ONE-DIMENSIONAL OPTIMIZATION ON NON-ARCHIMEDEAN FIELDS

KHODR SHAMSEDDINE AND VERA ZEIDAN

Abstract. One dimensional optimization on non-Archimedean fields is presented. We derive first and second order necessary and sufficient optimality conditions. For first order optimization, these conditions are similar to the corresponding real ones; but this is not the case for higher order optimization. This is due to the total disconnectedness of the given non-Archimedean field in the order topology, which renders the usual concept of differentiability weak. We circumvent this difficulty by using a stronger concept of differentiability based on the derivate approach, which entails a Taylor formula with remainder and hence a similar local behavior as in the real case.

1. Introduction

Let K be a totally ordered non-Archimedean field extension of $\mathbb{R}$. We introduce the following terminology.

Definition 1.1 ($\sim$, $\approx$, $\ll$, H, λ)**.** For $x, y \in K$, we say $x \sim y$ if there exist $n, m \in \mathbb{N}$ such that $n|x| > |y|$ and $m|y| > |x|$; for nonnegative $x, y \in K$, 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 H (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. It follows that $(H, +, <)$ is a totally ordered group, often referred to as the Hahn group or skeleton group, whose neutral element is the class $[1]$. It follows from the above that the projection λ from K to H is a valuation.

The theorem of Hahn [3] provides a complete classification of non-Archimedean extensions of $\mathbb{R}$ in terms of their skeleton groups. In fact, invoking the axiom of choice it is shown that the elements of any such field K can be written as formal power series over its skeleton group H with real coefficients, and the set of appearing "exponents" forms a well-ordered subset of H. The coefficient of the qth power in the Hahn representation of a given x will be denoted by $x[q]$, and the number d will be defined by $d[1] = 1$ and $d[q] = 0$ for $q \neq 1$. It is easy to check that $0 < d^q \ll 1$ 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 $x \neq 0$.

1991 *Mathematics Subject Classification.* 26E30, 12J25, 11D88, 78M50, 80M50.

Key words and phrases. Non-Archimedean calculus, Levi-Civita field, optimization, local extrema, convex functions.

From general properties of formal power series fields [7, 9], it follows that if H is divisible then K is real-closed. 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].

Throughout this paper, $\mathcal{N}$ will denote any totally ordered non-Archimedean field extension of $\mathbb{R}$ that is order complete and whose skeleton group G is Archimedean; i.e. a subgroup of $\mathbb{R}$. The smallest such field is the field of the formal Laurent series whose skeleton group is $\mathbb{Z}$; and the smallest such field that is also real-closed is the Levi-Civita field $\mathcal{R}$, first introduced in [5, 6]. In this case $H = \mathbb{Q}$, and for any element $x \in \mathcal{R}$, the set of exponents in the Hahn representation of x is a left-finite subset of $\mathbb{Q}$, i.e. below any rational bound r there are only finitely many exponents.

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 [1]. Having infinitely small numbers, the errors in classical numerical methods can be made infinitely small and hence irrelevant in all practical applications. One such application is the computation of derivatives of real functions representable on a computer [12], where both the accuracy of formula manipulators and the speed of classical numerical methods are achieved.

In the light of such practical usefulness of infinitely small numbers, it is natural to study optimization questions on non-Archimedean fields with the hope to extend the methods mentioned in the previous paragraph to find local extrema of functions as closely as allowed by machine precision. In this paper, we study general optimization questions and derive first and second order necessary and sufficient conditions for the existence of local maxima and minima of a function on a convex subset of a non-Archimedean field $\mathcal{N}$. We show that for first order optimization, the results are similar to the corresponding real ones. However, for second and higher order optimization, more work needs to be done: We show that conventional differentiability is not strong enough to just extend the real-case results; and a stronger differentiability concept will be used to solve that difficulty. We also characterize convex functions on convex sets of $\mathcal{N}$ in terms of first and second order derivatives. As we will see, some of the results known for the real case will be extended in the same format to our setting; however, other results will be proved under a mild extra assumption which is automatically satisfied in the special case of continuation of real functions.

Remark 1.2. *Throughout this paper, results whose proofs are very similar to those of the corresponding real-case ones will just be stated without proof.*

2. First Order Optimization

In this section, we prove first order optimization results similar to the real corresponding ones; and we show with examples that for higher order optimization, conventional differentiability is not sufficient to generalize the results from $\mathbb{R}$ to $\mathcal{N}$. In the following, $I(a, b)$ will denote any one of the intervals (a, b), $(a, b]$, $[a, b)$ or $[a, b]$. We have the following useful results.

Proposition 2.1. *Let $a < b$ in $\mathcal{N}$ be given, and let $f : I(a, b) \to \mathcal{N}$ be differentiable on $I(a, b)$ and have a local minimum (resp. local maximum) at $x_0 \in I(a, b)$. Then there exists $\delta > 0$ in $\mathcal{N}$ such that $f'(x_0)(x - x_0) \geq 0$ (resp. ≤ 0) for all $x \in I(a, b)$ satisfying $|x - x_0| < \delta$.*

Corollary 2.2. *Let $a < b$ in $\mathcal{N}$ be given, and let $f : I(a,b) \to \mathcal{N}$ be differentiable on $I(a,b)$ and have a local extremum (maximum or minimum) at $x_0 \in (a,b)$. Then $f'(x_0) = 0$.*

However, contrary to the real case, the following example shows that a function having a local minimum at a point $x_0 \in (a,b) \subset \mathcal{N}$ need not have the first nonvanishing derivative positive and of even order.

Example 2.3. Let $g : (-1,1) \to \mathcal{N}$ be given by $g(x)[q] = x[q/3]$ and let $f : (-1,1) \to \mathcal{N}$ be given by $f(x) = (g(x))^2 - x^7$. Then g is infinitely often differentiable on $(-1,1)$ with $g^{(j)}(x) = 0$ for all $j \in \mathbb{N}$ and for all $x \in (-1,1)$. Indeed, it suffices to show that g is differentiable on $(-1,1)$ with $g'(x) = 0$ for all $x \in (-1,1)$. We first observe that $g(x+y) = g(x) + g(y)$ for all $x, y \in (-1,1)$. Now let $x \in (-1,1)$ and $\epsilon > 0$ in $\mathcal{N}$ be given. Let $\delta = \min\{\epsilon, d\}$, and let $y \in (-1,1)$ be such that $0 < |y - x| < \delta$. Then

$$\left|\frac{g(y) - g(x)}{y - x}\right| = \left|\frac{g(y-x)}{y - x}\right| \sim (y-x)^2 \text{ since } g(y-x) \sim (y-x)^3.$$

Since $|y - x| < \min\{\epsilon, d\}$, we obtain that $(y-x)^2 \ll \epsilon$ both when ϵ is finite and when ϵ is infinitely small. Hence

$$\left|\frac{g(y) - g(x)}{y - x}\right| < \epsilon \text{ for all } y \in (-1,1) \text{ satisfying } 0 < |y - x| < \delta.$$

It follows that f is infinitely often differentiable on $(-1,1)$ with $f^{(j)}(0) = 0$ for all $j \in \{1, \ldots, 6\}$ and $f^{(7)}(0) = -7!$. We show that f has a relative minimum at 0. Let $x \in (-1,1)$ be such that $|x| \ll 1$. Then $(g(x))^2 \approx (x[\lambda(x)])^2 d^{6\lambda(x)}$ and $|x^7| \approx |x[\lambda(x)]|^7 d^{7\lambda(x)} \ll (x[\lambda(x)])^2 d^{6\lambda(x)}$. Hence $f(x) \approx (x[\lambda(x)])^2 d^{6\lambda(x)} > 0 = f(0)$. This is true for all $x \in (-1,1)$ satisfying $|x| \ll 1$. Hence f has a (strict) local minimum at 0.

Remark 2.4. *Even though $g'(x) = 0$ for all $x \in (-1,1)$, g is not constant; g is even strictly increasing on $(-1,1)$. Moreover, g is neither convex nor concave on $(-1,1)$ by Proposition 2.9 and the corresponding result for concave functions, since $g(d) = d^3 > 0 = g(0) + g'(0)d$ and $g(-d) = -d^3 < 0 = g(0) + g'(0)(-d)$. Finally, it is easy to check that g is a nontrivial order preserving field automorphism on $\mathcal{N}$; this situation is limited to the case of non-Archimedean fields since it is well-known that the only order preserving field automorphism on any given totally ordered Archimedean field is the identity mapping* [11].

Also, contrary to the real case, the following example shows that a function that is $2k$-times differentiable on an open interval (a,b) containing the point x_0, with $f^{(j)}(x_0) = 0$ for all $j \in \{1, \ldots, 2k-1\}$ and $f^{(2k)}(x_0) \neq 0$, need not have a local extremum at x_0.

Example 2.5. Let $g : (-1,1) \to \mathcal{N}$ be as in Example 2.3, and let $f : (-1,1) \to \mathcal{N}$ be given by $f(x) = g(x) - x^4$. Then f is four times differentiable on $(-1,1)$ with $f'(0) = f''(0) = f'''(0) = 0$ and $f^{(4)}(0) = -24$.

Now let $x \in (-1,1)$ be such that $|x| \ll 1$. Then $g(x) \approx x[\lambda(x)] d^{3\lambda(x)}$ and $x^4 \approx (x[\lambda(x)])^4 d^{4\lambda(x)} \ll |x[\lambda(x)]| d^{3\lambda(x)}$. Thus,

$$f(x) \approx x[\lambda(x)] d^{3\lambda(x)} \begin{cases} > f(0) & \text{if } 0 < x \ll 1 \\ < f(0) & \text{if } 0 < -x \ll 1 \end{cases};$$

and hence f has no local extremum at 0.

In Section 3, we introduce a stronger notion of differentiability based on the derivate concept [11, 2] and show that, for local extrema, this notion furnishes necessary and sufficient conditions analogous to those of the real case.

Definition 2.6 (Convex and Concave Functions)**.** Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$. Then we say that f is convex on $I(a,b)$ if $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ for all $x, y \in I(a,b)$ and for all $t \in [0,1]$. If the inequality is strict for all $x \neq y$ in $I(a,b)$ and for all $t \in (0,1)$, we say that f is strictly convex on $I(a,b)$. Finally, we say that f is concave on $I(a,b)$ if $-f$ is convex on $I(a,b)$.

The following result gives some geometric interpretation to Definition 2.6.

Proposition 2.7. *Let $a < b$ be given in $\mathcal{N}$, let $f : I(a,b) \to \mathcal{N}$, and let $S(a,b,f) \subset I(a,b) \times \mathcal{N}$ be defined by $S(a,b,f) = \{(x,r) : r \in \mathcal{N}$ and $f(x) \leq r\}$. Then f is convex on $I(a,b)$ if and only if $S(a,b,f)$ is convex.*

Proposition 2.8. *Let $a < b$ be given in $\mathcal{N}$. Then the following are true.*

1. *If $f_1, f_2 : I(a,b) \to \mathcal{N}$ are convex then so is $\alpha f_1 + \beta f_2$ for all $\alpha \geq 0$ and for all $\beta \geq 0$ in $\mathcal{N}$.*
2. *If $f : I(a,b) \to \mathcal{N}$ is convex then, for $c \in \mathcal{N}$, $\Gamma_c = \{x \in I(a,b) : f(x) \leq c\}$ is a convex subset of $I(a,b)$.*

Proposition 2.9. *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be differentiable. Then f is convex (resp. strictly convex) on $I(a,b)$ if and only if $f(y) \geq f(x) + f'(x)(y-x)$ (resp. $f(y) > f(x) + f'(x)(y-x)$) for all $x, y \in I(a,b)$.*

Corollary 2.10. *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be differentiable on $I(a,b)$. If f is convex (resp. strictly convex), then f' is nondecreasing (resp. strictly increasing) on $I(a,b)$.*

The converse of Corollary 2.10 is not true as the following example shows.

Example 2.11. Let $g : (-1,1) \to \mathcal{N}$ be as in Example 2.3 and let $f : (0,1) \to \mathcal{N}$ be given by $f(x) = -g(x) + x^4$. Then f is differentiable on $(0,1)$, with $f'(x) = 4x^3$ for all $x \in (0,1)$; and hence f' is strictly increasing on $(0,1)$. But $f(d) = -g(d) + d^4 = -d^3 + d^4 < 0 = f(0) + f'(0)d$, and hence f is not convex on $(0,1)$ by Proposition 2.9.

Theorem 2.12. *Let $a < b$ be given in $\mathcal{N}$, let $f : I(a,b) \to \mathcal{N}$ be convex and differentiable on $I(a,b)$, and let $x_0 \in I(a,b)$ be given. Then the following are equivalent:*

1. *f has a local minimum at x_0.*
2. *f has a global minimum at x_0.*
3. *$f'(x_0)(x - x_0) \geq 0$ for all $x \in I(a,b)$.*

Proposition 2.13. *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be twice differentiable on $I(a,b)$. If f is convex, then $f''(x) \geq 0$ for all $x \in I(a,b)$.*

Proof. Suppose that f is convex on $I(a,b)$; then f' is nondecreasing on $I(a,b)$ by Corollary 2.10. Now we show that $f''(x) \geq 0$ for all $x \in I(a,b)$. Suppose not. Then there exists $x_0 \in I(a,b)$ such that $f''(x_0) < 0$. Thus there exists $\delta > 0$ in $\mathcal{N}$ such that

$$\left| \frac{f'(x) - f'(x_0)}{x - x_0} - f''(x_0) \right| < \frac{-f''(x_0)}{2};$$

from which we infer that

$$\frac{f'(x)-f'(x_0)}{x-x_0}<0 \text{ for all } x\in I(a,b)\cap(x_0-\delta,x_0+\delta).$$

This contradicts the fact that f' is nondecreasing on $I(a,b)$. □

The converse is not true as the following example shows.

Example 2.14. Let $f:(-1,1)\to\mathcal{N}$ be as in Example 2.11. Then f is twice differentiable on $(-1,1)$ with $f''(x)=12x^2\geq 0$ for all $x\in(-1,1)$. But f is not convex on $(-1,1)$ as was shown in Example 2.11.

In the following section, we show that under the stronger concept of derivate differentiability and an additional mild condition, the converses of Corollary 2.10 and Proposition 2.13 will be true.

3. Derivate Continuity and Differentiability

In this section, we employ stronger concepts of continuity and differentiability based on which we obtain, in Theorem 3.13, a Taylor formula with remainder [11, 2]. This in turn will allow us to prove high order optimization results similar to the corresponding ones in the real case.

Definition 3.1. Let $a<b$ be given in $\mathcal{N}$ and let $f:I(a,b)\to\mathcal{N}$. Then we say that f is derivate continuous on $I(a,b)$ if there exists $M\in\mathcal{N}$, called a Lipschitz constant of f on $I(a,b)$, such that

$$\left|\frac{f(y)-f(x)}{y-x}\right|\leq M \text{ for all } x\neq y \text{ in } I(a,b).$$

It follows immediately from Definition 3.1 that if $f:I(a,b)\to\mathcal{N}$ is derivate continuous on $I(a,b)$ then f is uniformly continuous and bounded on $I(a,b)$.

If we define, $r_0(x,y)=(f(y)-f(x))/(y-x)$ for $y\neq x$ and $r_0(x,x)=0$, we readily obtain the following useful result [11, 2].

Lemma 3.2 (Remainder Formula 0). *Let $a<b$ in $\mathcal{R}$ and let $f:I(a,b)\to\mathcal{N}$ be derivate continuous on $I(a,b)$ with Lipschitz constant M. Then for all $x,y\in I(a,b)$, we have that $f(y)=f(x)+r_0(x,y)(y-x)$, with $\lambda(r_0(x,y))\geq\lambda(M)$.*

The proofs of Proposition 3.3 and Proposition 3.4 are straightforward; see [11].

Proposition 3.3. *Let $a<b$ be given in $\mathcal{N}$, let $f,g:I(a,b)\to\mathcal{N}$ be derivate continuous on $I(a,b)$, and let $\alpha\in\mathcal{N}$. Then $f+\alpha g$ and $f\cdot g$ are derivate continuous on $I(a,b)$.*

Proposition 3.4. *Let $a<b$ and $c<e$ in $\mathcal{N}$ be given, and let $f:I_1(a,b)\to\mathcal{N}$ and $g:I_2(c,e)\to\mathcal{N}$ be such that $f(I_1(a,b))\subset I_2(c,e)$, f is derivate continuous on $I_1(a,b)$ and g derivate continuous on $I_2(c,e)$. Then $g\circ f$ is derivate continuous on $I_1(a,b)$.*

Theorem 3.5. *Let $a<b$ be given in $\mathcal{N}$ and let $f:I(a,b)\to\mathcal{N}$ be derivate continuous on $I(a,b)$. Then there exists a unique function $g:[a,b]\to\mathcal{N}$, derivate continuous on $[a,b]$, such that $g|_{I(a,b)}=f$.*

Proof. We may assume that $I(a,b) \neq [a,b]$. First assume that $I(a,b) = (a,b]$. Let $f_0 = \lim_{n\to\infty} f(a + d^n(b-a))$, which exists [11] because f is uniformly continuous on $(a,b]$ and $\mathcal{N}$ is order complete. Define $g : [a,b] \to \mathcal{N}$ by $g(x) = f(x)$ if $x \in (a,b]$ and $g(a) = f_0$. It remains to show that g is derivate continuous on $[a,b]$. Let M be a Lipschitz constant of f on $(a,b]$ and let $x \neq y$ in $[a,b]$ be given. Without loss of generality, we may assume that $x < y$. First assume that $a < x$, then $x, y \in (a,b]$, and hence

$$\left|\frac{g(y) - g(x)}{y - x}\right| = \left|\frac{f(y) - f(x)}{y - x}\right| \leq M < 2M.$$

Now assume that $x = a$. There exists $N \in \mathbb{N}$ such that $d^N(b-a) < y - a$ and $\left|f\left(a + d^N(b-a)\right) - g(a)\right| \leq M(y-a)$. Then it follows that $0 < y - \left(a + d^N(b-a)\right) < y - a$; and hence

$$\begin{aligned}\left|\frac{g(y) - g(x)}{y - x}\right| &\leq \left|\frac{f(y) - f\left(a + d^N(b-a)\right)}{y - a}\right| + \left|\frac{f\left(a + d^N(b-a)\right) - g(a)}{y - a}\right| \\ &\leq \left|\frac{f(y) - f\left(a + d^N(b-a)\right)}{y - (a + d^N(b-a))}\right| + M \leq M + M = 2M.\end{aligned}$$

Hence g is derivate continuous on $[a,b]$ with Lipschitz constant $2M$, and $g|_{(a,b]} = f$.

Similarly, we can show that the result is true for the cases when $I(a,b) = [a,b)$ and $I(a,b) = (a,b)$. □

Using an argument similar to that in the proof of Theorem 3.5, we obtain the following result.

Corollary 3.6. *Let $a < b$ be given in $\mathcal{N}$, let $x_0 \in I(a,b)$, and let $f : I(a,b)\backslash\{x_0\} \to \mathcal{N}$ be derivate continuous on $I(a,b) \setminus \{x_0\}$. Then there exists a unique function $g : I(a,b) \to \mathcal{N}$, derivate continuous on $I(a,b)$, such that $g|_{I(a,b)\backslash\{x_0\}} = f$.*

Definition 3.7. Let $a < b$ be given in $\mathcal{N}$, let $f : I(a,b) \to \mathcal{N}$ be derivate continuous on $I(a,b)$, and let I_D denote the identity function on $I(a,b)$. Then we say that f is derivate differentiable on $I(a,b)$ if for all $x \in I(a,b)$, the function $\frac{f - f(x)}{I_D - x} : I(a,b) \setminus \{x\} \to \mathcal{N}$ is derivate continuous on $I(a,b) \setminus \{x\}$. In this case the unique continuation of $\frac{f-f(x)}{I_D - x}$ to $I(a,b)$, obtained by Corollary 3.6, will be called the first derivate function (or simply the derivate function) of f at x and will be denoted by $F_{1,x}$; moreover, the function value $F_{1,x}(x)$ will be called the derivative of f at x and will be denoted by $f'(x)$.

It follows immediately from Definition 3.7 that if $f : I(a,b) \to \mathcal{N}$ is derivate differentiable then f is differentiable in the conventional sense; moreover, the two derivatives at any given point of $I(a,b)$ agree. Also, using Definition 3.7 and Lemma 3.2, we obtain the following result [11, 2].

Lemma 3.8 (Remainder Formula 1)**.** *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be derivate differentiable on $I(a,b)$. Let $x \in I(a,b)$ be given, let $F_{1,x}$ be the derivate function of f at x, and let $M_{1,x}$ be a Lipschitz constant of $F_{1,x}$ on $I(a,b)$. Then for all $y \in I(a,b)$, we have that $f(y) = f(x) + f'(x)(y-x) + r_1(x,y)(y-x)^2$, with $\lambda(r_1(x,y)) \geq \lambda(M_{1,x})$.*

The proof of the following theorem is similar to that of the corresponding real case result; see [11].

Theorem 3.9. *Let $a < b$ be given in $\mathcal{N}$, let $f, g : I(a,b) \to \mathcal{N}$ be derivate differentiable on $I(a,b)$, and let $\alpha \in \mathcal{N}$. Then $f + \alpha g$ and $f \cdot g$ are derivate differentiable on $I(a,b)$, with derivatives $(f + \alpha g)' = f' + \alpha g'$ and $(f \cdot g)' = f' \cdot g + f \cdot g'$.*

Theorem 3.10 (Chain Rule)**.** *Let $a < b$ and $c < e$ in $\mathcal{N}$, and let $f : I_1(a,b) \to \mathcal{N}$ and $g : I_2(c,e) \to \mathcal{N}$ be such that $f(I_1(a,b)) \subset I_2(c,e)$, f is derivate differentiable on $I_1(a,b)$ and g derivate differentiable on $I_2(c,e)$. Then $g \circ f$ is derivate differentiable on $I_1(a,b)$, with derivative $(g \circ f)' = (g' \circ f) \cdot f'$.*

Proof. Let $x \in I(a,b)$ be given, and let $H_x : D \to \mathcal{N}$ be given by

$$H_x(y) = \begin{cases} \frac{(g\circ f)(y) - (g\circ f)(x)}{y-x} & \text{if } y \neq x \\ \\ g'(f(x))\, f'(x) & \text{if } y = x \end{cases}.$$

Then

$$\begin{aligned} H_x(y) &= \begin{cases} \frac{g(f(y)) - g(f(x))}{f(y)-f(x)} \frac{f(y)-f(x)}{y-x} & \text{if } y \neq x \text{ and } f(y) \neq f(x) \\ 0 & \text{if } y \neq x \text{ and } f(y) = f(x) \\ g'(f(x))\, f'(x) & \text{if } y = x \end{cases} \\ &= G_{1,f(x)}(f(y))\, F_{1,x}(y), \end{aligned}$$

where $F_{1,x}$ is the derivate function of f at x, and $G_{1,f(x)}$ the derivate function of g at $f(x)$. Hence $H_x = \left(G_{1,f(x)} \circ f\right) \cdot F_{1,x}$. Since f is derivate continuous on $I_1(a,b)$ and since $G_{1,f(x)}$ is continuous on $I_2(c,e)$, we have by Proposition 3.4 that $\left(G_{1,f(x)} \circ f\right)$ is continuous on $I_1(a,b)$. Since $F_{1,x}$ is continuous on $I_1(a,b)$, so is $\left(G_{1,f(x)} \circ f\right) \cdot F_{1,x} = H_x$ by Proposition 3.3. Hence $(g \circ f)$ is derivate differentiable on $I_1(a,b)$, with derivative $(g \circ f)'(x) = H_x(x) = g'(f(x))\, f'(x) = ((g' \circ f) \cdot f')(x)$ for all $x \in I_1(a,b)$. □

The following result provides a useful tool for checking the derivate differentiability of functions.

Theorem 3.11. *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be derivate continuous on $I(a,b)$. Suppose there exists $M \in \mathcal{N}$ and there exists a function $g : I(a,b) \to \mathcal{N}$ such that*

$$\left| \frac{f(y) - f(x)}{y - x} - g(x) \right| \leq M\, |y - x| \text{ for all } y \neq x \text{ in } I(a,b).$$

Then f is derivate differentiable on $I(a,b)$, with derivative $f' = g$.

Proof. It suffices to show that for all $x \in I(a,b)$, the function $F_{1,x} : I(a,b) \to \mathcal{N}$, given by

$$F_{1,x}(y) = \begin{cases} \frac{f(y)-f(x)}{y-x} & \text{if } y \neq x \\ \\ g(x) & \text{if } y = x \end{cases},$$

is derivate continuous on $I(a,b)$. It is sufficient to show that for all $x \in I(a,b)$, we have that $|F_{1,x}(y) - F_{1,x}(z)| \leq d^{-1} M |y - z|$ for all $y \neq z$ in $I(a,b)$. So let $x \in I(a,b)$ be given; and let $y \neq z$ in $I(a,b)$. Four cases are to be considered.

Case 1: Assume $y = x$. Then

$$\begin{aligned} |F_{1,x}(y) - F_{1,x}(z)| &= |F_{1,x}(x) - F_{1,x}(z)| = \left| g(x) - \frac{f(z) - f(x)}{z - x} \right| \\ &\leq M|x - z| = M|y - z| < d^{-1} M |y - z|. \end{aligned}$$

Case 2: Assume $z = x$. As in case 1, we have that $|F_{1,x}(y) - F_{1,x}(z)| \leq M|y-z| < d^{-1}M|y-z|$.

Case 3: Assume $y \neq x \neq z$ and $|y-z| \not\ll |y-x|$. Then $|y-z| \not\ll |z-x|$; for if $|y-z| \ll |z-x|$, then $|y-x| = |y-z+(z-x)| \approx |z-x| \gg |y-z|$, a contradiction. Thus,

$$\begin{aligned} |F_{1,x}(y) - F_{1,x}(z)| &\leq \left| \frac{f(y)-f(x)}{y-x} - g(x) \right| + \left| \frac{f(z)-f(x)}{z-x} - g(x) \right| \\ &\leq M|y-x| + M|z-x| < d^{-1}M|y-z| \end{aligned}$$

since $d^{-1}|y-z| \gg |y-x| + |z-x|$.

Case 4: Assume $y \neq x \neq z$ and $|y-z| \ll |y-x|$. Then $z-x = z-y+(y-x) \approx y-x$. Thus

$$\begin{aligned} |F_{1,x}(y) - F_{1,x}(z)| &= \left| \frac{f(y)-f(x)}{y-x} - \frac{f(z)-f(x)}{z-x} \right| \\ &= \left| \frac{y-z}{y-x} \right| \left| \frac{f(y)-f(z)}{y-z} - \frac{f(z)-f(x)}{z-x} \right| \\ &\leq \left| \frac{y-z}{y-x} \right| \left(\left| \frac{f(y)-f(z)}{y-z} - g(z) \right| + \left| \frac{f(z)-f(x)}{z-x} - g(z) \right| \right) \\ &\leq \left| \frac{y-z}{y-x} \right| M(|y-z| + |z-x|). \end{aligned}$$

Since $|z-x| \approx |y-x|$, we obtain that $|z-x| < d^{-1}|y-x|/2$. Also, since $|y-z| \ll |y-x|$, we obtain that $|y-z| < |y-x| < d^{-1}|y-x|/2$. Therefore,

$$|F_{1,x}(y) - F_{1,x}(z)| \leq \left| \frac{y-z}{y-x} \right| Md^{-1}|y-x| = d^{-1}M|y-z|;$$

which finishes the proof of the theorem. □

Definition 3.12 (n-times Derivate Differentiability)**.** Let $a < b$ be given in $\mathcal{N}$, and let $f : I(a,b) \to \mathcal{N}$. Let $n \geq 2$ be given in $\mathbb{N}$. Then we define n-times derivate differentiability of f on $I(a,b)$ inductively as follows: Having defined $(n-1)$-times derivate differentiability, we say that f is n-times derivate differentiable on $I(a,b)$ if f is $(n-1)$-times derivate differentiable on $I(a,b)$ and for all $x \in I(a,b)$, the $(n-1)$st derivate function $F_{n-1,x}$ is derivate differentiable on $I(a,b)$. For all $x \in I(a,b)$, the derivate function $F_{n,x}$ of $F_{n-1,x}$ at x will be called the nth derivate function of f at x, and the number $f^{(n)}(x) = n!F'_{n-1,x}(x)$ will be called the nth derivative of f at x and denoted by $f^{(n)}(x)$.

Using Lemma 3.8 and induction on n, we readily obtain the following result [11, 2].

Theorem 3.13 (Taylor Formula with Remainder)**.** *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be n-times derivate differentiable on $I(a,b)$. Let $x \in I(a,b)$ be given, let $F_{n,x}$ be the nth order derivate function of f at x, and let $M_{n,x}$ be a Lipschitz constant of $F_{n,x}$ on $I(a,b)$. Then for all $y \in I(a,b)$, we have that*

$$f(y) = f(x) + \sum_{j=1}^{n} \frac{f^{(j)}(x)}{j!}(y-x)^j + r_n(x,y)(y-x)^{n+1},$$

with $\lambda(r_n(x,y)) \geq \lambda(M_{n,x})$.

4. Higher Order Optimization

Theorem 4.1 (Necessary Conditions for Existence of Local Extrema)**.** *Let $a < b$ be given in $\mathcal{N}$, let $m \geq 2$, and let $f : I(a,b) \to \mathcal{N}$ be m-times derivate differentiable on $I(a,b)$. Suppose that f has a local extremum at $x_0 \in (a,b)$ and $l \leq m$ is the order of the first nonvanishing derivative of f at x_0. Then l is even. Moreover, $f^{(l)}(x_0)$ is positive if the extremum is a minimum and negative if the extremum is a maximum.*

Proof. Let M_{l,x_0} denote a Lipschitz constant for the lth derivate function of f at x_0. By the Taylor formula, Theorem 3.13, we have for all $x \in I(a,b)$ that

$$f(x) - f(x_0) = \frac{f^{(l)}(x_0)}{l!}(x-x_0)^l + r_l(x_0,x)(x-x_0)^{l+1},$$

with $\lambda\left(r_l(x_0,x)\right) \geq \lambda\left(M_{l,x_0}\right)$. Let $\delta > 0$ in $\mathcal{N}$ be such that

$$\lambda(\delta) > \lambda\left(f^{(l)}(x_0)\right) - \lambda\left(M_{l,x_0}\right) \text{ and } (x_0-\delta, x_0+\delta) \subset (a,b).$$

Then for all $x \in (x_0-\delta, x_0+\delta)$, we have that

$$\lambda\left(r_l(x_0,x)(x-x_0)\right) \geq \lambda\left(M_{l,x_0}\delta\right) = \lambda\left(M_{l,x_0}\right) + \lambda(\delta) > \lambda\left(f^{(l)}(x_0)\right).$$

Thus, $\lambda\left(r_l(x_0,x)(x-x_0)^{l+1}\right) > \lambda\left(f^{(l)}(x_0)(x-x_0)^l/(l!)\right)$; and hence

$$f(x) - f(x_0) \approx \frac{f^{(l)}(x_0)}{l!}(x-x_0)^l \text{ for all } x \neq x_0 \text{ in } (x_0-\delta, x_0+\delta)\,. \tag{4.1}$$

We show that j is even. Suppose not. Then from Equation (4.1), we obtain that $f(x) - f(x_0)$ changes sign as $x - x_0$ crosses 0, which contradicts the fact that f has an extremum at x_0. Thus, j is even. It follows, again from Equation (4.1), that $f(x) > f(x_0)$ for all $x \neq x_0$ in $(x_0-\delta, x_0+\delta)$ and hence f has a local minimum at x_0 if $f^{(j)}(x_0) > 0$. On the other hand, f has a local maximum at x_0 if $f^{(j)}(x_0) < 0$. □

Theorem 4.2 (Sufficient Conditions for Existence of Local Extrema)**.** *Let $a < b$ be given in $\mathcal{N}$, let $k \in \mathbb{N}$, and let $f : I(a,b) \to \mathcal{N}$ be $2k$-times derivate differentiable on $I(a,b)$. Let $x_0 \in (a,b)$ be such $f^{(j)}(x_0) = 0$ for all $j \in \{1,\ldots,2k-1\}$ and $f^{(2k)}(x_0) \neq 0$. Then f has a local minimum at x_0 if $f^{(2k)}(x_0) > 0$ and a local maximum if $f^{(2k)}(x_0) < 0$.*

Proof. The proof is done making use of the remainder formula, Theorem 3.13, just as in the proof of Theorem 4.1. □

The function f in Example 2.11 or Example 2.14 is twice derivate differentiable on $(-1,1)$. Hence the converses of Corollary 2.10 and Proposition 2.13 are not true even under the stronger concept of derivate differentiability. We note here that even though the function g in Example 2.3 was shown to be infinitely often differentiable in the conventional sense, it is only twice (but not three times) derivate differentiable. In the following, we show that under a mild condition on the remainder, the converses of Corollary 2.10 and Proposition 2.13 will be true.

Proposition 4.3. *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be derivate differentiable on $I(a,b)$ with $r_1(x,y)r_1(z,w) \geq 0$ (resp. > 0) for all $x,y,z,w \in I(a,b)$. Then f is convex (resp. strictly convex) if and only if f' is nondecreasing (resp. strictly increasing) on $I(a,b)$.*

Proof. If f is convex (resp. strictly convex) then f' is nondecreasing (resp. strictly increasing) on $I(a,b)$ by Corollary 2.10. Now suppose that f' is nondecreasing (resp. strictly increasing) on $I(a,b)$ and let $x \neq y$ in $I(a,b)$ be given. Then, using the remainder formula 3.8, $f(y) = f(x) + f'(x)(y-x) + r_1(x,y)(y-x)^2$ and $f(x) = f(y) + f'(y)(x-y) + r_1(y,x)(x-y)^2$. Adding the last two equalities, we get that $f'(y) - f'(x) = (r_1(x,y) + r_1(y,x))(y-x)$. Since f' is nondecreasing (resp. strictly increasing) on $I(a,b)$, we obtain that $r_1(x,y) + r_1(y,x) \geq 0$ (resp. > 0). Also, $r_1(x,y)r_1(y,x) \geq 0$ (resp. > 0) by the hypothesis of the proposition. Hence $r_1(x,y) \geq 0$ (resp. > 0) and $r_1(y,x) \geq 0$ (resp. > 0). It follows that $f(y) \geq f(x) + f'(x)(y-x)$ (resp. $f(y) > f(x) + f'(x)(y-x)$). Since this is true for all $x, y \in I(a,b)$, f is convex (resp. strictly convex) on $I(a,b)$ by Proposition 2.9. □

Remark 4.4. *The condition in Proposition 4.3 is automatically satisfied for functions that are continuations of real differentiable functions with nondecreasing (resp. strictly increasing) derivatives; and hence it is not a major restriction.*

Proposition 4.5. *Let $a < b$ be given in $\mathcal{N}$ and let $f : I(a,b) \to \mathcal{N}$ be twice derivate differentiable on $I(a,b)$ with $r_1(x,y)r_1(z,w) \geq 0$ for all $x, y, z, w \in I(a,b)$. Then f is convex on $I(a,b)$ if and only if $f''(x) \geq 0$ for all $x \in I(a,b)$.*

Proof. If f is convex then $f''(x) \geq 0$ for all $x \in I(a,b)$ by Proposition 2.13. Now suppose that $f''(x) \geq 0$ for all $x \in I(a,b)$. Then for all $x, y \in I(a,b)$, we have that $f(y) = f(x) + f'(x)(y-x) + r_1(x,y)(y-x)^2 = f(x) + f'(x)(y-x) + f''(x)(y-x)^2/2 + r_2(x,y)(y-x)^3$. Thus, $r_1(x,y) = f''(x)/2 + r_2(x,y)(y-x)$ for all $x, y \in I(a,b)$.

Let $x, y \in I(a,b)$ be given and let M_2 be a Lipschitz constant of the second derivate function of f at x. Then $|r_2(x,y)| \leq M_2$. Let $z \in I(a,b)$ be such that $|z-x| \ll |f''(x)|/M_2$. Then $|r_2(x,z)(z-x)| \ll |f''(x)|$; and hence $r_1(x,z) \approx f''(x)/2 \geq 0$. It follows then from $r_1(x,y)r_1(x,z) \geq 0$ that $r_1(x,y) \geq 0$; and hence $f(y) \geq f(x) + f'(x)(y-x)$. Therefore, f is convex on $I(a,b)$ by Proposition 2.9. □

5. Acknowledgments

The first author would like to thank his former PhD advisor, Professor Martin Berz, for introducing him to the exciting field of non-Archimedean Analysis and for many fruitful discussions about non-Archimedean fields in general and the Levi-Civita field in particular.

References

[1] M. Berz. Calculus and Numerics on Levi-Civita Fields. In M. Berz, C. Bischof, G. Corliss, and A. Griewank, editors, *Computational Differentiation: Techniques, Applications, and Tools*, pages 19–35, Philadelphia, 1996. SIAM.

[2] M. Berz. Nonarchimedean analysis and rigorous computation. *International Journal of Applied Mathematics*, 2:889–930, 2000.

[3] H. Hahn. Über die nichtarchimedischen Größensysteme. *Sitzungsbericht der Wiener Akademie der Wissenschaften Abt. 2a*, 117:601–655, 1907.

[4] Wolfgang Krull. Allgemeine Bewertungstheorie. *J. Reine Angew. Math.*, 167:160–196, 1932.

[5] Tullio Levi-Civita. Sugli infiniti ed infinitesimi attuali quali elementi analitici. *Atti Ist. Veneto di Sc., Lett. ed Art.*, 7a, 4:1765, 1892.

[6] Tullio Levi-Civita. Sui numeri transfiniti. *Rend. Acc. Lincei*, 5a, 7:91,113, 1898.

[7] S. MacLane. The universality of formal power series fields. *Bulletin American Mathematical Society*, 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] Paulo Ribenboim. Fields: Algebraically Closed and Others. *Manuscripta Mathematica*, 75:115–150, 1992.

[11] K. Shamseddine. *New Elements of Analysis on the Levi-Civita Field*. PhD thesis, Michigan State University, East Lansing, Michigan, USA, 1999. also Michigan State University report MSUCL-1147.

[12] K. Shamseddine and M. Berz. Exception Handling in Derivative Computation with Non-Archimedean Calculus. In M. Berz, C. Bischof, G. Corliss, and A. Griewank, editors, *Computational Differentiation: Techniques, Applications, and Tools*, pages 37–51, Philadelphia, 1996. SIAM.

Department of Mathematics, Michigan State University, East Lansing, MI 48824
E-mail address: khodr@math.msu.edu, zeidan@math.msu.edu