(s_i^0)\in\mathbb{R}$,

$$\begin{aligned}
h'_{j,l}(s_j^0) &= \lim_{k\to\infty}\frac{h_{j,l}(s_j^k)-h_{j,l}(s_j^0)}{s_j^k-s_j^0}\\
&= \lim_{k\to\infty}\frac{h_{i,l}(h_{j,i}(s_j^k))-h_{i,l}(h_{j,i}(s_j^0))}{h_{j,i}(s_j^k)-h_{j,l}(s_j^0)}\lim_{k\to\infty}\frac{h_{j,i}(s_j^k)-h_{j,l}(s_j^0)}{s_j^k-s_j^0}\\
&= \lim_{k\to\infty}\frac{h_{i,l}(s_i^k)-h_{i,l}(s_i^0)}{s_i^k-s_i^0}\lim_{k\to\infty}\frac{h_{j,i}(s_j^k)-h_{j,l}(s_j^0)}{s_j^k-s_j^0}\\
&= h'_{i,l}(s_i^0)\ h'_{j,i}(s_j^0)=h'_{i,l}(s_i^0)\ 0\in\mathbb{R}.
\end{aligned}$$

In other words, $S_j\subset S_i\setminus\{j\}\subsetneq S_i$, contradicting the minimality if S_i. □

Notice that the negative parts of Theorems 6 and 9 give us also the bounds for Theorem 12.

Theorem 13. (Generalization of Theorem 10) *For every $n \geq 2$ there exists a Baire class 1 function $h\colon \mathbb{R} \to \mathbb{R}^{n-1}$ such that $T(h)$-continuity of $f\colon \mathbb{R}^n \to \mathbb{R}$ implies its continuity.*

Proof. Let h_0 be given by (6) and p be a continuous (Peano-like space filling) function from $[-1,1]$ onto $[-1,1]^{n-1}$ such that $p(0) = \langle 0,\ldots,0\rangle$. Define function h on $[a,b]$ as $p \circ h_0 \restriction [a,b]$ and extend it to $\mathbb{R}$ in such a way that $h \restriction \mathbb{R} \setminus (a,b)$ is continuous and its graph contains n mutually perpendicular line segments. Then such an h is as required. The argument is a variation of that for [14, corollary 14] and it goes as follows:

Clearly, h is of Baire class 1. To see the main implication, choose a $T(h)$-continuous $f\colon \mathbb{R}^n \to \mathbb{R}$. We need to show that f is continuous.

The perpendicular segments in h insure that $f \circ r$ is separately continuous for some isometry r of $\mathbb{R}^n$. In particular, the set D of points of continuity of f is dense in $\mathbb{R}^n$. (See e.g. [12]. Actually, this is closely related to a 1932 result of of Wacław Sierpiński (1882-1969) from [47] that *A separately continuous function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is uniquely determined by its values on the dense subset of its domain.* See also [41].) It is enough to show (see e.g. [14, fact 3]) that $f(s^0) = \lim_k f(s^k)$ for every sequence $\bar{s} = \langle s^k \rangle_{k=1}^{\infty}$ in D convergent to an $s^0 \in \mathbb{R}^n$ for which the limit $\lim_k f(s^k)$ (finite or infinite) exists. We can also assume that $\bar{s}$ has diameter less than 1. Let $s^k = \langle s^k_1,\ldots,s^k_n\rangle$ for $k = 0,1,2,\ldots$.

Since P is "fat," there exists a subsequence $\langle s^{k_i}\rangle_{i=1}^{\infty}$ such that $\langle s_1^{k_i}\rangle_{i=1}^{\infty}$ is contained in a translate $t_1 + P$ of P. So, $s^0_1 \in t_1 + P$ and for $t = \langle t_1, s^0_2,\ldots,s^0_n\rangle$ we have $s^0 \in t + h \restriction P$, since $s_0 - t = \langle s^0_1 - t_1, 0,\ldots,0\rangle \in P \times \{0\}^{n-1} \subset h \restriction P$. Notice that, for every i, $s^{k_i} \in (t_1+P) \times [-1,1]^{n-1}$, as $s_1^{k_i} \in t_1 + P$ and $\bar{s}$ has the diameter less than 1. Since each s^{k_i} is a point of continuity of f and $(t_1+P)\times[-1,1]^{n-1}$ is contained in the closure of $t+h$, we can choose $\hat{s}^{k_i} \in t+h$ such that $|f(\hat{s}^{k_i}) - f(s^{k_i})| \leq 2^{-i}$ and $\|\hat{s}^{k_i} - s^{k_i}\| \leq 2^{-i}$. Hence, as $f \restriction t+h$ is continuous, $\lim_k f(s^k) = \lim_i f(s^{k_i}) = \lim_i f(\hat{s}^{k_i}) = f(\lim_i \hat{s}^{k_i}) = f(s^0)$, as required. □

4 Genocci-Peano type of examples for $n > 2$

When generalizing function $g(x,y)$ from (2) to higher dimensions, we need to decide whether to treat lines in $\mathbb{R}^2$ as the objects of dimension 1, or rather as hyperplanes, that is, objects of co-dimension 1. The first of these options leads the notion of *linearly continuous functions*. That is, functions $f\colon \mathbb{R}^n \to \mathbb{R}$ with continuous restrictions to any straight line. The second, to the notion of *hyperplane continuous functions*. That is, functions $f\colon \mathbb{R}^n \to \mathbb{R}$ with continuous restrictions to any (co-dimension one) hyperplane in $\mathbb{R}^n$. The class of linearly continuous functions on $\mathbb{R}^n$ has been studied for several decades, as we shall discuss in the next section. On the other hand, the class of hyperplane continuous functions on $\mathbb{R}^n$, for $n>2$, seem to first appear in the literature only in 2014, in the paper [12].

A simple example of discontinuous linearly continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ for any $n \geq 2$ is given by the following modification of (2):

$$h(x,y,z,\ldots) = g(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{for } \langle x,y\rangle \neq \langle 0,0\rangle \\ 0 & \text{for } \langle x,y\rangle = \langle 0,0\rangle. \end{cases} \tag{7}$$

Indeed, clearly $h = g\circ\pi_{xy}$ is discontinuous, where π_{xy} is a projection onto xy-plane. Also, for any straight line ℓ in $\mathbb{R}^n$, the map $h \restriction \ell = (g\restriction \pi_{xy}[\ell])\circ(\pi_{xy}\restriction \ell)$ is continuous, as so is $g\restriction \pi_{xy}[\ell]$, $\pi_{xy}[\ell]$ being either a line or a point. However, for $n>2$, the map h is discontinuous on the hyperplane $\{\langle x,y,z,\ldots\rangle \in \mathbb{R}^n\colon y=z\}$, leading to a question: *Do there exist simple, Genocci-Peano like, examples of discontinuous, hyperplane continuous functions on $\mathbb{R}^n$ for any $n>2$?* The answer to this question, which comes from the paper [13] by the current authors, is given in terms of the rational functions $g\colon \mathbb{R}^n \to \mathbb{R}$ of the form

$$g(x_1,x_2,\ldots,x_n) = \begin{cases} \frac{x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}}{x_1^{\beta_1}+x_2^{\beta_2}+\cdots+x_n^{\beta_n}} & \text{when } \langle x_1,x_2,\ldots,x_n\rangle \neq \langle 0,0,\ldots,0\rangle, \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $\alpha_i, \beta_i \in \mathbb{N}$ for all $i \in \{1,2,\ldots,n\}$.

Theorem 14. (Ciesielski & MIller 2014) *Let g be given by a formula (8) and let $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$ be even. Then*

(i) g is discontinuous if, and only if, $\sum_{i=1}^n \frac{\alpha_i}{\beta_i} \leq 1$.

(ii) g has a continuous restriction to every hyperplane in $\mathbb{R}^n$ if, and only if, $\sum_{i=1}^n \frac{\alpha_i}{\beta_i} - \frac{\alpha_k}{\beta_k} + \frac{\alpha_k}{\beta_{k-1}} > 1$ for every $k \in \{2,3,\ldots,n\}$.

For example, the following function

$$g(x_1,x_2,x_3,x_4) = \begin{cases} \frac{x_1x_2x_3^2x_4^3}{x_1^4+x_2^6+x_3^8+x_4^{10}} & \text{when } \langle x_1,x_2,x_3,x_4\rangle \neq \langle 0,0,0,0\rangle, \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

is discontinuous and hyperplane continuous, as it satisfies Theorem 14. Indeed, condition (i) holds, since $\frac{1}{4}+\frac{1}{6}+\frac{2}{8}+\frac{3}{10}=\frac{29}{30}<1$, while (ii) is justified by $\frac{1}{4}+\frac{1}{6}+\frac{2}{8}+\frac{3}{10}-\frac{1}{6}+\frac{1}{4}=\frac{21}{20}>1$, $\frac{1}{4}+\frac{1}{6}+\frac{2}{8}+\frac{3}{10}-\frac{2}{8}+\frac{2}{6}=\frac{21}{20}>1$, and $\frac{1}{4}+\frac{1}{6}+\frac{2}{8}+\frac{3}{10}-\frac{3}{10}+\frac{3}{8}=\frac{25}{24}>1$.

Also, Theorem 14 immediately implies the following corollary, which provides a "canonical" example for discontinuous and hyperplane continuous functions for any $n \geq 2$.

Corollary 15. (Ciesielski & Miller 2014) *For every $n > 1$, if $g_n \colon \mathbb{R}^n \to \mathbb{R}$ is defined as*

$$g_n(x_1, x_2, \ldots, x_n) = \begin{cases} \frac{x_1 x_2 \cdots x_{n-1} x_n^2}{x_1^2 + x_2^4 + \cdots + x_{n-1}^{2^{n-1}} + x_n^{2^n}} & \text{for } \langle x_1, x_2, \ldots, x_n \rangle \neq \langle 0, 0, \ldots, 0 \rangle, \\ 0 & \text{otherwise,} \end{cases}$$

then g_n is discontinuous but hyperplane continuous.

In particular, $g_2(x, y) = \frac{xy^2}{x^2+y^4}$ is the original Genocci-Peano example from (2), $g_3(x, y, z) = \frac{xyz^2}{x^2+y^4+z^8}$, and $g_4(x, y, z, t) = \frac{xyzt^2}{x^2+y^4+z^8+t^{16}}$.

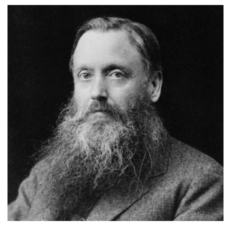
The direct proof that g_3 is discontinuous and hyperplane continuous, constituting a typical case of the general proof of Theorem 14 presented in [13], goes as follows. Function g_3 is discontinuous along the path $x = t^4$, $y = t^2$, and $z = t$, where $t \in \mathbb{R}$, since $g_3(t^4, t^2, t) = \frac{1}{3}$ for any $t \neq 0$. To see that g_3 is hyperplane continuous, we first note that $|g_3(x, y, z)| = \frac{|x|}{d^{\frac{1}{2}}} \frac{|y|}{d^{\frac{1}{4}}} \left(\frac{|z|}{d^{\frac{1}{8}}} \right)^2$, where $d = x^2 + y^4 + z^8$. Notice that each of the three quotients is bounded above by 1. Moreover, for $z = ax + by$ we have $\frac{|z|}{d^{\frac{1}{8}}} \leq |a| \frac{|x|}{d^{\frac{1}{8}}} + |b| \frac{|y|}{d^{\frac{1}{8}}} \to 0$ as $d \to 0$. So, g_3 is continuous at $\langle 0, 0, 0 \rangle$ on this hyperplane. Similarly, for $y = ax$ we have $\frac{|y|}{d^{\frac{1}{4}}} \to 0$ as $d \to 0$.

5 Sets of discontinuity points of separately, linearly, and hyperplane continuous functions

For a function $f \colon X \to \mathbb{R}$, let $D(f) \subset X$ be the set of points of discontinuity of f. Recall, that $D(f)$ is always an F_σ-set. That is, a countable union of sets closed in X. More specifically, $D(f) = \bigcup_{k \in \mathbb{N}} F_k$, where F_k is a the set of all $x \in X$ with oscillation greater than $1/k$. That is, such that for every open set U containing x there exist $y, z \in U$ with $|f(y) - f(z)| > 1/k$. It is easy to see that each set F_k is closed.

The goal of this section is to describe the study of the size and, more generally, the structure of the sets $D(f)$, where $f \colon \mathbb{R}^n \to \mathbb{R}$ varies either over the entire class $SC(\mathbb{R}^n)$ of separate continuous functions, or over some of its subclass, like the class of linearly or hyperplane continuous functions. Of course, by the remark above, any $D(f)$ must be an F_σ-set.






H. Hahn R.B. Kershner W.H. Young G.C. Young

5.1 Separate continuity

It appears that the first step toward understanding these sets came from Baire who showed, in his 1899 thesis [1], that *for any separately continuous function* $f\colon \mathbb{R}^2 \to \mathbb{R}$*, there exist the sets* $A, B \subset \mathbb{R}$ *of first Baire category (i.e., countable unions of nowhere dense sets) such that* $D(f) \subset A \times B$ and that *there is a separately continuous function* $f\colon \mathbb{R}^3 \to \mathbb{R}$ *such that* $D(f)$ *contains a line segment*. This last result was expanded, in 1919 [26], by Hans Hahn (1879-1934) who observed that the function $h(x, y, z, \ldots) = \frac{xy}{x^2+y^2}$ is discontinuous on the hyperplane $\{\langle x, y, z, \ldots\rangle \in \mathbb{R}^n \colon y = z\}$. The problem of characterizing the sets $D(f)$, for $f \in SC(\mathbb{R}^n)$, was settled in 1943 [28] by Richard B. Kershner[7] (1913-1982), by proving the following theorem. (Also, in 1949 G. Tolstoff [50] constructed a separately continuous function $f\colon \mathbb{R}^2 \to \mathbb{R}$ with $D(f)$ of positive Lebesgue measure. However, the existence of such example follows immediately from Theorem 16.)

Theorem 16. (Kershner 1943) *For any set* $D \subset \mathbb{R}^n$*,* $D = D(f)$ *for some separately continuous function* $f\colon \mathbb{R}^n \to \mathbb{R}$ *if, and only if,* D *is an* F_σ*-set and every orthogonal projection of* D *onto a coordinate hyperplane is of first category.*

The proof of a necessity of the condition from Theorem 16 is based on one of the early results on separately continuous functions, which, according to Baire [1], was first noticed by Vito Volterra (1860-1940). It said that every separately continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *quasi-continuous*. That is, has the property that for every $x \in \mathbb{R}^n$ and every pair of open sets $U \subset \mathbb{R}^n$ and

[7]Incidentally, Kershner seems to be better known for his contribution to the satellite navigation system than for his mathematical contributions.

Vito Volterra

$W \subset \mathbb{R}$, with $x \in U$ and $f(x) \in W$, there is a nonempty open subset $V \subset U$ such that $f[V] \subset W$. Note that this definition differs from the standard definition of continuity only in the fact that the point x need not be in V.

The other tool needed for the proof is the following version of a 1976 result [2, corollary 3.8] of J.C. Breckenridge and Togo Nishiura. (Compare also [12, lemma 4.1].)

Lemma 17. *If $Z = \mathbb{R}^k$ and $f\colon Z \times \mathbb{R}^m \to \mathbb{R}$ is such that $f(\cdot, y)$ is quasi-continuous for every $y \in \mathbb{R}^m$ and the map $\mathbb{R}^k \ni x \mapsto f(z, x) \in \mathbb{R}$ is separately continuous for every $z \in Z$, then the projection of $\pi_{\mathbb{R}^m}[D(f)]$ of $D(f)$ along Z is of first category in $\mathbb{R}^m$.*

Then, a necessity of the condition from Theorem 16 is obtained by applying the lemma to f treated as a function from $Z \times \mathbb{R}^1$, where Z is a fixed coordinate hyperplane in $\mathbb{R}^n$. (The map $f(\cdot, y)$ is quasi-continuous by the result of Volterra.)

To see that the condition from Theorem 16 is sufficient, notice that $D = \bigcup_{k=1}^{\infty} D_k$, where each D_k is compact. Fix a $k \in \mathbb{N}$ and notice that the orthogonal projection of D_k onto each coordinate hyperplane is nowhere dense, as it is compact and of first category. By easy induction, construct a family $\{B_i : i \in \mathbb{N}\}$ of open balls, each B_i centered at c_i, such that:

- for every distinct $i, j \in \mathbb{N}$, the projections, onto each coordinate hyperplane, of the sets B_i, B_j, and D_k are pairwise disjoint;
- D_k is the set of accumulation points of $\{c_i : i \in \mathbb{N}\}$.

For each i, let $g_i\colon \mathbb{R}^n \to [0, 4^{-i}]$ be continuous such that $g_i(c_i) = 1$ and $g_i(x) = 0$ for every $x \in \mathbb{R}^n \setminus B_i$. Define, similarly as in (5), the function $f_k = \sum_{i=1}^{\infty} g_i$. Notice that $f_k\colon \mathbb{R}^n \to [0, 4^{-i}]$ is separately continuous and

$D(f_k) = D_k$. Then, the function $f = \sum_{k=1}^{\infty} f_k$ is separately continuous (as uniform limit of such functions) and $D(f) = D$.

More on $D(f)$ for separately continuous f Recall also the following theorem of Karl Bögel [3]. (See also [27].)

Theorem 18. (Bögel 1926) *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be separately continuous. If $f(x, \cdot)$ is differentiable for every $y \in \mathbb{R}$, then $D(f)$ is nowhere dense.*

Note that a set of the first category, that is guaranteed without the assumption that functions $f(x, \cdot)$ are differentiable, may be strictly larger than a nowhere dense set.

5.2 Higher-dimensional versions of separate continuity

Let $0 < k < n$. Recall that a k-flat is a subset of $\mathbb{R}^n$ isometric to $\mathbb{R}^k$. We use a term *right k-flat* for any k-flat parallel to a vector subspace of $\mathbb{R}^n$ spanned by k-many coordinate vectors. The family of all k-flats is denoted as $\mathcal{F}_k$, while $\mathcal{F}_k^+$ will stand for the family of all right k-flats. In this notation, the class of $\mathcal{F}_1^+$-continuous ($\mathcal{F}_1$-continuous) $f : \mathbb{R}^n \to \mathbb{R}$ is identical with the class of separately (linearly, respectively) continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$. Also, the class of $\mathcal{F}_{n-1}$-continuous functions coincides with the class of hyperplane continuous functions f discussed in Section 4.

The natural higher-dimensional generalizations of separate continuity are the classes of $\mathcal{F}_k^+$-continuous functions $f : \mathbb{R}^n \to \mathbb{R}$. It is perhaps instructive to think of $\mathcal{F}_k^+$-continuous functions as those which are continuous when looked at in any k variables separately. The $\mathcal{F}_k^+$-continuous functions on $\mathbb{R}^n$ are fairly well documented in the literature. They have been studied in connection with the theory of Sobolev spaces, see e.g. [2].

Clearly, Theorem 16 gives a full characterization of the sets $D(f)$ for $\mathcal{F}_1^+$-continuous functions $f : \mathbb{R}^n \to \mathbb{R}$. The case of $\mathcal{D}_{n-1,n}^+$-continuous functions was settled in [2], by the result:

> $D = D(f)$ for some $\mathcal{D}_{n-1,n}^+$-continuous function $f : \mathbb{R}^n \to \mathbb{R}$ if, and only if, D is an F_σ-set and $D \subset D_1 \times D_2 \times \ldots \times D_n$, where each D_i is a first category subset of $\mathbb{R}$.

These results were generalized to the general case of $\mathcal{D}_{k,n}^+$-continuous functions by the first author and T. Glatzer [12, theorem 2.1] as follows.

Theorem 19. (Ciesielski & Glatzer 2014) *For every $0 < k < n$, $D = D(f)$ for some $\mathcal{D}_{k,n}^+$-continuous function $f : \mathbb{R}^n \to \mathbb{R}$ if, and only if, D is an F_σ-set whose orthogonal projection on any right $(n-k)$-flat is of first category.*

5.3 Linear continuity

Studies on the structure of $D(f)$ for linearly continuous functions $f\colon \mathbb{R}^n \to \mathbb{R}$ exist, but they are not as penetrating as those for separately continuous functions. One of the earliest such results can be found in a 1910 paper [53] of the husband and wife team William Henry Young (1863-1942) and Grace Chisholm Young (1868-1944)—the first woman that received a doctorate in any field in Germany (degree granted in 1895)—who gave an example of a linearly continuous function $f\colon [0,1]\times[0,1] \to \mathbb{R}$ for which $D(f)$ is uncountable in every non-empty open subset of its domain.

Linearly continuous functions were studied little after Young and Young's paper, until 1944, when Alexander S. Kronrod (1921–1986) took a course on the theory of real functions from N. Luzin (by then a rare opportunity). The course influenced him to begin a research program toward the development of a geometric theory of real functions (see [32]). This research program led to Kronrod asking for a description of the set of discontinuities of linearly continuous functions defined on the plane. The first partial answer to Kronrod's challenge came in 1976 from Semen G. Slobodnik [48], who gave the following necessary condition for a set D to be a discontinuity set of some linearly continuous function.

Theorem 20. (Slobodnik 1976) *If $D \subset \mathbb{R}^n$ is the set of discontinuity points of some linearly continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$, then D admits a representation $D = \bigcup_{i=1}^{\infty} D_i$, where each D_i is isometric to the graph of a Lipschitz function $\phi_i\colon K_i \to \mathbb{R}$ with K_i being a compact nowhere dense subset of $\mathbb{R}^{n-1}$.*

Recall, that the graph of a Lipschitz function defined on an $(n-1)$-dimensional space has the n-dimensional Lebesgue measure 0. (See e.g. [20].) In particular, the set $D(f)$ has Lebesgue measure 0 for every linearly continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$. This stands in contrast with the case of separately continuous functions. Indeed, by Theorem 19, there exists not only separately continuous, but even $\mathcal{D}^{+}_{n-1,n}$-continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ with $D(f) = P^n$, where $P \subset \mathbb{R}$ is a nowhere dense set of positive Lebesgue measure.

Slobodnik's response to Kronrod's challenge would be the only voice in this direction for twenty years. However, in 1997, E.E. Shnol' [46], recalling some conversations with Kronrod and evidently unaware of Slobodnik's results (or, for that matter, Kershner's), proved a necessary condition very much like Slobodnik's theorem (incorrectly listed as necessary and sufficient in the English translation) for functions on $\mathbb{R}^2$. (See also [24].)

The above theorem describes a necessary structure of sets $D(f)$ for linearly continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$. Some sufficient conditions for these sets, published in 2013 by the first author and T. Glatzer [11, theorem 2.3], are given below.

Theorem 21. (Ciesielski & Glatzer 2013) *If $D \subset \mathbb{R}^n$ can be written as $D = \bigcup_{i=1}^{\infty} D_i$, where each D_i is isometric to the graph of a function $f_i \restriction H_i$, with function $f_i \colon \mathbb{R}^{n-1} \to \mathbb{R}$ being convex and H_i being a closed nowhere dense subset of $\mathbb{R}^{n-1}$, then there exists a linearly continuous function $f \colon \mathbb{R}^n \to \mathbb{R}$, such that $D(f) = D$.*

For $n = 2$ the results remains true when the functions f_i are required to be continuously twice differentiable, instead of being convex.

Although Theorems 20 and 21 describe quite well the structures of the sets $D(f)$ for linearly continuous function $f \colon \mathbb{R}^n \to \mathbb{R}$, they do not provide a characterization for such sets. Actually, as it was pointed out in [11, corollary 2.4 and proposition 2.5], neither the necessary condition from Theorem 20 nor the sufficient condition from Theorem 21 characterize the sets $D(f)$ for linearly continuous function $f \colon \mathbb{R}^n \to \mathbb{R}$. Therefore, Kronrod's challenge discussed above is not fully solved, leading to the following problem:

Problem 4. (Kronrod) Find a non-trivial nice characterization of the sets $D(f)$ for linearly continuous function $f \colon \mathbb{R}^n \to \mathbb{R}$.

Note that in [12, theorem 2.3] the first author and T. Glatzer provide an answer for this problem for $n = 2$. However, this characterization is in terms of the topology of lines in $\mathbb{R}^2$ and is quite complicated. Finding a nicer characterization would be useful.

5.4 Higher-dimensional versions of linear continuity

Of course, we have in mind here the $\mathcal{F}_k$-continuity for the family $\mathcal{F}_k$ of all k-flats in $\mathbb{R}^n$. The following theorem of the first author and T. Glatzer, [12, theorem 2.2], is a generalization of Theorem 20. It provides an upper bound on the size of sets $D(f)$.

Theorem 22. (Ciesielski & Glatzer 2014) *Let $0 < k < n$ and $f \colon \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{F}_k$-continuous. Then, there exists a sequence $\langle f_i \rangle_{i<\omega}$ of Lipschitz functions from $\mathbb{R}^{n-k}$ to $\mathbb{R}^k$ such that $D(f)$ is covered by a union of isometric copies of the graphs of functions f_i.*

Notice, that this result immediately implies that the considered sets $D(f)$ must have Hausdorff dimension at most $n - k$. A lower bound for the sets $D(f)$ is given the next theorem, [12, proposition 2.4], which can be viewed as a higher-dimensional counterpart of Theorem 21.

Theorem 23. (Ciesielski & Glatzer 2014) *For every $0 < k < n$ and compact nowhere dense $K \subset \mathbb{R}$, there exists an $\mathcal{F}_k$-continuous function $f : \mathbb{R}^n \to \mathbb{R}$ with $D(f) = \{0\}^k \times K \times \mathbb{R}^{n-k-1}$. In particular, $D(f)$ may have positive $(n-k)$-Hausdorff measure.*

Moreover, if D is a countable union of arbitrary isometric copies of the sets of the form $\{0\}^k \times K \times \mathbb{R}^{n-k-1}$ as above, then $D = D(f)$ for some $\mathcal{F}_k$-continuous function $f : \mathbb{R}^n \to \mathbb{R}$.

For $0 < k < n$, let $\mathcal{D}^+_{k,n}$ (respectively $\mathcal{D}_{k,n}$) be the family of all sets $D(f)$ for $\mathcal{F}^+_k$-continuous ($\mathcal{F}_k$-continuous, respectively) functions $f : \mathbb{R}^n \to \mathbb{R}$. Then, we obviously have the inclusions indicated in the following chart

$$\begin{array}{ccccc} \mathcal{D}_{n-1,n} & \subset & \cdots & \subset & \mathcal{D}_{1,n} \\ \cap & & & & \cap \\ \mathcal{D}^+_{n-1,n} & \subset & \cdots & \subset & \mathcal{D}^+_{1,n}. \end{array}$$

Moreover, the measure theoretical consideration mentioned above shows that all indicated inclusions are proper. (For more details, see [12].) In particular, all classes of $\mathcal{F}^+_k$-continuous and $\mathcal{F}_k$-continuous functions are distinct.

It would be good to have an answer for a higher-dimensional analogue of Kronrod problem.

Problem 5. For $0 < k < n$, find a non-trivial nice characterization of the families $\mathcal{D}_{k,n}$.

There is also a higher-dimensional analogue of Problem 2.

Problem 6. For every $0 < k < n$, find the smallest number m^k_n such that every $\mathcal{F}_k$-continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is of Baire class m^k_n.

Notice, that

$$1 \leq m^k_n < \lceil n/k \rceil \quad \text{for every } 0 < k < n, \tag{10}$$

where $\lceil n/k \rceil$ is the smallest integer greater than or equal to n/k.

Indeed, the inequality $m_n \geq 1$ is justified by the existence of discontinuous hyperplane continuous functions. The other inequality holds, as for $m = \lceil n/k \rceil$ we have $\mathbb{R}^n = X_1 \times \cdots \times X_m$, where each $X_i = \mathbb{R}^{k_i}$ for some $k_i \leq k$. Since every $\mathcal{F}_k$-continuous function $f : \mathbb{R}^n \to \mathbb{R}$ can be treated as a map $f : X_1 \times \cdots \times X_m \to \mathbb{R}$, which is separately continuous, f is in the Baire class $m - 1$ by a version of Theorem 4(i) for the functions defined on the product of metric spaces.

Clearly (10) implies that $m^k_n = 2$ for $\lceil n/k \rceil = 2$, that is, for $k \geq n/2$. So, Problem 6 is of interest only for $0 < k < n/2$.

5.5 $(D^2)^*$-continuity

Theorem 22 implies that even a function satisfying the strongest among the $\mathcal{F}_k$-continuity conditions—the hyperplane (i.e., $\mathcal{F}_{n-1}$-) continuity—can have an uncountable set of points of discontinuity. Is this a general rule that for any family $\mathcal{H}$ of subsets of $\mathbb{R}^n$, the following dichotomy holds: either $\mathcal{H}$-continuity implies continuity, or there exists an $\mathcal{H}$-continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ with an uncountable $D(f)$?

As it might be expected, this is not the case. For example, if $\mathcal{H}$ consists of all graphs of continuous functions $g\colon \mathbb{R} \to \mathbb{R}$ with $g(0) \neq 0$, then any $\mathcal{H}$-continuous function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is continuous at every $\langle x, y\rangle \neq \langle 0, 0\rangle$. However, the characteristic function $\chi_{\{\langle 0,0\rangle\}}$ of the singleton $\{\langle 0, 0\rangle\}$ is $\mathcal{H}$-continuous with $D\left(\chi_{\{\langle 0,0\rangle\}}\right) = \{\langle 0, 0\rangle\}$.

Nevertheless, the dichotomy is true for generalized continuities considered in the other parts of this paper. Indeed, the strongest among all such continuities which still does not imply continuity is that of $(D^2)^*$-continuity. At the same time, the first author and T. Glatzer [10] constructed the following example.

Theorem 24. **(Ciesielski & Glatzer 2012)** *There exists a $(D^2)^*$-continuous function $f\colon \mathbb{R}^2 \to \mathbb{R}$ for which $D(f)$ is a perfect set of arbitrarily large 1-Hausdorff measure.*

Problem 7. Is there an analogue of an example from Theorem 24 for functions of more than two variables?

6 Concluding remarks

In this paper we described a long history and the current research concerning the continuity of real valued functions on several variables, emphasizing an interplay between the functions of one variable and those of more than one variable. Many of the discussed topics are so fundamental, that they first appeared in calculus textbooks, rather than in research papers. This includes Theorem X of Cauchy from [9], the examples $\frac{xy}{x^2+y^2}$ and $\frac{xy^2}{x^2+y^4}$ from the Genocchi-Peano text [23], and the earlier version of these examples included in the Thomae's book [49]. Some other results, like Theorem 8 of Luzin or examples of discontinuous hyperplane continuous functions $\frac{xyz^2}{x^2+y^4+z^8}$ or $\frac{x_1x_2x_3^2x_4^3}{x_1^4+x_2^6+x_3^8+x_4^{10}}$, perhaps should be included in a standard calculus curriculum. One might be even tempted to include in a multivariable calculus text a version of Theorem 6 of A. Rosenthal 1955, or at least some version of the example (4).

Perhaps going beyond the mathematics textbooks, some elements of the presented discussion of Theorem X of Cauchy can bring us back to a fundamental question of *Why the Archimedean set of real numbers should be considered more suitable for doing analysis, rather than a non-Archimedean one?* Taking under consideration the implications of Theorem X—the easiness of representing the continuity of functions of many real variables in terms of functions of one variable—maybe a non-standard analysis deserves a more central place in analysis that it currently holds? Perhaps, a position more similar to one that complex analysis enjoys? But, certainly, such a shift of paradigm would require considerably more study of non-standard analysis that currently exists. After all, complex analysis is one of the most applicable parts of mathematics, while, so far, there seem to be relatively few similar results for the analysis on the non-Archimedean set(s) of reals.

References

[1] R. Baire, *Sur les fonctions des variables réelles*, Annali di Matematica Pura ed Applicata, **3** (1899), 1–122.

[2] J. C. Breckenridge and T. Nishiura, *Partial Continuity, Quasi-Continuity, and Baire Spaces*, Bull. Inst. Math. Acad. Sinica, **4** (1976), 191–203.

[3] K. Bögel, *Über partiell differenzierbare Funktionen*, Math. Z., **25** (1926), 490–498.

[4] J. Boman, *Differentiability of a function and of its compositions with functions of one variable*, Math. Scand., **20** (1967), 249–268.

[5] A. Borovik and M. G. Katz, *Who Gave You the Cauchy-Weierstrass Tale? The Dual History of Rigorous Calculus*, Found. Sci., **17** (2012), 245–276.

[6] R. E. Bradley and C. E. Sandifer, *Cauchy's Cours d'analyse. An Annotated Translation*, Springer, 2009.

[7] A. M. Bruckner, *Differentiation of Real Functions*, CRM Monograph Series 5, American Mathematical Society, 1994.

[8] J. Bourgain, D. H. Fremlin, and M. Talagrand, *Pointwise compact sets of Baire measurable functions*, Amer. J. Math., **100** (1978), 845–886.

[9] L. A. Cauchy, *Cours D'analyse de l'Ecole Royale Polytechnique*, Paris, 1821.

[10] K. C. Ciesielski and T. Glatzer, *Functions continuous on twice differentiable curves, discontinuous on large sets*, Real Anal. Exchange, **37**(2) (2012), 353–361.

[11] K. C. Ciesielski and T. Glatzer, *Sets of Discontinuities of Linearly Continuous Functions*, Real Anal. Exchange, **38**(2) (2013), 337–389.

[12] K. C. Ciesielski and T. Glatzer, *Sets of discontinuities for functions continuous on flats*, Real Anal. Exchange, **39**(1) (2014), 117–138.

[13] K. C. Ciesielski and D. Miller, *On a Genocchi-Peano example*, submitted.

[14] K. C. Ciesielski and J. Rosenblatt, *Restricted continuity and a theorem of Luzin*, Colloq. Math., **135**(2) (2014), 211–225.

[15] J. Činčura, T. Šalát, and T. Visnyai, *On separately continuous function* $f\colon \ell^2 \to \mathbb{R}$, Acta Academie Paedagogicae Agriensis, **31** (2004), 11–18.

[16] J. Cleave, *The concept of 'variable' in nineteenth century analysis*, The British Journal for the Philosophy of Science, **30**(3) (1979), 266–278.

[17] R. O. Davies and J. Dravecký, *On the measurability of functions of two variables*, Mat. Časopis Sloven. Akad. Vied, **23** (1973), 285–289.

[18] O. P. Dzagnidze, *Separately continuous functions in a new sense are continuous*, Real Analysis Exchange, **24**(2) (1998–99), 695–702.

[19] P. Ehrlich, *The rise of non-Archimedean mathematics and the roots of a misconception. I. The emergence of non-Archimedean systems of magnitudes*, Arch. Hist. Exact Sci., **60**(1) (2006), 1–121.

[20] K. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge, 1985.

[21] W. Felscher, *Bolzano, Cauchy, epsilon, delta*, American Mathematical Monthly, **107**(9) (2000), 844–862.

[22] C. Freiling, *On the problem of characterizing derivatives*, Real Anal. Exchange, **23**(2) (1997–98), 805–812.

[23] A. Genocchi and G. Peano, *Calcolo differentiale e principii di Calcolo*, Torino, 1884.

[24] T. Glatzer, *Continuities on Subspaces*, Ph.D. Thesis (written under supervision of K. C. Ciesielski), West Virginia University, 2013.

[25] Z. Grande, *La mesurabilité des fonctions de deux variables*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys., **22** (1974), 657–661.

[26] H. Hahn, *Über Funktionen mehrer Veränderlichen die nach jeder einzelnen Veränderlichen stetig sind*, Math Zeit., **4** (1919), 306–313.

[27] M. Jarnicki and P. Pflug, *Directional vs. Joint Regularity*, Notices Amer. Math. Soc., **58** (2011), 896–904.

[28] R. Kershner, *The Continuity of Functions of Many Variables*, Trans. Amer. Math. Soc., **53** (1943), 83–100.

[29] M. Kline, *Mathematical Thought From Ancient to Modern Times*, Oxford University Press, 1972.

[30] R. L. Kruse and J. J. Deely, *Joint Continuity of Monotonic Functions*, Amer. Math. Monthly, **76**(1) (1969), 74–76.

[31] M. Laczkovich and A. Miller, *Measurability of functions with approximately continuous vertical sections and measurable horizontal sections*, Colloq. Math., **69**(2) (1995), 299–308.

[32] E. M. Landis and I. M. Iaglom, *Remembering A. S. Kronrod*, Mathematical Intelligencer, **24**(1) (2002), 22–30.

[33] H. Lebesgue, *Sur l'approximation des fonctions*, Bull. Sciences Math, **22** (1898), 278–287.

[34] H. Lebesgue, *Sur les fonctions representable analytiquement*, J. Math. Pure Appl., **6** (1905), 139–216.

[35] N. N. Luzin, *Teoria funktsii deistvitel'nogo peremennogo*, 2nd ed., National Educational Publishers, Moscow, 1948.

[36] E. Marczewski and C. Ryll-Nardzewski, *Sur la measurabilite des fonctions de plusieurs variables*, Ann. Soc. Polon. Math., **25** (1952), 145–154.

[37] L. Mou and M. McAsey, *Remark on the continuity of multivariate Functions*, Real Anal. Exchange, **37** (2011), 167–176.

[38] O. Nanyes, *Limits of Functions of Two Variables*, College Math. J., **36**(4) (2005), 326–329.

[39] O. Nanyes, *Continuously differentiable curves detect limits of functions of two variables*, College Math. J., **45**(1) (2014), 54–56.

[40] Z. Piotrowski, *Separate and joint continuity*, Real Anal. Exchange, **11** (1985/86), 293–322.

[41] Z. Piotrowski, *Separate and joint continuity, II*, Real Anal. Exchange, **15** (1989/90), 248–258.

[42] Z. Piotrowski, *The genesis of separate versus joint continuity*, Tatra Mountains Math. Publ., **8** (1996), 113–126.

[43] A. Robinson, *Non-standard Analysis*, North-Holland, Amsterdam 1966.

[44] A. Rosenthal, *On the Continuity of Functions of Several Variables*, Math. Zeitschr., **63** (1955), 31-38.

[45] L. Scheeffer, *Theorie der Maxima und Minima einer Function von zwei Variabeln*, Math. Ann., **35** (1890), 541–576.

[46] E. E. Shnol', *Functions of two variables continuous along straight lines*, Mat. Zametki, **62**(2) (1997). English translation available in Math Notes, **62** (1997), 255–259.

[47] W. Sierpiński, *Sur une propriété de fonctions de deux variables réelles, continues par rapport à chacune de variables*, Publ. Math. Univ. Belgrade, **1** (1932), 125–128.

[48] S. G. Slobodnik, *An Expanding System of Linearly Closed Sets*, Mat. Zametki, **19** (1976), 67–84; English translation Math Notes, **19** (1976), 39–48.

[49] J. Thomae, *Abriss einer Theorie der complexen Funktionen und der Thetafunktionen einer Veränderlichen*, Halle, 1873. (First edition published in 1870.)

[50] G. P. Tolstoff, *On partial derivatives*, (Russian) Izv. Akad. Nauk SSSR Ser. Mat. vol. **13** (1949), 425–446; Amer. Math. Soc. Transl. No. **69**.

[51] T. Visnyai, *Strongly separately continuous and separately quasicontinuous functions* $f\colon \ell^2 \to \mathbb{R}$, Real Anal. Exchange, **38** (2013), 499–511.

[52] W. H. Young, *A note on monotone functions*, The Quarterly Journal of Pure and Applied Mathematics (Oxford Ser.), **41** (1910), 79–87.

[53] W. H. Young and G. C. Young, *Discontinuous functions continuous with respect to every straight line*, Quart. J. Math., Oxford (Series I) **41** (1910), 87–93.

[54] Z. Zahorski, *Sur l'ensamble des points singuliére d'une fonction d'une variable réele admettand des dérivées des tous orders*, Fund. Math., **34** (1947), 183–245.