ISSN 2070-0466, p-Adic Numbers, Ultrametric Analysis and Applications, 2021, Vol. 13, No. 2, pp. 148–165. © Pleiades Publishing, Ltd., 2021.

RESEARCH ARTICLES

Taylor's Theorem, the Inverse Function Theorem and the Implicit Function Theorem for Weakly Locally Uniformly Differentiable Functions on Non-Archimedean Spaces*

Khodr Shamseddine**

Department of Physics and Astronomy, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

Received October 10, 2020; in final form, December 14, 2020; accepted December 16, 2020

Abstract—Let $\mathcal{N}$ be a non-Archimedean ordered field extension of the real numbers that is real closed and Cauchy complete in the topology induced by the order. In this paper, we first review the properties of weakly locally uniformly differentiable (WLUD) functions [1] at a point or on an open subset of $\mathcal{N}$. WLUD functions are C^1 and they form an $\mathcal{N}$-algebra that is closed under composition and contains all polynomial functions. Moreover, they satisfy an inverse function theorem, a local intermediate value theorem and a local mean value theorem. We define k times weakly locally uniformly differentiable (WLUD^k) functions from $\mathcal{N}$ to $\mathcal{N}$, then we state and prove a Taylor theorem with remainder for WLUD^k functions on $\mathcal{N}$. Finally, we generalize the concept of weak local uniform differentiability to functions from $\mathcal{N}^n$ to $\mathcal{N}^m$ with $m, n \in \mathbb{N}$, then we formulate and prove the inverse function theorem for WLUD functions from $\mathcal{N}^n$ to $\mathcal{N}^n$ and the implicit function theorem for WLUD functions from $\mathcal{N}^n$ to $\mathcal{N}^m$ with $m < n$ in $\mathbb{N}$.

DOI: 10.1134/S2070046621020035

Key words: *non-Archimedean analysis, non-Archimedean valued fields, intermediate value theorem, inverse function theorem, mean value theorem, implicit function theorem.*

1. INTRODUCTION

We start this section by reviewing some basic terminology and facts about non-Archimedean fields. So let F be an ordered non-Archimedean field extension of the field of real numbers $\mathbb{R}$. We introduce the following terminology.

Definition 1.1 ($\sim$, $\approx$, $\ll$, S_F, λ). *For $x, y \in F^* := F \setminus \{0\}$, we say that x is of the same order as y and write $x \sim y$ if there exist $n, m \in \mathbb{N}$ such that $n|x| > |y|$ and $m|y| > |x|$, where $|\cdot|$ denotes the ordinary absolute value on F: $|x| = \max\{x, -x\}$.*
For nonnegative $x, y \in F$, we say that x is infinitely smaller than y and write $x \ll y$ if $nx < y$ for all $n \in \mathbb{N}$, and we say that x is infinitely small if $x \ll 1$ and x is finite if $x \sim 1$; finally, we say that x is approximately equal to y and write $x \approx y$ if $x \sim y$ and $|x - y| \ll |x|$. We also set $\lambda(x) = [x]$, the class of x under the equivalence relation $\sim$.

The set of equivalence classes S_F (under the relation $\sim$) is naturally endowed with an addition via $[x] + [y] = [x \cdot y]$ and an order via $[x] < [y]$ if $|y| \ll |x|$ (or $|x| \gg |y|$), both of which are readily checked to be well-defined. Note that we use $+$ instead of $\cdot$ for the operation in S_F because, for the fields discussed in this paper, S_F is isomorphic to an additive subgroup of $\mathbb{R}$. It follows that $(S_F, +, <)$ is an ordered group, often referred to as the Hahn group or skeleton group, whose neutral element is [1], the class of 1. It follows from the above that the projection λ from F^* to S_F is a valuation.

The theorem of Hahn [3] provides a complete classification of non-Archimedean ordered field extensions of $\mathbb{R}$ in terms of their skeleton groups. In fact, invoking the axiom of choice it is shown that

*The text was submitted by the author in English.
**E-mail: khodr.shamseddine@umanitoba.ca

the elements of any such ordered field F can be written as (generalized) formal power series (also called Hahn series) over its skeleton group S_F with real coefficients, and the set of appearing exponents forms a well-ordered subset of S_F. That is, for all $x \in F$, we have that

$$x = \sum_{q \in S_F} a_q d^q; \tag{1.1}$$

with $a_q \in \mathbb{R}$ for all q, d a positive infinitely small element of F, and the support of x, given by

$$\text{supp}(x) := \{q \in S_F : a_q \neq 0\},$$

forming a well-ordered subset of S_F. With the representation of elements of F as in Equation (1.1), it follows that for $x \neq 0$ in F,

$$\lambda(x) = \min\left(\text{supp}(x)\right),$$

which exists since supp(x) is well-ordered. Moreover, we set $\lambda(0) = \infty$.

Addition, multiplication and order on the Hahn series are defined as follows. Given $x = \sum_{q \in \text{supp}(x)} a_q d^q$ and $y = \sum_{t \in \text{supp}(y)} b_t d^t$, then

$$x + y = \sum_{r \in \text{supp}(x) \cup \text{supp}(y)} (a_r + b_r) d^r; \text{ and}$$

$$x \cdot y = \sum_{r \in \text{supp}(x) \oplus \text{supp}(y)} \left(\sum_{\substack{q \in \text{supp}(x),\, t \in \text{supp}(y) \\ q + t = r}} a_q \cdot b_t \right) d^r. \tag{1.2}$$

Note that, since supp(x) and supp(y) are well-ordered, only finitely many terms contribute to the sum

$$\sum_{\substack{q \in \text{supp}(x),\, t \in \text{supp}(y) \\ q + t = r}} a_q \cdot b_t$$

in Equation (1.2) for each $r \in \text{supp}(x) \oplus \text{supp}(y)$.

Given a nonzero $x = \sum_{q \in \text{supp}(x)} a_q d^q$, then $x > 0$ if and only if $a_{\lambda(x)} > 0$.

From general properties of formal power series fields [7, 9], it follows that if S_F is divisible then F is real closed; that is, every positive element of F is a square in F and every polynomial of odd degree over F has at least one root in F. For a general overview of the algebraic properties of formal power series fields, we refer to the comprehensive overview by Ribenboim [10], and for an overview of the related valuation theory the book by Krull [4]. A thorough and complete treatment of ordered structures can also be found in [8]. A more comprehensive survey of all non-Archimedean fields can be found in [2].

Throughout this paper, we will denote by $\mathcal{N}$ any totally ordered non-Archimedean field extension of $\mathbb{R}$ that is real closed and complete in the order topology and whose skeleton group $S_{\mathcal{N}}$ is Archimedean, i.e. a subgroup of $\mathbb{R}$. The coefficient a_q of the qth power in the Hahn representation of a given x will be denoted by $x[q]$, and hence the number d is given by $d[1] = 1$ and $d[q] = 0$ for $q \neq 1$. It is easy to check that, for $q \in S_{\mathcal{N}}$, $0 < d^q \ll 1$ if and only if $q > 0$ and $d^q \gg 1$ if and only if $q < 0$; moreover, $x \approx x[\lambda(x)] d^{\lambda(x)}$ for all $x \neq 0$.

The smallest such field $\mathcal{N}$ is the Levi-Civita field $\mathcal{R}$, first introduced in [5, 6]. In this case $S_{\mathcal{R}} = \mathbb{Q}$, and for any element $x \in \mathcal{R}$, supp(x) is a left-finite subset of $\mathbb{Q}$, i.e. below any rational bound r there are only finitely many exponents in the Hahn representation of x. The Levi-Civita field $\mathcal{R}$ is of particular interest because of its practical usefulness. Since the supports of the elements of $\mathcal{R}$ are left-finite, it is possible to represent these numbers on a computer. Having infinitely small numbers allows for many computational applications; one such application is the computation of derivatives of real functions representable on a computer [14, 15], where both the accuracy of formula manipulators and the speed of classical numerical methods are achieved. For a review of the Levi-Civita field $\mathcal{R}$, see [13] and references therein.

In the wider context of valuation theory, it is interesting to note that the topology induced by the order on $\mathcal{N}$ is the same as that introduced via the valuation λ, as shown in Remark 1.2 below. It follows therefore that the field $\mathcal{N}$ is just a special case of the class of fields discussed in [12].

Remark 1.2. *The mapping $\Lambda : \mathcal{N} \times \mathcal{N} \to \mathbb{R}$, given by $\Lambda(x, y) = \exp(-\lambda(x - y))$, is an ultrametric distance (and hence a metric); the valuation topology it induces is equivalent to the order topology (we will use τ_v to denote either one of the two topologies in this paper). For if A is an open set in the order topology and $a \in A$, then there exists $r > 0$ in $\mathcal{N}$ such that, for all $x \in \mathcal{N}$, $|x - a| < r \Rightarrow x \in A$. Let $l = \exp(-\lambda(r))$, then we also have that, for all $x \in \mathcal{N}$, $\Lambda(x, a) < l \Rightarrow x \in A$; and hence A is open with respect to the valuation topology. The other direction of the equivalence of the topologies follows analogously.*

It follows from Remark 1.2 that $\mathcal{N}$ which is complete in the order topology is also complete in the valuation topology τ_v induced by the ultrametric Λ.

Remark 1.3. *Like the field $^*\mathbb{R}$ of Nonstandard Analysis [11, 17], the field $\mathcal{N}$ is a sequentially complete non-Archimedean ordered field extension of the field of real numbers $\mathbb{R}$; and the embedding of $\mathbb{R}$ in $\mathcal{N}$ is compatible with the orders in $\mathbb{R}$ and $\mathcal{N}$. However, while in Nonstandard Analysis there is a generally valid transfer principle that allows the transformation of known results of conventional analysis, here all relevant calculus theorems are developed separately. Moreover, besides being non-Archimedeanly valued, the fact that the field $\mathcal{N}$ has a total order (which is also non-Archimedean) gives the field a richer structure, thus opening up new possibilities of study, like monotonicity, which are not available in other non-Archimedean valued fields like the p-adic fields for example [12]. This makes $\mathcal{N}$ an outstanding example, worthy to be studied in detail in its own right.*

The following results were proved in [16]; they show that the topological structure of $\mathcal{N}$ is different from that of $\mathbb{R}$ or $\mathbb{C}$, and that makes doing Calculus on the field more difficult.

- $(\mathcal{N}, \tau_v)$ is a totally disconnected topological space. It is Hausdorff and nowhere locally compact. There are no countable bases. The topology induced to $\mathbb{R}$ is the discrete topology. As an immediate consequence of the fact that $(\mathcal{N}, \tau_v)$ is totally disconnected, it follows that, for any $x_0 \in \mathcal{N}$, the connected component of x_0 is $\{x_0\}$; moreover, the topology is zero-dimensional, that is, there is a base of clopen sets for the topology.

- If we view $\mathcal{N}$ as an infinite dimensional vector space over $\mathbb{R}$ then τ_v is not a vector topology; that is, $(\mathcal{N}, \tau_v)$ is not a linear topological space.

- If A is compact in $(\mathcal{N}, \tau_v)$ then A is closed and bounded and it has an empty interior in $(\mathcal{N}, \tau_v)$, that is,
$$\mathrm{int}(A) := \{a \in A : \exists r > 0 \text{ in } \mathcal{N} \ni (a - r, a + r) \subset A\} = \emptyset.$$
The converse is not true: the set $A = [0, 1] \cap \mathbb{Q}$ is a (countably infinite) closed and bounded subset of $\mathcal{N}$ with an empty interior; but A is not compact in $(\mathcal{N}, \tau_v)$ [16].

- Given a sequence (x_n) of elements of $\mathcal{N}$, the series $\sum_{n=1}^{\infty} x_n$ converges if and only if the sequence (x_n) converges to zero.

2. WEAK LOCAL UNIFORM DIFFERENTIABILITY

As hinted to in the Introduction above, because of the total disconnectedness of the field $\mathcal{N}$ in the order topology, the standard theorems of real calculus like the intermediate value theorem, the inverse function theorem and the mean value theorem require stronger smoothness criteria of the functions involved in order for the theorems to hold. In this section we will present one such criterion: the so-called 'weak local uniform differentiability', we will review recent work based on that smoothness criterion and then present new results.

2.1. Review of Recent Results

In [1], we focus our attention on $\mathcal{N}$-valued functions of one variable. We study the properties of weakly locally uniformly differentiable (WLUD) functions at a point $x_0 \in \mathcal{N}$ or on an open subset A of $\mathcal{N}$. In particular, we show that WLUD functions are C^1, they include all polynomial functions, and they are closed under addition, multiplication and composition. Then we generalize the definition of weak local uniform differentiability to any order. In particular, we study the properties of WLUD^2 functions at a point $x_0 \in \mathcal{N}$ or on an open subset A of $\mathcal{N}$; and we show that WLUD^2 functions are C^2, they include all polynomial functions, and they are closed under addition, multiplication and composition. Finally, we formulate and prove an inverse function theorem as well as a local intermediate value theorem and a local mean value theorem for these functions. Here we only recall the main definitions and results (without proofs) and refer the reader to [1] for the details.

Definition 2.1. *Let $A \subseteq \mathcal{N}$ be open, let $f : A \to \mathcal{N}$, and let $x_0 \in A$ be given. We say that f is weakly locally uniformly differentiable (abbreviated as WLUD) at x_0 if f is differentiable in a neighbourhood Ω of x_0 in A and if for every $\epsilon > 0$ in $\mathcal{N}$ there exists $\delta > 0$ in $\mathcal{N}$ such that for every $x, y \in (x_0 - \delta, x_0 + \delta) \cap \Omega$ we have that $|f(y) - f(x) - f'(x)(y - x)| \leq \epsilon |y - x|$. Moreover, we say that f is WLUD on A if f is WLUD at every point in A.*

We extend the WLUD concept to higher orders of differentiability and we define WLUD^n as follows.

Definition 2.2. *Let $A \subseteq \mathcal{N}$ be open, let $f : A \to \mathcal{N}$, let $x_0 \in A$, and let $n \in \mathbb{N}$ be given. We say that f is WLUD^n at x_0 if f is n times differentiable in a neighbourhood Ω of x_0 in A and if for every $\epsilon > 0$ in $\mathcal{N}$ there exists $\delta > 0$ in $\mathcal{N}$ such that for every $x, y \in (x_0 - \delta, x_0 + \delta) \cap \Omega$ we have that*

$$\left| f(y) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y - x)^k \right| \leq \epsilon |y - x|^n .$$

Moreover, we say that f is WLUD^n on A if f is WLUD^n at every point in A.

Definition 2.3. *Let $A \subseteq \mathcal{N}$ be open, let $f : A \to \mathcal{N}$, and let $x_0 \in A$ be given. We say that f is WLUD$^\infty$ at x_0 if f is WLUD^n at x_0 for every $n \in \mathbb{N}$. Moreover, we say that f is WLUD$^\infty$ on A if f is WLUD$^\infty$ at every point in A.*

Theorem 2.4 (Inverse Function Theorem). *Let $A \subseteq \mathcal{N}$ be open, let $f : A \to \mathcal{N}$ be WLUD on A, and let $x_0 \in A$ be such that $f'(x_0) \neq 0$. Then there exists a neighborhood Ω of x_0 in A such that*

(i) $f|_\Omega$ is one-to-one;

(ii) $f(\Omega)$ is open; and

(iii) f^{-1} exists and is WLUD on $f(\Omega)$ with $(f^{-1})' = 1/\left(f' \circ f^{-1}\right)$.

Theorem 2.5 (Local Intermediate Value Theorem). *Let $A \subseteq \mathcal{N}$ be open, let $f : A \to \mathcal{N}$ be WLUD on A, and let $x_0 \in A$ be such that $f'(x_0) \neq 0$. Then there exists a neighborhood Ω of x_0 in A such that for any $a < b$ in $f(\Omega)$ and for any $c \in (a, b)$, there is an*

$$x \in \left(\min\left\{f^{(-1)}(a), f^{(-1)}(b)\right\}, \max\left\{f^{(-1)}(a), f^{(-1)}(b)\right\}\right)$$

such that $f(x) = c$.

Theorem 2.6 (Local Mean Value Theorem). *Let $A \subseteq \mathcal{N}$ be open, let $f : A \to \mathcal{N}$ be WLUD^2 on A, and let $x_0 \in A$ be such that $f''(x_0) \neq 0$. Then there exists a neighborhood Ω of x_0 in A such that f has the mean value property on Ω. That is, for every $a, b \in \Omega$ with $a < b$, there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

As in the real case, the mean value property can be used to prove other important results. In particular, while L'Hôpital's rule does not hold for differentiable functions on $\mathcal{N}$, we prove the result under similar conditions to those of the local mean value theorem. To do this we first prove the local equivalent of the Cauchy mean value theorem (Lemma 2.7). The proof is obtained from the mean value property the same way as in the real case.

Lemma 2.7. *Let $A \subset \mathcal{N}$ be open, let $f, g : A \to \mathcal{N}$ be WLUD^2 on A, and let $x_0 \in A$ be such that $f''(x_0) \neq 0$ and $g''(x_0) \neq 0$. Then there exists a neighborhood Ω of x_0 in A such that for every $a, b \in \Omega$ with $a < b$, there exists $c \in (a, b)$ such that*

$$f'(c)\,(g(b) - g(a)) = g'(c)\,(f(b) - f(a))\,.$$

Theorem 2.8 (L'Hôpital's Rule)**.** *Let $A \subset \mathcal{N}$ be open, let $f, g : A \to \mathcal{N}$ be WLUD^2 on A, and let $a \in A$ be such that $f''(a) \neq 0$ and $g''(a) \neq 0$. Furthermore, suppose that $f(a) = g(a) = 0$, that there exists a neighborhood Ω of a in A such that $g'(x) \neq 0$ for every $x \in \Omega \setminus \{a\}$, and that $\lim\limits_{x \to a} f'(x)/g'(x)$ exists. Then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Proof. By Lemma 2.7, there exists a neighborhood U of a in A such that, for every $x \in U$, there exists c between x and a such that

$$f'(c)(g(x) - g(a)) = g'(c)(f(x) - f(a)).$$

Let $\delta_1 > 0$ in $\mathcal{N}$ be such that $(a - \delta_1, a + \delta_1) \subseteq U \cap \Omega$, let $L = \lim_{x \to a} f'(x)/g'(x)$, and let $\epsilon > 0$ in $\mathcal{N}$ be given. Then there exists $\delta_2 > 0$ in $\mathcal{N}$ such that for all $x \in (a - \delta_2, a + \delta_2)$ we have that

$$\left| \frac{f'(x)}{g'(x)} - L \right| < \epsilon.$$

Let $\delta = \min\{\delta_1, \delta_2\}$, and let $x \neq a$ in $(a - \delta, a + \delta)$ be given such that $g(x) \neq 0$. Then, there exists c between x and a [thus $c \in (a - \delta, a + \delta)$] such that

$$\frac{f'(c)}{g'(c)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)}.$$

Since $c \in (a - \delta, a + \delta) \subset (a - \delta_2, a + \delta_2)$, it follows that

$$\left| \frac{f(x)}{g(x)} - L \right| = \left| \frac{f'(c)}{g'(c)} - L \right| < \epsilon.$$

□

In this paper, we will formulate and prove a Taylor theorem with remainder for WLUD^n functions. Then we will extend the concept of WLUD to functions from $\mathcal{N}^n$ to $\mathcal{N}^m$ with $m, n \in \mathbb{N}$ and study the properties of those functions as we did for functions from $\mathcal{N}$ to $\mathcal{N}$. Then we will formulate and prove the inverse function theorem for WLUD functions from $\mathcal{N}^n$ to $\mathcal{N}^n$ and the implicit function theorem for WLUD functions from $\mathcal{N}^n$ to $\mathcal{N}^m$ with $m < n$ in $\mathbb{N}$.

2.2. Taylor Theorem with Remainder

In this section, we use the concept of WLUD^k to formulate and prove a Taylor theorem with remainder. As in the real case, the proof of the theorem uses the mean value theorem. However, in the non-Archimedean setting, stronger conditions on the function are needed than in the real case.

Theorem 2.9. *(Taylor's Theorem with Remainder) Let $A \subseteq \mathcal{N}$ be open, let $n \in \mathbb{N}$ be given, and let $f : A \to \mathcal{N}$ be WLUD^{n+2} on A. Assume further that $f^{(m)}$ is WLUD^2 on A for $0 \le m \le n$. Then, for every $x \in A$, there exists a neighborhood U of x in A such that, for any $y \in U$, there exists $c \in [\min(y,x), \max(y,x)]$ such that*

$$R_n(y) := f(y) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k = \frac{f^{(n+1)}(c)}{(n+1)!}(y-x)^{n+1}. \qquad (2.1)$$

Proof. Let $x \in A$ be given. First note that Equation (2.1) holds trivially for $y = x$. For $y \neq x$ in A, define $F : A \to \mathcal{N}$ by

$$F(t) = f(y) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(y-t)^k.$$

Then $F(x) = R_n(y)$ and $F'(t) = -\frac{f^{(n+1)}(t)}{n!}(y-t)^n$ for all $t \in A$.

Now let $G : A \to \mathcal{N}$ be given by

$$G(t) = F(t) - \left(\frac{y-t}{y-x}\right)^{n+1} F(x).$$

Then $G(t)$ is WLUD^2 on A.

Case I: $G''(x) \neq 0$. Then we can apply the mean value theorem, Theorem 2.6, to G. Note that $G(x) = G(y) = 0$. Applying the mean value theorem to G, there exists a neighborhood U of x in A, such that for every $y \neq x$ in U we can find $c \in [\min(y,x), \max(y,x)]$ such that

$$\begin{aligned} 0 &= \frac{G(y)-G(x)}{y-x} = G'(c) \\ &= F'(c) + (n+1)\frac{(y-c)^n}{(y-x)^{n+1}}F(x) \\ &= -\frac{f^{(n+1)}(c)}{n!}(y-c)^n + (n+1)\frac{(y-c)^n}{(y-x)^{n+1}}F(x). \end{aligned}$$

It follows that

$$R_n(y) = F(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(y-x)^{n+1}.$$

Case II: $G''(x) = 0$. Note that, for all $t \in A$, we have that

$$\begin{aligned} G''(t) &= F''(t) - \frac{n(n+1)}{(y-x)^2}\left(\frac{y-t}{y-x}\right)^{n-1} F(x) \\ &= \frac{f^{(n+1)}(t)}{(n-1)!}(y-t)^{n-1} - \frac{f^{(n+2)}(t)}{n!}(y-t)^n - \frac{(n+1)n}{(y-x)^2}\left(\frac{y-t}{y-x}\right)^{n-1} F(x). \end{aligned}$$

Since $G''(x) = 0$, we obtain that

$$\frac{f^{(n+1)}(x)}{(n-1)!}(y-x)^{n-1} - \frac{f^{(n+2)}(x)}{n!}(y-x)^n = \frac{(n+1)n}{(y-x)^2}F(x)$$

from which we obtain that

$$f(y) - \sum_{k=0}^{n+2} \frac{f^{(k)}(x)}{k!}(y-x)^k = -\frac{2n+2}{n}\frac{f^{(n+2)}(x)}{(n+2)!}(y-x)^{n+2}. \qquad (2.2)$$

We will show that, in this case (Case II), $f^{(n+2)}(x) = 0$. Assume, to the contrary, that $f^{(n+2)}(x) \neq 0$, and let $\epsilon' = \left|\frac{f^{(n+2)}(x)}{(n+2)!}\right|$. Then $\epsilon' > 0$. Since f is WLUD^{n+2} at x, there exists a neighborhood U' of x in A such that, for every $y \in U'$, we have that

$$\left| f(y) - \sum_{k=0}^{n+2} \frac{f^{(k)}(x)}{k!} (y-x)^k \right| \leq \epsilon' \left| y - x \right|^{n+2},$$

which contradicts Equation (2.2) above. Thus, $f^{(n+2)}(x) = 0$. Then, it follows from Equation (2.2) that

$$f(y) - \sum_{k=0}^{n+1} \frac{f^{(k)}(x)}{k!} (y-x)^k = 0$$

and hence

$$R_n(y) = \frac{f^{(n+1)}(x)}{(n+1)!} (y-x)^{n+1}.$$

Note that in Case II, the result holds for any $y \in A$ and that $c = x$ in this case. □

3. LINEAR TRANSFORMATIONS FROM $\mathcal{N}^n$ TO $\mathcal{N}^m$

In this section we review the properties of linear transformations from $\mathcal{N}^n$ into $\mathcal{N}^m$, which are similar to those of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Proposition 3.1. *Let $\boldsymbol{L} : \mathcal{N}^n \to \mathcal{N}^m$ be a linear transformation. Then $\{|\boldsymbol{L}(\boldsymbol{t})| : |\boldsymbol{t}| \leq 1\}$ is bounded.*

Proof. Let $\underline{\boldsymbol{L}} = \begin{pmatrix} L_{11} & L_{12} & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{m1} & L_{m2} & \dots & L_{mn} \end{pmatrix}$ denote the matrix of the linear transformation $\boldsymbol{L}$, and let $\alpha = \max\{|L_{ij}| : i = 1, \dots, m; j = 1, \dots, n\}$. Then, for $|\boldsymbol{t}| \leq 1$, we have that

$$\begin{aligned} |\boldsymbol{L}(\boldsymbol{t})| = |\underline{\boldsymbol{L}}\boldsymbol{t}| &= \left| \begin{pmatrix} L_{11} & L_{12} & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{m1} & L_{m2} & \dots & L_{mn} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} \right| = \left| \begin{pmatrix} \sum_{j=1}^n L_{1j} t_j \\ \vdots \\ \vdots \\ \sum_{j=1}^n L_{mj} t_j \end{pmatrix} \right| \\ &= \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^n L_{ij} t_j \right)^2} \leq \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^n |L_{ij}||t_j| \right)^2} \\ &\leq \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^n \alpha \cdot 1 \right)^2} = \sqrt{\sum_{i=1}^m (n\alpha)^2} = \sqrt{m} n \alpha. \end{aligned}$$

Thus, $\{|\boldsymbol{L}(\boldsymbol{t})| : |\boldsymbol{t}| \leq 1\}$ is bounded above by $\sqrt{m} n \alpha$. □

Corollary 3.2. *Let $\boldsymbol{L} : \mathcal{N}^n \to \mathcal{N}^m$ be a linear transformation and let $\mathfrak{L}$ be an upper bound for $\{|\boldsymbol{L}(\boldsymbol{t})| : |\boldsymbol{t}| \leq 1\}$. Then $|\boldsymbol{L}(\boldsymbol{t})| \leq \mathfrak{L}|\boldsymbol{t}|$ for all $\boldsymbol{t} \in \mathcal{N}^n$.*

Proof. Let $\boldsymbol{t} \in \mathcal{N}^n$. If $\boldsymbol{t} = \boldsymbol{0}$, then $|\boldsymbol{L}(\boldsymbol{t})| = 0 = \mathfrak{L}|\boldsymbol{t}|$ and we are done. Otherwise, let $c = |\boldsymbol{t}|^{-1}$; then $|c\boldsymbol{t}| = 1$, and so $c|\boldsymbol{L}(\boldsymbol{t})| = |c\boldsymbol{L}(\boldsymbol{t})| = |\boldsymbol{L}(c\boldsymbol{t})| \leq \mathfrak{L}$. Thus,

$$|\boldsymbol{L}(\boldsymbol{t})| \leq \frac{1}{c}\mathfrak{L} = \mathfrak{L}|\boldsymbol{t}|.$$

□

Corollary 3.3. *Let $\boldsymbol{L} : \mathcal{N}^n \to \mathcal{N}^n$ be an invertible linear transformation and let $\bar{\mathfrak{L}}$ be an upper bound for $\{|\boldsymbol{L}^{-1}(\boldsymbol{t})| : |\boldsymbol{t}| \leq 1\}$. Then $|\boldsymbol{L}(\boldsymbol{t})| \geq \dfrac{|\boldsymbol{t}|}{\bar{\mathfrak{L}}}$ for all $\boldsymbol{t} \in \mathcal{N}^n$.*

Proof. First we note that, since $\boldsymbol{L}^{-1}$ is invertible, $\boldsymbol{L}^{-1}(\boldsymbol{t}) = \boldsymbol{0}$ only if $\boldsymbol{t} = \boldsymbol{0}$; and hence $\bar{\mathfrak{L}} > 0$. Now let $\boldsymbol{t} \in \mathcal{N}^n$ be given. Then $|\boldsymbol{t}| = |\boldsymbol{L}^{-1}(\boldsymbol{L}(\boldsymbol{t}))| \leq \bar{\mathfrak{L}}|\boldsymbol{L}(\boldsymbol{t})|$; and hence $|\boldsymbol{L}(\boldsymbol{t})| \geq \dfrac{|\boldsymbol{t}|}{\bar{\mathfrak{L}}}$. □

Lemma 3.4. *Let $\boldsymbol{g} : \mathcal{N}^n \to \mathcal{N}^m$ be C^1; and let $\mathfrak{L}$ be an upper bound for $\{|\boldsymbol{Dg}(\boldsymbol{x_0})(\boldsymbol{x})| : |\boldsymbol{x}| \leq 1\}$, where $\boldsymbol{Dg}(\boldsymbol{x_0})$ denotes the linear map from $\mathcal{N}^n$ to $\mathcal{N}^m$ defined by the $m \times n$ Jacobian matrix of $\boldsymbol{g}$ at $\boldsymbol{x_0}$:*

$$\begin{pmatrix} \boldsymbol{g}_1^1(\boldsymbol{x_0}) & \boldsymbol{g}_2^1(\boldsymbol{x_0}) & \dots & \boldsymbol{g}_n^1(\boldsymbol{x_0}) \\ \boldsymbol{g}_1^2(\boldsymbol{x_0}) & \boldsymbol{g}_2^2(\boldsymbol{x_0}) & \dots & \boldsymbol{g}_n^2(\boldsymbol{x_0}) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{g}_1^m(\boldsymbol{x_0}) & \boldsymbol{g}_2^m(\boldsymbol{x_0}) & \dots & \boldsymbol{g}_n^m(\boldsymbol{x_0}) \end{pmatrix}$$

with $\boldsymbol{g}_j^i(\boldsymbol{x_0}) = \frac{\partial g_i}{\partial x_j}(\boldsymbol{x_0})$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. Then, for all $\epsilon > 0$ in $\mathcal{N}$, there exists $\delta > 0$ in $\mathcal{N}$ such that $|\boldsymbol{Dg}(\boldsymbol{y})(\boldsymbol{x})| \leq (\mathfrak{L} + \epsilon)|\boldsymbol{x}|$ for all $\boldsymbol{y} \in B_\delta(\boldsymbol{x_0})$ and for all $\boldsymbol{x} \in \mathcal{N}^n$.

Proof. Let $\epsilon > 0$ in $\mathcal{N}$ be given. Since $\boldsymbol{g}$ is C^1 then for every i and j, there exists $\delta_j^i > 0$ in $\mathcal{N}$ such that $|\boldsymbol{g}_j^i(\boldsymbol{x_0}) - \boldsymbol{g}_j^i(\boldsymbol{y})| < \dfrac{\epsilon}{n\sqrt{m}}$ whenever $\boldsymbol{y} \in B_{\delta_j^i}(\boldsymbol{x_0})$. Let $\delta = \min\{\delta_j^i : i = 1, \dots, m; j = 1, \dots, n\}$. Then, using the proof of Proposition 3.1, for all $\boldsymbol{y} \in B_\delta(\boldsymbol{x_0})$ we have that $\{|\boldsymbol{Dg}(\boldsymbol{x_0})(\boldsymbol{x}) - \boldsymbol{Dg}(\boldsymbol{y})(\boldsymbol{x})| : |\boldsymbol{x}| \leq 1\}$ is bounded above by ϵ. Thus, by Corollary 3.2, we have that $|\boldsymbol{Dg}(\boldsymbol{x_0})(\boldsymbol{x}) - \boldsymbol{Dg}(\boldsymbol{y})(\boldsymbol{x})| \leq \epsilon|\boldsymbol{x}|$ for all $\boldsymbol{y} \in B_\delta(\boldsymbol{x_0})$ and for all $\boldsymbol{x} \in \mathcal{N}^n$. Therefore $|\boldsymbol{Dg}(\boldsymbol{y})(\boldsymbol{x})| \leq \epsilon|\boldsymbol{x}| + |\boldsymbol{Dg}(\boldsymbol{x_0})(\boldsymbol{x})|$; and hence $|\boldsymbol{Dg}(\boldsymbol{y})(\boldsymbol{x})| \leq (\epsilon + \mathfrak{L})|\boldsymbol{x}|$ for all $\boldsymbol{y} \in B_\delta(\boldsymbol{x_0})$ and for all $\boldsymbol{x} \in \mathcal{N}^n$. □

4. WLUD FUNCTIONS FROM $\mathcal{N}^n$ TO $\mathcal{N}^m$

In the rest of the paper, let A denote an open subset of $\mathcal{N}^n$; consequently, whenever we speak of a ball $B_\delta(\boldsymbol{x})$ around a point $\boldsymbol{x}$ in A, it is assumed that $\delta > 0$ is small enough so that $B_\delta(\boldsymbol{x}) \subset A$.

Definition 4.1 (Uniformly Differentiable). *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be differentiable on A. Then we say that $\boldsymbol{f}$ is uniformly differentiable on A if for all $\epsilon > 0$ in $\mathcal{N}$, there exists $\delta > 0$ in $\mathcal{N}$ such that whenever $\boldsymbol{x}, \boldsymbol{y} \in A$ and $|\boldsymbol{y} - \boldsymbol{x}| < \delta$ we have that $|\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| \leq \epsilon|\boldsymbol{y} - \boldsymbol{x}|$.*

Definition 4.2 (Weakly Locally Uniformly Differentiable). *Let $A \subset \mathcal{N}^n$ be open, let $\boldsymbol{f} : A \to \mathcal{N}^m$, and let $\boldsymbol{x_0} \in A$ be given. Then we say that $\boldsymbol{f}$ is weakly locally uniformly differentiable (WLUD) at $\boldsymbol{x_0}$ if $\boldsymbol{f}$ is differentiable in a neighborhood Ω of $\boldsymbol{x_0}$ in A and if for every $\epsilon > 0$ in $\mathcal{N}$ there exists $\delta > 0$ in $\mathcal{N}$ such that for all $\boldsymbol{x}, \boldsymbol{y} \in B_\delta(\boldsymbol{x_0}) \cap \Omega$, we have that*

$$|\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| \leq \epsilon|\boldsymbol{y} - \boldsymbol{x}|.$$

Moreover, we say that $\boldsymbol{f}$ is WLUD on A if $\boldsymbol{f}$ is WLUD at every point in A.

Remark 4.3. *It is clear from the two definitions above that if $\boldsymbol{f}$ is uniformly differentiable on A then $\boldsymbol{f}$ is WLUD at every point in A and hence $\boldsymbol{f}$ is WLUD on A.*

Proposition 4.4. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be differentiable at $\boldsymbol{x} \in A$. Then $\boldsymbol{f}$ is continuous at $\boldsymbol{x}$.*

Proof. Let $\mathfrak{L}_\mathfrak{r} > 0$ be an upper bound for $\{|\boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y})| : |\boldsymbol{y}| \leq 1\}$. Since $\boldsymbol{f}$ is differentiable at $\boldsymbol{x}$, there exists $\delta_0 > 0$ in $\mathcal{N}$ such that whenever $\boldsymbol{y} \in B_{\delta_0}(\boldsymbol{x})$, we have that $|\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| \leq \mathfrak{L}_\mathfrak{r}|\boldsymbol{y} - \boldsymbol{x}|$. Let $\epsilon > 0$ in $\mathcal{N}$ be given. Let $\delta = \min\{\delta_0, \frac{\epsilon}{2\mathfrak{L}_\mathfrak{r}}\}$. Then for $\boldsymbol{y} \in B_\delta(\boldsymbol{x})$ we have that

$$\begin{aligned}
|\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})| &= |\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}) + \boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| \\
&\leq |\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| + |\boldsymbol{Df}(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| \\
&\leq \mathfrak{L}_\mathfrak{r}|\boldsymbol{y} - \boldsymbol{x}| + \mathfrak{L}_\mathfrak{r}|\boldsymbol{y} - \boldsymbol{x}| \\
&= 2\mathfrak{L}_\mathfrak{r}|\boldsymbol{y} - \boldsymbol{x}| \\
&< 2\mathfrak{L}_\mathfrak{r}\delta \\
&\leq \epsilon.
\end{aligned}$$

□

Corollary 4.5. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be differentiable on A. Then $\boldsymbol{f}$ is continuous on A.*

Theorem 4.6. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be WLUD at $\boldsymbol{x_0} \in A$. Then $\boldsymbol{f}$ is C^1 at $\boldsymbol{x_0}$.*

Proof. Let $\epsilon > 0$ in $\mathcal{N}$ be given. Then there exists $\delta_1 > 0$ in $\mathcal{N}$ such that $\boldsymbol{f}$ is differentiable on $B_{\delta_1}(\boldsymbol{x_0})$ and, for $\boldsymbol{s}, \boldsymbol{t} \in B_{\delta_1}(\boldsymbol{x_0})$, we have that

$$|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{Df}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \leq \frac{\epsilon}{4}|\boldsymbol{s} - \boldsymbol{t}|.$$

Let $\delta_2 = \delta_1/2$ and let $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ be given. Then, for any $\boldsymbol{x} \in B_{\delta_2}(\boldsymbol{x_0})$, we have that

$$|f^i(\boldsymbol{x} + \delta_2\hat{\boldsymbol{e_j}}) - f^i(\boldsymbol{x}) - \boldsymbol{D}f^i(\boldsymbol{x})(\delta_2\hat{\boldsymbol{e_j}})| \leq \frac{\epsilon\delta_2}{4}.$$

That is,

$$|f^i(\boldsymbol{x} + \delta_2\hat{\boldsymbol{e_j}}) - f^i(\boldsymbol{x}) - f^i_j(\boldsymbol{x})\delta_2| \leq \frac{\epsilon\delta_2}{4}.$$

Now, since f^i is continuous on $B_{\delta_1}(\boldsymbol{x_0})$, there exists $\delta_3 > 0$ such that, for all $\boldsymbol{s} \in B_{\delta_3}(\boldsymbol{x_0})$, we have that

$$|f^i(\boldsymbol{s}) - f^i(\boldsymbol{x_0})| \leq \frac{\epsilon\delta_2}{4}.$$

Additionally, there exists $\delta_4 > 0$ such that, for all $\boldsymbol{s} \in B_{\delta_4}(\boldsymbol{x_0} + \delta_2\hat{\boldsymbol{e_j}})$, we have that

$$|f^i(\boldsymbol{s}) - f^i(\boldsymbol{x_0} + \delta_2\hat{\boldsymbol{e_j}})| \leq \frac{\epsilon\delta_2}{4}.$$

Let $\delta = \min\{\delta_2, \delta_3, \delta_4\}$ and let $\boldsymbol{y} \in B_\delta(\boldsymbol{x_0})$ be given. Then we have that

$$\begin{aligned}
|f^i_j(\boldsymbol{y})\delta_2 - f^i_j(\boldsymbol{x_0})\delta_2| &= |f^i(\boldsymbol{x_0} + \delta_2\hat{\boldsymbol{e_j}}) - f^i(\boldsymbol{x_0}) - f^i_j(\boldsymbol{x_0})\delta_2 \\
&\quad + f^i_j(\boldsymbol{y})\delta_2 - f^i(\boldsymbol{y} + \delta_2\hat{\boldsymbol{e_j}}) + f^i(\boldsymbol{y}) \\
&\quad + f^i(\boldsymbol{x_0}) - f^i(\boldsymbol{y}) - f^i(\boldsymbol{x_0} + \delta_2\hat{\boldsymbol{e_j}}) + f^i(\boldsymbol{y} + \delta_2\hat{\boldsymbol{e_j}})| \\
&\leq |f^i(\boldsymbol{x_0} + \delta_2\hat{\boldsymbol{e_j}}) - f^i(\boldsymbol{x_0}) - f^i_j(\boldsymbol{x_0})\delta_2| \\
&\quad + |f^i(\boldsymbol{y} + \delta_2\hat{\boldsymbol{e_j}}) - f^i(\boldsymbol{y}) - f^i_j(\boldsymbol{y})\delta_2| \\
&\quad + |f^i(\boldsymbol{x_0}) - f^i(\boldsymbol{y})| + |f^i(\boldsymbol{x_0} + \delta_2\hat{\boldsymbol{e_j}}) - f^i(\boldsymbol{y} + \delta_2\hat{\boldsymbol{e_j}})| \\
&\leq \frac{\epsilon\delta_2}{4} + \frac{\epsilon\delta_2}{4} + \frac{\epsilon\delta_2}{4} + \frac{\epsilon\delta_2}{4} \\
&= \epsilon\delta_2.
\end{aligned}$$

Thus, for all $\boldsymbol{y} \in B_\delta(\boldsymbol{x_0})$, we have that

$$|f_j^i(\boldsymbol{y}) - f_j^i(\boldsymbol{x_0})| \le \epsilon.$$

Thus, f_j^i is continuous at $\boldsymbol{x_0}$ for all $i \in \{1, \ldots, m\}$ and for all $j \in \{1, \ldots, n\}$; and hence $\boldsymbol{f}$ is C^1 at $\boldsymbol{x_0}$. □

Corollary 4.7. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be WLUD on A. Then $\boldsymbol{f}$ is C^1 on A.*

Remark 4.8. *Theorem 4.6 and Corollary 4.7 show that the class of WLUD functions at a point $\boldsymbol{x_0}$ (respectively on an open set A) is a subset of the class of C^1 functions at $\boldsymbol{x_0}$ (respectively on A). However, this is still large enough to include all polynomial functions as Corollary 4.19 and Corollary 4.20 below will show.*

Lemma 4.9. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be WLUD at $\boldsymbol{x_0} \in A$. Then*

$$\forall \epsilon > 0\ \exists \delta > 0 \ni (\boldsymbol{s}, \boldsymbol{t} \in B_\delta(\boldsymbol{x_0}) \Rightarrow |\boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{f}(\boldsymbol{s}) - D\boldsymbol{f}(\boldsymbol{x_0})(\boldsymbol{t} - \boldsymbol{s})| \le \epsilon |\boldsymbol{t} - \boldsymbol{s}|). \tag{4.1}$$

Proof. Let $\epsilon > 0$ in $\mathcal{N}$ be given. Since $\boldsymbol{f}$ is WLUD at $\boldsymbol{x_0}$, there exists $\delta_0 > 0$ in $\mathcal{N}$ such that $\boldsymbol{f}$ is differentiable on $B_{\delta_0}(\boldsymbol{x_0})$ and, for all $\boldsymbol{s}, \boldsymbol{t} \in B_{\delta_0}(\boldsymbol{x_0})$, we have that

$$|\boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{s})(\boldsymbol{t} - \boldsymbol{s})| \le \frac{\epsilon}{2} |\boldsymbol{t} - \boldsymbol{s}|.$$

Moreover, since $\boldsymbol{f}$ is C^1 at $\boldsymbol{x_0}$ by Theorem 4.6, there exists $\delta > 0$ in $\mathcal{N}$, $\delta \le \delta_0$, such that for all $\boldsymbol{s} \in B_\delta(\boldsymbol{x_0})$, we have that

$$|\boldsymbol{Df}(\boldsymbol{s})(\boldsymbol{t} - \boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{x_0})(\boldsymbol{t} - \boldsymbol{s})| \le \frac{\epsilon}{2} |\boldsymbol{t} - \boldsymbol{s}|.$$

Thus, for all $\boldsymbol{s}, \boldsymbol{t} \in B_\delta(\boldsymbol{x_0}) \subseteq B_{\delta_0}(\boldsymbol{x_0})$, we have that

$$\begin{aligned}
|\boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{x_0})(\boldsymbol{t} - \boldsymbol{s})| &= |\boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{s})(\boldsymbol{t} - \boldsymbol{s}) \\
&\quad + \boldsymbol{Df}(\boldsymbol{s})(\boldsymbol{t} - \boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{x_0})(\boldsymbol{t} - \boldsymbol{s})| \\
&\le |\boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{s})(\boldsymbol{t} - \boldsymbol{s})| \\
&\quad + |\boldsymbol{Df}(\boldsymbol{s})(\boldsymbol{t} - \boldsymbol{s}) - \boldsymbol{Df}(\boldsymbol{x_0})(\boldsymbol{t} - \boldsymbol{s})|. \\
&\le \frac{\epsilon}{2} |\boldsymbol{t} - \boldsymbol{s}| + \frac{\epsilon}{2} |\boldsymbol{t} - \boldsymbol{s}| \\
&= \epsilon |\boldsymbol{t} - \boldsymbol{s}|.
\end{aligned}$$

Thus $\boldsymbol{f}$ satisfies (4.1). □

Proposition 4.10. *Let $\boldsymbol{L} : \mathcal{N}^n \to \mathcal{N}^m$ be a linear transformation. Then $\boldsymbol{L}$ is uniformly differentiable, and hence WLUD, on $\mathcal{N}^n$.*

Proof. As in the real case, $\boldsymbol{L}$ is differentiable with $\boldsymbol{DL}(\boldsymbol{x}) = \boldsymbol{L}$ for all $\boldsymbol{x} \in \mathcal{N}^n$. Let $\epsilon > 0$ in $\mathcal{N}$ be given. Then for any $\boldsymbol{s}, \boldsymbol{t} \in \mathcal{N}^n$ we have that

$$\begin{aligned}
|\boldsymbol{L}(\boldsymbol{s}) - \boldsymbol{L}(\boldsymbol{t}) - \boldsymbol{DL}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| &= |\boldsymbol{L}(\boldsymbol{s}) - \boldsymbol{L}(\boldsymbol{t}) - \boldsymbol{L}(\boldsymbol{s} - \boldsymbol{t})| \\
&= |\boldsymbol{L}(\boldsymbol{s} - \boldsymbol{t}) - \boldsymbol{L}(\boldsymbol{s} - \boldsymbol{t})| = 0.
\end{aligned}$$

Thus $\boldsymbol{L}$ is uniformly differentiable on $\mathcal{N}^n$. □

Proposition 4.11. *Let $\boldsymbol{f}, \boldsymbol{g} : A \to \mathcal{N}^m$ be WLUD at $\boldsymbol{x_0} \in A$; and let $\alpha \in \mathcal{N}$ be given. Then $\alpha\boldsymbol{f} + \boldsymbol{g}$ is WLUD at $\boldsymbol{x_0}$. That is, any linear combination of WLUD functions at $\boldsymbol{x_0}$ is again WLUD at $\boldsymbol{x_0}$.*

Proof. If $\alpha = 0$ then there is nothing to prove; so without loss of generality we may assume $\alpha \neq 0$. Since $\boldsymbol{f}$ and $\boldsymbol{g}$ are both WLUD at $\boldsymbol{x_0}$, there exists $\delta_0 >$ in $\mathcal{N}$ such that $B_{\delta_0}(\boldsymbol{x_0}) \subset A$ and such that $\boldsymbol{f}$ and $\boldsymbol{g}$ are differentiable on $B_{\delta_0}(\boldsymbol{x_0})$. It follows that $\alpha\boldsymbol{f} + \boldsymbol{g}$ is differentiable on $B_{\delta_0}(\boldsymbol{x_0})$, with $\boldsymbol{D}(\alpha\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{x}) = \alpha\boldsymbol{D}\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{D}\boldsymbol{g}(\boldsymbol{x})$ for all $\boldsymbol{x} \in B_{\delta_0}(\boldsymbol{x_0})$.

Now let $\epsilon > 0$ in $\mathcal{N}$ be given. Then there exists $\delta_{\boldsymbol{f}} > 0$ in $\mathcal{N}$, $\delta_{\boldsymbol{f}} \leq \delta_0$, such that for all $\boldsymbol{s}, \boldsymbol{t} \in B_{\delta_{\boldsymbol{f}}}(\boldsymbol{x_0})$, we have that

$$|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{D}\boldsymbol{f}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \leq \frac{\epsilon}{2|\alpha|}|\boldsymbol{s} - \boldsymbol{t}|.$$

Also, there exists $\delta_{\boldsymbol{g}} > 0$ in $\mathcal{N}$, $\delta_{\boldsymbol{g}} \leq \delta_0$, such that for all $\boldsymbol{s}, \boldsymbol{t} \in B_{\delta_{\boldsymbol{g}}}(\boldsymbol{x_0})$, we have that

$$|\boldsymbol{g}(\boldsymbol{s}) - \boldsymbol{g}(\boldsymbol{t}) - \boldsymbol{D}\boldsymbol{g}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \leq \frac{\epsilon}{2}|\boldsymbol{s} - \boldsymbol{t}|.$$

Let $\delta = \min\{\delta_{\boldsymbol{f}}, \delta_{\boldsymbol{g}}\}$. Then $0 < \delta \leq \delta_0$ and, for all $\boldsymbol{s}, \boldsymbol{t} \in B_\delta(\boldsymbol{x_0})$, we have that

$$\begin{aligned}
&|(\alpha\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{s}) - (\alpha\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{t}) - \boldsymbol{D}(\alpha\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \\
=&|\alpha\boldsymbol{f}(\boldsymbol{s}) + \boldsymbol{g}(\boldsymbol{s}) - (\alpha\boldsymbol{f}(\boldsymbol{t}) + \boldsymbol{g}(\boldsymbol{t})) - \alpha\boldsymbol{D}\boldsymbol{f}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t}) - \boldsymbol{D}\boldsymbol{g}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \\
\leq&|\alpha\left[\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{D}\boldsymbol{f}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})\right]| + |\boldsymbol{g}(\boldsymbol{s}) - \boldsymbol{g}(\boldsymbol{t}) - \boldsymbol{D}\boldsymbol{g}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \\
\leq&|\alpha|\frac{\epsilon}{2|\alpha|}|\boldsymbol{s} - \boldsymbol{t}| + \frac{\epsilon}{2}|\boldsymbol{s} - \boldsymbol{t}| \\
=&\epsilon|\boldsymbol{s} - \boldsymbol{t}|.
\end{aligned}$$

□

Corollary 4.12. *Let $\boldsymbol{f}, \boldsymbol{g} : A \to \mathcal{N}^m$ be WLUD on A; and let $\alpha \in \mathcal{N}$ be given. Then $\alpha\boldsymbol{f} + \boldsymbol{g}$ is WLUD on A.*

Theorem 4.13. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be WLUD at $\boldsymbol{x_0} \in A$ and let $\boldsymbol{g} : C \to \mathcal{N}^p$ be WLUD at $\boldsymbol{f}(\boldsymbol{x_0}) \in C$, where A is an open subset of $\mathcal{N}^n$, C an open subset of $\mathcal{N}^m$ and $\boldsymbol{f}(A) \subseteq C$. Then $\boldsymbol{g} \circ \boldsymbol{f}$ is WLUD at $\boldsymbol{x_0}$.*

Proof. There exists $\delta_1 > 0$ in $\mathcal{N}$ such that $B_{\delta_1}(\boldsymbol{x_0}) \subset A$ and $\boldsymbol{f}$ is differentiable on $B_{\delta_1}(\boldsymbol{x_0})$, and there exists $\delta_2 > 0$ in $\mathcal{N}$ such that $B_{\delta_2}(\boldsymbol{f}(\boldsymbol{x_0})) \subset C$ and $\boldsymbol{g}$ is differentiable on $B_{\delta_2}(\boldsymbol{f}(\boldsymbol{x_0}))$. Since $\boldsymbol{f}$ is continuous at $\boldsymbol{x_0}$ by Proposition 4.4, we may assume that δ_1 is small enough so that $\boldsymbol{f}(B_{\delta_1}(\boldsymbol{x_0})) \subset B_{\delta_2}(\boldsymbol{f}(\boldsymbol{x_0}))$. As in the real case, it follows that $\boldsymbol{g} \circ \boldsymbol{f}$ is differentiable on $B_{\delta_1}(\boldsymbol{x_0})$, with

$$\boldsymbol{D}(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = \boldsymbol{D}\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) \circ \boldsymbol{D}\boldsymbol{f}(\boldsymbol{x})$$

for all $\boldsymbol{x} \in B_{\delta_1}(\boldsymbol{x_0})$.

Let $\mathfrak{L}_{\boldsymbol{f}} > 0$ be an upper bound for $\{|\boldsymbol{D}\boldsymbol{f}(\boldsymbol{x_0})(\boldsymbol{y})| : |\boldsymbol{y}| \leq 1\}$ and let $\mathfrak{L}_{\boldsymbol{g}} \geq 1$ be an upper bound for $\{|\boldsymbol{D}\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x_0}))(\boldsymbol{y})| : |\boldsymbol{y}| \leq 1\}$. By Lemma 3.4, there exists $\delta_3 > 0$ in $\mathcal{N}$, $\delta_3 \leq \delta_1$, such that whenever $\boldsymbol{s} \in B_{\delta_3}(\boldsymbol{x_0})$ we have that $|\boldsymbol{D}\boldsymbol{f}(\boldsymbol{s})(\boldsymbol{y})| \leq 2\mathfrak{L}_{\boldsymbol{f}}|\boldsymbol{y}|$ for all $\boldsymbol{y} \in \mathcal{N}^n$. Similarly, there exists $\delta_4 > 0$ in $\mathcal{N}$, $\delta_4 \leq \delta_2$, such that whenever $\boldsymbol{u} \in B_{\delta_4}(\boldsymbol{f}(\boldsymbol{x_0}))$ we have that $|\boldsymbol{D}\boldsymbol{g}(\boldsymbol{u})(\boldsymbol{v})| \leq 2\mathfrak{L}_{\boldsymbol{g}}|\boldsymbol{v}|$ for all $\boldsymbol{v} \in \mathcal{N}^m$.

Now let $\epsilon > 0$ in $\mathcal{N}$ be given. By definition, there exists $\delta_{\boldsymbol{g}} > 0$ in $\mathcal{N}$, $\delta_g \leq \delta_4 \leq \delta_2$, such that

$$|\boldsymbol{g}(\boldsymbol{u}) - \boldsymbol{g}(\boldsymbol{v}) - \boldsymbol{D}\boldsymbol{g}(\boldsymbol{v})(\boldsymbol{u} - \boldsymbol{v})| \leq \frac{\epsilon}{2(\epsilon + 2\mathfrak{L}_{\boldsymbol{f}})}|\boldsymbol{u} - \boldsymbol{v}|$$

for all $\boldsymbol{u}, \boldsymbol{v} \in B_{\delta_g}(\boldsymbol{f}(\boldsymbol{x_0}))$. Also, there exists $\delta > 0$ in $\mathcal{N}$, $\delta \leq \delta_3 \leq \delta_1$, such that

$$|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{D}\boldsymbol{f}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \leq \frac{\epsilon}{4\mathfrak{L}_{\boldsymbol{g}}}|\boldsymbol{s} - \boldsymbol{t}|$$

for all $\boldsymbol{s}, \boldsymbol{t} \in B_\delta(\boldsymbol{x_0})$. Again, by the continuity of $\boldsymbol{f}$ at $\boldsymbol{x_0}$, we may assume that δ is small enough so that $\boldsymbol{f}(B_\delta(\boldsymbol{x_0})) \subset B_{\delta_{\boldsymbol{g}}}(\boldsymbol{f}(\boldsymbol{x_0}))$.

Let $\boldsymbol{s}, \boldsymbol{t} \in B_\delta(\boldsymbol{x_0})$ be given. Then we have that

$$|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t})| \leq \frac{\epsilon}{4\mathfrak{L}_{\boldsymbol{g}}}|\boldsymbol{s} - \boldsymbol{t}| + |\boldsymbol{D}\boldsymbol{f}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \leq (\epsilon + 2\mathfrak{L}_{\boldsymbol{f}})|\boldsymbol{s} - \boldsymbol{t}|.$$

Also, we have that $\boldsymbol{f}(\boldsymbol{s}), \boldsymbol{f}(\boldsymbol{t}) \in B_{\delta_{\boldsymbol{g}}}(\boldsymbol{f}(\boldsymbol{x_0})) \subseteq B_{\delta_4}(\boldsymbol{f}(\boldsymbol{x_0}))$; and hence

$$|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{s})) - \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{t})) - \boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t}))(\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}))| \leq \frac{\epsilon}{2(\epsilon + 2\mathfrak{L}_{\boldsymbol{f}})}|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t})|.$$

Therefore, for all $\boldsymbol{s}, \boldsymbol{t} \in B_\delta(\boldsymbol{x_0})$, we have that

$$\begin{aligned}
&|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{s})) - \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{t})) - \boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t})) \circ \boldsymbol{Df}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \\
=&|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{s})) - \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{t})) - \boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t}))(\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t})) \\
&+ \boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t}))(\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t})) - \boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t})) \circ \boldsymbol{Df}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \\
\leq&|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{s})) - \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{t})) - \boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t}))(\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}))| \\
&+ |\boldsymbol{Dg}(\boldsymbol{f}(\boldsymbol{t}))(\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{Df}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t}))| \\
\leq&\frac{\epsilon}{2(\epsilon + 2\mathfrak{L}_{\boldsymbol{f}})}|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t})| + 2\mathfrak{L}_{\boldsymbol{g}}|\boldsymbol{f}(\boldsymbol{s}) - \boldsymbol{f}(\boldsymbol{t}) - \boldsymbol{Df}(\boldsymbol{t})(\boldsymbol{s} - \boldsymbol{t})| \\
\leq&\frac{\epsilon}{2}|\boldsymbol{s} - \boldsymbol{t}| + 2\mathfrak{L}_{\boldsymbol{g}}\frac{\epsilon}{4\mathfrak{L}_{\boldsymbol{g}}}|\boldsymbol{s} - \boldsymbol{t}| \\
=&\frac{\epsilon}{2}|\boldsymbol{s} - \boldsymbol{t}| + \frac{\epsilon}{2}|\boldsymbol{s} - \boldsymbol{t}| \\
=&\epsilon|\boldsymbol{s} - \boldsymbol{t}|.
\end{aligned}$$

□

Corollary 4.14. *Let $\boldsymbol{f} : A \to \mathcal{N}^m$ be WLUD on A and let $\boldsymbol{g} : C \to \mathcal{N}^p$ be WLUD on C, with $\boldsymbol{f}(A) \subseteq C$. Then $\boldsymbol{g} \circ \boldsymbol{f}$ is WLUD on A.*

Lemma 4.15. *Let $h : \mathcal{N}^2 \to \mathcal{N}$ be given by $h(x_1, x_2) = x_1 x_2$. Then h is uniformly differentiable, and hence WLUD, on $\mathcal{N}^2$; with $\boldsymbol{D}h(x_1, x_2) = (x_2 \;\; x_1)$.*

Proof. Let $\epsilon > 0$ in $\mathcal{N}$ be given. Let $\delta = \epsilon$. Then for all $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ in $\mathcal{N}^2$ satisfying $|\boldsymbol{y} - \boldsymbol{x}| < \delta$, we have that

$$\begin{aligned}
|h(\boldsymbol{y}) - h(\boldsymbol{x}) - \boldsymbol{D}h(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})| &= |h(\boldsymbol{y}) - h(\boldsymbol{x}) - (x_2 \;\; x_1)(\boldsymbol{y} - \boldsymbol{x})| \\
&= |y_1 y_2 - x_1 x_2 - x_2(y_1 - x_1) - x_1(y_2 - x_2)| \\
&= |y_1 y_2 - x_2 y_1 - x_1 y_2 + x_1 x_2| \\
&= |(y_1 - x_1)(y_2 - x_2)| \\
&\leq |\boldsymbol{y} - \boldsymbol{x}|^2 \\
&\leq \epsilon|\boldsymbol{y} - \boldsymbol{x}|.
\end{aligned}$$

□

Proposition 4.16. *Let $f, g : A \to \mathcal{N}$ be WLUD at $\boldsymbol{x_0} \in A$ (where A is, as before, an open subset of $\mathcal{N}^n$). Then fg is WLUD at $\boldsymbol{x_0}$.*

Proof. Define $\boldsymbol{k} : A \to \mathcal{N}^2$ by

$$\boldsymbol{k}(\boldsymbol{x}) = \begin{pmatrix} f(\boldsymbol{x}) \\ g(\boldsymbol{x}) \end{pmatrix};$$

and let $h : \mathcal{N}^2 \to \mathcal{N}$ be as in Lemma 4.15 above. Then $\boldsymbol{k}$ is WLUD at $\boldsymbol{x_0}$ and h is WLUD at $\boldsymbol{k}(\boldsymbol{x_0})$. It follows from Theorem 4.13 that $fg = h \circ \boldsymbol{k}$ is WLUD at $\boldsymbol{x_0}$. □

Corollary 4.17. *Let $f, g : A \to \mathcal{N}$ be WLUD on A. Then fg is WLUD on A.*

Lemma 4.18. *For each $j \in \{1, \ldots, n\}$, the function $f_j : \mathcal{N}^n \to \mathcal{N}$, given by*

$$f_j(x_1, x_2, \ldots, x_n) = x_j,$$

is uniformly differentiable, and hence WLUD, on $\mathcal{N}^n$.

Proof. Let $j \in \{1, \ldots, n\}$ be given. Then for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{N}^n$ and for all $\alpha, \beta \in \mathcal{N}$, we have that

$$f_j(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = (\alpha \boldsymbol{x} + \beta \boldsymbol{y})_j = \alpha x_j + \beta y_j = \alpha f_j(\boldsymbol{x}) + \beta f_j(\boldsymbol{y}).$$

Hence f_j is a linear transformation from $\mathcal{N}^n$ to $\mathcal{N}$. It follows from Proposition 4.10 that f_j is uniformly differentiable on $\mathcal{N}^n$. □

Using the results of Proposition 4.16 and Lemma 4.18, we infer that any monomial function is WLUD on $\mathcal{N}^n$. It then follows from Proposition 4.11 that any polynomial function is WLUD on $\mathcal{N}^n$.

Corollary 4.19. *Let $f : \mathcal{N}^n \to \mathcal{N}$ be a polynomial function. Then f is WLUD on $\mathcal{N}^n$.*

Corollary 4.20. *Let $\boldsymbol{f} : \mathcal{N}^n \to \mathcal{N}^m$ be given by*

$$\boldsymbol{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$$

with f_i a polynomial function from $\mathcal{N}^n$ to $\mathcal{N}$ for all $i \in \{1, \ldots, m\}$. Then $\boldsymbol{f}$ is WLUD on $\mathcal{N}^n$.

4.1. Inverse Function Theorem

We start this section with some preliminary results needed to prove the inverse function theorem. Let $\delta_1 > 0$ in $\mathcal{N}$ be given, let $c \in \mathcal{N}$ be such that $0 < c \ll 1$, and let $\boldsymbol{\phi} : B_{\delta_1}(\boldsymbol{0}) \subset \mathcal{N}^n \to \mathcal{N}^n$ be such that

$$|\boldsymbol{\phi}(\boldsymbol{t})| \le c|\boldsymbol{t}| \text{ for all } \boldsymbol{t} \in B_{\delta_1}(\boldsymbol{0}). \tag{4.2}$$

Then $\boldsymbol{\phi}(B_{\delta_1}(\boldsymbol{0})) \subseteq B_{\delta_1}(\boldsymbol{0})$. For $m \in \mathbb{N}$ let $\boldsymbol{\phi}^{[m]} = \underbrace{\boldsymbol{\phi} \circ \cdots \circ \boldsymbol{\phi}}_{m \text{ times}}$ and set $\boldsymbol{\phi}^{[0]} = \boldsymbol{I}$ (the identity map on $\mathcal{N}^n$).

Using induction, it can be shown that, for all $m \in \mathbb{N}$, we have that

(a) $\boldsymbol{\phi}^{[m]}(B_{\delta_1}(\boldsymbol{0})) \subseteq B_{\delta_1}(\boldsymbol{0})$; and

(b) $|\boldsymbol{\phi}^{[m]}(\boldsymbol{t})| \le c^m |\boldsymbol{t}|$ for all $\boldsymbol{t} \in B_{\delta_1}(\boldsymbol{0})$.

Lemma 4.21. *Let $\delta_1 > 0$ in $\mathcal{N}$ be given, let $c \in \mathcal{N}$ be such that $0 < c \ll 1$, and let $\boldsymbol{\phi} : B_{\delta_1}(\boldsymbol{0}) \subset \mathcal{N}^n \to \mathcal{N}^n$ be continuous on $B_{\delta_1}(\boldsymbol{0})$ and satisfy (4.2). Let $\delta \in \mathcal{N}$ be such that $0 < \delta \le (1 - c)\delta_1$ and let $\boldsymbol{\psi}(\boldsymbol{t}) = \sum_{m=0}^{\infty} \boldsymbol{\phi}^{[m]}(\boldsymbol{t})$, for all $\boldsymbol{t} \in B_\delta(\boldsymbol{0})$. Then, for all $\boldsymbol{t} \in B_\delta(\boldsymbol{0})$, we have that*

- $|\boldsymbol{\psi}(\boldsymbol{t})| \le \dfrac{|\boldsymbol{t}|}{1 - c}$*; and*
- $\boldsymbol{\psi}(\boldsymbol{t}) - \boldsymbol{\phi}[\boldsymbol{\psi}(\boldsymbol{t})] = \boldsymbol{t}$.

Proof. Let $t \in B_\delta(\boldsymbol{0})$ be given. By (b) above, we have that $|\boldsymbol{\phi}^{[m]}(\boldsymbol{t})| \le c^m |\boldsymbol{t}|$ for all $m \in \mathbb{N}$. Also, we have that $\lim\limits_{m \to \infty} c^m = 0$ since $c \ll 1$ and the skeleton group of $\mathcal{N}$ is Archimedean. Thus, $\lim\limits_{m \to \infty} |\boldsymbol{\phi}^{[m]}(\boldsymbol{t})| = 0$ and hence $\sum_{m=0}^{\infty} \boldsymbol{\phi}^{[m]}(\boldsymbol{t})$ converges in $\mathcal{N}$. For each $r \in \mathbb{N}$, let $\boldsymbol{\psi}_r(\boldsymbol{t}) = \sum_{m=0}^{r} \boldsymbol{\phi}^{[m]}(\boldsymbol{t})$. Then

$$|\boldsymbol{\psi}_r(\boldsymbol{t})|, |\boldsymbol{\psi}(\boldsymbol{t})| \le \sum_{m=0}^{\infty} |\boldsymbol{\phi}^{[m]}(\boldsymbol{t})| \le |\boldsymbol{t}| \sum_{m=0}^{\infty} c^m = \frac{|\boldsymbol{t}|}{1 - c}.$$

Therefore, $\boldsymbol{\psi}_r(\boldsymbol{t}), \boldsymbol{\psi}(\boldsymbol{t}) \in B_{\delta_1}(\mathbf{0})$ for all $r \in \mathbb{N}$. Furthermore,

$$\boldsymbol{\psi}_r - \boldsymbol{\phi} \circ \boldsymbol{\psi}_r = \sum_{m=0}^{r} \boldsymbol{\phi}^{[m]} - \sum_{m=1}^{r+1} \boldsymbol{\phi}^{[m]} = \boldsymbol{I} - \boldsymbol{\phi}^{[r+1]};$$

and hence

$$\boldsymbol{\psi}_r(\boldsymbol{t}) - \boldsymbol{\phi}[\boldsymbol{\psi}_r(\boldsymbol{t})] = \boldsymbol{t} - \boldsymbol{\phi}^{[r+1]}(\boldsymbol{t}). \tag{4.3}$$

It is readily seen that $\lim_{r\to\infty} \boldsymbol{\psi}_r(\boldsymbol{t}) = \boldsymbol{\psi}(\boldsymbol{t})$, and so $\lim_{r\to\infty} \boldsymbol{\phi}[\boldsymbol{\psi}_r(\boldsymbol{t})] = \boldsymbol{\phi}[\boldsymbol{\psi}(\boldsymbol{t})]$ since $\boldsymbol{\phi}$ is continuous on $B_\delta(\mathbf{0})$. Also, $\lim_{r\to\infty} \boldsymbol{\phi}^{[r]}(\boldsymbol{t}) = \mathbf{0}$. Thus, by letting $r \to \infty$ on both sides of Equation (4.3), we obtain that $\boldsymbol{\psi}(\boldsymbol{t}) - \boldsymbol{\phi}[\boldsymbol{\psi}(\boldsymbol{t})] = \boldsymbol{t}$. □

Lemma 4.22. *Let $\boldsymbol{g} : A \to \mathcal{N}^n$ be WLUD at $\boldsymbol{t_1} \in A$, with $J\boldsymbol{g}(\boldsymbol{t_1}) \neq 0$, where $J\boldsymbol{g}(\boldsymbol{t_1})$ denotes the Jacobian (determinant) of $\boldsymbol{g}$ at $\boldsymbol{t_1}$; and let $\boldsymbol{x_1} = \boldsymbol{g}(\boldsymbol{t_1})$. Then there exist $\delta, \eta > 0$ and a function $\boldsymbol{F}$ defined on $B_\eta(\boldsymbol{x_1})$ such that:*

(i) $B_\delta(\boldsymbol{t_1}) \subseteq A$;

(ii) $\boldsymbol{g}|_{B_\delta(\boldsymbol{t_1})}$ is one-to-one;

(iii) $B_\eta(\boldsymbol{x_1}) \subseteq \boldsymbol{g}(B_\delta(\boldsymbol{t_1}))$ and $\boldsymbol{F}(B_\eta(\boldsymbol{x_1})) \subseteq B_\delta(\boldsymbol{t_1})$;

(iv) $\boldsymbol{g}[\boldsymbol{F}(\boldsymbol{x})] = \boldsymbol{x} \; \forall \boldsymbol{x} \in B_\eta(\boldsymbol{x_1})$; and

(v) $\boldsymbol{F}$ is WLUD at $\boldsymbol{x_1}$ with $\boldsymbol{DF}(\boldsymbol{x_1}) = [\boldsymbol{Dg}(\boldsymbol{t_1})]^{-1}$.

Proof. Without loss of generality, we may assume that $\boldsymbol{t}_1 = \mathbf{0}$ and $\boldsymbol{x}_1 = \mathbf{0}$; for if this is not the case then we can replace $\boldsymbol{g}(\boldsymbol{t})$ with $\tilde{\boldsymbol{g}}(\boldsymbol{t}) := \boldsymbol{g}(\boldsymbol{t} + \boldsymbol{t_1}) - \boldsymbol{x_1}$. Since $\boldsymbol{g}$ is WLUD at $\mathbf{0}$, there exists $\omega_0 > 0$ in $\mathcal{N}$ such that $B_{\omega_0}(\mathbf{0}) \subset A$ and $\boldsymbol{g}$ is differentiable on $B_{\omega_0}(\mathbf{0})$. Also, since $\boldsymbol{g}$ is C^1 at $\mathbf{0}$ by Theorem 4.6, there exists $\omega_1 > 0$ in $\mathcal{N}$ such that $B_{\omega_1}(\mathbf{0}) \subset A$ and $J\boldsymbol{g}(\boldsymbol{t}) \neq 0$ for all $\boldsymbol{t} \in B_{\omega_1}(\mathbf{0})$. Let $\omega = \min\{\omega_0, \omega_1\}$. By Lemma 4.9, $\boldsymbol{g}$ satisfies (4.1) at $\mathbf{0}$. Let $\boldsymbol{L} = \boldsymbol{Dg}(\mathbf{0})$; then $\boldsymbol{L}^{-1}$ exists since $J\boldsymbol{g}(\mathbf{0}) \neq 0$. Let $\boldsymbol{\phi} = \boldsymbol{I} - \boldsymbol{L}^{-1} \circ \boldsymbol{g}$. Then $\boldsymbol{\phi}$ is WLUD at $\mathbf{0}$ and differentiable on $B_\omega(\mathbf{0})$. Moreover, we have that $\boldsymbol{\phi}(\mathbf{0}) = \mathbf{0}$ and

$$\boldsymbol{D\phi}(\mathbf{0}) = \boldsymbol{D}(\boldsymbol{I} - \boldsymbol{L}^{-1} \circ \boldsymbol{g})(\mathbf{0}) = \boldsymbol{I} - \boldsymbol{L}^{-1} \circ \boldsymbol{Dg}(\mathbf{0}) = \boldsymbol{I} - \boldsymbol{L}^{-1} \circ \boldsymbol{L} = \mathbf{0}.$$

Let $c \in \mathcal{N}$ be such that $0 < c \ll 1$. Since $\boldsymbol{\phi}$ satisfies (4.1) at $\mathbf{0}$, there exists $\delta_0 > 0$ in $\mathcal{N}$ such that $B_{\delta_0}(\mathbf{0}) \subset A$ and, for all $\boldsymbol{s}, \boldsymbol{t} \in B_{\delta_0}(\mathbf{0})$, we have that

$$|\boldsymbol{\phi}(\boldsymbol{s}) - \boldsymbol{\phi}(\boldsymbol{t}) - \boldsymbol{D\phi}(\mathbf{0})(\boldsymbol{s} - \boldsymbol{t})| \leq c\,|\boldsymbol{s} - \boldsymbol{t}|\,.$$

Since $\boldsymbol{D\phi}(\mathbf{0}) = \mathbf{0}$, it follows that

$$|\boldsymbol{\phi}(\boldsymbol{s}) - \boldsymbol{\phi}(\boldsymbol{t})| \leq c\,|\boldsymbol{s} - \boldsymbol{t}| \text{ for all } \boldsymbol{s}, \boldsymbol{t} \in B_{\delta_0}(\mathbf{0}). \tag{4.4}$$

Let $\boldsymbol{s}, \boldsymbol{t} \in B_{\delta_0}(\mathbf{0})$ be such that $\boldsymbol{g}(\boldsymbol{s}) = \boldsymbol{g}(\boldsymbol{t})$. Then $\boldsymbol{\phi}(\boldsymbol{s}) - \boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{s} - \boldsymbol{t}$. Using (4.4) above, it follows that

$$|\boldsymbol{s} - \boldsymbol{t}| = |\boldsymbol{\phi}(\boldsymbol{s}) - \boldsymbol{\phi}(\boldsymbol{t})| \leq c\,|\boldsymbol{s} - \boldsymbol{t}|\,.$$

Since c is infinitely small, it follows that $\boldsymbol{s} = \boldsymbol{t}$ and hence $\boldsymbol{g}|_{B_{\delta_0}(\mathbf{0})}$ is one-to-one.

Let $\bar{\mathfrak{L}} > 0$ be an upper bound for $\{|\boldsymbol{L}^{-1}(\boldsymbol{t})| : |\boldsymbol{t}| \leq 1\}$. Since $\boldsymbol{g}$ satisfies (4.1) at $\mathbf{0}$, there exists $\delta_g > 0$ such that, for all $\boldsymbol{s}$,$\boldsymbol{t} \in B_{\delta_g}(\mathbf{0})$, we have that

$$|\boldsymbol{g}(\boldsymbol{s}) - \boldsymbol{g}(\boldsymbol{t}) - \boldsymbol{L}(\boldsymbol{s} - \boldsymbol{t})| \leq \frac{1}{2\bar{\mathfrak{L}}}\,|\boldsymbol{s} - \boldsymbol{t}|\,. \tag{4.5}$$

Also, since $\boldsymbol{g}$ is C^1 at $\mathbf{0}$, it follows from Lemma 3.4 that there exists $\delta_d > 0$ in $\mathcal{N}$ such that, for all $\boldsymbol{s} \in B_{\delta_d}(\mathbf{0})$ and for all $\boldsymbol{x} \in \mathcal{N}^n$, we have that $\left|(\boldsymbol{Dg}(\boldsymbol{s}))^{-1}\boldsymbol{x}\right| \leq 2\bar{\mathfrak{L}}\,|\boldsymbol{x}|$. Let $\delta = \min\{(1-c)\delta_0, \omega, \delta_g, \delta_d\}$. Then $B_\delta(\mathbf{0}) \subset B_{\delta_0}(\mathbf{0}) \subset A$ and hence $\boldsymbol{g}|_{B_\delta(\mathbf{0})}$ is one-to-one. This proves (i) and (ii).

By (4.4), with $\boldsymbol{t}=\boldsymbol{0}$, we have that $|\boldsymbol{\phi}(\boldsymbol{s})| \le c\,|\boldsymbol{s}|$ for all $\boldsymbol{s} \in B_\delta(\boldsymbol{0})$; thus, we have a function $\boldsymbol{\psi}$ with the properties of Lemma 4.21. Let $\eta = \frac{\delta}{\bar{\mathfrak{L}}}(1-c)$ and define $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{\psi}\left(\boldsymbol{L}^{-1}(\boldsymbol{x})\right)$ for all $\boldsymbol{x} \in B_\eta(\boldsymbol{0})$. Thus, for all $\boldsymbol{x} \in B_\eta(\boldsymbol{0})$, we have that

$$|\boldsymbol{F}(\boldsymbol{x})| = \left|\boldsymbol{\psi}\left(\boldsymbol{L}^{-1}(\boldsymbol{x})\right)\right| \le \frac{\left|\boldsymbol{L}^{-1}(\boldsymbol{x})\right|}{(1-c)} \le \frac{\bar{\mathfrak{L}}\,|\boldsymbol{x}|}{(1-c)} < \frac{\bar{\mathfrak{L}}\eta}{(1-c)} = \delta.$$

Hence $\boldsymbol{F}\left(B_\eta(\boldsymbol{0})\right) \subseteq B_\delta(\boldsymbol{0})$. Furthermore, $(\boldsymbol{I}-\boldsymbol{\phi})|_{B_\delta(\boldsymbol{0})} = \left(\boldsymbol{L}^{-1}\circ\boldsymbol{g}\right)\big|_{B_\delta(\boldsymbol{0})}$; and by Lemma 4.21, we have that

$$\left((\boldsymbol{I}-\boldsymbol{\phi})\circ\boldsymbol{\psi}\right)|_{B_\delta(\boldsymbol{0})} = \boldsymbol{I}|_{B_\delta(\boldsymbol{0})}\,.$$

Thus,

$$\left(\boldsymbol{L}^{-1}\circ\boldsymbol{g}\circ\boldsymbol{\psi}\right)\big|_{B_\delta(\boldsymbol{0})} = \boldsymbol{I}|_{B_\delta(\boldsymbol{0})}\,;$$

and hence

$$\boldsymbol{g}\left(\boldsymbol{\psi}(\boldsymbol{t})\right) = \boldsymbol{L}(\boldsymbol{t}) \text{ for all } \boldsymbol{t} \in B_\delta(\boldsymbol{0})\,.$$

Let $\boldsymbol{x} \in B_\eta(\boldsymbol{0})$ and set $\boldsymbol{t} = \boldsymbol{L}^{-1}(\boldsymbol{x})$. Then

$$|\boldsymbol{t}| \le \bar{\mathfrak{L}}\,|\boldsymbol{x}| \le \bar{\mathfrak{L}}\eta = (1-c)\delta < \delta.$$

Thus, $\boldsymbol{L}^{-1}(\boldsymbol{x}) \in B_\delta(\boldsymbol{0})$. It follows that

$$\boldsymbol{g}\left(\boldsymbol{F}(\boldsymbol{x})\right) = \boldsymbol{g}\left(\boldsymbol{\psi}\left(\boldsymbol{L}^{-1}(\boldsymbol{x})\right)\right) = \boldsymbol{L}\left(\boldsymbol{L}^{-1}(\boldsymbol{x})\right) = \boldsymbol{x} \text{ for all } \boldsymbol{x} \in B_\eta(\boldsymbol{0})$$

and hence $B_\eta(\boldsymbol{0}) \subseteq \boldsymbol{g}\left(B_\delta(\boldsymbol{0})\right)$, since $\boldsymbol{F}(\boldsymbol{x}) \in B_\delta(\boldsymbol{0})$ for all $\boldsymbol{x} \in B_\eta(\boldsymbol{0})$. This proves (iii) and (iv).

Claim: $|\boldsymbol{s}-\boldsymbol{t}| \le 2\bar{\mathfrak{L}}\,|\boldsymbol{g}(\boldsymbol{s})-\boldsymbol{g}(\boldsymbol{t})|$ for all $\boldsymbol{s},\boldsymbol{t} \in B_\delta(\boldsymbol{0})$.

Let $\boldsymbol{s},\boldsymbol{t} \in B_\delta(\boldsymbol{0})$. Then, by (4.5), $|\boldsymbol{g}(\boldsymbol{s})-\boldsymbol{g}(\boldsymbol{t})-\boldsymbol{L}(\boldsymbol{s}-\boldsymbol{t})| \le \frac{|\boldsymbol{s}-\boldsymbol{t}|}{2\bar{\mathfrak{L}}}$. It follows that

$$\begin{aligned} |\boldsymbol{g}(\boldsymbol{s})-\boldsymbol{g}(\boldsymbol{t})| &\ge |\boldsymbol{L}(\boldsymbol{s}-\boldsymbol{t})| - \frac{|\boldsymbol{s}-\boldsymbol{t}|}{2\bar{\mathfrak{L}}} \\ &\ge \frac{|\boldsymbol{s}-\boldsymbol{t}|}{\bar{\mathfrak{L}}} - \frac{|\boldsymbol{s}-\boldsymbol{t}|}{2\bar{\mathfrak{L}}}, \text{using Corollary 3.3} \\ &= \frac{|\boldsymbol{s}-\boldsymbol{t}|}{2\bar{\mathfrak{L}}}. \end{aligned}$$

This completes the proof of the claim.

Now let $\epsilon > 0$ in $\mathcal{N}$ be given. Since $\boldsymbol{g}|_{B_\delta(\boldsymbol{0})}$ is WLUD at $\boldsymbol{0}$, there exists $\delta_1 > 0$ in $\mathcal{N}$, $\delta_1 \le \delta$, such that, for all $\boldsymbol{s},\boldsymbol{t} \in B_{\delta_1}(\boldsymbol{0})$, we have that

$$|\boldsymbol{g}(\boldsymbol{s})-\boldsymbol{g}(\boldsymbol{t})-\boldsymbol{D}\boldsymbol{g}(\boldsymbol{t})(\boldsymbol{s}-\boldsymbol{t})| \le \frac{\epsilon|\boldsymbol{s}-\boldsymbol{t}|}{4\bar{\mathfrak{L}}^2}.$$

Let $\xi = \frac{\delta_1}{2\bar{\mathfrak{L}}}$. Then

$$0 < \xi \le \frac{\delta}{2\bar{\mathfrak{L}}} < \eta = \frac{\delta}{\bar{\mathfrak{L}}}(1-c);$$

and hence $B_\xi(\boldsymbol{0}) \subset B_\eta(\boldsymbol{0})$. Now let $\boldsymbol{x},\boldsymbol{y} \in B_\xi(\boldsymbol{0})$ be given. Since $B_\xi(\boldsymbol{0}) \subset B_\eta(\boldsymbol{0}) \subseteq \boldsymbol{g}\left(B_\delta(\boldsymbol{0})\right)$, then there exists $\boldsymbol{t_x},\boldsymbol{t_y} \in B_\delta(\boldsymbol{0})$ such that $\boldsymbol{g}(\boldsymbol{t_x}) = \boldsymbol{x}$, and $\boldsymbol{g}(\boldsymbol{t_y}) = \boldsymbol{y}$. Since $\boldsymbol{F}\left(B_\xi(\boldsymbol{0})\right) \subset \boldsymbol{F}\left(B_\eta(\boldsymbol{0})\right) \subseteq B_\delta(\boldsymbol{0})$ we get that $\boldsymbol{F}\left(\boldsymbol{g}(\boldsymbol{t_x})\right) = \boldsymbol{F}(\boldsymbol{x}) \in B_\delta(\boldsymbol{0})$. Thus, $\boldsymbol{g}\left(\boldsymbol{F}(\boldsymbol{x})\right) = \boldsymbol{g}\left(\boldsymbol{F}\left(\boldsymbol{g}(\boldsymbol{t_x})\right)\right) = \boldsymbol{g}(\boldsymbol{t_x})$ by (4.22). Since $\boldsymbol{g}$ is one-to-one on $B_\delta(\boldsymbol{0})$, it follows that $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{t_x}$. Similarly $\boldsymbol{F}(\boldsymbol{y}) = \boldsymbol{t_y}$. Moreover, we have that

$$|\boldsymbol{t_x}| = |\boldsymbol{t_x}-\boldsymbol{0}| \le 2\bar{\mathfrak{L}}\,|\boldsymbol{g}(\boldsymbol{t_x})-\boldsymbol{g}(\boldsymbol{0})| = 2\bar{\mathfrak{L}}|\boldsymbol{x}-\boldsymbol{0}| < 2\bar{\mathfrak{L}}\xi = \delta_1.$$

Thus, $\boldsymbol{t_x} \in B_{\delta_1}(\mathbf{0})$; and similarly, we show that $\boldsymbol{t_y} \in B_{\delta_1}(\mathbf{0})$. It follows that

$$|\boldsymbol{g}(\boldsymbol{t_y}) - \boldsymbol{g}(\boldsymbol{t_x}) - \boldsymbol{Dg}(\boldsymbol{t_x})(\boldsymbol{t_y} - \boldsymbol{t_x})| \leq \frac{\epsilon|\boldsymbol{t_y} - \boldsymbol{t_x}|}{4\bar{\mathfrak{L}}^2}.$$

Note that $(\boldsymbol{Dg}(\boldsymbol{t}))^{-1}$ exists since $J\boldsymbol{g}(\boldsymbol{t}) \neq 0$ for all $\boldsymbol{t} \in B_\omega(\mathbf{0}) \supseteq B_\delta(\mathbf{0})$. Now,

$$\begin{aligned}
\left|\boldsymbol{F}(\boldsymbol{y}) - \boldsymbol{F}(\boldsymbol{x}) - (\boldsymbol{Dg}(\boldsymbol{t_x}))^{-1}(\boldsymbol{y} - \boldsymbol{x})\right| &= \left|(\boldsymbol{Dg}(\boldsymbol{t_x}))^{-1}(\boldsymbol{y} - \boldsymbol{x} - \boldsymbol{Dg}(\boldsymbol{t_x})(\boldsymbol{F}(\boldsymbol{y}) - \boldsymbol{F}(\boldsymbol{x})))\right| \\
&\leq 2\bar{\mathfrak{L}}\,|\boldsymbol{y} - \boldsymbol{x} - \boldsymbol{Dg}(\boldsymbol{t_x})(\boldsymbol{F}(\boldsymbol{y}) - \boldsymbol{F}(\boldsymbol{x}))| \\
&= 2\bar{\mathfrak{L}}\,|\boldsymbol{g}(\boldsymbol{t_y}) - \boldsymbol{g}(\boldsymbol{t_x}) - \boldsymbol{Dg}(\boldsymbol{t_x})(\boldsymbol{t_y} - \boldsymbol{t_x})| \\
&\leq 2\bar{\mathfrak{L}}\left(\frac{\epsilon}{4\bar{\mathfrak{L}}^2}\right)|\boldsymbol{t_y} - \boldsymbol{t_x}| = \frac{\epsilon|\boldsymbol{t_y} - \boldsymbol{t_x}|}{2\bar{\mathfrak{L}}} \\
&\leq \frac{\epsilon}{2\bar{\mathfrak{L}}}(2\bar{\mathfrak{L}})\,|\boldsymbol{g}(\boldsymbol{t_y}) - \boldsymbol{g}(\boldsymbol{t_x})| \\
&= \epsilon|\boldsymbol{y} - \boldsymbol{x}|.
\end{aligned}$$

This shows that $\boldsymbol{F}$ is differentiable on $B_\xi(\mathbf{0}) \subset B_\eta(\mathbf{0})$ with $\boldsymbol{DF}(\boldsymbol{x}) = (\boldsymbol{Dg}(\boldsymbol{t}))^{-1}$ where $\boldsymbol{g}(\boldsymbol{t}) = \boldsymbol{x}$. Thus, $\boldsymbol{F}$ is WLUD at $\mathbf{0}$ with $\boldsymbol{DF}(\mathbf{0}) = (\boldsymbol{Dg}(\mathbf{0}))^{-1}$. □

Theorem 4.23 (Inverse Function Theorem). *Let $\boldsymbol{g} : A \to \mathcal{N}^n$ be WLUD on A and let $\boldsymbol{t_0} \in A$ be such that $J\boldsymbol{g}(\boldsymbol{t_0}) \neq 0$. Then there is a neighborhood Ω of $\boldsymbol{t_0}$ such that:*

(i) $\boldsymbol{g}|_\Omega$ is one-to-one;

(ii) $\boldsymbol{g}(\Omega)$ is open;

(iii) the inverse $\boldsymbol{f}$ of $\boldsymbol{g}|_\Omega$ is WLUD on $\boldsymbol{g}(\Omega)$; and $\boldsymbol{Df}(\boldsymbol{x}) = [\boldsymbol{Dg}(\boldsymbol{t})]^{-1}$ for $\boldsymbol{t} \in \Omega$ and $\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{t})$.

Proof. Using Lemma 4.22, we can find a neighborhood Ω_0 of $\boldsymbol{t_0}$ such that $\boldsymbol{g}|_{\Omega_0}$ is one-to-one. Also, since $\boldsymbol{g}$ is C^1 on A by Corollary 4.7 and since $J\boldsymbol{g}(\boldsymbol{t_0}) \neq 0$, there exists a neighborhood Ω_1 of $\boldsymbol{t_0}$ such that $J\boldsymbol{g}(\boldsymbol{t}) \neq 0$ for all $\boldsymbol{t} \in \Omega_1$. Let $\Omega \subseteq \Omega_0 \cap \Omega_1$ be a neighborhood of $\boldsymbol{t_0}$. Then $\boldsymbol{g}|_\Omega$ is one-to-one. Let $\boldsymbol{f} = (\boldsymbol{g}|_\Omega)^{-1}$ with domain $\boldsymbol{g}(\Omega)$. Let $\boldsymbol{t} \in \Omega$ and $\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{t})$. Lemma 4.22 applied to $\boldsymbol{g}|_\Omega$ at the point $\boldsymbol{t}$ gives us δ, η, and $\boldsymbol{F}$ as stated in that lemma. Since $B_\eta(\boldsymbol{x}) \subseteq \boldsymbol{g}(B_\delta(\boldsymbol{t})) \subseteq \boldsymbol{g}(\Omega)$ and $\boldsymbol{g}$ is one-to-one on Ω, it follows that

$$\boldsymbol{g}(\boldsymbol{F}(\boldsymbol{y})) = \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{y})) \text{ and hence } \boldsymbol{F}(\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{y}) \text{ for all } \boldsymbol{y} \in B_\eta(\boldsymbol{x}).$$

Since each $\boldsymbol{x} \in \boldsymbol{g}(\Omega)$ has such a neighborhood $B_\eta(\boldsymbol{x})$ in $\boldsymbol{g}(\Omega)$, it follows that $\boldsymbol{g}(\Omega)$ is open. Moreover, since $\boldsymbol{F}$ is WLUD at $\boldsymbol{x}$ for all $\boldsymbol{x} \in \boldsymbol{g}(\Omega)$ by Lemma 4.22, it follows that $\boldsymbol{f}$ is WLUD on $\boldsymbol{g}(\Omega)$. Finally, using Lemma 4.22 again, we have that

$$\boldsymbol{Df}(\boldsymbol{x}) = \boldsymbol{DF}(\boldsymbol{x}) = [\boldsymbol{Dg}(\boldsymbol{t})]^{-1} \text{ for } \boldsymbol{t} \in \Omega \text{ and } \boldsymbol{x} = \boldsymbol{g}(\boldsymbol{t}).$$

□

As in the real case, the inverse function theorem will be used to prove the implicit function theorem.

4.2. Implicit Function Theorem

We start this final section with some notations that will be useful in the statement and proof of the Implicit Function Theorem.

Notation 4.24. Let $A \subseteq \mathcal{N}^n$ be open and let $\boldsymbol{\Phi} : A \to \mathcal{N}^m$ be WLUD on A. For

$$\boldsymbol{t} = (t_1, ..., t_{n-m}, t_{n-m+1}, ..., t_n) \in A,$$

let

$$\hat{\boldsymbol{t}} = (t_1, ..., t_{n-m}) \text{ and } \tilde{J}\boldsymbol{\Phi}(\boldsymbol{t}) = \det\left(\frac{\partial(\Phi_1, ..., \Phi_m)}{\partial(t_{n-m+1}, ..., t_n)}\right).$$

Theorem 4.25. *Let* $\mathbf{\Phi} : A \to \mathcal{N}^m$ *be WLUD on A, where* $A \subseteq \mathcal{N}^n$ *is open and* $1 \le m < n$. *Let* $\boldsymbol{t_0} \in A$ *be such that* $\mathbf{\Phi}(\boldsymbol{t_0}) = \mathbf{0}$ *and* $\tilde{J}\mathbf{\Phi}(\boldsymbol{t_0}) \neq 0$. *Then there exist a neighborhood U of* $\boldsymbol{t_0}$, *a neighborhood* R *of* $\hat{\boldsymbol{t_0}}$ *and* $\boldsymbol{\phi} : R \to \mathcal{N}^m$ *that is WLUD on R such that*

$$\tilde{J}\mathbf{\Phi}(\boldsymbol{t}) \neq 0 \text{ for all } \boldsymbol{t} \in U,$$

and

$$\{\boldsymbol{t} \in U : \mathbf{\Phi}(\boldsymbol{t}) = \mathbf{0}\} = \{(\hat{\boldsymbol{t}}, \boldsymbol{\phi}(\hat{\boldsymbol{t}})) : \hat{\boldsymbol{t}} \in R\}.$$

Proof. Since $\mathbf{\Phi}$ is C^1 on A by Corollary 4.7 and since $\tilde{J}\mathbf{\Phi}(\boldsymbol{t_0}) \neq 0$, there exists a neighborhood U_0 of $\boldsymbol{t_0}$ such that $\tilde{J}\mathbf{\Phi}(\boldsymbol{t}) \neq 0$ for all $\boldsymbol{t} \in U_0$. Let $\boldsymbol{g} : U_0 \to \mathcal{N}^n$ be defined as

$$\begin{aligned} g_i(\boldsymbol{t}) &= t_i & 1 \le i \le n-m \\ g_{n-m+j}(\boldsymbol{t}) &= \Phi_j(\boldsymbol{t}) & 1 \le j \le m. \end{aligned}$$

Then $\boldsymbol{g}$ is WLUD on U_0 and has the Jacobian matrix

$$\left(\begin{array}{ccc|ccc} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline \Phi_1^1(\boldsymbol{t}) & \cdots & \Phi_{n-m}^1(\boldsymbol{t}) & \Phi_{n-m+1}^1(\boldsymbol{t}) & \cdots & \Phi_n^1(\boldsymbol{t}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Phi_1^m(\boldsymbol{t}) & \cdots & \Phi_{n-m}^m(\boldsymbol{t}) & \Phi_{n-m+1}^m(\boldsymbol{t}) & \cdots & \Phi_n^m(\boldsymbol{t}) \end{array}\right).$$

Thus, $J\boldsymbol{g}(\boldsymbol{t}) = \tilde{J}\mathbf{\Phi}(\boldsymbol{t}) \neq 0$ on U_0. Applying the Inverse Function Theorem (Theorem 4.23 above) to $\boldsymbol{g}$ at $\boldsymbol{t_0}$, we get a neighborhood U of $\boldsymbol{t_0}$ in U_0 such that $\boldsymbol{g}(U)$ is open, and $\boldsymbol{g}|_U$ is one-to-one. Additionally, $\boldsymbol{g}|_U$ has an inverse $\boldsymbol{f}$ which is WLUD on $g(U)$. Let $R = \{\hat{\boldsymbol{t}} \in \mathcal{N}^{n-m} : (\hat{\boldsymbol{t}}, \mathbf{0}) \in \boldsymbol{g}(U)\}$. Then R is open since $\boldsymbol{g}(U)$ is open. Let $\boldsymbol{\phi} : R \to \mathcal{N}^m$ be defined as

$$\phi_l(\hat{\boldsymbol{t}}) = f_{n-m+l}(\hat{\boldsymbol{t}}, \mathbf{0}) \qquad 1 \le l \le m.$$

Then

$$\boldsymbol{t} \in U \text{ and } \mathbf{\Phi}(\boldsymbol{t}) = \mathbf{0} \iff \hat{\boldsymbol{t}} \in R \text{ and } \boldsymbol{g}(\boldsymbol{t}) = (\hat{\boldsymbol{t}}, \mathbf{0}).$$

Moreover, since $\boldsymbol{g}|_U$ and $\boldsymbol{f}$ are inverses, it follows that

$$\boldsymbol{g}(\boldsymbol{t}) = (\hat{\boldsymbol{t}}, \mathbf{0}) \iff \boldsymbol{t} = \boldsymbol{f}(\hat{\boldsymbol{t}}, \mathbf{0}).$$

Thus ,

$$\begin{aligned} \{\boldsymbol{t} \in U : \mathbf{\Phi}(\boldsymbol{t}) = \mathbf{0}\} &= \{\boldsymbol{t} \in U : \boldsymbol{g}(\boldsymbol{t}) = (\hat{\boldsymbol{t}}, \mathbf{0}), \hat{\boldsymbol{t}} \in R\} \\ &= \{\boldsymbol{t} \in U : \boldsymbol{t} = \boldsymbol{f}(\hat{\boldsymbol{t}}, \mathbf{0}), \hat{\boldsymbol{t}} \in R\} \\ &= \{(\hat{\boldsymbol{t}}, \boldsymbol{\phi}(\hat{\boldsymbol{t}})) : \hat{\boldsymbol{t}} \in R\}. \end{aligned}$$

□

ACKNOWLEDGEMENTS

The author would like to thank Zhenghang Du and Yudong Chen for their help with the proof of the Taylor theorem with remainder, Theorem 2.9, during their MITACS Globalink internship in my research group from June to August, 2019.

FUNDING

This research was funded by the Natural Sciences and Engineering Council of Canada (NSERC, Grant # RGPIN/4965-2017).

REFERENCES

1. G. Bookatz and K. Shamseddine, "Calculus on a non-archimedean field extension of the real numbers: inverse function theorem, intermediate value theorem and mean value theorem," Contemp. Math. **704**, 49–67 (2018).
2. A. Barría Comicheo and K. Shamseddine, "Summary on non-archimedean valued fields," Contemp. Math. **704**, 1–36 (2018).
3. H. Hahn, "Über die nichtarchimedischen Größensysteme," Sitzungsbericht der Wiener Akademie der Wissenschaften Abt. 2a **117**, 601–655 (1907).
4. W. Krull, "Allgemeine Bewertungstheorie," J. Reine Angew. Math. **167**, 160–196 (1932).
5. T. Levi-Civita, "Sugli infiniti ed infinitesimi attuali quali elementi analitici," Atti Ist. Veneto di Sc., Lett. ed Art. 7a **4**, 1765 (1892).
6. T. Levi-Civita, "Sui numeri transfiniti," Rend. Acc. Lincei 5a **7**, 91, 113 (1898).
7. S. MacLane, "The universality of formal power series fields," Bull. Amer. Math. Soc. **45**, 888 (1939).
8. S. Priess-Crampe, *Angeordnete Strukturen: Gruppen, Körper, projektive Ebenen* (Springer, Berlin, 1983).
9. F. J. Rayner, "Algebraically closed fields analogous to fields of Puiseux series," J. London Math. Soc. **8**, 504–506 (1974).
10. P. Ribenboim, "Fields: Algebraically closed and others," Manuscr. Math. **75**, 115–150 (1992).
11. A. Robinson, *Non-Standard Analysis* (North-Holland, 1974).
12. W. H. Schikhof, *Ultrametric Calculus: An Introduction to p-Adic Analysis* (Cambridge University Press, 1985).
13. K. Shamseddine, "Analysis on the Levi-Civita field and computational applications," Appl. Math. Comput. **255**, 44–57 (2015).
14. K. Shamseddine and M. Berz, "Exception handling in derivative computation with non-Archimedean calculus," in *Computational Differentiation: Techniques, Applications, and Tools*, pp. 37–51 (SIAM, Philadelphia, 1996).
15. K. Shamseddine and M. Berz, "The differential algebraic structure of the Levi-Civita field and applications," Int. J. Appl. Math. **3**, 449–465 (2000).
16. K. Shamseddine and T. Sierens, "On locally uniformly differentiable functions on a complete non-Archimedean ordered field extension of the real numbers," ISRN Math. Anal., Art. ID 387053 (2012).
17. N. Vakil, *Real Analysis through Modern Infinitesimals* (Cambridge University Press, 2011).