mathematics

Article

Generalized Approach to Differentiability

Nikola Koceić-Bilan * and Snježana Braić

Faculty of Science, University of Split, 21000 Split, Croatia
* Correspondence: koceic@pmfst.hr; Tel.: +385-21619202

Abstract: In the traditional approach to differentiability, found in almost all university textbooks, this notion is considered only for interior points of the domain of function or for functions with an open domain. This approach leads to the fact that differentiability has usually been considered only for functions with an open domain in $\mathbb{R}^n$, which severely limits the possibility of applying the potential techniques and tools of differential calculus to a broader class of functions. Although there is a great need for generalization of the notion of differentiability of a function in various problems of mathematical analysis and other mathematical branches, the notion of differentiability of a function at the non-interior points of its domain has almost not been considered or successfully defined. In this paper, we have generalized the differentiability of scalar and vector functions of several variables by defining it at non-interior points of the domain of the function, which include not only boundary points but also all points at which the notion of linearization is meaningful (points admitting nbd rays). This generalization allows applications in all areas where standard differentiability can be applied. With this generalized approach to differentiability, some unexpected phenomena may occur, such as a function discontinuity at a point where a function is differentiable, the non-uniqueness of differentials... However, if one reduces this theory only to points with some special properties (points admitting a linearization space with dimension equal to the dimension of the ambient Euclidean space of the domain and admitting a raylike neighborhood, which includes the interior points of a domain), then all properties and theorems belonging to the known theory of differentiability remain valid in this extended theory. For generalized differentiability, the corresponding calculus (differentiation techniques) is also provided by matrices—representatives of differentials at points. In this calculus the role of partial derivatives (which in general cannot exist for differentiable functions at some points) is taken by directional derivatives.

Keywords: differentiability; partial derivatives; derivatives in the direction; set of linear contribution; linearization space; neighbourhood ray; raylike neighbourhood

MSC: 26B05; 26B12

check for updates

Citation: Koceić-Bilan, N.; Braić, S. Generalized Approach to Differentiability. *Mathematics* **2022**, *10*, 3085. https://doi.org/10.3390/math10173085

Academic Editor: Andrey Amosov

Received: 21 July 2022
Accepted: 22 August 2022
Published: 27 August 2022

1. Introduction and Motivation

One of the basic ideas of differential calculus is to better approximate a given function $f : X \to \mathbb{R}^m, X \subseteq \mathbb{R}^n$, locally by an affine function, i.e., to linearize it at a point $P_0 \in X$. For this to be possible, the function must be differentiable at this point which means that there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that the limit

$$\lim_{H\to 0} \frac{f(P_0 + H) - f(P_0) - A(H)}{||H||} \tag{1}$$

exists and is equal to $0 \in \mathbb{R}^m$. For practical reasons, differentiability in mathematical analysis has been defined and considered almost only for functions $f : \Omega \to \mathbb{R}^m$ with an open domain $\Omega \subseteq \mathbb{R}^n$ [1–4]. Since every point of an open set $\Omega \subseteq \mathbb{R}^n$ is an accumulation point of Ω [1] then for every point $P_0 \in \Omega$ it holds that $0 \in \mathbb{R}^n$ is the accumulation point of the domain D of the function

$$H \mapsto \frac{f(P_0 + H) - f(P_0) - A(H)}{||H||}$$

for a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$. Indeed, there exists $r > 0$ such that the open ball $B(P_0, r)$ is contained in Ω and consequently $B(0, r) \subseteq D$ and the limit from the definition of differentiability (1) is reasonable to consider. (Recall that the limit of the function can be considered only at an accumulation point of the domain.)

However, reducing differentiability only to an open domain, i.e., to the interior points of a domain, has, in addition to many successful applications and advantages, some obvious deficiencies. For example, for the function $f : [0, \infty\rangle \to \mathbb{R}\ \ f(x) = \sqrt{x^3}$ it holds

$$\lim_{h \to 0^+} \frac{\sqrt{(0+h)^3} - \sqrt{0^3} - 0h}{|h|} = 0,$$

so this function can be well approximated by the zero operator, i.e., it could be linearized at the point $0 \in \mathbb{R}$ inside the natural domain of f, but due to the conditions from the definition of differentiability (that a point must belong to the interior of the domain [2]), the differentiability of the function at this point is usually not considered at all. Even though this issue can be overcome by extending the definition of differentiability (derivability) of a real function of a real variable to the endpoints of the given domain using one side limits [5], for a function of several variables the problem of differentiability at non-interior points of the domain remains current. For example, the differentiability of the function

$$f : D \to \mathbb{R}\ \ f(x, y) = \sqrt{y - x^3},\ D = \left\{(x, y) \in \mathbb{R}^2 \mid y \geq x^3\right\}$$

cannot be considered in all boundary points $(x, x^3), x \in \mathbb{R}$, although it can be well linearized locally by the zero operator in those points. Similarly, because of the reduction to open sets, the question of the existence of tangents [6] and tangent planes [4] of a function $f : \mathrm{Cl}\,\Omega \to \mathbb{R}$, $\Omega \subseteq \mathbb{R}$ or $\mathbb{R}^2$, at points $(x, f(x))$, $x \in \mathrm{Fr}\,\Omega$, remains open. For example, due to this reduction we cannot obtain the tangent of the function $x \mapsto \sqrt{x^3}$ at the point $O = (0, 0)$ although it is obvious that for points $T_x = (x, f(x))$, $x \in \langle 0, \infty\rangle$, the secants OT_x tend to the line $y = 0$ as x tends to 0, and the line $y = 0$ should be the tangent of this function at the point O. Moreover, the study of the local conditional extreme of a scalar function is reduced to the study of a function whose domain is not necessarily an open set, so that the problem of finding a conditional extreme cannot be clarified or fully studied if differentiability is studied only on open sets. Furthermore, a differentiable function would lose the property of differentiability at many points if differentiability at boundary points is not considered when switching from one Cartesian coordinate system to other non-affine coordinate systems (or vice versa).

These are some of the reasons that indicate that the notion of differentiability should be generalized by observing differentiability not only at interior points of sets, but much more broadly, at points of any domain $X \subseteq \mathbb{R}^n$ of a function $f : X \to \mathbb{R}^m$ in which the notion of differentiability and linearization is meaningful. John W. Milnor mentioned this problem in his famous series of lectures on differential topology which dates back to 1965 [7]. We will show that this extension is meaningful for all points $P_0 \in X$ for which there is at least one point $Q \in X \backslash \{P_0\}$ such that the line segment $\overline{P_0 Q}$ is contained in X. Indeed, this is the most general case in which a linear operator can linearize a function at a point (at least on a line segment to which this point belongs). The linearization space is then a one-dimensional vector subspace of $\mathbb{R}^n$ which is also the smallest vector subspace on which it is interesting to consider and specify a linear operator.

In the history of modern mathematics one can find some other issues or (overlooked) problems of mathematical analysis like this one [8], where we take for granted some traditional approaches, common requirements and (sometimes wrong) conclusions. Concerning differentiability, one can find in the literature some generalizations of differentiability (derivability) such as the fractional derivative [9] or the derivative at the endpoints of a segment [5]. In this paper, we provide a natural generalization of differentiability of a function by defining it at some non-interior points of the domain of function. These points include not only the boundary points of the domain, but also all points in which the notion of differentiability and linearization is meaningful. For this generalized case, a corresponding calculus (techniques of differentiation) is also provided.

2. Preliminaries

In this paper we use the notation $(|)$ for the Euclidean scalar product on $\mathbb{R}^n$, the notation $|| \; ||$ for the Euclidean norm and the notation d for the Euclidean metric. We use the notation O for the point $(0, \ldots, 0) \in \mathbb{R}^n$ or we simply write $0 \in \mathbb{R}^n$.

Let $\Omega \subseteq \mathbb{R}^n$ be an open set, $f : \Omega \to \mathbb{R}^m$ a function, and $P_0 = \left(x_1^0, \ldots, x_n^0\right)$ an arbitrary point in Ω. To approximate the function f on the open ball $B(P_0, r) \subseteq \Omega$, $r > 0$, at the point P_0 with the special affine function $\alpha : \mathbb{R}^n \to \mathbb{R}^m \;\; \alpha(H) = f(P_0) + A(H)$ means to find a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ [10] such that $f(P_0 + H) \sim \alpha(H)$ for any $H \in B(O, r) \subseteq \mathbb{R}^n$. Geometrically interpreted, in the case of $m = 1$ this means that we want to replace the part of the graph of the function f at the point $(P_0, f(P_0))$ by the part of the graph of the affine function

$$\alpha(x_1, \ldots, x_n) = f(P_0) + a_1 x_1 + \cdots + a_n x_n, a_i \in \mathbb{R}, i = 1, \ldots n,$$

i.e., the part of the hyperplane in $\mathbb{R}^{n+1}$. The desirable property of such an approximation is that it is as accurate as possible at points closer to the point P_0, i.e., that the error

$$r(H) := f(P_0 + H) - \alpha(H) = f(P_0 + H) - f(P_0) - A(H)$$

tends to zero as H tends to zero. However, if f is a continuous function, then the error $r(H)$ always tends to zero as H tends to zero (because every linear operator acting between finite-dimensional vector spaces is continuous). This would mean that there is an adequate local replacement by the affine function of any continuous mapping, which is not the case. For example, if we consider the function $f : \mathbb{R}^2 \to \mathbb{R} \;\; f(x, y) = \sqrt{x^2 + y^2}$, it is easy to see that on the open ball $B((0, 0), \varepsilon)$ we cannot approximate this function by an affine function, i.e., we cannot replace its graph well enough by a part of the plane passing through the origin $O \in \mathbb{R}^3$, although this is perfectly possible on all rays starting in O. Thus, it is not only necessary that the error $r(H)$ can be made arbitrarily small (because every continuous function has this property), but even more so that the relative error $\frac{r(H)}{||H||}$ can be made arbitrarily small, which leads us to the definition of differentiability of the function f at the point P_0, which is as follows [3]:

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. A function $f : \Omega \to \mathbb{R}^m$ is differentiable at a point $P_0 \in \Omega$ if there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that the limit

$$\lim_{H \to 0} \frac{f(P_0 + H) - f(P_0) - A(H)}{||H||}$$

exists and is equal to $0 \in \mathbb{R}^m$. We then call the linear operator A the differential of the function$f$ at the point P_0, it is unique and we denote it by $df(P_0)$.

A linear operator A is the differential of the function f at the point P_0 if and only if

$$f(P_0 + H) - f(P_0) = A(H) + r(H)$$

where $r : B(0, \varepsilon) \to \mathbb{R}^m$ is the error function with the property

$$\lim_{H \to 0} \frac{r(H)}{||H||} = 0 \text{ and } B(P_0, \varepsilon) \subseteq \Omega.$$

3. Linearization of Function

Definition 1. *Let $X \subseteq \mathbb{R}^n$ be a set and $P_0 \in X$. We say that the point P_0* ***admits a neighborhood ray*** *(or simply admits a nbd ray)* ***in*** *X if there exists $H \in \mathbb{R}^n \backslash \{0\}$ such that the line segment $\overline{P_0P_0 + H}$ is contained in X.*

This notion is of particular importance to us because we will consider the linearization of a function exactly at points in a domain that admit at least one nbd ray in the domain (and not, as before, only at points from its interior).

Example 1.
- (a) *No point of a sphere admits a nbd ray in it.*
- (b) *Every point of a non-trivial convex set admits a nbd ray in that set.*

Since every line segment is a convex set and every nontrivial convex set contains the line segment between any two of its points, a point $P_0 \in X \subseteq \mathbb{R}^n$ admits a nbd ray in X if and only if there exists a nontrivial convex set $K \subseteq \mathbb{R}^n$ such that $P_0 \in K \subseteq X$.

Definition 2. *Let $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point admitting nbd ray in X. The set*

$$\Delta_{X,P_0} := \left\{H \in \mathbb{R}^n \backslash \{0\} \mid \overline{P_0P_0 + H} \subseteq X\right\}$$

is called the ***set of linear contributions at*** *P_0* ***in*** *X, and its linear hull [10]*

$$\Sigma_{X,P_0} := \left[\Delta_{X,P_0}\right]$$

is said to be the ***linearization space at*** *P_0* ***with respect to*** *X.*

For a function $f : X \to \mathbb{R}^m$ and a point $P_0 \in X$ we say that Σ_{X,P_0} is the ***linearization space of the function*** *f* ***at the point*** *P_0.*

Example 2.
- (a) *Let the points $P, Q, R \in \mathbb{R}^n$ be in general position, i.e., let them be the three non-collinear points. Then it holds*

$$\begin{aligned} \Delta_{\overline{PQ},P} &= \{t(Q-P) \mid t \in \langle 0,1]\} \text{ and} \\ \Sigma_{\overline{PQ},P} &= \{t(Q-P) \mid t \in \mathbb{R}\}, \\ \Delta_{\overline{PQ}\cup\overline{PR},P} &= \left(\overline{O(Q-P)} \cup \overline{O(R-P)}\right)\backslash\{O\} \text{ and} \\ \Sigma_{\overline{PQ}\cup\overline{PR},P} &= \{\alpha(Q-P) + \beta(R-P) \mid \alpha,\beta \in \mathbb{R}\}. \end{aligned}$$

- (b) *Let $D = \{P \in \mathbb{R}^n \mid ||P - P_0|| \le r\}$ and $Q \in D$. Then it holds*

$$\Delta_{D,Q} = \{P - Q \mid P \in D\backslash\{Q\}\} \text{ and}$$

$$\Sigma_{D,Q} = \mathbb{R}^n.$$

Let us now generalize the notion of differentiability of a function to points admitting nbd ray in the domain. This will allow us to consider differentiability at points where this was not possible so far.

Definition 3. *Let $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point admitting nbd ray in X. We say that a function $f : X \to \mathbb{R}^m$ is* ***differentiable at*** *P_0 if there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that the limit*

$$\lim_{\substack{H\to 0 \\ H\in\Delta_{X,P_0}}} \frac{f(P_0 + H) - f(P_0) - A(H)}{||H||} \tag{2}$$

exists and is equal to $0 \in \mathbb{R}^m$*. If such a linear operator exists, we call it the* ***differential of the function*** f ***at the point*** P_0.

The function f *is* ***differentiable on*** X *if* f *is differentiable at every point of* X.

Notice that if a point $P_0 \in X$ admits nbd ray in X, then P_0 is an accumulation point of the set X and then $0 \in \mathbb{R}^n$ is an accumulation point of the set Δ_{X,P_0}. Indeed, if $\overline{P_0P_0+H} \subseteq X$ then every nbd of P_0 contains some points of the line segment $\overline{P_0P_0+H} \subseteq X$ and consequently every nbd of 0 intersects Δ_{X,P_0}. Therefore, the limit from the previous definition makes sense to consider. Furthermore, the natural domain $D \subseteq \mathbb{R}^n$ of the function

$$H \mapsto \frac{f(P_0+H) - f(P_0) - A(H)}{||H||} \tag{3}$$

could be in general a superset of Δ_{X,P_0}, so it is necessary to emphasize that the limit (2) is considered only on the set Δ_{X,P_0} (this is the limit of the restriction of the function (3) to the set Δ_{X,P_0} at 0). Otherwise, the values of the above function at points that do not belong to Δ_{X,P_0} but are in D and near 0 may affect the existence of the limit of the function (3) at 0, which we do not want to allow. But, if $P_0 \in \operatorname{Int} X$ then it holds

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||} = \lim_{H\to 0} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||},$$

which is a consequence of the following theorem:

Theorem 1. *Let* $f : X \to \mathbb{R}^m$*,* $X \subseteq \mathbb{R}^n$*,* $Y \subseteq X$ *and* P_0 *be an accumulation point of the set* Y*. Let* U *be an open neighborhood of the point* P_0 *in* $\mathbb{R}^n$ *such that* $(U \backslash \{P_0\}) \cap X \subseteq Y$*. If the restriction* $f|_Y : Y \to \mathbb{R}^m$ *has the limit at the point* P_0*, then* f *has the limit at* P_0 *and they are equal, i.e.,* $\lim\limits_{P_0}(f|_Y) = \lim\limits_{P_0} f$*.*

Proof. Let $Q_0 := \lim\limits_{P_0}(f|_Y)$ and let $B(Q_0, \varepsilon)$ be an open ball in $\mathbb{R}^m$. Then there exists an open neighborhood V of P_0 in $\mathbb{R}^n$ such that $V \subseteq U$ and $f|_Y(V \cap (Y \backslash \{P_0\})) \subseteq B(Q_0, \varepsilon)$. Hence,

$$f(V \cap (X \backslash \{P_0\})) = f(V \cap (Y \backslash \{P_0\})) = f|_Y(V \cap (Y \backslash \{P_0\})) \subseteq B(Q_0, \varepsilon)$$

which implies that $\lim\limits_{P_0} f = Q_0$. □

Therefore, in the above definition of differentiability of a function at a point, we can omit the notation of the restriction in the limit if this point belongs to the interior of the domain. In this case, the above definition coincides with the previously known definition of this notion. Thus, the Definition 3 is a natural generalization of the notion of differentiability and this generalization brings many advantages and solves many contentious issues and problems (e.g., at the boundary points of a domain...), which we will explain hereinafter with several various examples.

From the definition of differentiability, it follows that the linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ is the differential of a function $f : X \to \mathbb{R}^m$ at a point $P_0 \in X \subseteq \mathbb{R}^n$ if and only if

$$f(P_0+H) - f(P_0) = A(H) + r(H),$$

where $r : \Delta_{X,P_0} \to \mathbb{R}^m$ is the error function with the property $\lim\limits_{H\to 0} \frac{r(H)}{||H||} = 0$. Notice that the above equation makes sense only on Δ_{X,P_0}, i.e., only for a sufficiently small neighborhood U of the point $0 \in \mathbb{R}^n$ [2] we can write

$$f(P_0+H) - f(P_0) \sim A(H)$$

for $H \in U \cap \Delta_{X,P_0}$. Likewise, the linear operator A, although defined on $\mathbb{R}^n$, has its true meaning from the point of view of approximating the function f only on the linearization space $\Sigma_{X,P_0} \subseteq \mathbb{R}^n$.

It is important to notice that the differential of a function at a point need not be unique (which could not be the case so far). Indeed, if the linearization space Σ_{X,P_0} is a proper subset of $\mathbb{R}^n$ and if there exists a differential $A : \mathbb{R}^n \to \mathbb{R}^m$ of the function f at P_0, then every linear operator $B : \mathbb{R}^n \to \mathbb{R}^m$ that coincides with A on the subspace Σ_{X,P_0} (and there are infinitely many of them) is the differential of the function f at P_0 because it satisfies the conditions of the definition of differentiability. Let us formalize this consideration by the following statement:

Proposition 1. *Let $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point admitting a nbd ray in X. If the function f is differentiable at P_0 and $A : \mathbb{R}^n \to \mathbb{R}^m$ is the differential of the function f at the point P_0 then every linear operator $B : \mathbb{R}^n \to \mathbb{R}^m$ which agrees with A on the vector space Σ_{X,P_0} is also the differential of the function f at the point P_0.*

Proof. Using equality

$$A|_{\Sigma_{X,P_0}} = B|_{\Sigma_{X,P_0}}$$

it is easy to check that

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||} = 0 = \lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - B(H)}{||H||}$$

holds. □

Now, we will show that all differentials of f at P_0 are equal on the linearization space Σ_{X,P_0}.

Theorem 2. *Let $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point admitting a nbd ray in X. If the differential of the function f exists at the point P_0 then it is unique on the vector space Σ_{X,P_0}.*

Proof. Suppose that $A, B : R^n \to \mathbb{R}^m$ are two linear operators for which

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||} = 0 = \lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - B(H)}{||H||}.$$

Then it holds

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{B(H) - A(H)}{||H||} = 0.$$

Every vector $H \in \Sigma_{X,P_0}$ can be written as a linear combination of vectors from Δ_{X,P_0}. Therefore, $A|_{\Sigma_{X,P_0}} = B|_{\Sigma_{X,P_0}}$ if and only if $A(H) = B(H)$ for every $H \in \Delta_{X,P_0}$. If $H \in \Delta_{X,P_0}$ then $tH \in \Delta_{X,P_0}$ for every $t \in \langle 0, 1]$ and

$$\lim_{t\to 0^+} \frac{B(tH) - A(tH)}{||tH||} = 0,$$

so it follows

$$0 = \lim_{t\to 0^+} \frac{(B-A)(tH)}{||tH||} = \lim_{t\to 0^+} \frac{t(B-A)(H)}{|t|||H||} = \frac{(B-A)(H)}{||H||}.$$

Therefore, $A(H) = B(H)$ for every $H \in \Delta_{X,P_0}$ and it holds

$$A|_{\Sigma_{X,P_0}} = B|_{\Sigma_{X,P_0}}.$$

□

Remark 1. *One might think that the cases where the linearization space Σ_{X,P_0} is a proper subset of $\mathbb{R}^n$ and the differential of the function f exists at the point P_0 cause certain difficulties because the differential is not unique, but it is unique where it should be, i.e., on the linearization space Σ_{X,P_0}. According to the previous theorem, all differentials of the function f at the point P_0 coincide in the space Σ_{X,P_0} and this is the only thing that is important for us because only in this space we can use the differential to approximate the function f at the point P_0.*

Corollary 1. *Let $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, $P_0 \in X$ be a point admitting nbd ray in X and $\Sigma_{X,P_0} = \mathbb{R}^n$. If the differential of the function f exists at the point P_0, then it is unique.*

Proof. This follows from Theorem 2 and Proposition 1. □

If the differential of a function $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, exists at a point $P_0 \in X$ and is unique, we denote it by $df(P_0)$.

If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator then f is differentiable on $\mathbb{R}^n$ and $df(P) = f$ at any point $P \in \mathbb{R}^n$. In particular, the projection map $p_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, n$, is a linear operator and $dp_i(P) = p_i$ for every $P \in \mathbb{R}^n$. Usually $dp_i(P)$ is denoted by dx_i.

An affine mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ $f(P) = P_0 + A(P)$, where $P_0 \in \mathbb{R}^m$ and $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator, is differentiable on $\mathbb{R}^n$ and $df(P) = A$ for every point $P \in \mathbb{R}^n$.

Example 3. *Let $H = (1, 0) \in \mathbb{R}^2$ and $f : \overline{OH} \to \mathbb{R}$ $f(x, y) = 3x$. Since f is the restriction of the linear operator $A : \mathbb{R}^2 \to \mathbb{R}$ $A(x, y) = 3x$ on the convex set $\overline{OH}$, f is differentiable at any point of the domain and the differential at any point is equal to A. The linearization space of the function f at any point of the domain is $\Sigma = \mathbb{R}\times\{0\}$ and since it is a 1-dimensional subspace of $\mathbb{R}^2$, the differential of f is not unique. Moreover, all linear operators $\mathbb{R}^2 \to \mathbb{R}$, represented by a matrix $\begin{bmatrix} 3 & p \end{bmatrix}$, $p \in \mathbb{R}$, are all its different differentials. However, according to the previous theorem, the restriction of all these differentials on Σ is the same.*

Notice that according to the traditional definition of differentiability, this function would not be differentiable at any point in its domain. On the other hand, the function f is perfectly linearized since its graph is $\overline{OT} \subseteq \mathbb{R}^3$, $T = (1, 0, 3)$, and it would be incorrect to say that it cannot be linearized (since its graph is perfectly linearized by the part of the line OT). However, since for functions whose domain is a subset of $\mathbb{R}^2$ the graph is linearized by part of the plane, we can do that in infinitely many ways, since the entire pencil of planes passes through the line OT, so its linearization is not unique. However, if we take the set $\overline{OH} \cup \overline{OH_1}$, $H_1 = (0, 1) \in \mathbb{R}^2$ for the domain of the function f, then $\Sigma_{\overline{OH}\cup\overline{OH_1},O} = \mathbb{R}^2$, and by the previous theorem the differential of the function f at $O \in \mathbb{R}^2$ is unique, i.e., its linearization is the part of the unique plane passing through the line OT and OT_1, $T_1 = (0, 1, 0)$, $O \in \mathbb{R}^3$ (the graph of the function f is the set $\overline{OT} \cup \overline{OT_1}$, which is a part of this plane).

Corollary 2. *Let $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}$ and $x_0 \in X$ be a point admitting nbd ray in X. If a differential of the function f exists at the point x_0, then it is unique.*

Proof. Since $\Sigma_{X,x_0} = \mathbb{R}$, the statement follows from the previous corollary. □

Example 4. *Let us consider the function $f : D \to \mathbb{R}$ $f(x) = \sqrt{y - x^3}$, $D = \left\{(x, y) \in \mathbb{R}^2 \mid y \geq x^3\right\}$ from the introduction. Since for $(0, 0) \in \mathbb{R}^2$ it holds*

$$\Delta_{D,(0,0)} = \left(\left\{(x, y) \in \mathbb{R}^2 \mid x \leq 0,\ y \geq 0\right\} \cup \left\{(x, y) \in \mathbb{R}^2 \mid y \geq x^3,\ x > 0\right\}\right) \setminus \{(0, 0)\}$$

and

$$\lim_{\substack{(h_1,h_2)\to(0,0)\\(h_1,h_2)\in\Delta_{D,(0,0)}}} \frac{\sqrt{h_2 - h_1^3} - \sqrt{0} - O(h_1, h_2)}{\|(h_1, h_2)\|} = 0,$$

where $O : \mathbb{R}^2 \to \mathbb{R}$ denotes the zero operator, f is differentiable at the point 0. Moreover, it follows from $\Sigma_{D,0} = \mathbb{R}^2$ that the zero operator is the unique differential of f at the point $(0,0)$, i.e., $df(0,0) = 0$.

Definition 4. *Let $P_0 \in X \subseteq \mathbb{R}^n$ and $V \in \mathbb{R}^n \backslash \{0\}$. We say that a point P_0 **admits a neighborhood ray in** X **in the direction of** V if there exists $\lambda_0 \in \mathbb{R}^+$ such that $\overline{P_0 P_0 + \lambda_0 V} \subseteq X$.*

Proposition 2. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Every point $P_0 \in \Omega$ admits a nbd ray in Ω in the direction of all vectors $H \in \mathbb{R}^n \backslash \{0\}$ and $\Sigma_{\Omega,P_0} = \mathbb{R}^n$.*

Proof. Let $P_0 \in \Omega$ and let $H \in \mathbb{R}^n \backslash \{0\}$ be arbitrary. Since Ω is open, there exists a ball $B(P_0, r) \subseteq \Omega$, and since a ball is a convex set, $\overline{P_0 P_0 + \frac{r}{2}\frac{H}{||H||}} \subseteq B(P_0, r)$ holds. Therefore, P_0 admits nbd ray in $B(P_0, r)$ in the direction of H and then admits it in Ω. Furthermore, from $\Delta_{B(P_0,r),P_0} \subseteq \Delta_{\Omega,P_0}$ it follows that

$$\Sigma_{B(P_0,r),P_0} \subseteq \Sigma_{\Omega,P_0}$$

and $\frac{r}{2}\frac{H}{||H||} \in \Delta_{B(P_0,r),P_0}$ implies $H \in \Sigma_{B(P_0,r),P_0}$ for every $H \in \mathbb{R}^n \backslash \{0\}$. So,

$$\Sigma_{B(P_0,r),P_0} = \mathbb{R}^n$$

and then $\Sigma_{\Omega,P_0} = \mathbb{R}^n$. □

Corollary 3. *If $\Omega \subseteq \mathbb{R}^n$ is an open set and $f : \Omega \to \mathbb{R}^m$ is differentiable at a point $P_0 \in \Omega$, then the differential of the function f at the point P is unique.*

Proof. This follows from the previous proposition and corollary 1. □

Remark 2. *We have already mentioned that the new definition of differentiability (Definition 3) coincides with the well-known definition of this notion when the domain of a function is an open set. It follows that in this particular case all previously known properties of differentials hold, including the property of uniqueness. However, the new theory induced by the extended definition of differentiability provides the proof of the uniqueness of the differential of an open domain function without relying on prior general knowledge of it.*

Proposition 3. *Let $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, $Y \subseteq X$ and $P_0 \in Y$ be a point admitting nbd ray in Y. If f is differentiable at P_0 then $f|_Y$ is differentiable at P_0 and the differentials of the functions f and $f|_Y$ at P_0 coincide on Σ_{Y,P_0}.*

Proof. Since the function f is differentiable at P_0, there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that $\lim\limits_{\substack{H \to 0 \\ H \in \Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||} = 0$. Now, $\Delta_{Y,P_0} \subseteq \Delta_{X,P_0}$ implies

$$0 = \lim_{\substack{H \to 0 \\ H \in \Delta_{Y,P_0}}} \frac{f(P_0 + H) - f(P_0) - A(H)}{||H||} = \lim_{\substack{H \to 0 \\ H \in \Delta_{Y,P_0}}} \frac{f|_Y(P_0 + H) - f|_Y(P_0) - A(H)}{||H||}$$

from which it follows that the function $f|_Y$ is differentiable at P_0 and that the linear operator A is its differential at P_0. Now, by the Theorem 2, we conclude that every other differential at P_0 coincides with A on Σ_{Y,P_0}. □

The converse does not hold, i.e., if the restriction of a function $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, to a subset $Y \subseteq X$ is differentiable at a point $P_0 \in Y$, then in general the function f need not be differentiable at that point. This will be shown by the following counterexample. But we will prove that if Y is open in X, then differentiability on Y implies differentiability on X.

Example 5. *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ $f(x,y) = \sqrt{x^2+y^2}$. *Let us consider the restrictions of the function* f *to sets*

$$X_1 = \{(0,y) \mid y \in [0,\infty\rangle\} \text{ and } X_2 = \{(0,y) \mid y \in \langle -\infty, 0]\}.$$

For functions

$$f_1 := f|_{X_1} \; f_1(0,y) = y \text{ and } f_2 := f|_{X_2} \; f_2(0,y) = -y,$$

the linearization space at each point in their domains is $\Sigma = \{(0,y) \mid y \in \mathbb{R}\}$. *The function* f_1 *is the restriction of the linear operator* p_2 *to the convex set* X_1, *so* p_2 *is the differential of the function* f_1 *at any point in* X_1 *(not unique because the dimension of* Σ *is less than* 2, *but they all coincide on* Σ*). Similarly, the differential of the function* f_2 *at every point in* X_2 *is* $-p_2$. *If the function* f *were differentiable at the point* $0 \in \mathbb{R}^2$, *then, by Proposition 3 and Theorem 2, the differentials of the functions* f_1 *and* f_2 *at the point* 0 *would coincide on* Σ *which is obviously not the case.*

Theorem 3. *Let* $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, $P_0 \in X$ *and* U *be a neighborhood of the point* P_0 *in* $\mathbb{R}^n$. *If* P_0 *admits a nbd ray in* $U \cap X$ *and if* $f|_{U\cap X}$ *is differentiable at* P_0 *then* f *is also differentiable at* P_0.

Proof. Since U is a neighborhood of the point P_0 in $\mathbb{R}^n$, there exists $r \in \mathbb{R}^+$ such that $B(P_0,r) \subseteq U$ and then $B(P_0,r) \cap X \subseteq U \cap X$. Therefore, $B(0,r) \cap \Delta_{X,P_0} \subseteq \Delta_{U\cap X,P_0}$. Due to differentiability of the function $f|_{U\cap X}$, there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$0 = \lim_{\substack{H\to 0\\ H\in\Delta_{U\cap X,P_0}}} \frac{f|_{U\cap X}(P_0+H) - f|_{U\cap X}(P_0) - A(H)}{||H||} = \lim_{\substack{H\to 0\\ H\in\Delta_{U\cap X,P_0}}} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||}.$$

Since $\Delta_{U\cap X,P_0} \subseteq \Delta_{X,P_0}$ and $B(0,r) \cap \Delta_{X,P_0} \subseteq \Delta_{X,P_0}$, by Theorem 1, it follows

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0) - A(H)}{||H||} = 0,$$

which implies that f is differentiable at P_0. □

The following statement follows from the previous theorem.

Corollary 4. *Let* $f : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, $\Omega \subseteq X$ *be an open set in* $\mathbb{R}^n$ *and* $P_0 \in \Omega$. *If* $f|_\Omega$ *is differentiable at* P_0 *then* f *is also differentiable at* P_0 *and* $df(P_0) = df|_\Omega(P_0)$.

We will show now that differentiability does not imply continuity in general (which cannot be the case for a function with an open domain).

Example 6. *Let* $P_n = \left(0, \frac{1}{n}\right)$, $Q_n = \left(1, \frac{1}{n}\right) \in \mathbb{R}^2$, $n \in \mathbb{N}$, *and*

$$X = \bigcup_{n\in\mathbb{N}} \overline{P_nQ_n} \cup \overline{(0,0)(1,0)} \subseteq \mathbb{R}^2.$$

Let us consider the function $f : X \to \mathbb{R}$

$$f(P) = \begin{cases} n, & P \in \overline{P_nQ_n} \\ 0, & P \in \overline{(0,0)(1,0)}. \end{cases}$$

For the point $O = (0,0) \in X$ *it holds*

$$\Delta_{X,O} = \langle 0,1] \times \{0\} \text{ and } \Sigma_{X,O} = \mathbb{R}\times\{0\}.$$

The function f *is differentiable at the point* O *(it is differentiable at every point of its domain and the zero operator is one of its differentials), but* f *is discontinuous at all points of the line segment* $\overline{(0,0)(1,0)}$, *so it is discontinuous at* O.

In this example, the dimension of the linearization space $\Sigma_{X,O}$ is less than the dimension of the whole space $\mathbb{R}^2$. However, even if the dimension of a linearization space is equal to the dimension of the whole space $\mathbb{R}^n$, a function need not be continuous. This is shown by the following counterexample.

Example 7. *Let $\mathbb{S}^1 \subseteq \mathbb{R}^2$ be the 1-sphere and $P_0, Q, R \in \mathbb{S}^1$ three distinct points on it. Consider the union of two circular arcs and their corresponding chords $X = \widehat{P_0Q} \cup \widehat{P_0R} \cup \overline{P_0Q} \cup \overline{P_0R}$. The function $f : X \to \mathbb{R}$*

$$f(x,y) = \begin{cases} 0, \ (x,y) \in \overline{P_0Q} \cup \overline{P_0R} \\ 1, \ (x,y) \in \left(\widehat{P_0Q} \cup \widehat{P_0R}\right) \setminus \{Q, R, P_0\} \end{cases}$$

is differentiable at P_0. Namely, $\Delta_{X,P_0} = \left(\overline{O(Q-P_0)} \cup \overline{O(R-P_0)}\right) \setminus \{O\}$ and

$$\lim_{\substack{H \to 0 \\ H \in \Delta_{X,P_0}}} \frac{f(P_0+H) - f(P_0)}{||H||} = 0.$$

Since $\Sigma_{X,P_0} = \mathbb{R}^2$, the differential is unique and $df(P_0)$ is the zero operator. But the function f is discontinuous at P_0 because $\lim\limits_{\substack{P \to P_0 \\ P \in \overline{P_0Q} \cup \overline{P_0R}}} f(P) = 0$ and $\lim\limits_{\substack{P \to P_0 \\ P \in \widehat{P_0Q} \cup \widehat{P_0R}}} f(P) = 1$ hold.

To ensure that differentiability of a function at a point implies continuity at that point, we need an additional condition, which is introduced in the following definition.

Definition 5. *Let $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point admitting nbd ray in X. A neighborhood U of the point P_0 in X is said to be* ***raylike neighborhood of the point** P_0 **in** X provided $\overline{P_0P} \subseteq U$ holds for every $P \in U$. If there exists at least one raylike nbd in X of the point P_0, we say that the point P_0* ***admits raylike nbd in** X.*

It is easy to see that every point of a non-trivial convex set admits raylike nbd in that set, and then every point of the open set $\Omega \subseteq \mathbb{R}^n$ admits raylike nbd in Ω.

Theorem 4. *Let a point $P_0 \in X \subseteq \mathbb{R}^n$ admits raylike nbd in X. If $f : X \to \mathbb{R}^m$ is differentiable at P_0 then it is also continuous at P_0.*

Proof. Let U be a raylike nbd of the point P_0 in X. Then $U - P_0 = \{P - P_0 : P \in U\}$ is a neighborhood of the point $O \in \mathbb{R}^n$ in $\Delta_{X,P_0} \cup \{O\}$. To prove that f is continuous at P_0 it suffices to prove that $f|_U$ is continuous at P_0, i.e., that $\lim\limits_{\substack{H \to 0 \\ H \in U - P_0}} f(P_0 + H) = f(P_0)$. By the assumed differentiability, there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(P_0 + H) - f(P_0) = A(H) + r(H) \tag{4}$$

where $r : \Delta_{X,P_0} \to \mathbb{R}^m$ is an error function with the property $\lim\limits_{H \to 0} \frac{r(H)}{||H||} = 0$. Then $\lim\limits_{H \to 0} r(H) = 0$. Since $U - P_0 \subseteq \Delta_{X,P_0} \cup \{0\}$,

$$\lim_{\substack{H \to 0 \\ H \in U - P_0}} r(H) = \lim_{H \to 0} r(H) = 0.$$

Every linear operator operating between finite dimensional vectorial spaces is continuous, therefore

$$0 = A(0) = \lim_{H \to 0} A(H) = \lim_{\substack{H \to 0 \\ H \in U - P_0}} A(H).$$

Hence, by (4), it follows

$$\lim_{\substack{H \to 0 \\ H \in U - P_0}} f(P_0 + H) = f(P_0).$$

□

Obviously, if $\Omega \subseteq \mathbb{R}^n$ is an open set, $f : \Omega \to \mathbb{R}^m$, $P_0 \in \Omega$ and f is differentiable at P_0, then f is also continuous at P_0. Thus, if the domain of a function is an open set, differentiability implies continuity. The same is true for any convex domain.

4. Partial and Directional Derivatives

Definition 6. *Let $X \subseteq \mathbb{R}^n$, $n \geq 2$, $V \in \mathbb{R}^n \backslash \{0\}$ and $P_0 \in X$ be a point admitting nbd ray in X in the direction of the vector V. The set*

$$\Delta_{V(X,P_0)} := \{tV \mid t \in \mathbb{R}\} \cap \Delta_{X,P_0}$$

*is called the **set of linear contributions at** P_0 **in the direction of** V **into** X**.***

Let us denote

$$\begin{aligned} h_1 &:= \inf\{t \in \mathbb{R} \mid tV \in \Delta_{X,P_0}\}, \\ h_2 &:= \sup\{t \in \mathbb{R} \mid tV \in \Delta_{X,P_0}\}, \end{aligned} \tag{5}$$

where $h_1 = -\infty$ ($h_2 = \infty$), provided the set $\{t \in \mathbb{R} \mid tV \in \Delta_{X,P_0}\}$ is not bounded from below (above). Let X_{V,P_0} denotes the largest convex subset of the set $X \cap \overline{P_0P_0 + V}$ containing the point P_0 ($\overline{P_0P_0 + V}$ denotes the line passing through the points P_0 and $P_0 + V$). If $h_1, h_2 \in \mathbb{R}$ then $X_{V,P_0} = X \cap \overline{(P_0 + h_1V)(P_0 + h_2V)}$ (it is the line segment with or without boundary points), otherwise X_{V,P_0} is the half-line or the line $\overline{P_0P_0 + V}$.

Definition 7. *Let a point $P_0 \in X \subseteq \mathbb{R}^n$ admit nbd ray in X in the direction of $V \in \mathbb{R}^n \backslash \{0\}$. We say that a function $f : X \to \mathbb{R}$ has the **derivative at** P_0 **in the direction of** V if there exists*

$$\lim_{\substack{h \to 0 \\ hV \in \Delta_{V(X,P_0)}}} \frac{f(P_0 + hV) - f(P_0)}{h}.$$

*This limit, if it exists, is denoted by $\partial_V f(P_0)$ and is called the **derivative of** f **at** P_0 **in the direction of** V.*

*The derivative at P_0 in the direction of e_i (e_i is the i-th basis vector of the standard ordered basis for $\mathbb{R}^n$) is called **i-th partial derivative of** f **at** P_0 and is denoted by $\partial_i f(P_0)$.*

Notice that, for $P_0 = (x_1^0, \ldots, x_n^0)$,

$$\partial_i f(P_0) = \lim_{\substack{h \to 0 \\ h \in \langle h_1, h_2 \rangle \backslash \{0\}}} \frac{f(x_1^0, \ldots, x_{i-1}^0, x_i^0 + h, x_{i+1}^0, \ldots, x_n^0) - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0)}{h}$$

holds. If $h_1 < 0$, by Theorem 1, it follows

$$\partial_i f(P_0) = \lim_{h \to 0} \frac{f(x_1^0, \ldots, x_{i-1}^0, x_i^0 + h, x_{i+1}^0, \ldots, x_n^0) - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0)}{h}.$$

Theorem 5. *Let $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$, $V \in \mathbb{R}^n \backslash \{0\}$ and $P_0 \in X$ be a point admitting nbd ray in X in the direction of V. The function f has the derivative at P_0 in the direction of V if and only if its restriction $f|_{X_{V,P_0}}$ is differentiable at P_0. The value of each differential of the function $f|_{X_{V,P_0}}$ at P_0, at V is $\partial_V f(P_0)$.*

Proof. Suppose f has the derivative at P_0 in the direction of V. If $\{V, v_1, \ldots, v_{n-1}\}$ is some basis for $\mathbb{R}^n$, then any linear operator $A : \mathbb{R}^n \to \mathbb{R}$, given by

$$A(V) = \partial_V f(P_0),\ A(v_i) = a_i$$

for some real numbers a_i, $i = 1, \ldots, n-1$, is a differential of the function $f|_{X_{V,P_0}}$. Namely,

$$\lim_{\substack{h\to 0^+\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0) - h\partial_V f(P_0)}{|h|}$$

$$= \lim_{\substack{h\to 0^+\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f(P_0+hV) - f(P_0) - \partial_V f(P_0)h}{h} =$$

$$\lim_{\substack{h\to 0^+\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f(P_0+hV) - f(P_0)}{h} - \partial_V f(P_0) =$$

$$\lim_{\substack{h\to 0\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f(P_0+hV) - f(P_0)}{h} - \partial_V f(P_0) = 0.$$

Similarly,

$$\lim_{\substack{h\to 0^-\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0) - h\partial_V f(P_0)}{|h|} = 0$$

which implies

$$\lim_{\substack{h\to 0\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0) - h\partial_V f(P_0)}{|h|} = 0.$$

Therefore,

$$\lim_{\substack{H\to 0\\ H\in\Delta_{V(X,P_0)}}} \frac{f|_{X_{V,P_0}}(P_0+H) - f|_{X_{V,P_0}}(P_0) - A(H)}{||H||} =$$

$$= \lim_{\substack{h\to 0\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0) - A(hV)}{|h|||V||} =$$

$$\frac{1}{||V||} \lim_{\substack{h\to 0\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0) - h\partial_V f(P_0)}{|h|} = 0.$$

Hence, $f|_{X_{V,P_0}}$ is differentiable at the point P_0.

Now assume that $f|_{X_{V,P_0}}$ is differentiable at the point P_0, i.e., that there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}$ such that

$$\lim_{\substack{H\to 0\\ H\in\Delta_{V(X,P_0)}}} \frac{f|_{X_{V,P_0}}(P_0+H) - f|_{X_{V,P_0}}(P_0) - A(H)}{||H||} = 0.$$

Then

$$0 = \lim_{\substack{h\to 0\\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0) - hA(V)}{|h|||V||}$$

from which follows

$$
\begin{aligned}
A(V) &= \lim_{\substack{h\to 0 \\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f|_{X_{V,P_0}}(P_0+hV) - f|_{X_{V,P_0}}(P_0)}{h} \\
&= \lim_{\substack{h\to 0 \\ h\in\langle h_1,h_2\rangle\setminus\{0\}}} \frac{f(P_0+hV) - f(P_0)}{h} = \partial_V f(P_0).
\end{aligned}
$$

Thus, f has the derivative at P_0 in the direction of V and it is equal to $A(V)$. □

Corollary 5. *Let $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$, $V \in \mathbb{R}^n\setminus\{0\}$ and $P_0 \in X$ be a point admitting nbd ray in X in the direction of V and let f be differentiable at P_0. Then f has the derivative at P_0 in the direction of V and the value of each differential of the function f at P_0, at V is equal to $\partial_V f(P_0)$. If $\Sigma_{X,P_0} = \mathbb{R}^n$ then $\partial_V f(P_0) = df(P_0)(V)$.*

Proof. The statement follows from the previous theorem and Proposition 3. □

The converse of this corollary does not hold, i.e., a function f at a point P_0 can have directional derivatives and need not be differentiable at P_0, as shown in the following example.

Example 8. *The function*

$$
f : \mathbb{R}^2 \to \mathbb{R} \quad f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}
$$

has both partial derivatives at the point $(0,0)$ but f is not continuous at this point, so by Theorem 4 f is not differentiable at $(0,0)$. It is interesting to note that this function has the derivative at the point $(0,0)$ in the direction of any $V = (v_1, v_2) \in \mathbb{R}^2\setminus\{0\}$, and

$$
\partial_V f(0,0) = \lim_{h\to 0} \frac{hv_1 hv_2}{h^2 v_1^2 + h^2 v_2^2} = \frac{v_1 v_2}{v_1^2 + v_2^2}.
$$

5. Differentiable Functions

5.1. Properties of Differentials

We will now prove the following important results which hold for differentiable functions.

Proposition 4. *Let $X \subseteq \mathbb{R}^n$, $P_0 \in X$ be a point admitting nbd ray in X, $f, g : X \to \mathbb{R}^m$ be differentiable functions at P_0, and $A, B : \mathbb{R}^n \to \mathbb{R}^m$ be differentials of the functions f and g at P_0, respectively. Then the function $\lambda f + \mu g : X \to \mathbb{R}^m$ is differentiable at P_0 for any $\lambda, \mu \in \mathbb{R}$ and $\lambda A + \mu B$ is its differential at P_0.*

Proof. By the differentiability of the functions f and g at P_0 it holds

$$
f(P_0 + H) - f(P_0) = A(H) + r_1(H)
$$

and

$$
g(P_0 + H) - g(P_0) = B(H) + r_2(H),
$$

for every $H \in \Delta_{X,P_0} \subseteq \mathbb{R}^n$, where $r_1, r_2 : \Delta_{X,P_0} \to \mathbb{R}^m$ are the functions with the property

$$
\lim_{H\to 0} \frac{r_1(H)}{||H||} = 0 = \lim_{H\to 0} \frac{r_2(H)}{||H||}.
$$

Now, it follows

$$
(\lambda f + \mu g)(P_0 + H) - (\lambda f + \mu g)(P_0) = \lambda(f(P_0 + H) - f(P_0)) + \mu(g(P_0 + H) - g(P_0))
$$

$$= \lambda(A(H) + r_1(H)) + \mu(B(H) + r_2(H)) =$$
$$= \lambda A(H) + \mu B(H) + \lambda r_1(H) + \mu r_2(H) =$$
$$= (\lambda A + \mu B)(H) + \lambda r_1(H) + \mu r_2(H),$$

for every $H \in \Delta_{X,P_0}$. Since for the function $r(H) = \lambda r_1(H) + \mu r_2(H)$ it holds $\lim\limits_{H\to 0} \frac{r(H)}{||H||} = 0$, the function $\lambda f + \mu g$ is differentiable at P_0 and the linear operator $\lambda A + \mu B$ is its differential at this point. □

Proposition 5. *Let $X \subseteq \mathbb{R}^n$, $P_0 \in X$ be a point admitting nbd ray in X, $\alpha : X \to \mathbb{R}$ and $f : X \to \mathbb{R}^m$ be differentiable functions at P_0, and $A : \mathbb{R}^n \to \mathbb{R}$ and $B : \mathbb{R}^n \to \mathbb{R}^m$ be differentials of the functions α and f at P_0, respectively. Then the function $\alpha \cdot f : X \to \mathbb{R}^m$ $(\alpha \cdot f)(P) = \alpha(P)f(P)$ is differentiable at P_0 and $\alpha(P_0)B + f(P_0)A : \mathbb{R}^n \to \mathbb{R}^m$ $(\alpha(P_0)B + f(P_0)A)(H) = \alpha(P_0)B(H) + A(H)f(P_0)$ is its differential at P_0.*

Proof. By the differentiability of the functions α and f at P_0, it holds

$$\alpha(P_0 + H) - \alpha(P_0) = A(H) + r_1(H)$$

and

$$f(P_0 + H) - f(P_0) = B(H) + r_2(H),$$

for every $H \in \Delta_{X,P_0} \subseteq \mathbb{R}^n$, where $r_1 : \Delta_{X,P_0} \to \mathbb{R}$ and $r_2 : \Delta_{X,P_0} \to \mathbb{R}^m$ are the functions with the properties

$$\lim_{H\to 0} \frac{r_1(H)}{||H||} = 0 \text{ and } \lim_{H\to 0} \frac{r_2(H)}{||H||} = 0.$$

Hence $\lim\limits_{H\to 0} r_1(H) = 0$ and $\lim\limits_{H\to 0} r_2(H) = 0$. Now we infer that

$$(\alpha \cdot f)(P_0 + H) - (\alpha \cdot f)(P_0) = \alpha(P_0)B(H) + f(P_0)A(H)+$$
$$\alpha(P_0)r_2(H) + r_1(H)f(P_0) + (A(H) + r_1(H))(B(H) + r_2(H)),$$

holds, for every $H \in \Delta_{X,P_0}$. Therefore, it is sufficient to prove that $\lim\limits_{H\to 0} \frac{r(H)}{||H||} = 0$ for the function $r : \Delta_{X,P_0} \to \mathbb{R}^m$

$$r(H) = \alpha(P_0)r_2(H) + r_1(H)f(P_0) + (A(H) + r_1(H))(B(H) + r_2(H)).$$

From the properties of the functions r_1 and r_2 it follows that

$$\lim_{H\to 0} \frac{\alpha(P_0)r_2(H)}{||H||} = 0 = \lim_{H\to 0} \frac{r_1(H)f(P_0)}{||H||}.$$

Furthermore, by the boundedness of a linear operator, there exists $\lambda > 0$ such that $||A(H)|| \leq \lambda||H||$, for every $H \in \mathbb{R}^n$ [11,12]. Therefore, since the linear operator is continuos and its value at zero is equal to zero, by the properties of the errors functions r_1 and r_2, it follows

$$\begin{aligned} 0 &\leq \left\| \lim_{H\to 0} \frac{(A(H) + r_1(H))(B(H) + r_2(H))}{||H||} \right\| \\ &= \left\| \lim_{H\to 0} \left(\frac{A(H)}{||H||} + \frac{r_1(H)}{||H||} \right) (B(H) + r_2(H)) \right\| \\ &\leq \left\| \lim_{H\to 0} \left(\lambda + \frac{r_1(H)}{||H||} \right) (B(H) + r_2(H)) \right\| = 0. \end{aligned}$$

This implies

$$\lim_{H\to 0}\frac{(A(H)+r_1(H))(B(H)+r_2(H))}{||H||}=0.$$

Therefore, the function αf is differentiable at P_0 and the linear operator $\alpha(P_0)B+f(P_0)A$ is its differential at this point. □

Proposition 6. *Let $X\subseteq\mathbb{R}^n$, $P_0\in X$ be a point admitting nbd ray in X, $f:X\to\mathbb{R}$ be a differentiable function at P_0, and $A:\mathbb{R}^n\to\mathbb{R}$ be the differential of the function f. If f is continuous at P_0 and $f(P_0)\neq 0$, then the function $\frac{1}{f}$ is differentiable at P_0 and $-\frac{1}{(f(P_0))^2}A:\mathbb{R}^n\to\mathbb{R}$ is its differential at P_0.*

Proof. By the continuity of the function f at P_0, there exists an open neighborhood O of the point P_0 in X such that $f(P)\neq 0$ for every $P\in O$. Since $O=U\cap X$ for some neighborhood U of P_0 in $\mathbb{R}^n$, it suffices to prove that the restriction function $\frac{1}{f}|_O$ is differentiable at P_0 (Theorem 3). By the assumption,

$$f(P_0+H)-f(P_0)=A(H)+r(H)$$

holds for every $H\in\Delta_{O,P_0}\subseteq\mathbb{R}^n$, where $r:\Delta_{O,P_0}\to\mathbb{R}$ is the function with the property

$$\lim_{H\to 0}\frac{r(H)}{||H||}=0.$$

It follows

$$\frac{1}{f(P_0+H)}-\frac{1}{f(P_0)}=-\frac{f(P_0+H)-f(P_0)}{f(P_0)f(P_0+H)}=-\frac{A(H)+r(H)}{f(P_0)f(P_0+H)}=$$

$$=-\frac{A(H)}{(f(P_0))^2}+\frac{A(H)(f(P_0+H)-f(P_0))-f(P_0)r(H)}{(f(P_0))^2f(P_0+H)},$$

for every $H\in\Delta_{O,P_0}$. Thus, it is sufficient to prove the equality $\lim\limits_{H\to 0}\frac{r_1(H)}{||H||}=0$ for the function $r_1:\Delta_{O,P_0}\to\mathbb{R}$

$$r_1(H)=\frac{A(H)(f(P_0+H)-f(P_0))-f(P_0)r(H)}{(f(P_0))^2f(P_0+H)}.$$

Notice that

$$\lim_{H\to 0}\left(\frac{r(H)}{||H||}\frac{f(P_0)}{(f(P_0))^2f(P_0+H)}\right)=0$$

holds. By the boundedness of a linear operator, there exists $\lambda>0$ such that $||A(H)||\leq\lambda||H||$, for every $H\in\mathbb{R}^n$ [11,12]. It implies

$$0\leq\left|\left|\frac{A(H)(f(P_0+H)-f(P_0))}{||H||(f(P_0))^2f(P_0+H)}\right|\right|\leq\lambda\left|\left|\frac{f(P_0+H)-f(P_0)}{(f(P_0))^2f(P_0+H)}\right|\right|.$$

Since f is continuous at P_0 it holds

$$\lim_{H\to 0}\left|\left|\frac{f(P_0+H)-f(P_0)}{(f(P_0))^2f(P_0+H)}\right|\right|=0.$$

Now, we infer

$$\lim_{H\to 0}\left|\left|\frac{A(H)(f(P_0+H)-f(P_0))}{||H||(f(P_0))^2f(P_0+H)}\right|\right|=0$$

and finally

$$\lim_{H\to 0} \frac{A(H)(f(P_0+H)-f(P_0))}{||H||(f(P_0))^2 f(P_0+H)} = 0.$$

This proves that the function $-\frac{A}{(f(P_0))^2}$ is the differential of the function $\frac{1}{f}$ at the point P_0. $\square$

Corollary 6. *Let $X \subseteq \mathbb{R}^n$, $P_0 \in X$ be a point admitting nbd ray in X, $\alpha : X \to \mathbb{R}$ and $f : X \to \mathbb{R}^m$ be differentiable functions at P_0, and $A : \mathbb{R}^n \to \mathbb{R}$ and $B : \mathbb{R}^n \to \mathbb{R}^m$ be differentials of the functions α and f at P_0, respectively. If α is continuous at P_0 and $\alpha(P_0) \neq 0$, then the function $\frac{1}{\alpha} f$ is differentiable at P_0 and*

$$\frac{\alpha(P_0)B - f(P_0)A}{(\alpha(P_0))^2} : \mathbb{R}^n \to \mathbb{R}^m$$

Proof. This follows from the two previous propositions. $\square$

Let us now prove that the composition of differentiable functions is differentiable.

Theorem 6. *Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$, $f : X \to \mathbb{R}^m$, $f(X) \subseteq Y$, and $g : Y \to \mathbb{R}^p$. Let $P_0 \in X$ be a point admitting raylike nbd in X, and let $Q_0 = f(P_0)$ be the point admitting raylike nbd in Y. If f is differentiable at P_0 and g is differentiable at Q_0, then the composition $g \circ f : X \to \mathbb{R}^p$ is differentiable at P_0 and $B \circ A$ is its differential at the point P_0, where $A : \mathbb{R}^n \to \mathbb{R}^m$ and $B : \mathbb{R}^m \to \mathbb{R}^p$ are differentials at the points P_0 and Q_0 of the functions f and g, respectively.*

Proof. Since f is differentiable at the point P_0, there exists a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(P) - f(P_0) = A(P - P_0) + r_1(P - P_0)$$

for each $P \in X$ such that $P - P_0 \in \Delta_{X,P_0}$, and for each $r_1 : \Delta_{X,P_0} \to \mathbb{R}^m$ such that

$$\lim_{H\to 0} \frac{r_1(H)}{||H||} = \lim_{P\to P_0} \frac{r_1(P-P_0)}{||P-P_0||} = 0. \tag{6}$$

Similarly, there exists a linear operator $B : \mathbb{R}^m \to \mathbb{R}^p$ such that

$$g(Q) - g(Q_0) = B(Q - Q_0) + r_2(Q - Q_0)$$

for each $Q \in Y$ such that $Q - Q_0 \in \Delta_{Y,Q_0}$, and for each $r_2 : \Delta_{Y,Q_0} \to \mathbb{R}^p$ such that

$$\lim_{H\to 0} \frac{r_2(H)}{||H||} = \lim_{Q\to Q_0} \frac{r_2(Q-Q_0)}{||Q-Q_0||} = 0.$$

By the assumption there exists a raylike nbd O of P_0 in X. Since $O = U \cap X$ for some neighborhood U of the point P_0 in $\mathbb{R}^n$, by Theorem 3, it suffices to prove that $g \circ f|_O$ is differentiable at P_0. Now,

$$\begin{aligned}(g\circ f)(P) - (g\circ f)(P_0) &= B(f(P) - f(P_0)) + r_2(f(P) - f(P_0)) \\ &= B(A(P-P_0) + r_1(P-P_0)) + r_2(f(P) - f(P_0)) \\ &= B\circ A(P-P_0) + B(r_1(P-P_0)) + r_2(f(P) - f(P_0))\end{aligned}$$

for every $P \in O$. Therefore, we have to prove that

$$\lim_{H\to 0} \frac{r(H)}{||H||} = \lim_{P\to P_0} \frac{r(P-P_0)}{||P-P_0||} = 0 \tag{7}$$

for the function $r : (O - P_0)\backslash\{0\} \to \mathbb{R}^p$ defined by

$$r(H) = B(r_1(H)) + r_2(f(P_0+H) - f(P_0)).$$

By the continuity of a linear operator, it follows that

$$\lim_{P\to P_0}\frac{B(r_1(P-P_0))}{||P-P_0||}=\lim_{P\to P_0}B\left(\frac{r_1(P-P_0)}{||P-P_0||}\right)=B\left(\lim_{P\to P_0}\frac{r_1(P-P_0)}{||P-P_0||}\right)=B(0)=0.$$

It remains to prove

$$\lim_{P\to P_0}\frac{r_2(f(P)-f(P_0))}{||P-P_0||}=0. \tag{8}$$

Since a linear operator is linearly bounded [11,12], there exists $\lambda>0$ such that

$$||A(P-P_0)||\leq\lambda||P-P_0||$$

for every $P\in\mathbb{R}^n$. Let $\varepsilon>0$. By the equality (7) we infer that there exists $\delta'>0$ such that $B(Q_0,\delta')\cap Y$ is a raylike nbd of Q_0 in Y and

$$||r_2(Q-Q_0)||\leq\frac{\varepsilon}{2\lambda}||Q-Q_0||$$

holds for every $Q\in B(Q_0,\delta')\cap Y$. Furthermore, by the condition (6) and the continuity of the function f at the point P_0, there exists $\delta>0$ such that $B(P_0,\delta)\cap X\subseteq O$, $f(B(P_0,\delta)\cap X)\subseteq B(Q_0,\delta')\cap Y$ and

$$\frac{||r_1(P-P_0)||}{||P-P_0||}<\lambda$$

for every $P\in(B(P_0,\delta)\backslash\{P_0\})\cap X$. Hence,

$$d\left(\frac{r_2(f(P)-f(P_0))}{||P-P_0||},0\right)=\frac{||r_2(f(P)-f(P_0))||}{||P-P_0||}\leq$$

$$\leq\frac{\frac{\varepsilon}{2\lambda}||f(P)-f(P_0)||}{||P-P_0||}\leq\frac{\varepsilon}{2\lambda}\frac{||A(P-P_0)||+||r_1(P-P_0)||}{||P-P_0||}\leq\frac{\varepsilon}{2\lambda}(\lambda+\lambda)=\varepsilon$$

for every $P\in(B(P_0,\delta)\backslash\{P_0\})\cap X$. Thus we have proved (8). □

Corollary 7. *Let $\Omega_1\subseteq\mathbb{R}^n$ and $\Omega_2\subseteq\mathbb{R}^m$ be open sets, and $f:\Omega_1\to\mathbb{R}^m$and $g:\Omega_2\to\mathbb{R}^p$ such that $f(\Omega_1)\subseteq\Omega_2$. If f is differentiable at $P_0\in\Omega_1$ and g is differentiable at $f(P_0)\in\Omega_2$, then the composition $g\circ f:\Omega_1\to\mathbb{R}^p$ is differentiable at P_0 and $d(g\circ f)(P_0)=dg(f(P_0))\circ df(P_0)$.*

Proof. Since every point of an open set admits raylike nbd in that set, the statement follows from the previous theorem, the Proposition 2 and the Corollary 1. □

Proposition 7. *Let $X\subseteq\mathbb{R}^n$, $Y\subseteq\mathbb{R}^m$ and $f:X\to Y$ be a bijection. Let the points $P_0\in X$ and $Q_0=f(P_0)\in Y$ admit a raylike nbd in X and Y, respectively, and let $\Sigma_{X,P_0}=\mathbb{R}^n$ and $\Sigma_{Y,Q_0}=\mathbb{R}^m$. If the function f is differentiable at P_0 and if its inverse $f^{-1}:Y\to X$ is differentiable at $Q_0\in Y$, then $m=n$, the differential $df(P_0):\mathbb{R}^n\to\mathbb{R}^n$ is a regular operator and $d(f^{-1})(Q_0)=(df(P_0))^{-1}$.*

Proof. Since $f^{-1}\circ f=1_X$ and $f\circ f^{-1}=1_Y$, by the previous theorem, it follows that

$$d\left(f^{-1}\right)(Q_0)\circ df(P_0)=d(1_X)(P_0)\text{ and }df(P_0)\circ d\left(f^{-1}\right)(Q_0)=d(1_Y)(Q_0).$$

Furthermore, an identity is a linear operator, so that the differentials of all restrictions of the identity are equal to that identity. Now from

$$d\left(f^{-1}\right)(Q_0)\circ df(P_0)=1_{\mathbb{R}^n}\text{ and }df(P_0)\circ d\left(f^{-1}\right)(Q_0)=1_{\mathbb{R}^m}$$

it follows that the linear operators $df(P_0) : \mathbb{R}^n \to \mathbb{R}^m$ and $d(f^{-1})(Q_0) : \mathbb{R}^m \to \mathbb{R}^n$ are bijections, i.e., isomorphisms, so that $n = m$ [10]. Then $d(f^{-1})(Q_0) = (df(P_0))^{-1}$. $\square$

5.2. Differentiability of Real Functions of One Variable

Let $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}$, and $x_0 \in X$ be a point admitting nbd ray in X. Then $\Sigma_{X,x_0} = \mathbb{R}$, and if f is differentiable at x_0, then the differential of f at x_0 is unique and

$$\lim_{\substack{h\to 0\\ h\in\Delta_{X,x_0}}} \frac{f(x_0+h) - f(x_0) - df(x_0)(h)}{|h|} = 0.$$

Moreover, since $df(x_0)$ is a linear operator, there exists a unique $a \in \mathbb{R}$ such that

$$\lim_{\substack{h\to 0\\ h\in\Delta_{X,x_0}}} \frac{f(x_0+h) - f(x_0) - ah}{|h|} = 0.$$

Therefore, it follows that

$$a = \lim_{\substack{h\to 0\\ h\in\Delta_{X,x_0}}} \frac{f(x_0+h) - f(x_0)}{h}.$$

This limit is denoted by $f'(x_0)$ and is called the **derivative of the function f at the point x_0**.

In general, if $f'(x_0)$ exists for a function $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}$, at a point $x_0 \in X$, then the function f is said to be **derivable at** x_0. If the function f is derivable at every point of X, then we say that f **is derivable**. If X is an open set then

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

Notice that differentiability of the function f at x_0 implies derivability and

$$df(x_0)(h) = f'(x_0)h.$$

Also, derivability of the function f at x_0 (existence of the number $f'(x_0)$) implies differentiability at x_0, i.e., it holds following theorem:

Theorem 7. *Let $f : X \to \mathbb{R}$ and $x_0 \in X \subseteq \mathbb{R}$ be a point admitting nbd ray in X. The function f is differentiable at x_0 if and only if it is derivable at x_0.*

Generalizing the notion of derivability (differentiability) of a real function of a real variable to points admitting an nbd ray in the domain of this function and not belonging to the interior of this domain allows the phenomenon of derivable (differentiable) but discontinuous functions at a given point, as shown in the following example.

Example 9. *Let $X = \bigcup_{n\in\mathbb{N}} \left[-\frac{1}{2n}, -\frac{1}{2n+1}\right] \cup [0,1]$ and $f : X \to \mathbb{R}$ be the function defined by*

$$f(x) = \begin{cases} n, & x \in \left[-\frac{1}{2n}, -\frac{1}{2n+1}\right] \\ 0, & x \in [0,1] \end{cases}.$$

The function f is derivable (differentiable) at every point of its domain X, and $f'(x) = 0$ for every $x \in X$, but f is discontinuous at 0.

However, if a point $x_0 \in X \subseteq \mathbb{R}$ of a function $f : X \to \mathbb{R}$ admits raylike nbd in X and if f is derivable at x_0, then f is continuous at x_0, which follows from Theorem 4 and the previous theorem.

Let us now consider the question of the tangent to the graph of a function $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}$. Since the number $\frac{f(x_0+h)-f(x_0)}{h}$ is the slope of the secant passing through the points $(x_0, f(x_0))$ and $(x_0+h, f(x_0+h))$, the number $\lim\limits_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0)$, if it exists, is the slope of the tangent line (the limiting position of secant) to the graph of the function f at the point $(x_0, f(x_0))$. Hence, we distinguish two cases:

(a) If f is derivable at the point x_0, then $f'(x_0)$ exists and we define **the tangent to the graph of the function** f **at the point** $(x_0, f(x_0))$ as the line passing through the point $(x_0, f(x_0))$ whose the slope is $f'(x_0)$, so that its equation is

$$y - f(x_0) = f'(x_0)(x - x_0).$$

(b) If the function f at x_0 is not derivable, but is continuous and

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = \infty \text{ or } \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = -\infty,$$

then the line $x = x_0$ is the limiting position of the secant and we call it **tangent at the point** $(x_0, f(x_0))$ **to the graph of the function** f. For example, since $\lim\limits_{h\to 0} \frac{\sqrt[3]{h}-0}{h} = \infty$, the line $x = 0$ is tangent to the graph of the function $x \mapsto \sqrt[3]{x}$ at the point $(0, 0)$.

Likewise, the line $x = x_0$ is called **the tangent to the graph of the function** f **at** $(x_0, f(x_0))$ provided that

$$\lim_{h\to 0^+} \frac{f(x_0+h)-f(x_0)}{h} = \infty(-\infty) \text{ and } \lim_{h\to 0^-} \frac{f(x_0+h)-f(x_0)}{h} = -\infty(\infty).$$

Thus, for the function $x \mapsto \sqrt[3]{x^2}$, since $\lim\limits_{h\to 0^+} \frac{\sqrt[3]{h^2}-0}{h} = \infty$ and $\lim\limits_{h\to 0^-} \frac{\sqrt[3]{h^2}-0}{h} = -\infty$, it follows that the line $x = 0$ is the tangent to the graph of this function at the point $(0, 0)$.

5.3. Differentiability of Functions of Several Real Variables

Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$ be differentiable at a point $P_0 \in X$ that admits nbd ray in X in the direction of k linearly independent vectors $V_1, \ldots, V_k$, $k \le n$, and let these vectors form the basis of the space Σ_{X,P_0}. The differential $A : \mathbb{R}^n \to \mathbb{R}$ of the function f at P_0 is uniquely determined on Σ_{X,P_0} (Theorem 2) by values $\partial_{V_i} f(P_0) = A(V_i)$, $i = 1, \ldots, k$ (Corollary 5), so that, for every $H \in \Sigma_{X,P_0}$, there exist numbers h_i, $i = 1, \ldots, k$, such that $H = h_1 V_1 + \cdots + h_k V_k$ and $A(H) = \sum\limits_{i=1}^{k} \partial_{V_i} f(P_0) h_i$.

If $k = n$, then the vectors $V_1, \ldots, V_n$ form the basis of the space $\Sigma_{X,P_0} = \mathbb{R}^n$, therefore, the linear operator $df(P_0)$ is unique (Corollary 1) and is uniquely determined by the values $\partial_{V_i} f(P_0) = df(P_0)(V_i)$, $i = 1, \ldots, n$. For each $H \in \mathbb{R}^n$ there exist numbers h_i, $i = 1, \ldots, n$, such that $H = h_1 V_1 + \cdots + h_n V_n$ and $df(P_0)(H) = \sum\limits_{i=1}^{n} \partial_{V_i} f(P_0) h_i$. This proves the following theorem.

Theorem 8. *Let $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point admitting nbd ray in X in the direction of n linearly independent vectors $V_1, \ldots, V_n$. If $f : X \to \mathbb{R}$ is differentiable at P_0, then f has the derivatives at P_0 in the direction of $V_i, \ldots, V_n$, and for any choice of vector $H = h_1 V_1 + \cdots + h_n V_n \in \mathbb{R}^n$ holds*

$$df(P_0)(H) = \sum_{i=1}^{n} \partial_{V_i} f(P_0) h_i.$$

If $\Omega \subseteq \mathbb{R}^n$ is an open set and a function $f : \Omega \to \mathbb{R}$ is differentiable at $P_0 \in \Omega$ then all of its partial derivatives at P_0 exist and

$$df(P_0)(H) = \sum_{i=1}^{n} \partial_i f(P_0) h_i = (\operatorname{grad} f(P_0)|H)$$

for every $H = (h_1, \ldots, h_n) \in \mathbb{R}^n$.

Notice that the existence of the differential of a function $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$, at $P_0 \in X$ now no longer depends on the existence of the partial derivatives at this point (which was the case so far). If the function f is differentiable at P_0 and if the point P_0 does not admit nbd ray in the direction of e_i for some $i \in \{1, \ldots, n\}$, but admits nbd rays in X in the direction of n linearly independent vectors $V_1, \ldots, V_n$, then the role of the partial derivatives of f at P_0 is taken over in the differential $df(P_0)$ by derivatives in the direction of $V_1, \ldots, V_n$ and $df(P_0)$ is represented by the matrix

$$\begin{bmatrix} \partial_{V_1} f(P_0) & \cdots & \partial_{V_n} f(P_0) \end{bmatrix}$$

in the pair of ordered bases $(V_1, \ldots, V_n)$ and (e_1), $e_1 = 1$. The converse of the previous theorem is not true. Namely, at some point derivatives of f in the direction of all vectors in $\mathbb{R}^n \backslash \{0\}$ can exist and the function f need not be differentiable at that point, as we have shown in the Example 8.

5.4. Differentiability of Vector Functions

The question of differentiability of a vector function $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$ is equivalent to the question of differentiability of its coordinate functions $f_i = p_i \circ f$, $i = 1, \ldots, m$.

Theorem 9. *Let $X \subseteq \mathbb{R}^n$, $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$, and $P_0 \in X$ be a point admitting nbd ray in X. The function f is differentiable at P_0 if and only if f_i is differentiable at P_0 for every $i = 1, \ldots, m$. A linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ is the differential of the function f at P_0 if and only if $p_i \circ A : \mathbb{R}^n \to \mathbb{R}$ is the differential of the function f_i at P_0, for $i = 1, \ldots, m$.*

Proof. The differential of the function f at P_0 is a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ with the property

$$f(P_0 + H) - f(P_0) = A(H) + r(H)$$

for every $H \in \Delta_{X,P_0}$, where $r = (r_1, \ldots, r_m) : \Delta_{X,P_0} \to \mathbb{R}^m$ is the function such that $\lim\limits_{H \to 0} \frac{r(H)}{||H||} = 0$. It follows that

$$f_i(P_0 + H) - f_i(P_0) = A_i(H) + r_i(H),$$

for every $H \in \Delta_{X,P_0}$, and $\lim\limits_{H \to 0} \frac{r_i(H)}{||H||} = 0$, for every $i \in \{1, \ldots, m\}$, where A_i denotes the i-th coordinate function $p_i \circ A : \mathbb{R}^n \to \mathbb{R}$, which is also a linear operator as A is. Thus, the function $f_i : X \to \mathbb{R}$ is differentiable at P_0 for every $i = 1, \ldots, m$, i.e., the i-th coordinate function of the differential of the function f at P_0 is the differential of the i-th coordinate function f_i of f.

In the same way it can be shown that the converse statement is also valid, i.e., that the differentiability of the coordinate functions f_i at P_0 implies the differentiability of the function $f = (f_1, \ldots, f_m)$. □

Corollary 8. *Let $X \subseteq \mathbb{R}^n$ and $P_0 \in X$ be a point that admits nbd rays in X in the direction of n linearly independent vectors $V_1, \ldots, V_n$. If $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$ is differentiable at P_0 then there exists $\partial_{V_i} f(P_0) := (\partial_{V_i} f_1(P_0), \ldots, \partial_{V_i} f_m(P_0)) \in \mathbb{R}^m$ for $i = 1, \ldots, n$ and*

$$df(P_0)(H) = \sum_{i=1}^{n} \partial_{V_i} f(P_0) h_i, \; H = h_1 V_1 + \cdots + h_n V_n \in \mathbb{R}^n.$$

Proof. The statement follows from the previous theorem, Corollary 1 and Theorem 8. □

In particular, if a function $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}^n$, is differentiable at a point $P_0 \in X$ and if the point P_0 admits nbd ray in X in the direction of $e_1, \ldots, e_n$, then the numbers $\partial_j f_i(P_0)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, uniquely determine the differential $df(P_0) : \mathbb{R}^n \to \mathbb{R}^m$ of f at P_0. This linear operator in the pair of standard bases is represented by the well-known Jacobi matrix. But if the point P_0 does not admit nbd ray in X in the direction of e_i for some $i \in \{1, \ldots, n\}$, but admits nbd ray in X in the direction of n linearly independent vectors $V_1, \ldots, V_n$, then the role of partial derivatives of the functions f_i, $i = 1, \ldots, m$, at P_0 is taken over in the differential $df(P_0)$ by derivatives in the direction of $V_1, \ldots, V_n$ and $df(P_0)$ is represented by the matrix

$$\begin{bmatrix} \partial_{V_1} f_1(P_0) & \ldots & \partial_{V_n} f_1(P_0) \\ & \vdots & \\ \partial_{V_1} f_m(P_0) & \ldots & \partial_{V_n} f_m(P_0) \end{bmatrix}$$

in the pair of ordered bases $(V_1, \ldots, V_n)$ and $(e_1, \ldots, e_m)$. Let us show an application of this generalized calculus by the following simple example.

Example 10. *Let $V_1, V_2 \in \mathbb{R}^2 \backslash \{0\}$ be two linear independent vectors and let $P_0 \in X \subseteq \mathbb{R}^2$ be a point admitting raylike nbd in X and admitting nbd ray in X in the direction of V_1 and V_2. Let $f(x, y) = (u, v, w) : X \to \mathbb{R}^3$ and $g(u, v, w) : \mathbb{R}^3 \to \mathbb{R}$ be differentiable functions at $P_0 = (x_0, y_0)$ and $Q_0 = f(P_0) = (u_0, v_0, w_0)$, respectively. By Theorem 6 the function $g \circ f$ is differentiable at P_0 and by Corollary 5 there exist directional derivatives $\partial_{V_1}(g \circ f)(P_0)$, $\partial_{V_2}(g \circ f)(P_0)$, $\partial_{V_1} f(P_0) = \big(\partial_{V_1} u(P_0), \partial_{V_1} v(P_0), \partial_{V_1} w(P_0)\big)$ and $\partial_{V_2} f(P_0) = \big(\partial_{V_2} u(P_0), \partial_{V_2} v(P_0), \partial_{V_2} w(P_0)\big)$. The differential $df(P_0)$ is represented by the matrix*

$$\begin{bmatrix} \partial_{V_1} u(P_0) & \partial_{V_2} u(P_0) \\ \partial_{V_1} v(P_0) & \partial_{V_2} v(P_0) \\ \partial_{V_1} w(P_0) & \partial_{V_2} w(P_0) \end{bmatrix}$$

in the pair of ordered bases (V_1, V_2) and (e_1, e_2, e_3). The differential $d(g \circ f)(P_0)$ is represented by the matrix

$$\begin{bmatrix} \partial_{V_1}(g \circ f)(P_0) & \partial_{V_2}(g \circ f)(P_0) \end{bmatrix}$$

in the pair of ordered basis (V_1, V_2) and (e_1), $e_1 = 1$. Now, the equality

$$d(g \circ f)(P_0) = dg(Q_0) \circ df(P_0)$$

induces the matrix equation

$$\begin{bmatrix} \partial_{V_1}(g \circ f)(P_0) & \partial_{V_2}(g \circ f)(P_0) \end{bmatrix} =$$

$$= \begin{bmatrix} \partial_1 g(Q_0) & \partial_2 g(Q_0) & \partial_3 g(Q_0) \end{bmatrix} \begin{bmatrix} \partial_{V_1} u(P_0) & \partial_{V_2} u(P_0) \\ \partial_{V_1} v(P_0) & \partial_{V_2} v(P_0) \\ \partial_{V_1} w(P_0) & \partial_{V_2} w(P_0) \end{bmatrix}$$

which implies the following formulas:

$$\partial_{V_i}(g \circ f)(P_0) = \partial_{V_i} u(P_0) \cdot \partial_1 g(Q_0) + \partial_{V_i} v(P_0) \cdot \partial_2 g(Q_0) + \partial_{V_i} w(P_0) \cdot \partial_3 g(Q_0),$$

$i = 1, 2$.

5.5. Differentiability of Vector Functions of One Variable

Let $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$, $X \subseteq \mathbb{R}$, be a vector function of one variable. If f is differentiable at a point $x_0 \in X$ then the differential is unique (Corollary 2). Moreover, by Theorem 9, functions f_i are differentiable at x_0 for $i = 1, \ldots, n$, and

$$df(x_0)(h) = \big(f_1'(x_0)h, \ldots, f_m'(x_0)h\big).$$

The vector $\big(f_1'(x_0), \ldots, f_m'(x_0)\big)$ is called the **derivative of the vector function** f **at the point** x_0 and is denoted by $f'(x_0)$. Obviously, $df(x_0)(h) = hf'(x_0)$. From Theorems 7 and 9 it follows that the differentiability of f at x_0 is equivalent to the derivability of its coordinate functions $f_1, \ldots, f_m$ at x_0, consequently instead of differentiability of f we often use the term derivability of f.

Definition 8. *Let $x_0 \in X \subseteq \mathbb{R}$ admit raylike nbd in X. We say that $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$ is* ***regular at*** *x_0 if f is derivable at x_0 and $f'(x_0) \neq 0$.*

Definition 9. *Let $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$ be a regular at $x_0 \in X \subseteq \mathbb{R}$ and let any $x \in f^{-1}(\{f(x_0)\})$ admits raylike nbd in X. We say that f is* ***geometrically smooth at*** *x_0 provided f is derivable on $f^{-1}(\{f(x_0)\})$ and the vectors $f'(x_0)$ and $f'(x)$ are collinear for every $x \in f^{-1}(\{f(x_0)\})$. If f is geometrically smooth at all points in its domain, then we say that f is* ***geometrically smooth****.*

Let $m > 1$ and $f = (f_1, \ldots, f_m) : X \to \mathbb{R}^m$ be a geometrically smooth at a point $x_0 \in X$. The vector $\frac{f(x_0+h)-f(x_0)}{h}$ is the direction vector of the secant passing through the points $f(x_0)$ and $f(x_0 + h)$, and when $h \to 0$ we obtain the vector $f'(x_0)$ which is the direction vector of the tangent to the image $f(X) \subseteq \mathbb{R}^m$ of the function f at the point $f(x_0)$, i.e., the **tangent to the image of the function** f **at the point** $f(x_0)$ is the line passing through the point $f(x_0)$ and its direction vector is $f'(x_0)$. Therefore its equation is

$$\frac{y_1 - f_1(x_0)}{f_1'(x_0)} = \cdots = \frac{y_m - f_m(x_0)}{f_m'(x_0)}.$$

Notice that at the point $f(x_0) = (x_0, g(x_0))$ the terms tangent to the image of the function $f = (1_X, g) : X \to \mathbb{R}^2$, $X \subseteq \mathbb{R}$ and the tangent to the graph of or a function $g : X \to \mathbb{R}$ coincide.

The tangent to the image of a vector function f of one variable makes sense only at points where the function f is geometrically smooth. This means, first of all, that we consider only points at which the function f is derivable. In addition to derivability, the regularity of f is also required, since the direction vector of each line is different from the zero vector. Furthermore, the tangent to the image of the function f at the point $f(x_0)$ only makes sense if there exist derivatives of f at all points $x \in f^{-1}(\{f(x_0)\})$ and these derivatives are collinear vectors. Otherwise, we would get different tangents at $f(x_0)$ depending on which point $x \in f^{-1}(\{f(x_0)\})$ we choose. Therefore, this condition is also included in the definition of geometrically smooth function f at a point x_0, since only at these points the tangent is uniquely determined. For example the function

$$f : \mathbb{R} \to \mathbb{R}^2 \ f(t) = (\cos t, \sin t)$$

is a geometrically smooth function, since it is derivable and regular at every point in the domain. The image of f is the circle $\mathbb{S}^1$ and for an arbitrary point $f(t_0)$ the derivatives of f at all points in the set $f^{-1}(\{\, f(t_0)\}) = \{t_0 + 2k\pi,\ k \in \mathbb{Z}\}$ are equal. On the other hand, the function

$$g : \left[0, \frac{\pi}{3}\right] \to \mathbb{R}^2 \ g(t) = (\sin 3t \cos t, \sin 3t \sin t)$$

is not geometrically smooth at the points 0 and $\frac{\pi}{3}$. Namely, $g(0) = (0,0) = g\left(\frac{\pi}{3}\right)$, the function g is derivable and

$$g'(t) = (3\cos 3t\cos t - \sin 3t\sin t, 3\cos 3t\sin t + \sin 3t\cos t),$$

but vectors

$$g'(0) = (3,0) \text{ and } g'\left(\frac{\pi}{3}\right) = \left(-\frac{3}{2}, -3\frac{\sqrt{3}}{2}\right)$$

are not collinear. Therefore, the tangent to the image of the function g at the point $(0,0)$ is not defined.

Notation 1. *Although the notion of a geometrically smooth function allows the correct definition of a tangent to the image of the function at a point (a tangent cannot be a line without direction and must be unique at the point in the image of the function), the problem of the dependence of the tangent on the observed function remains, i.e., a tangent to the image of a function depends directly on the observed function and not only on its image. Indeed, two functions f and g may have the same image Γ and for a point $P \in \Gamma$ the function f need not be geometrically smooth at x and g geometrically smooth at y for every $x \in f^{-1}(\{P\})$ and every $y \in g^{-1}(\{P\})$, nor need the vectors $f'(x)$ and $g'(y)$, if they exist, be collinear. For example, the image of the functions*

$$f,\ g : [-1,1] \to \mathbb{R}^2\ f(t) = (t,t), g(t) = \left(t^3, t^3\right)$$

is the line segment $\overline{(-1,-1)(1,1)}$. The function f is geometrically smooth at the point 0, $f'(0) = (1,1)$ and the tangent to the image of f at $(0,0)$ is the line $y = x$. On the other hand, the function g is not regular at the point 0, so the tangent to its image (the same line segment) is not defined. If we want to define a tangent to a set which is the image of a function of one variable, but does not depend on the function itself, we should consider curves, i.e., smooth 1-parameterizable sets, which is beyond the scope of this paper.

6. Tangent Plane

Let $X \subseteq \mathbb{R}^n$, $n \geq 2$, and $f : X \to \mathbb{R}$. Let

$$S = \{(x_1, \dots, x_n) \in X \mid f(x_1, \dots, x_n) = 0\}$$

be a nonempty set, and $P_0 = \left(x_1^0, \dots, x_n^0\right) \in S$ be a point admitting nbd ray in X in the direction of $e_1, \dots, e_n$ and admitting raylike nbd in X. Let f be differentiable at P_0 and $\nabla f(P_0) \neq 0$. For a continuous function $r = (r_1, \dots, r_n) : [a,b] \to S$, $[a,b] \subseteq \mathbb{R}$, which is differentiable and geometrically smooth at a point $t_0 \in [a,b]$ and for which $r(t_0) = P_0$, the direction vector of the tangent to the image $r([a,b])$ of r at the point P_0 is $r'(t_0)$. Since $f \circ r = 0$ then $d(f \circ r)(t_0) = 0$ and, by Theorem 6,

$$0 = d(f \circ r)(t_0) = df(P_0) \circ dr(t_0),$$

i.e.,

$$0 = \begin{bmatrix} \partial_1 f(P_0) & \cdots & \partial_n f(P_0) \end{bmatrix} \begin{bmatrix} r_1'(t_0) \\ \vdots \\ r_n'(t_0) \end{bmatrix} = \left(\nabla f(P_0) | r'(t_0)\right) \tag{9}$$

therefore the vectors $\nabla f(P_0)$ and $r'(t_0)$ are orthogonal. Thus, the direction vector of the tangent to the image of any function $r : [a,b] \to S$ with the above properties at the point P_0 is orthogonal to $\nabla f(P_0)$ which implies that all these tangents lie in the same plane; we call this hyperplane the **tangent plane to the set S at the point P_0**. Since $\nabla f(P_0)$ is its normal vector, its equation is

$$\partial_1 f(P_0)(x_1 - x_1^0) + \cdots + \partial_n f(P_0)(x_n - x_n^0) = 0.$$

Let F be defined by $F(x_1,\dots,x_{n+1}) = x_{n+1} - f(x_1,\dots,x_n)$ for $(x_1,\dots,x_n) \in X$. The tangent plane to the set

$$\begin{aligned}\Gamma &= \{(x_1,\dots,x_{n+1}) \in \mathbb{R}^{n+1} \mid F(x_1,\dots,x_{n+1}) = 0,\ (x_1,\dots,x_n) \in X\} \\ &= \Big\{(x_1,\dots,x_n, f(x_1,\dots,x_n)) \in \mathbb{R}^{n+1} \mid (x_1,\dots,x_n) \in X\Big\}\end{aligned}$$

at the point $(P_0, f(P_0))$ is called **tangent plane to the graph of the function f at the point** $(P_0, f(P_0))$. The vector $(-\partial_1 f(P_0),\dots,-\partial_n f(P_0), 1)$ is a normal vector of this plane, so its equation is

$$x_{n+1} - f(P_0) = \sum_{i=1}^{n} \partial_i f(P_0)\Big(x_i - x_i^0\Big) = df(P_0)\Big(x_1 - x_1^0,\dots,x_n - x_n^0\Big).$$

Let us now define the tangent plane to the graph of a scalar function in an even more general case, i.e., at points which do not admit nbd ray in X in the direction of some of the vectors $e_1,\dots,e_n$ but in the direction of some n linearly independent vectors.

Definition 10. *Let $f : X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$, be differentiable at a point $P_0 = \big(x_1^0,\dots,x_n^0\big) \in X$ that admits raylike nbd in X, and let $\Sigma_{X,P_0} = \mathbb{R}^n$. The hyperplane*

$$x_{n+1} - f(P_0) = df(P_0)\Big(x_1 - x_1^0,\dots,x_n - x_n^0\Big)$$

*is called the **tangent plane to the graph of the function f at the point** $(P_0, f(P_0))$.*

Remark 3. *Since the coordinates of the vector $P - P_0$ are given in the standard basis of $\mathbb{R}^n$, it is assumed that the operator $df(P_0)$ is defined in the pair of ordered basis $(e_1,\dots,e_n)$ and $(e_1 = 1)$. The numbers $df(P_0)(e_i)$, $i = 1,\dots,n$, are partial derivatives of the function f only if P_0 admits nbd ray in the direction of e_i, for $i = 1,\dots,n$.*

Let us show the justification of the previously defined notion. Let $f : X \to \mathbb{R}$ fulfill conditions of the previous definition and let P_0 admits nbd ray in the direction of a vector $V = (v_1,\dots,v_n) \in \mathbb{R}^n \backslash \{0\}$ in X such that $\overline{P_0 P_0 + V} \subseteq X$. Let us consider the parametrization of the line segment $\overline{P_0 P_0 + V}$, i.e., the function $r = (r_1,\dots,r_n) : [0,1] \to \mathbb{R}^n$ $r(t) = P_0 + tV$. Due to the assumed differentiability of the function f at P_0, there exists the derivative $\partial_V f(P_0)$ in the direction of V (Corollary 5) and then the function $f \circ r$ is derivable at 0 because

$$\begin{aligned}(f \circ r)'(0) &= \lim_{\substack{h \to 0 \\ h \in \langle 0,1 \rangle}} \frac{f(P_0 + hV) - f(P_0)}{h} \\ &= \lim_{\substack{h \to 0 \\ hV \in \Delta_{V(X,P_0)}}} \frac{f(P_0 + hV) - f(P_0)}{h} = \partial_V f(P_0).\end{aligned}$$

The image of the function $(r_1,\dots,r_n, f \circ r) : [0,1] \to \mathbb{R}^{n+1}$ belongs to the graph of f, passes through the point $(P_0, f(P_0))$ and the direction vector of the tangent to the image of the function $(r_1,\dots,r_n, f \circ r)$ at the point $(P_0, f(P_0))$ is $(v_1,\dots,v_n, \partial_V f(P_0))$. We will now show that this tangent lies in the tangent plane to the graph of the function f at $(P_0, f(P_0))$, i.e., that the point

$$\Big(x_1^0 + v_1,\dots,x_n^0 + v_n, f(P_0) + \partial f_V(P_0)\Big)$$

belongs to this plane. According to Corollary 5, $df(P_0)(V) = \partial_V f(P_0)$ holds. This implies

$$(f(P_0) + \partial_V f(P_0)) - f(P_0) = df(P_0)\Big(x_1^0 + v_1 - x_1^0,\dots,x_n^0 + v_n - x_n^0\Big)$$

which proves that all points of the tangent belong to the tangent plane. This means that all previously described tangents lie in the same hyperplane and the normal vector of this hyperplane is orthogonal to the vector $(v_1, \ldots, v_n, \partial_V f(P_0))$ where which justifies the previous definition.

Remark 4. *For a scalar function $f : D \to \mathbb{R}$ of two variables which fulfils conditions of the previous definition at a point $P_0 = (x_0, y_0) \in D$ admitting nbd rays in the direction of two non-collinear vectors $V = (v_1, v_2)$ and $V' = (v_1', v_2')$ in D, the vector of the direction of the normal line of the tangent plane to the graph of the function f at $(x_0, y_0, f(P_0))$ is $(v_1, v_2, \partial_V f(P_0)) \times (v_1', v_2', \partial_{V'} f(P_0))$. This vector is orthogonal to the vector $\left(\overline{v_1}, \overline{v_2}, \partial_{(\overline{v_1},\overline{v_2})} f(P_0)\right)$ where $(\overline{v_1}, \overline{v_2})$is an arbitrary vector such that the point P_0 admits nbd ray in D in the direction of $(\overline{v_1}, \overline{v_2})$. In particular, at a point $P = (x, y) \in \operatorname{Int} D$, i.e., at a point that admits nbd rays in D in the direction of vectors e_1 and e_2, the normal vector of the tangent plane in $(x, y, f(P))$ can be calculated as $(1, 0, \partial_x f(P)) \times (0, 1, \partial_y f(P)) = (-\partial_x f(P), -\partial_y f(P), 1)$.*

7. Differentiability in Different Coordinate Systems

7.1. Affine Coordinate Systems

Let $(O', (e'))$ be an affine coordinate system [10] where $O' \in \mathbb{R}^n$ is its origin and $(e') = (e_1', \ldots, e_n')$ is an ordered basis for the vector space $\mathbb{R}^n$. For any $P \in \mathbb{R}^n$ the numbers $x_1', \ldots, x_n' \in \mathbb{R}$ for which $P - O' = x_1'e_1' + \cdots + x_n'e_n'$ are the coordinates of the point P in that affine system and we write $P = (x_1', \ldots, x_n')$. Obviously, $O' = (0, \ldots, 0)$. If $(e) = (e_1, \ldots, e_n)$ is the standard ordered basis and $O \in \mathbb{R}^n$, $(O, (e))$ is said to be the **standard affine coordinate system.**

Let $P \in \mathbb{R}^n$ has notation $(x_1, \ldots, x_n)$ in the standard affine coordinate system $(O, (e))$ and let $(x_1', \ldots, x_n')$ be its notation in the affine coordinate system $(O', (e'))$. Then the linear operator $A_{(e)(e')} : \mathbb{R}^n \to \mathbb{R}^n$ $A(e_i) = e_i'$, $i = 1, \ldots, n$, is an isomorphism and its corresponding affine isomorphism $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ give us an analytical relation between the coordinates in the both coordinate systems according to the rule

$$(x_1, \ldots, x_n) = \sigma(x_1', \ldots, x_n') := \left(x_1^0, \ldots, x_n^0\right) + A_{(e)(e')}(x_1', \ldots, x_n'), \tag{10}$$

where $(x_1^0, \ldots, x_n^0)$ is the notation of the point O' in the standard affine coordinate system $(O, (e))$. The affine isomorphism σ is called the **function of the transition from the standard affine coordinate system to the affine coordinate system** $(O', (e'))$. As an affine isomorphism, the function σ is differentiable, as well as its inverse σ^{-1}.

Definition 11. *Let $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}^m$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ be the transition function (10) and $X' := \sigma^{-1}(X)$. The function $f_\sigma := f \circ \sigma : X' \to \mathbb{R}^m$ is called the **representation of the function f in the affine coordinate system** $(O', (e'))$.*

We will now prove the following important statements which hold for affine coordinate systems:

Theorem 10. *Let $X \subseteq \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ be the function of the transition from the standard affine coordinate system to an affine coordinate system $(O', (e'))$ and let $P_0 \in X$ be a point admitting nbd ray in X. Let $f : X \to \mathbb{R}^m$ be a function, $X' := \sigma^{-1}(X)$ and $f_\sigma : X' \to \mathbb{R}^m$ be its representation in the affine coordinate system $(O', (e'))$. Then it holds:*

(i) The point $P_0' = \sigma^{-1}(P_0)$ admits nbd ray in X' and $\Delta_{X',P_0'} = A^{-1}_{(e)(e')}(\Delta_{X,P_0})$.

(ii) If f is differentiable at P_0, then the function f_σ is differentiable at $\sigma^{-1}(P_0)$.

(iii) If $\Sigma_{X,P_0} = \mathbb{R}^n$ and f is differentiable at P_0, then the function f_σ is also differentiable at $P_0' = \sigma^{-1}(P_0)$ and $df_\sigma(P_0') = df(P_0) \circ A_{(e)(e')}$.

Proof.

(i) Since an affine function maps a line segment to the line segment whose endpoints are the image of the endpoints of the original line segment, the transition function σ^{-1} maps the line segment $\overline{P_0P_0+H}$ to the line segment $\overline{P_0'P_0'+A^{-1}_{(e),(e')}(H)}$, from which follows the statement.

(ii) Let $\Delta'_{X,P_0} = A^{-1}(\Delta_{X,P_0})$. By assumption, there exists a linear operator $B : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H)-f(P_0)-B(H)}{||H||} = 0.$$

Let us show that the linear operator $B\circ A : \mathbb{R}^n \to \mathbb{R}^m$ is the differential of the function f_σ at the point $P_0' = \sigma^{-1}(P_0)$. It holds

$$\lim_{\substack{H'\to 0\\ H'\in\Delta'_{X,P_0}}} \frac{f\circ\sigma(P_0'+H')-f\circ\sigma(P_0')-B\circ A(H')}{||H'||} =$$

$$\lim_{\substack{H'\to 0\\ H'\in\Delta'_{X,P_0}}} \frac{f(O'+A(P_0')+A(H'))-f(P_0)-B(A(H'))}{||H'||} =$$

$$= \lim_{\substack{H'\to 0\\ H'\in\Delta'_{X,P_0}}} \frac{f(P_0+A(H'))-f(P_0)-B(A(H'))}{||A^{-1}\circ A(H')||} =$$

$$\lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H)-f(P_0)-B(H)}{||A^{-1}(H)||} =$$

$$= \lim_{\substack{H\to 0\\ H\in\Delta_{X,P_0}}} \frac{f(P_0+H)-f(P_0)-B(H)}{||H)||\cdot\frac{||A^{-1}(H)||}{||H)||}}$$

By the Lemma 1, there exists $m\in\mathbb{R}^+$ such that

$$0 \le \frac{||f(P_0+H)-f(P_0)-B(H)||}{||H||\cdot\frac{||A^{-1}(H)||}{||H)||}} \le \frac{||f(P_0+H)-f(P_0)-B(H)||}{||H||\cdot m},$$

for every $H\in\Delta_{X,P_0}$, thus from the provious identities it follows

$$\lim_{\substack{H'\to 0\\ H'\in\Delta'_{X,P_0}}} \frac{f_\sigma(P_0'+H')-f_\sigma(P_0')-B\circ A(H')}{||H'||} = 0.$$

(iii) It follows from the previous two statements.

□

Lemma 1. *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be an isomorphism. Then there exists $m\in\mathbb{R}^+$ such that*

$$\frac{||A(H)||}{||H||} \ge m, \text{ for every } H\in\mathbb{R}^n.$$

Proof. Assume the contrary that there exists a sequence (H_k) in $\mathbb{R}^n$ and a real sequence (m_k) converging to 0 such that

$$\frac{||A(H_k)||}{||H_k||} < m_k, \text{ for every } k\in\mathbb{N}.$$

Consider a sequence of points (P_k), $P_k = \frac{H_k}{||H_k||}$, $k \in \mathbb{N}$, in $\mathbb{R}^n$. Since the sphere $\mathbb{S}^{n-1}$ is a compact set, the sequence (P_k), contained in it, has a certain convergent subsequence $\left(P_{k_l}\right)$ whose limit P_0 belongs to the sphere [3]. Therefore, $||P_0|| = 1$. From $\left(P_{k_l}\right) \to P_0$ follows $\left(\left|\left|A\left(P_{k_l}\right)\right|\right|\right) \to ||A(P_0)||$. By the property of a norm and a linear operator, it holds

$$\left|\left|A\left(P_{k_l}\right)\right|\right| = \left|\left|A\left(\frac{H_k}{||H_k||}\right)\right|\right| = \frac{||A(H_k)||}{||H_k||} < m_k, \text{ for every } k \in \mathbb{N}.$$

This implies $\lim\left(\left|\left|A\left(P_{k_l}\right)\right|\right|\right) = 0$, from which it follows $||A(P_0)|| = 0$ and consequently $A(P_0) = 0$. But since A is an isomorphism, this implies $P_0 = 0$ which contradicts the equality $||P_0|| = 1$. $\square$

This theorem shows us that the notion of differentiability of a function does not depend on the chosen affine coordinate system. We will now show that it is not true for some other coordinate systems.

7.2. Polar, Elliptical, Cylindrical and Spherical Coordinate Systems

For each point $T = (x, y) \in \mathbb{R}^2$ in the standard affine coordinate system we define the coordinates r, ϕ in the standard polar coordinate system (so-called polar coordinates) by the following formulas

$$r = \sqrt{x^2 + y^2},$$

$$\phi := \arg(T) = \begin{cases} \arctan\left(\frac{y}{x}\right), & x > 0,\ y \geq 0 \\ \arctan\left(\frac{y}{x}\right) + \pi, & x < 0,\ y \leq 0 \text{ or } x < 0,\ y \geq 0 \\ \arctan\left(\frac{y}{x}\right) + 2\pi, & x > 0,\ y \leq 0 \\ \frac{\pi}{2}, & x = 0,\ y > 0 \\ \frac{3\pi}{2}, & x = 0,\ y < 0 \\ 0, & x = 0,\ y = 0. \end{cases}$$

This defines bijection

$$\rho : \Theta_0 \to \mathbb{R}^2,\ \Theta_0 := \langle 0, \infty\rangle \times [0, 2\pi\rangle \cup \{(0,0)\}$$

$$(x, y) = \rho(r, \phi) = (r\cos\phi, r\sin\phi).$$

which we call the **transition function from the standard affine coordinate system to the standard polar coordinate system**. The polar coordinate grid consists of "lines" $\phi = \phi_0$ and "lines" $r = r_0$ for $(r_0, \phi_0) \in \Theta_0$. This polar coordinate grid is mapped to a "spider web" in an affine coordinate system consisting of all concentric circles with center 0 and all half lines with 0 as endpoint.

Notice that the function ρ is not a homeomorphism [2] since its inverse has discontinuity at all points $(x, 0)$, $x \in [0, \infty\rangle$. For this reason, in applications and transitions from the polar to the affine coordinate system, the restriction of the transition function is used

$$\rho|_{\langle 0,\infty\rangle \times \langle 0,2\pi\rangle} : \langle 0, \infty\rangle \times \langle 0, 2\pi\rangle \to \mathbb{R}^2 \backslash p_0,\ p_0 := \{(x, 0) \mid x \geq 0\},$$

which is a homeomorphism. With the transition function ρ, the rectangle

$$\{(r, \phi) \mid r_1 \leq r \leq r_2,\ \phi_1 \leq \phi \leq \phi_2\},\ 0 < r_1 < r_2, 0 \leq \phi_1 < \phi_2 < 2\pi,$$

is mapped to an area bounded by corresponding circles and half-lines, therefore the polar coordinate system is more suitable to consider such sections than any other system in the plane. Every point of the set Θ_0 admits nbd ray in it and the function ρ is differentiable at every point $(r, \phi) \in \Theta_0$, and the differential is represented by a matrix

$$\begin{bmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{bmatrix}.$$

Let $X \subseteq \mathbb{R}^2$, $f : X \to \mathbb{R}^m$ be a function and $X' = \rho^{-1}(X) \subseteq \Theta_0$. Then, for the function $f_\rho := f \circ \rho : X' \to \mathbb{R}^m$, we say that the **representation of the function f is in the polar coordinate system or in polar coordinates**.

Example 11. *The representation of a function $f : \mathbb{R}^2 \to \mathbb{R}$ $f(x,y) = \sqrt{x^2+y^2}$ in the polar coordinate system is $f_\rho : \Theta_0 \to \mathbb{R}$ $f_\rho(r,\phi) = r$. Notice that the function f is not differentiable at the point O (Example 5) but the function f_ρ is differentiable at the point $(0,0) = \rho^{-1}(O)$ and it holds $df_\rho(0,0) = p_1$.*

Therefore, we conclude that the notion of differentiability of a function depends on the chosen coordinate system (affine or polar) in which it is represented. Such a phenomenon is not possible in the transition from one affine coordinate system to another (Theorem 10). Thus, in the transition from polar coordinates to affine coordinates (or vice versa), the differentiability of the function need not be preserved, nor does the notion of admissibility of nbd rays. To avoid such undesirable phenomena, we will consider only functions f (in affine coordinates) whose domains are open sets in $\mathbb{R}^2\backslash p_0$, and functions f_ρ (in polar coordinates) whose domains are open sets in $\langle 0,\infty\rangle \times \langle 0,2\pi\rangle$. Since $\rho|_{\langle 0,\infty\rangle \times \langle 0,2\pi\rangle}$ is a diffeomorphism (differentiable bijection which inverse is also differentiable), the following holds according to the Theorem 6.

Corollary 9. *Let $\Omega \subseteq \mathbb{R}^2\backslash p_0$ be an open set, $f : \Omega \to \mathbb{R}^m$ a function and $f_\rho : \Omega' \to \mathbb{R}^m$ representation of the function f in polar coordinates, where $\Omega' = \rho^{-1}(\Omega)$. If f is differentiable at a point $(x,y) \in \Omega$ then the function $f_\rho : \Omega' \to \mathbb{R}^m$, is differentiable at $(r,\phi) = \rho^{-1}(x,y)$ and it holds $d\big(f_\rho\big)(r,\phi) = df(x,y) \circ d\rho(r,\phi)$.*

Notation 2. *Differentiability of a function represented in polar coordinates should be considered only on a formal level. Indeed, the idea of linearization of a function, i.e., its local approximation by an affine function, makes sense only in affine coordinates. For example, for the scalar function $z = f_\rho(r,\phi)$ in polar coordinates, the differential at the point$(r_0,\phi_0) \neq (0,0)$ is the linear operator A, $A(r,\phi) = \partial_r f_\rho(r_0,\phi_0)r + \partial_\phi f_\rho(r_0,\phi_0)\phi$, which is no longer a linear operator in affine coordinates. Differentiability of a function represented in polar coordinates in the context of an affine coordinate system not be understood as a possibility of local approximation by an affine function, but by a linear combination of the functions $z = \sqrt{x^2+y^2}$ and $z = \arg(x,y)$ since these can be considered special in polar coordinates. For functions with a conic graph (like those in the previous example), such an approximation is more appropriate than linearization, and differentiation of functions in polar coordinates has exactly this meaning. Thus, the function from the previous example becomes a linear operator in polar coordinates and its differential is equal to itself at every point and in this context it is a perfect approximation.*

8. Conclusions

In this work we have presented some results obtained in the research conducted during COVID epidemic. It was motivated by some issues and shortcomings which occur in some applications of the traditional approach to differentiability. These problems were noticed by the first author, who has many years of experience in giving classes of various courses in mathematical analysis to students from University of Split, Croatia. In the traditional approach to differentiability, which is featured in almost all university textbooks, this notion is considered only for interior points of domain of function or for functions with an open domain. We have generalized differentiability of scalar and vector functions of several variables by defining it at non-interior points of the domain of function, which include not only boundary points but also all points where a notion of linearization is meaningful (points admitting nbd rays). This generalization allows applications in many fields of mathematics and engineering or, in short, in all areas where standard differentiability can be applied. With this generalized approach to differentiability, some unexpected phenomena may occur, such as the non-uniqueness of the differential in some special cases, a function discontinuity at a point where a function is differentiable, which is possible only for points that do not admit raylike nbd in a domain of function. But if one reduces this theory only to points with some special properties (points admitting a linearization space with dimension equal to the dimension of the ambient Euclidean space of the domain and admitting a raylike neighborhood, which includes the interior points of a domain), then all properties and theorems belonging to the known theory of differentiability remain valid in this extended theory. For generalized differentiability, the corresponding calculus (differentiation techniques) is also provided by matrices—representatives of differentials at points. In this calculus, the role of partial derivatives (which generally may not exist for differentiable functions at some points) is taken by directional derivatives. The results presented open the possibility for further research and examination of known theorems on standard differentiability in a new context.

Author Contributions: Conceptualization, N.K.-B.; investigation, N.K.-B.; writing—original draft preparation, N.K.-B. and S.B.; writing—review and editing, N.K.-B. and S.B.; supervision, N.K.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aksoy, A.; Khamsi, M. *A Problem Book in Real Analysis*; Springer: New York, USA, 2010; pp. 197–198.
2. Dugundji, J. *Topology*, 12th ed.; Allyn and Bacon, Inc.: Boston, MA, USA, 1978; pp. 62–97.
3. Rudin, W. *Principles of Mathematical Analysis*, 3rd ed.; McGraw-Hill, Inc.: New York, NY, USA, 1976; pp. 51–52, 212–213.
4. Zorich, V. *Mathematical Analysis I*, 4th ed.; Springer: Berlin, Germany, 2004; Chapters 5 and 8.
5. Wade, W. *An Introduction to Analysis*, 4th ed.; Pearson Education, Inc.: Hoboken, NJ, USA, 2010; Chapters 3 and 4.
6. Wrede, R.; Spiegel, M. *Theory and Problems of Advanced Calculus*, 2nd ed.; McGraw-Hill, Inc.: New York, NY, USA, 1976; pp. 65–89.
7. Differencial Topology with Prof. John W. Milnor. Available online: https://www.youtube.com/watch?v=u5C0GKmMHQ4&list=PLS8dWbmb9L0fPQclKNhgYddkNr8QWMGzX (accessed on 20 June 2022).
8. Koceić Bilan, N.; Jelić, I. On intersetions of exponential and logarithmic curves. *Ann. Math. Inform.* **2014**, *43*, 159–170.
9. Kukushkin, M. Abstract Fractional Calculus for m-Accretive Operators. *Int. J. Appl. Math.* **2021**, *34*, 1–41. [CrossRef]
10. Hogben, L. *Handbook of Linear Algebra*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2017.
11. Bühler, T.; Salamon, D. *Functional Analysis*; American Mathematical Society: Providence, RI, USA, 2018; pp. 22–26.
12. Appendix Functional Analysis and Operators. Available online: https://link.springer.com/content/pdf/bbm%3A978-3-030-34949-3%2F1 (accessed on 20 July 2022).