Applied Mathematics and Computation 255 (2015) 44–57

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Analysis on the Levi-Civita field and computational applications

Khodr Shamseddine

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

ARTICLE INFO

Keywords:
Non-Archimedean analysis
Levi-Civita fields
Power series
Measure theory and integration
Optimization
Computational applications

This paper is dedicated to the loving memory of my brother Saïd Shamseddine (1968–2013).

ABSTRACT

In this paper, we present an overview of some of our research on the Levi-Civita fields $\mathcal{R}$ and $\mathcal{C}$. $\mathcal{R}$ (resp. $\mathcal{C}$) is the smallest non-Archimedean field extension of the real (resp. complex) numbers that is Cauchy-complete and real closed (resp. algebraically closed); in fact, $\mathcal{R}$ is small enough to allow for the calculus on the field to be implemented on a computer and used in applications such as the fast and accurate computation of the derivatives of real functions as "differential quotients" up to very high orders. We summarize the convergence and analytical properties of power series, showing that they have the same smoothness behavior as real and complex power series; we present a Lebesgue-like measure and integration theory on the Levi-Civita field $\mathcal{R}$; we discuss solutions to one-dimensional and multi-dimensional optimization problems based on continuity and differentiability concepts that are stronger than the topological ones; and we give a brief summary of the results of our ongoing work on developing a non-Archimedean operator theory on a Banach space over $\mathcal{C}$.

1. Introduction

An overview of recent research on the Levi-Civita fields $\mathcal{R}$ and $\mathcal{C}$ will be presented. We recall that the elements of $\mathcal{R}$ and its complex counterpart $\mathcal{C}$ are functions from $\mathbb{Q}$ to $\mathbb{R}$ and $\mathbb{C}$, respectively, with left-finite support (denoted by supp). That is, below every rational number q, there are only finitely many points where the given function does not vanish. For the further discussion, it is convenient to introduce the following terminology.

Definition 1.1 ($\lambda, \sim, \approx$). For $x \neq 0$ in $\mathcal{R}$ or $\mathcal{C}$, we let $\lambda(x) = \min(\mathrm{supp}(x))$, which exists because of the left-finiteness of $\mathrm{supp}(x)$; and we let $\lambda(0) = +\infty$. Moreover, we denote the value of x at $q \in \mathbb{Q}$ with brackets like $x[q]$.

Given $x,\ y \neq 0$ in $\mathcal{R}$ or $\mathcal{C}$, we say $x \sim y$ if $\lambda(x) = \lambda(y)$; and we say $x \approx y$ if $\lambda(x) = \lambda(y)$ and $x[\lambda(x)] = y[\lambda(y)]$.

At this point, these definitions may feel somewhat arbitrary; but after having introduced an order on $\mathcal{R}$, we will see that λ describes orders of magnitude, the relation $\approx$ corresponds to agreement up to infinitely small relative error, while $\sim$ corresponds to agreement of order of magnitude.

The sets $\mathcal{R}$ and $\mathcal{C}$ are endowed with formal power series multiplication and componentwise addition, which make them into fields [5] in which we can isomorphically embed $\mathbb{R}$ and $\mathbb{C}$ (respectively) as subfields via the map $\Pi : \mathbb{R},\ \mathbb{C} \to \mathcal{R},\ \mathcal{C}$ defined by

$$\Pi(x)[q] = \begin{cases} x & \text{if } q = 0 \\ 0 & \text{else} \end{cases}. \tag{1.1}$$

E-mail address: Khodr.Shamseddine@umanitoba.ca

http://dx.doi.org/10.1016/j.amc.2014.04.108
0096-3003/

Definition 1.2 (*Order in* $\mathcal{R}$). Let $x, y \in \mathcal{R}$ be given. Then we say that $x > y$ (or $y < x$) if $x \neq y$ and $(x-y)[\lambda(x-y)] > 0$; and we say $x \geqslant y$ (or $y \leqslant x$) if $x = y$ or $x > y$.

It follows that the relation $\geqslant$ (or $\leqslant$) defines a total order on $\mathcal{R}$ which makes it into an ordered field. Note that, given $a < b$ in $\mathcal{R}$, we define the $\mathcal{R}$-interval $[a,b] = \{x \in \mathcal{R} : a \leqslant x \leqslant b\}$, with the obvious adjustments in the definitions of the intervals $[a,b[,]a,b]$, and $]a,b[$. Moreover, the embedding Π in Eq. (1.1) of $\mathbb{R}$ into $\mathcal{R}$ is compatible with the order.

The order leads to the definition of an ordinary absolute value on $\mathcal{R}$:

$$|x| = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0; \end{cases}$$

which induces the same topology on $\mathcal{R}$ (called the order topology or valuation topology) as that induced by the ultrametric absolute value:

$$|x|_u = e^{-\lambda(x)},$$

as was shown in [36]. Moreover, two corresponding absolute values are defined on $\mathcal{C}$ in the natural way:

$$|x+iy| = \sqrt{x^2+y^2}; \text{and} |x+iy|_u = e^{-\lambda(x+iy)} = \max\{|x|_u, |y|_u\}.$$

Thus, $\mathcal{C}$ is topologically isomorphic to $\mathcal{R}^2$ provided with the product topology induced by $|\cdot|$ (or $|\cdot|_u$) in $\mathcal{R}$.

We note in passing here that $|\cdot|_u$ is a non-Archimedean valuation on $\mathcal{R}$ (resp. $\mathcal{C}$); that is, it satisfies the following properties

(1) $|v|_u \geqslant 0$ for all $v \in \mathcal{R}$ (resp. $v \in \mathcal{C}$) and $|v|_u = 0$ if and only if $v = 0$;
(2) $|vw|_u = |v|_u |w|_u$ for all $v, w \in \mathcal{R}$ (resp. $v, w \in \mathcal{C}$); and
(3) $|v+w|_u \leqslant \max\{|v|_u, |w|_u\}$ for all $v, w \in \mathcal{R}$ (resp. $v, w \in \mathcal{C}$): the strong triangle inequality.

Thus, $(\mathcal{R}, |\cdot|)$ and $(\mathcal{C}, |\cdot|)$ are non-Archimedean valued fields.

Besides the usual order relations on $\mathcal{R}$, some other notations are convenient.

Definition 1.3 ($\ll, \gg$). Let $x, y \in \mathcal{R}$ be non-negative. We say x is infinitely smaller than y (and write $x \ll y$) if $nx < y$ for all $n \in \mathbb{N}$; we say x is infinitely larger than y (and write $x \gg y$) if $y \ll x$. If $x \ll 1$, we say x is infinitely small; if $x \gg 1$, we say x is infinitely large. Infinitely small numbers are also called infinitesimals or differentials. Infinitely large numbers are also called infinite. Non-negative numbers that are neither infinitely small nor infinitely large are also called finite.

Definition 1.4 (*The Number d*). Let d be the element of $\mathcal{R}$ given by $d[1] = 1$ and $d[q] = 0$ for $q \neq 1$.

It is easy to check that $d^q \ll 1$ if $q > 0$ and $d^q \gg 1$ if $q < 0$. Moreover, for all $x \in \mathcal{R}$ (resp. $\mathcal{C}$), the elements of $\text{supp}(x)$ can be arranged in ascending order, say $\text{supp}(x) = \{q_1, q_2, \ldots\}$ with $q_j < q_{j+1}$ for all j; and x can be written as $x = \sum_{j=1}^{\infty} x[q_j]d^{q_j}$, where the series converges in the valuation topology [5].

Altogether, it follows that $\mathcal{R}$ (resp. $\mathcal{C}$) is a non-Archimedean field extension of $\mathbb{R}$ (resp. $\mathbb{C}$). For a detailed study of these fields, we refer the reader to [5,32,26,6,33,34,40,7,35,41,36,37,28,38,39,30,1,31]. In particular, it is shown that $\mathcal{R}$ and $\mathcal{C}$ are complete with respect to the natural (valuation) topology.

It follows therefore that the fields $\mathcal{R}$ and $\mathcal{C}$ are just special cases of the class of fields discussed in [20]. For a general overview of the algebraic properties of formal power series fields in general, we refer the reader to the comprehensive overview by Ribenboim [19], and for an overview of the related valuation theory to the books by Krull [11], Schikhof [20] and Alling [3]. A thorough and complete treatment of ordered structures can also be found in [18].

Besides being the smallest ordered non-Archimedean field extension of the real numbers that is both complete in the order topology and real closed, 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 [5]; and having infinitely small numbers in the field allows for many computational applications similar to those obtained with the numerical system employed by Sergeyev in [21–25]. One such application is the computation of derivatives of real functions representable on a computer [32], where both the accuracy of formula manipulators and the speed of classical numerical methods are achieved.

In the following sections, we present a brief overview of recent research done on $\mathcal{R}$ and $\mathcal{C}$; and we refer the interested reader to the respective papers for a more detailed study of any of the research topics summarized below.

2. Calculus on $\mathcal{R}$

The following examples show that functions on a finite interval of $\mathcal{R}$ behave in a way that is different from (and even opposite to) what we would expect under similar conditions in $\mathbb{R}$.

Example 2.1. Let $f_1 : [0,1] \to \mathcal{R}$ be given by

$$f_1(x) = \begin{cases} d^{-1} & \text{if } 0 \leqslant x < d \\ d^{-1/\lambda(x)} & \text{if } d \leqslant x \ll 1 \\ 1 & \text{if } x \sim 1. \end{cases}$$

Then f_1 is continuous on [0,1]; but for $d \leqslant x \ll 1,\ f_1(x)$ grows without bound.

Example 2.2. Let $f_2 : [-1, 1] \to \mathcal{R}$ be given by

$$f_2(x) = x - x[0].$$

Then f_2 is continuous on $[-1,1]$. However, f_2 assumes neither a maximum nor a minimum on $[-1,1]$. The set $f_2([-1,1])$ is bounded above by any positive real number and below by any negative real number; but it has neither a least upper bound nor a greatest lower bound.

Example 2.3. Let $f_3 : [0,1] \to \mathcal{R}$ be given by

$$f_3(x) = \begin{cases} 1 & \text{if } x \sim 1 \\ 0 & \text{if } x \ll 1 \end{cases}.$$

Then f_3 is continuous on [0,1] and differentiable on (0,1), with $f_3'(x) = 0$ for all $x \in (0,1)$. We have that $f_3(0) = 0$ and $f_3(1) = 1$; but $f_3(x) \neq 1/2$ for all $x \in [0,1]$. Moreover, f_3 is not constant on [0,1] even though $f_3'(x) = 0$ for all $x \in (0,1)$.

Example 2.4. Let $f_4 : [-1,1] \to \mathcal{R}$ be given by

$$f_4(x) = x[0] + \sum_{\nu=1}^{\infty} x_\nu d^{3q_\nu} \quad \text{when } x = x[0] + \sum_{\nu=1}^{\infty} x_\nu d^{q_\nu}.$$

Then $f_4'(x) = 0$ for all $x \in (-1,1)$. But f_4 is obviously not constant on $[-1,1]$.

Remark 2.5. The extension f of f_4 to $\mathcal{R}$, that is $f : \mathcal{R} \to \mathcal{R}$ given by $f(x)[q] = x[q/3]$, is differentiable on all of $\mathcal{R}$ with vanishing derivative everywhere. Moreover, f is an example of a nontrivial order preserving field automorphism on $\mathcal{R}$ [29]; in $\mathbb{R}$ (or any other ordered Archimedean field) the identity map is the only order preserving field automorphism.

Example 2.6. Let $f_5 : [-1,1] \to \mathcal{R}$ be given by

$$f_5(x) = -f_4(x) + x^4,$$

where f_4 is the function from Example 2.4. Then $f_5'(x) = 4x^3$ for all $x \in (-1,1)$. Thus, $f_5' > 0$ on (0,1); but f_5 is not increasing on $(0,1) : f_5(d^2) > f_5(d)$ even though $d^2 < d$. Also f_5' is strictly increasing and $f_5'' \geqslant 0$ on $(-1,1)$; but f_5 is not convex on $(-1,1)$ since $f_5(d) = -d^3 + d^4 < 0 = f_5(0) + f_5'(0)d$.

Example 2.7. Let $f_6 : [-1,1] \to \mathcal{R}$ be given by

$$f_6(x) = -(f_4(x))^2 + x^8,$$

where f_4 is again the function from Example 2.4. Then f_6 is infinitely often differentiable on $(-1,1)$ with $f_6^{(j)}(0) = 0$ for $1 \leqslant j \leqslant 7$ and $f_6^{(8)}(0) = 8! > 0$. But f_6 has a relative maximum at 0.

The difficulties embodied in the examples above are not specific to $\mathcal{R}$, but are common to all non-Archimedean ordered fields; and they result from the fact that $\mathcal{R}$ is disconnected in the topology induced by the order. This makes developing Analysis on the field more difficult than in the real case; for example, the existence of nonconstant functions whose derivatives vanish everywhere on an interval (as in Example 2.4) makes integration much harder and renders the solutions of the simplest initial value problems (e.g. $y' = 0;\ y(0) = 0$) not unique. To circumvent such difficulties, different approaches have been employed. For example, by imposing stronger conditions on the function than in the real case, we obtain versions of the intermediate value theorem, the inverse function theorem and the implicit function theorem [38,39]; by carefully defining a measure on $\mathcal{R}$ in [35,30], we succeed in developing an integration theory with similar properties to those of the Lebesgue integral of Real Analysis; and by using a stronger concept of continuity and differentiability than in the real case, one-dimensional and multi-dimensional optimization results similar to those from Real Analysis have been obtained for $\mathcal{R}$-valued functions [40,41].

3. Review of power series and $\mathcal{R}$-analytic functions

Power series on the Levi-Civita field $\mathcal{R}$ have been studied in details in [26,33,36,37,28]; work prior to that had been mostly restricted to power series with real coefficients. In [13,14,17,12], they could be studied for infinitely small arguments only, while in [5], using the newly introduced weak topology (see Definition 3.4 below), also finite arguments were possible. Moreover, power series over complete valued fields in general have been studied by Schikhof [20], Alling [3] and others in valuation theory, but always in the valuation topology.

In [33], we study the general case when the coefficients in the power series are Levi-Civita numbers (i.e. elements of $\mathcal{R}$ or $\mathcal{C}$), using the weak convergence. We derive convergence criteria for power series which allow us to define a radius of convergence η such that the power series converges weakly for all points whose distance from the center is smaller than η by a finite amount and it converges in the order topology for all points whose distance from the center is infinitely smaller than η.

In [36] it is shown that, within their radius of convergence, power series are infinitely often differentiable and the derivatives to any order are obtained by differentiating the power series term by term. Also, power series can be re-expanded around any point in their domain of convergence and the radius of convergence of the new series is equal to the difference between the radius of convergence of the original series and the distance between the original and new centers of the series. We then study a class of functions that are given locally by power series (which we call $\mathcal{R}$-analytic functions) and show that they are closed under arithmetic operations and compositions and they are infinitely often differentiable with the derivative functions of all orders being $\mathcal{R}$-analytic themselves.

In [37], we focus on the proof of the intermediate value theorem for the $\mathcal{R}$-analytic functions. Given a function f that is $\mathcal{R}$-analytic on an interval $[a,b]$ and a value S between $f(a)$ and $f(b)$, we use iteration to construct a sequence of numbers in $[a,b]$ that converges strongly to a point $c \in [a,b]$ such that $f(c) = S$. The proof is quite involved, making use of many of the results proved in [33,36] as well as some results from Real Analysis.

Finally, in [28], we state and prove necessary and sufficient conditions for the existence of relative extrema. Then we use that as well as the intermediate value theorem and its proof to prove the extreme value theorem, the mean value theorem, and the inverse function theorem for functions that are $\mathcal{R}$-analytic on an interval $[a,b]$, thus showing that such functions behave as nicely as real analytic functions.

In the following, we summarize some of the key results in [33,36,37,28]. We start with a brief review of the convergence of sequences in two different topologies.

Definition 3.1. A sequence (s_n) in $\mathcal{R}$ or $\mathcal{C}$ is called regular if the union of the supports of all members of the sequence is a left-finite subset of $\mathbb{Q}$.

Definition 3.2. We say that a sequence (s_n) converges strongly in $\mathcal{R}$ or $\mathcal{C}$ if it converges in the valuation topology.

It is shown in [4] that the fields $\mathcal{R}$ and $\mathcal{C}$ are complete with respect to the valuation topology; and a detailed study of strong convergence can be found in [26,33].

Since power series with real (complex) coefficients do not converge strongly for any nonzero real (complex) argument, it is advantageous to study a new kind of convergence. We do that by defining a family of semi-norms on $\mathcal{R}$ or $\mathcal{C}$, which induces a topology weaker than the topology induced by the absolute value and called weak topology [5,26,33,27].

Definition 3.3. Given $r \in \mathbb{R}$, we define a mapping $|\cdot|_r : \mathcal{R}$ or $\mathcal{C} \to \mathbb{R}$ as follows: $|x|_r = \max\{|x[q]| : q \in \mathbb{Q}$ and $q \leqslant r\}$.

The maximum in Definition 3.3 exists in $\mathbb{R}$ since, for any $r \in \mathbb{R}$, only finitely many of the $x[q]$'s considered do not vanish.

Definition 3.4. A sequence (s_n) in $\mathcal{R}$ (resp. $\mathcal{C}$) is said to be weakly convergent if there exists $s \in \mathcal{R}$ (resp. $\mathcal{C}$), called the weak limit of the sequence (s_n), such that for all $\epsilon > 0$ in $\mathbb{R}$, there exists $N \in \mathbb{N}$ such that $|s_m - s|_{1/\epsilon} < \epsilon$ for all $m \geqslant N$.

It is shown [5] that $\mathcal{R}$ and $\mathcal{C}$ are not Cauchy complete with respect to the weak topology and that strong convergence implies weak convergence to the same limit. A detailed study of weak convergence is found in [5,26,33,27].

3.1. Power series

In the following, we review strong and weak convergence criteria for power series, Theorems 3.5 and 3.6, the proofs of which are given in [33]. We also note that Theorem 3.5 is a special case of the result on page 59 of [20].

Theorem 3.5 (Strong Convergence Criterion for Power Series). *Let (a_n) be a sequence in $\mathcal{R}$ (resp. $\mathcal{C}$), and let*

$$\lambda_0 = \limsup_{n\to\infty}\left(\frac{-\lambda(a_n)}{n}\right) \text{ in } \mathbb{R}\cup\{-\infty,\infty\}.$$

Let $x_0 \in \mathcal{R}$ (resp. $\mathcal{C}$) be fixed and let $x \in \mathcal{R}$ (resp. $\mathcal{C}$) be given. Then the power series $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ converges strongly if $\lambda(x-x_0) > \lambda_0$ and is strongly divergent if $\lambda(x-x_0) < \lambda_0$ or if $\lambda(x-x_0) = \lambda_0$ and $-\lambda(a_n)/n > \lambda_0$ for infinitely many n.

Theorem 3.6 (Weak Convergence Criterion for Power Series). *Let (a_n) be a sequence in $\mathcal{R}$ (resp. $\mathcal{C}$), and let $\lambda_0 = \limsup_{n\to\infty}(-\lambda(a_n)/n) \in \mathbb{Q}$. Let $x_0 \in \mathcal{R}$ (resp. $\mathcal{C}$) be fixed, and let $x \in \mathcal{R}$ (resp. $\mathcal{C}$) be such that $\lambda(x - x_0) = \lambda_0$. For each $n \geqslant 0$, let $b_n = a_n d^{n\lambda_0}$. Suppose that the sequence (b_n) is regular and write $\bigcup_{n=0}^{\infty} supp(b_n) = \{q_1, q_2, \ldots\}$; with $q_{j_1} < q_{j_2}$ if $j_1 < j_2$. For each n, write $b_n = \sum_{j=1}^{\infty} b_{n_j} d^{q_j}$, where $b_{n_j} = b_n[q_j]$. Let*

$$\eta = \frac{1}{\sup\left\{\limsup_{n\to\infty}|b_{n_j}|^{1/n} : j \geqslant 1\right\}} \quad \text{in } \mathbb{R}\cup\{\infty\}, \tag{3.1}$$

with the conventions $1/0 = \infty$ and $1/\infty = 0$. Then $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converges absolutely weakly if $|(x - x_0)[\lambda_0]| < \eta$ and is weakly divergent if $|(x - x_0)[\lambda_0]| > \eta$.

Remark 3.7. The number η in Eq. (3.1) is referred to as the radius of weak convergence of the power series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$.

As an immediate consequence of Theorem 3.6, we obtain the following result which allows us to extend real and complex functions representable by power series to the Levi-Civita fields $\mathcal{R}$ and $\mathcal{C}$. This result is of particular interest for the application [32] mentioned in Section 1 above and discussed in Section 6 below.

Corollary 3.8. (Power Series with Purely Real or Complex Coefficients). *Let $\sum_{n=0}^{\infty} a_n X^n$ be a power series with purely real (resp. complex) coefficients and with classical radius of convergence equal to η. Let $x \in \mathcal{R}$ (resp. $\mathcal{C}$), and let $A_n(x) = \sum_{j=0}^{n} a_j x^j \in \mathcal{R}$ (resp. $\mathcal{C}$). Then, for $|x| < \eta$ and $|x| \neg\approx \eta$, the sequence $(A_n(x))$ converges absolutely weakly. We define the limit to be the continuation of the power series to $\mathcal{R}$ (resp. $\mathcal{C}$).*

Definition 3.9 (*The Functions Exp, Cos, Sin, Cosh, and Sinh*). By Corollary 3.8, the series

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}, \quad \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}, \quad \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}, \quad \sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}, \quad \text{and} \sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}$$

converge absolutely weakly in $\mathcal{R}$ (resp. $\mathcal{C}$) for any $x \in \mathcal{R}$ (resp. $\mathcal{C}$), at most finite in (ordinary) absolute value (that is, for $\lambda(x) \geq 0$). For any such x, define

$$\exp(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!};$$
$$\cos(x) = \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!};$$
$$\sin(x) = \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!};$$
$$\cosh(x) = \sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!};$$
$$\sinh(x) = \sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}.$$

A detailed study of the transcendental functions introduced on $\mathcal{R}$ in Definition 3.9 can be found in [26]. In particular, we show that addition theorems similar to the real ones hold, which is essential for the implementation of these functions on a computer (see Section 1.5 in [26]).

3.2. $\mathcal{R}$-analytic functions

In this section, we review the algebraic and analytical properties of a class of functions that are given locally by power series and we refer the reader to [36,37,28] for a more detailed study.

Definition 3.10. Let $a < b$ in $\mathcal{R}$ be given and let $f : [a, b] \to \mathcal{R}$. Then we say that f is expandable or $\mathcal{R}$-analytic on $[a, b]$ if for all $x \in [a, b]$ there exists a positive $\delta \sim b - a$ in $\mathcal{R}$, and there exists a regular sequence $(a_n(x))$ in $\mathcal{R}$ such that, under weak convergence, $f(y) = \sum_{n=0}^{\infty} a_n(x)(y - x)^n$ for all $y \in (x - \delta, x + \delta) \cap [a, b]$.

It is shown in [36] that if f is $\mathcal{R}$-analytic on $[a, b]$ then f is bounded on $[a, b]$; also, if g is $\mathcal{R}$-analytic on $[a, b]$ and $\alpha \in \mathcal{R}$ then $f + \alpha g$ and $f \cdot g$ are $\mathcal{R}$-analytic on $[a, b]$. Moreover, the composition of $\mathcal{R}$-analytic functions is $\mathcal{R}$-analytic. Furthermore, using the fact that power series on $\mathcal{R}$ are infinitely often differentiable within their domain of convergence and the derivatives to any order are obtained by differentiating the power series term by term [36], we obtain the following result.

Theorem 3.11. *Let $a < b$ in $\mathcal{R}$ be given, and let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$. Then f is infinitely often differentiable on $[a,b]$, and for any positive integer m, we have that $f^{(m)}$ is $\mathcal{R}$-analytic on $[a,b]$. Moreover, if f is given locally around $x_0 \in [a,b]$ by $f(x) = \sum_{n=0}^{\infty} a_n(x_0)(x-x_0)^n$, then $f^{(m)}$ is given by*

$$f^{(m)}(x) = \sum_{n=m}^{\infty} n(n-1)\dots(n-m+1)a_n(x_0)(x-x_0)^{n-m}.$$

In particular, we have that $a_m(x_0) = f^{(m)}(x_0)/m!$ for all $m = 0,1,2,\dots$.

In [37], we prove the intermediate value theorem for $\mathcal{R}$-analytic functions on an interval $[a,b]$.

Theorem 3.12 (Intermediate Value Theorem). *Let $a < b$ in $\mathcal{R}$ be given and let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$. Then f assumes on $[a,b]$ every intermediate value between $f(a)$ and $f(b)$.*

Since Theorem 3.12 is a central result in the study of power series and $\mathcal{R}$-analytic functions, we present in the following the key steps of the proof and refer the reader to [37] for the detailed (lengthy) proof.

- Without loss of generality, we may assume that f is not constant on $[a,b]$. Let $F : [0,1] \to \mathcal{R}$ be given by

$$F(x) = f((b-a)x + a) - \frac{f(a)+f(b)}{2}.$$

Then F is $\mathcal{R}$-analytic on $[0,1]$; and f assumes on $[a,b]$ every intermediate value between $f(a)$ and $f(b)$ if and only if F assumes on $[0,1]$ every intermediate value between $F(0) = (f(a)-f(b))/2$ and $F(1) = (f(b)-f(a))/2 = -F(0)$. So without loss of generality, we may assume that $a = 0,\ b = 1$, and $f = F$. Also, since scaling the function by a constant factor does not affect the existence of intermediate values, we may assume that

$$i(f) := \min\{\mathrm{supp}(f(x)) : x \in [0,1]\} = 0.$$

- We define $f_R : [0,1] \cap \mathbb{R} \to \mathbb{R}$ by $f_R(X) = f(X)[0]$. Then f_R is a real-valued analytic function on the real interval $[0,1] \cap \mathbb{R}$. Let S be between $f(a) = f(0)$ and $f(b) = f(1)$; and let $S_R = S[0]$. Then S_R is a real value between $f_R(0)$ and $f_R(1)$. We use the classical intermediate value theorem to find a real point $X_0 \in [0,1]$ such that $f_R(X_0) = S_R$.
- We use iteration to construct a convergent sequence (x_n) such that $\lambda(x_n) > 0$ and $\lambda(x_{n+2} - x_{n+1}) > \lambda(x_{n+1} - x_n)$ for all $n \in \mathbb{N}$. Let $x = \lim_{n\to\infty} x_n$; then $\lambda(x) > 0$, and we show that

$$X_0 + x \in [0,1] \quad \text{and } f(X_0 + x) = S.$$

A close look at that proof shows that if f is not constant on $[a,b]$ and S is between $f(a)$ and $f(b)$ then there are only finitely many points c in $[a,b]$ such that $f(c) = S$. This is crucial for the proof of the extreme value theorem for the $\mathcal{R}$-analytic functions in [28].

In [28], we complete the study of $\mathcal{R}$-analytic functions: we state and prove necessary and sufficient conditions for the existence of relative extrema; then we prove the extreme value theorem, the mean value theorem and the inverse function theorem for these functions, thus showing that $\mathcal{R}$-analytic functions have all the nice properties of real analytic functions.

Theorem 3.13. *Let $a < b$ in $\mathcal{R}$ be given; let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$; let $x_0 \in (a,b)$ and let $m \in \mathbb{N}$ be the order of the first nonvanishing derivative of f at x_0. Then f has a relative extremum at x_0 if and only if m is even. In that case (m is even), the extremum is a minimum if $f^{(m)}(x_0) > 0$ and a maximum if $f^{(m)}(x_0) < 0$.*

Theorem 3.14 (Extreme Value Theorem). *Let $a < b$ in $\mathcal{R}$ be given and let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$. Then f assumes a maximum and a minimum on $[a,b]$.*

Using the intermediate value theorem and the extreme value theorem, then the following results become easy to prove.

Corollary 3.15. *Let $a < b$ in $\mathcal{R}$ be given and let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$. Then there exist $m, M \in \mathcal{R}$ such that $f([a,b]) = [m,M]$.*

Corollary 3.16 (Mean Value Theorem). *Let $a < b$ in $\mathcal{R}$ be given and let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$. Then there exists $c \in (a,b)$ such that*

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

Corollary 3.17. *Let $a < b$ in $\mathcal{R}$ be given, and let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$. Then the following are true.*

(i) *If $f'(x) \neq 0$ for all $x \in (a,b)$ then either $f'(x) > 0$ for all $x \in (a,b)$ and f is strictly increasing on $[a,b]$, or $f'(x) < 0$ for all $x \in (a,b)$ and f is strictly decreasing on $[a,b]$.*
(ii) *If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant on $[a,b]$.*

Corollary 3.18 (*Inverse Function Theorem*). *Let $a < b$ in $\mathcal{R}$ be given, let $f : [a,b] \to \mathcal{R}$ be $\mathcal{R}$-analytic on $[a,b]$, and let $x_0 \in (a,b)$ be such that $f'(x_0) > 0$ (resp. $f'(x_0) < 0$). Then there exists $\delta > 0$ in $\mathcal{R}$ such that*

(i) *$f' > 0$ and f is strictly increasing (resp. $f' < 0$ and f is strictly decreasing) on $[x_0 - \delta, x_0 + \delta]$.*
(ii) *$f([x_0 - \delta, x_0 + \delta]) = [m, M]$ where $m = f(x_0 - \delta)$ and $M = f(x_0 + \delta)$ (resp. $m = f(x_0 + \delta)$ and $M = f(x_0 - \delta)$).*
(iii) *$\exists g : [m, M] \to [x_0 - \delta, x_0 + \delta]$, strictly increasing (resp. strictly decreasing) on $[m, M]$, such that*
  – *g is the inverse of f on $[x_0 - \delta, x_0 + \delta]$;*
  – *g is differentiable on $[m, M]$; and for all $y \in [m, M]$,*

$$g'(y) = \frac{1}{f'(g(y))}.$$

Remark 3.19. Since power series over $\mathcal{R}$ are $\mathcal{R}$-analytic on any interval within their domain of convergence, all the results of Section 3.2 hold as well for power series on any interval in which the series converges.

4. Measure theory and integration

Using the nice smoothness properties of power series summarized above, we developed a Lebesgue-like measure and integration theory on $\mathcal{R}$ in [35,30] that uses the power series as the family of simple functions instead of the step functions used in the real case. This was possible in particular because the family $\mathcal{S}(a,b)$ of power series (that converge weakly) on a given interval $I(a,b) \subset \mathcal{R}$ (where $I(a,b)$ denotes any one of the intervals $[a,b]$, $(a,b]$, $[a,b)$ or (a,b)) satisfies the following crucial properties.

(1) $\mathcal{S}(a,b)$ is an algebra that contains the identity function;
(2) for all $f \in \mathcal{S}(a,b)$, f is Lipschitz on $I(a,b)$ and there exists an anti-derivative F of f in $\mathcal{S}(a,b)$, which is unique up to a constant;
(3) for all differentiable $f \in \mathcal{S}(a,b)$, if $f' = 0$ on (a,b) then f is constant on $I(a,b)$; moreover, if $f' \geqslant 0$ on (a,b) then f is non-decreasing on $I(a,b)$.

Definition 4.1. Let $A \subset \mathcal{R}$ be given. Then we say that A is measurable if for every $\epsilon > 0$ in $\mathcal{R}$, there exist a sequence of mutually disjoint intervals (I_n) and a sequence of mutually disjoint intervals (J_n) such that $\cup_{n=1}^{\infty} I_n \subset A \subset \cup_{n=1}^{\infty} J_n$, $\sum_{n=1}^{\infty} l(I_n)$ and $\sum_{n=1}^{\infty} l(J_n)$ converge in $\mathcal{R}$, and $\sum_{n=1}^{\infty} l(J_n) - \sum_{n=1}^{\infty} l(I_n) \leqslant \epsilon$.

Given a measurable set A, then for every $k \in \mathbb{N}$, we can select a sequence of mutually disjoint intervals (I_n^k) and a sequence of mutually disjoint intervals (J_n^k) such that $\sum_{n=1}^{\infty} l(I_n^k)$ and $\sum_{n=1}^{\infty} l(J_n^k)$ converge in $\mathcal{R}$ for all k,

$$\cup_{n=1}^{\infty} I_n^k \subset \cup_{n=1}^{\infty} I_n^{k+1} \subset A \subset \cup_{n=1}^{\infty} J_n^{k+1} \subset \cup_{n=1}^{\infty} J_n^k \quad \text{and} \quad \sum_{n=1}^{\infty} l\left(J_n^k\right) - \sum_{n=1}^{\infty} l\left(I_n^k\right) \leqslant d^k$$

for all $k \in \mathbb{N}$. Since $\mathcal{R}$ is Cauchy-complete in the order topology, it follows that $\lim_{k\to\infty} \sum_{n=1}^{\infty} l(I_n^k)$ and $\lim_{k\to\infty} \sum_{n=1}^{\infty} l(J_n^k)$ both exist and they are equal. We call the common value of the limits the measure of A and we denote it by $m(A)$. Thus,

$$m(A) = \lim_{k\to\infty} \sum_{n=1}^{\infty} l(I_n^k) = \lim_{k\to\infty} \sum_{n=1}^{\infty} l(J_n^k).$$

We prove in [35] that the measure defined above has similar properties to those of the Lebesgue measure of Real Analysis. Then we define a measurable function on a measurable set $A \subset \mathcal{R}$ using Definition 4.1 and simple functions (convergent power series).

Definition 4.2. Let $A \subset \mathcal{R}$ be a measurable subset of $\mathcal{R}$ and let $f : A \to \mathcal{R}$ be bounded on A. Then we say that f is measurable on A if for all $\epsilon > 0$ in $\mathcal{R}$, there exists a sequence of mutually disjoint intervals (I_n) such that $I_n \subset A$ for all n, $\sum_{n=1}^{\infty} l(I_n)$ converges in $\mathcal{R}$, $m(A) - \sum_{n=1}^{\infty} l(I_n) \leqslant \epsilon$ and f is simple on I_n for all n.

In [35], we derive a simple characterization of measurable functions and we show that they form an algebra. Then we show that a measurable function is differentiable almost everywhere and that a function measurable on two measurable subsets of $\mathcal{R}$ is also measurable on their union and intersection.

We define the integral of a simple function over an interval $I(a,b)$ and we use that to define the integral of a measurable function f over a measurable set A.

Definition 4.3. Let $a < b$ in $\mathcal{R}$, let $f : I(a,b) \to \mathcal{R}$ be simple on $I(a,b)$, and let F be a simple anti-derivative of f on $I(a,b)$. Then the integral of f over $I(a,b)$ is the $\mathcal{R}$ number

$$\int_{I(a,b)} f = \lim_{x \to b} F(x) - \lim_{x \to a} F(x).$$

The limits in Definition 4.3 account for the case when the interval $I(a,b)$ does not include one or both of the end points; and these limits exist since F is Lipschitz on $I(a,b)$.

Now let $A \subset \mathcal{R}$ be measurable, let $f : A \to \mathcal{R}$ be measurable and let M be a bound for $|f|$ on A. Then for every $k \in \mathbb{N}$, there exists a sequence of mutually disjoint intervals $(I_n^k)_{n \in \mathbb{N}}$ such that $\cup_{n=1}^{\infty} I_n^k \subset A$, $\sum_{n=1}^{\infty} l(I_n^k)$ converges, $m(A) - \sum_{n=1}^{\infty} l(I_n^k) \leqslant d^k$, and f is simple on I_n^k for all $n \in \mathbb{N}$. Without loss of generality, we may assume that $I_n^k \subset I_n^{k+1}$ for all $n \in \mathbb{N}$ and for all $k \in \mathbb{N}$. Since $\lim_{n \to \infty} l(I_n^k) = 0$, and since $\left| \int_{I_n^k} f \right| \leqslant M l(I_n^k)$ (proved in [35] for simple functions), it follows that

$$\lim_{n \to \infty} \int_{I_n^k} f = 0 \quad \text{for all } k \in \mathbb{N}.$$

Thus, $\sum_{n=1}^{\infty} \int_{I_n^k} f$ converges in $\mathcal{R}$ for all $k \in \mathbb{N}$ [33].

We show that the sequence $\left(\sum_{n=1}^{\infty} \int_{I_n^k} f \right)_{k \in \mathbb{N}}$ converges in $\mathcal{R}$; and we define the unique limit as the integral of f over A.

Definition 4.4. Let $A \subset \mathcal{R}$ be measurable and let $f : A \to \mathcal{R}$ be measurable. Then the integral of f over A, denoted by $\int_A f$, is given by

$$\int_A f = \lim_{\substack{\sum_{n=1}^{\infty} l(I_n) \to m(A) \\ \cup_{n=1}^{\infty} I_n \subset A \\ (I_n) \text{ are mutually disjoint} \\ f \text{ is simple on } I_n \ \forall\ n}} \sum_{n=1}^{\infty} \int_{I_n} f.$$

It turns out that the integral in Definition 4.4 satisfies similar properties to those of the Lebesgue integral on $\mathbb{R}$ [35]. In particular, we prove the linearity property of the integral and that if $|f| \leqslant M$ on A then $\left| \int_A f \right| \leqslant M m(A)$, where $m(A)$ is the measure of A. We also show that the sum of the integrals of a measurable function over two measurable sets is equal to the sum of its integrals over the union and the intersection of the two sets.

In [30], which is a continuation of the work done in [35] and complements it, we show, among other results, that the uniform limit of a sequence of convergent power series on an interval $I(a,b)$ is again a power series that converges on $I(a,b)$. Then we use that to prove the uniform convergence theorem in $\mathcal{R}$.

Theorem 4.5. *Let $A \subset \mathcal{R}$ be measurable, let $f : A \to \mathcal{R}$, for each $k \in \mathbb{N}$ let $f_k : A \to \mathcal{R}$ be measurable on A, and let the sequence (f_k) converge uniformly to f on A. Then f is measurable on A, $\lim_{k \to \infty} \int_A f_k$ exists, and*

$$\lim_{k \to \infty} \int_A f_k = \int_A f.$$

5. Optimization

In [40], we consider unconstrained one-dimensional optimization on $\mathcal{R}$. 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 $\mathcal{R}$. We show that for first order optimization, the results are similar to the corresponding real ones. However, for second and higher order optimization, we show that conventional differentiability is not strong enough to just extend the real-case results (see Examples 2.6 and 2.7); and a stronger concept of differentiability, the so-called derivate differentiability (see Definition 5.4 below), is used to solve that difficulty. We also characterize convex functions on convex sets of $\mathcal{R}$ in terms of first and second order derivatives.

In the following, we review the definitions of derivate continuity and differentiability in one dimension, as well as some related results and we refer the interested reader to [6,26,31] for a more detailed study. Throughout this section, $I(a,b)$ will denote any one of the intervals $]a,b[$, $]a,b]$, $[a,b[$ or $[a,b]$.

Definition 5.1. Let $a < b$ be given in $\mathcal{R}$ and let $f : I(a,b) \to \mathcal{R}$. Then we say that f is derivate continuous on $I(a,b)$ if there exists $M \in \mathcal{R}$, called a Lipschitz constant of f on $I(a,b)$, such that

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

Note that the $|\cdot|$ in Definition 5.1 is the ordinary absolute value defined on $\mathcal{R}$ in the Introduction. It follows immediately from Definition 5.1 that if $f : I(a,b) \to \mathcal{R}$ is derivate continuous on $I(a,b)$ then f is uniformly continuous and bounded on $I(a,b)$.

Remark 5.2. It is clear that the concept of derivate continuity in Definition 5.1 coincides with that of Lipschitz continuity when restricted to $\mathbb{R}$. We chose to call it derivate continuity here so that, after having defined derivate differentiability in Definition 5.4 and higher order derivate differentiability in Definition 5.6, we can think of derivate continuity as derivate differentiability of "order zero", just as is the case for continuity in $\mathbb{R}$.

Remark 5.3. Definition 5.1 can be generalized in the obvious way to functions on any countable unions of intervals of $\mathcal{R}$.

Definition 5.4. Let $a < b$ be given in $\mathcal{R}$, let $f : I(a,b) \to \mathcal{R}$ 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 $\frac{f-f(x)}{I_d-x} : I(a,b)\setminus\{x\} \to \mathcal{R}$ is derivate continuous on $I(a,b)\setminus\{x\}$. In this case the unique continuation of $\frac{f-f(x)}{I_d-x}$ to $I(a,b)$ 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 5.4 that if $f : I(a,b) \to \mathcal{R}$ is derivate differentiable then f is differentiable in the conventional sense; moreover, the two derivatives at any given point of $I(a,b)$ agree. As for derivate continuity, the definition of derivate differentiability can be generalized to functions on countable unions of intervals of $\mathcal{R}$.

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

Theorem 5.5. *Let $a < b$ be given in $\mathcal{R}$ and let $f : I(a,b) \to \mathcal{R}$ be derivate continuous on $I(a,b)$. Suppose there exists $M \in \mathcal{R}$ and there exists a function $g : I(a,b) \to \mathcal{R}$ such that*

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

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

Definition 5.6. (*n-times Derivate Differentiability*). Let $a < b$ be given in $\mathcal{R}$, and let $f : I(a,b) \to \mathcal{R}$. Let $n \geqslant 2$ 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 $f^{(n)}(x)$.

One of the most useful consequences of the derivate differentiability concept is that it gives rise to a Taylor formula with remainder while the conventional (topological) differentiability does not. We only state the result here and refer the reader to [6,26,31] for its proof. We also note that, as an immediate result of Theorem 5.7, we obtain local expandability in Taylor series around $x_0 \in I(a,b)$ of a given function that is infinitely often derivate differentiable on $I(a,b)$ [6,26,31].

Theorem 5.7. (Taylor Formula with Remainder). *Let $a < b$ be given in $\mathcal{R}$ and let $f : I(a,b) \to \mathcal{R}$ 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 $F_{n,x}$. 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 $\lambda(r_n(x,y)) \geqslant \lambda(M_{n,x})$.

Using Theorem 5.7, we are able to generalize in [40] most of one-dimensional optimization results of Real Analysis. For example, we obtain the following two results which state necessary and sufficient conditions for the existence of local (relative) extrema.

Theorem 5.8 (Necessary Conditions for Existence of Local Extrema). *Let $a < b$ be given in $\mathcal{R}$, let $m \geqslant 2$, and let $f : I(a,b) \to \mathcal{R}$ be m-times derivate differentiable on $I(a,b)$. Suppose that f has a local extremum at $x_0 \in (a,b)$ and $l \leqslant m$ 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.*

Theorem 5.9 (Sufficient Conditions for Existence of Local Extrema). *Let $a < b$ be given in $\mathcal{R}$, let $k \in \mathbb{N}$, and let $f : I(a,b) \to \mathcal{R}$ 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 $f^{(2k)}(x_0) < 0$.*

In [41,31], we generalize the concepts of derivate continuity and differentiability to higher dimensions; and this yields a Taylor Formula with a bounded remainder term for C^m functions (in the derivate sense) from an open subset of $\mathcal{R}^n$ to $\mathcal{R}$.

Theorem 5.10 (Taylor Formula for Functions of Several Variables). *Let $D \subset \mathcal{R}^n$ be open, let $\vec{x}_0 \in D$ be given and let $f : D \to \mathcal{R}$ be C^q on D. Then there exist $M, \delta > 0$ in $\mathcal{R}$ such that $B_\delta(\vec{x}_0) \subset D$ and, for all $\vec{x} \in B_\delta(\vec{x}_0)$, we have that*

$$f(\vec{x}) = f(\vec{x}_0) + \sum_{j=1}^{q}\left(\frac{1}{j!}\sum_{l_1,\ldots,l_j=1}^{n}\left(\partial_{l_1}\ldots\partial_{l_j}f(\vec{x}_0)\pi_{k=1}^{j}\left(x_{l_k} - x_{0,l_k}\right)\right)\right) + R_{q+1}(\vec{x}_0,\vec{x}),$$

where $\left|R_{q+1}(\vec{x}_0,\vec{x})\right| \leqslant M|\vec{x} - \vec{x}_0|^{q+1}$.

Then we use that to derive necessary and sufficient conditions of second order for the existence of a minimum of an $\mathcal{R}$-valued function on $\mathcal{R}^n$ subject to equality and inequality constraints. More specifically, we solve the problem of minimizing a function $f : \mathcal{R}^n \to \mathcal{R}$, subject to the following set of constraints:

$$\begin{cases} h_1(\vec{x}) & = & 0 \\ & \vdots & \\ h_m(\vec{x}) & = & 0 \end{cases} \text{and} \begin{cases} g_1(\vec{x}) & \leqslant & 0 \\ & \vdots & \\ g_p(\vec{x}) & \leqslant & 0 \end{cases}, \tag{5.1}$$

where all the functions in Eq. (5.1) are from $\mathcal{R}^n$ to $\mathcal{R}$. A point $\vec{x}_0 \in \mathcal{R}^n$ is said to be a feasible point if it satisfies the constraints in Eq. (5.1).

Definition 5.11. Let $\vec{x}_0$ be a feasible point for the constraints in Eq. (5.1) and let $I(\vec{x}_0) = \{l \in \{1,\ldots,p\} : g_l(\vec{x}_0) = 0\}$. Then we say that $\vec{x}_0$ is regular for the constraints if $\{\nabla h_j(\vec{x}_0) : j = 1,\ldots,m; \nabla g_l(\vec{x}_0) : l \in I(\vec{x}_0)\}$ forms a linearly independent subset of vectors in $\mathcal{R}^n$.

The following theorem provides *necessary conditions* of second order for a local minimizer $\vec{x}_0$ of a function f subject to the constraints in Eq. (5.1). The result is a generalization of the corresponding real result [15,10] and the proof (see [41]) is similar to that of the latter; but one essential difference is the form of the remainder formula. In the real case, the remainder term is related to the second derivative at some intermediate point, while here that is not the case. However, the concept of derivate differentiability puts a bound on the remainder term; and this is instrumental in the proof of the theorem.

Theorem 5.12. *Suppose that f, $\{h_j\}_{j=1}^m$, $\{g_l\}_{l=1}^p$ are C^2 on some open set $D \subset \mathcal{R}^n$ containing the point $\vec{x}_0$ and that $\vec{x}_0$ is a regular point for the constraints in Eq. (5.1). If $\vec{x}_0$ is a local minimizer for f under the given constraints, then there exist $\alpha_1,\ldots,\alpha_m,\beta_1,\ldots,\beta_p \in \mathcal{R}$ such that*

(i) $\beta_l \geqslant 0$ *for all* $l \in \{1,\ldots,p\}$,
(ii) $\beta_l g_l(\vec{x}_0) = 0$ *for all* $l \in \{1,\ldots,p\}$,
(iii) $\nabla f(\vec{x}_0) + \sum_{j=1}^m \alpha_j \nabla h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l \nabla g_l(\vec{x}_0) = \vec{0}$, *and*
(iv) $\vec{y}^T\left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^m \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l \nabla^2 g_l(\vec{x}_0)\right)\vec{y} \geqslant 0$ *for all* $\vec{y} \in \mathcal{R}^n$ *satisfying* $\nabla h_j(\vec{x}_0)\vec{y} = 0$ *for all* $j \in \{1,\ldots,m\}$, $\nabla g_l(\vec{x}_0)\vec{y} = 0$ *for all* $l \in L = \{k \in I(\vec{x}_0) : \beta_k > 0\}$ *and* $\nabla g_l(\vec{x}_0)\vec{y} \leqslant 0$ *for all* $l \in I(\vec{x}_0) \setminus L$.

In the following theorem, we present second order *sufficient conditions* for a feasible point $\vec{x}_0$ to be a local minimum of a function f subject to the constraints in Eq. (5.1). It is a generalization of the real result [10] and reduces to it, when restricted to functions from $\mathbb{R}^n$ to $\mathbb{R}$. In fact, since ϵ in condition (iv) below is allowed to be infinitely small, the condition $|\nabla h_j(\vec{x}_0)\vec{y}| < \epsilon$ would reduce to $\nabla h_j(\vec{x}_0)\vec{y} = 0$, when restricted to $\mathbb{R}$. Similarly, one can readily see that the other conditions are mere generalizations of the corresponding real ones. However, the proof (see [41]) is different than that of the real result since the supremum principle does not hold in $\mathcal{R}$.

Theorem 5.13. *Suppose that f, $\{h_j\}_{j=1}^m$, $\{g_l\}_{l=1}^p$ are C^2 on some open set $D \subset \mathcal{R}^n$ containing the point $\vec{x}_0$ and that $\vec{x}_0$ is a feasible point for the constraints in Eq. (5.1) such that, for some $\alpha_1,\ldots,\alpha_m,\beta_1,\ldots,\beta_p \in \mathcal{R}$ and for some ϵ, $\gamma > 0$ in $\mathcal{R}$, we have that*

(i) $\beta_l \geqslant 0$ *for all* $l \in \{1,\ldots,p\}$,
(ii) $\beta_l g_l(\vec{x}_0) = 0$ *for all* $l \in \{1,\ldots,p\}$,
(iii) $\nabla f(\vec{x}_0) + \sum_{j=1}^m \alpha_j \nabla h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l \nabla g_l(\vec{x}_0) = \vec{0}$, *and*
(iv) $\vec{y}^T\left(\nabla^2 f(\vec{x}_0) + \sum_{j=1}^m \alpha_j \nabla^2 h_j(\vec{x}_0) + \sum_{l=1}^p \beta_l \nabla^2 g_l(\vec{x}_0)\right)\vec{y} \geqslant \gamma$ *for all* $\vec{y} \in \mathcal{R}^n$ *satisfying* $|\vec{y}| = 1$, $|\nabla h_j(\vec{x}_0)\vec{y}| < \epsilon$ *for all* $j \in \{1,\ldots,m\}$, $|\nabla g_l(\vec{x}_0)\vec{y}| < \epsilon$ *for all* $l \in L = \{k : \beta_k > 0\}$ *and* $\nabla g_l(\vec{x}_0)\vec{y} < \epsilon$ *for all* $l \in I(\vec{x}_0) \setminus L$, *where* $I(\vec{x}_0) = \{k : g_k(\vec{x}_0) = 0\}$.

Then $\vec{x}_0$ is a strict local minimum for f under the constraints of Eq. (5.1).

6. Computation of derivatives of real functions

The general question of efficient differentiation is at the core of many parts of the work on perturbation and aberration theories relevant in Physics and Engineering; for an overview, see for example [8]. In this case, derivatives of highly complicated functions have to be computed to high orders. However, even when the derivative of the function is known to exist at the given point, numerical methods fail to give an accurate value of the derivative; the error increases with the order, and for orders greater than three, the errors often become too large for the results to be practically useful.

On the other hand, while formula manipulators like Mathematica are successful in finding low-order derivatives of simple functions, they fail for high-order derivatives of very complicated functions. Moreover, they fail to find the derivatives of certain functions at given points even though the functions are differentiable at the respective points. This is generally connected to the occurrence of non-differentiable parts that do not affect the differentiability of the end result as well as the occurrence of branch points in coding as in IF-ELSE structures.

Using calculus on $\mathcal{R}$ and the fact that the field has infinitely small numbers represents a new method for computational differentiation that avoids the well-known accuracy problems of numerical differentiation tools. It also avoids the often rather stringent limitations of formula manipulators that restrict the complexity of the function that can be differentiated, and the orders to which differentiation can be performed.

By a computer function, we denote any real-valued function that can be typed on a computer. The $\mathcal{R}$ numbers as well as the continuations to $\mathcal{R}$ of the intrinsic functions (and hence of all computer functions) have all been implemented for use on a computer, using the code COSY INFINITY [9,16]. Using the calculus on $\mathcal{R}$, we formulate a necessary and sufficient condition for the derivatives of a computer function to exist, and show how to find these derivatives whenever they exist [32,34]. The new technique of computing the derivatives of computer functions, which we summarize below, achieves results that combine the accuracy of formula manipulators with the speed of classical numerical methods, that is the best of both worlds. The method is much faster than Mathematica and other formula manipulators since no symbolic differentiation is required before the numerical evaluation of the derivatives. Moreover, the results obtained are accurate up to machine precision-the error is infinitely small and hence it does not mix with the real derivative; this represents a clear advantage over traditional numerical differentiation methods in which case finite errors result from digit cancelation in the floating point representation and for high orders the errors usually become too large for the results to be of any practical use.

Lemma 6.1. *Let f be a computer function. Then f is defined at x_0 if and only if $f(x_0)$ can be computed on a computer.*

This lemma hinges on a careful implementation of the intrinsic functions and operations, in particular in the sense that they should be executable for any floating point number in the domain of definition that produces a result within the range of allowed floating point numbers.

Lemma 6.2. *Let f be a computer function, and let x_0 be such that $f(x_0-d), f(x_0)$, and $f(x_0+d)$ are all defined. Then f is continuous at x_0 if and only if*

$$f(x_0-d) =_0 f(x_0) =_0 f(x_0+d).$$

If $f(x_0)$ and $f(x_0+d)$ are defined, but $f(x_0-d)$ is not, then f is right-continuous at x_0 if and only if $f(x_0+d) =_0 f(x_0)$. Finally, if $f(x_0)$ and $f(x_0-d)$ are defined, but $f(x_0+d)$ is not, then f is left-continuous at x_0 if and only if $f(x_0-d) =_0 f(x_0)$.

Theorem 6.3. *Let f be a computer function that is continuous at x_0, and let $f(x_0-d)$ and $f(x_0+d)$ be both defined. Then f is differentiable at x_0 if and only if*

$$\frac{f(x_0+d)-f(x_0)}{d} \quad \text{and} \quad \frac{f(x_0)-f(x_0-d)}{d}$$

are both at most finite in absolute value, and their real parts agree. In this case,

$$\frac{f(x_0+d)-f(x_0)}{d} =_0 f'(x_0) =_0 \frac{f(x_0)-f(x_0-d)}{d}.$$

If f is differentiable at x_0, then f is twice differentiable at x_0 if and only if

$$\frac{f(x_0+2d)-2f(x_0+d)+f(x_0)}{d^2} \quad \text{and} \quad \frac{f(x_0)-2f(x_0-d)+f(x_0-2d)}{d^2}$$

are both at most finite in absolute value, and their real parts agree. In this case

$$\frac{f(x_0+2d)-2f(x_0+d)+f(x_0)}{d^2} =_0 f^{(2)}(x_0) =_0 \frac{f(x_0)-2f(x_0-d)+f(x_0-2d)}{d^2}.$$

In general, if f is $(n-1)$ times differentiable at x_0, then f is n times differentiable at x_0 if and only if

$$\frac{\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}f(x_0+jd)}{d^n} \quad \text{and} \quad \frac{\sum_{j=0}^{n}(-1)^{j}\binom{n}{j}f(x_0-jd)}{d^n}$$

are both at most finite in absolute value, and their real parts agree. In this case,

$$\frac{\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}f(x_0+jd)}{d^n}=_0 f^{(n)}(x_0)=_0\frac{\sum_{j=0}^{n}(-1)^{j}\binom{n}{j}f(x_0-jd)}{d^n}.$$

Since knowledge of $f(x_0-d)$ and $f(x_0+d)$ gives us all the information about a computer function f in a real positive radius σ around x_0, we have the following result which states that, from the mere knowledge of $f(x_0-d)$ and $f(x_0+d)$, we can find at once the order of differentiability of f at x_0 and the accurate values of all existing derivatives.

Theorem 6.4. *Let f be a computer function that is defined at x_0; and let $n\in\mathbb{N}$ be given. Then f is n times differentiable at x_0 if and only if $f(x_0-d)$ and $f(x_0+d)$ are both defined and can be written as*

$$f(x_0-d)=_n f(x_0)+\sum_{j=1}^{n}(-1)^j\alpha_j d^j \quad \text{and } f(x_0+d)=_n f(x_0)+\sum_{j=1}^{n}\alpha_j d^j,$$

where the α_j's are real numbers. Moreover, in this case $f^{(j)}(x_0)=j!\,\alpha_j$ for $1\leqslant j\leqslant n$.

Now consider, as an example, the function

$$g(x)=\frac{\sin\left(x^3+2x+1\right)+\dfrac{3+\cos\left(\sin\left(\ln|1+x|\right)\right)}{\exp\left(\tanh\left(\sinh\left(\cosh\left(\frac{\sin\left(\cos\left(\tan\left(\exp(x)\right)\right)\right)}{\cos\left(\sin\left(\exp\left(\tan\left(x+2\right)\right)\right)\right)}\right)\right)\right)\right)}}{2+\sin\left(\sinh\left(\cos\left(\tan^{-1}\left(\ln\left(\exp(x)+x^2+3\right)\right)\right)\right)\right)}. \tag{6.1}$$

Table 1
$g^{(n)}(0),\ 0\leqslant n\leqslant 10$, computed with $\mathcal{R}$ calculus.

Order n	$g^{(n)}(0)$	CPU Time
0	1.004845319007115	1.820 msec
1	0.4601438089634254	2.070 msec
2	−5.266097568233224	3.180 msec
3	−52.82163351991485	4.830 msec
4	−108.4682847837855	7.700 msec
5	16451.44286410806	11.640 msec
6	541334.9970224757	18.050 msec
7	7948641.189364974	26.590 msec
8	−144969388.2104904	37.860 msec
9	−15395959663.01733	52.470 msec
10	−618406836695.3634	72.330 msec

Table 2
$g^{(n)}(0),\ 0\leqslant n\leqslant 6$, computed with Mathematica.

Order n	$g^{(n)}(0)$	CPU Time
0	1.004845319007116	0.11 sec
1	0.4601438089634254	0.17 sec
2	−5.266097568233221	0.47 sec
3	−52.82163351991483	2.57 sec
4	−108.4682847837854	14.74 sec
5	16451.44286410805	77.50 sec
6	541334.9970224752	693.65 sec

Table 3
$g^{(n)}(0),\ 1\leqslant n\leqslant 10$, computed numerically.

Order n	$g^{(n)}(0)$	Relative Error
1	0.4601437841866840	54×10^{-9}
2	−5.266346392944456	47×10^{-6}
3	−52.83767867680922	30×10^{-5}
4	−87.27214664649106	0.20
5	19478.29555909866	0.18
6	633008.9156614641	0.17
7	−12378052.73279768	2.6
8	−1282816703.632099	7.8
9	83617811421.48561	6.4
10	91619495958355.24	149

Using the $\mathcal{R}$ calculus, we find $g^{(n)}(0)$ for $0 \leqslant n \leqslant 10$. These numbers are listed in Table 1; we note that, for $0 \leqslant n \leqslant 10$, we list the CPU time needed to obtain all derivatives of g at 0 up to order n and not just $g^{(n)}(0)$. For comparison purposes, we give in Table 2 the function value and the first six derivatives computed with Mathematica. Note that the respective values listed in Tables 1 and 2 agree. However, Mathematica used much more CPU time to compute the first six derivatives, and it failed to find the seventh derivative as it ran out of memory. We also list in Table 3 the first ten derivatives of g at 0 computed numerically using the numerical differentiation formulas

$$g^{(n)}(0) = (\Delta x)^{-n}\left(\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}g(j\Delta x)\right), \quad \Delta x = 10^{-16/(n+1)},$$

for $1 \leqslant n \leqslant 10$, together with the corresponding relative errors obtained by comparing the numerical values with the respective exact values computed using $\mathcal{R}$ calculus.

On the other hand, formula manipulators fail to find the derivatives of certain functions at given points even though the functions are differentiable at the respective points. For example, the functions

$$g_1(x) = |x|^{5/2} \cdot g(x) \quad \text{and } g_2(x) = \begin{cases} \frac{1-\exp(-x^2)}{x} \cdot g(x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases},$$

where $g(x)$ is the function given in Eq. (6.1), are both differentiable at 0; but the attempt to compute their derivatives using formula manipulators fails. This is not specific to g_1 and g_2, and is generally connected to the occurrence of non-differentiable parts that do not affect the differentiability of the end result, of which case g_1 is an example, as well as the occurrence of branch points in coding as in IF-ELSE structures, of which case g_2 is an example.

In the following section we give a brief summary of our work on developing a non-Archimedean operator theory on a Banach space over the complex Levi-Civita field $\mathcal{C}$ which is the subject of an ongoing collaboration with José Aguayo (Universidad de Concepción, Chile) and Miguel Nova (Universidad Católica de la Santísima, Concepción, Chile). For lack of space, we will omit all the details here and refer the interested reader to [1,2].

7. Non-Archimedean operator theory

Let c_0 denote the space of all null sequences of elements in $\mathcal{C}$. The natural inner product on c_0 induces the sup-norm of c_0. In [1], we show that c_0 is not orthomodular then we characterize those closed subspaces of c_0 with an orthonormal complement with respect to the inner product. Such a subspace, together with its orthonormal complement, defines a special kind of projection, the normal projection. We present characterizations of normal projections as well as other kinds of operators, the self-adjoint and compact operators on c_0. In [2], we work on some B*-algebras of operators, including those mentioned above; and we define an inner product on such algebras that induces the usual norm of operators. Finally, in a paper currently in preparation, we study the properties of positive operators on c_0, which we then use to introduce a partial order on the B*-algebra of compact and self-adjoint operators on c_0 and study the properties of that partial order.

References

[1] J. Aguayo, M. Nova, K. Shamseddine, Characterization of compact and self-adjoint operators on free Banach spaces of countable type over the complex Levi-Civita field, J. Math. Phys. 54 (2) (2013).
[2] J. Aguayo, M. Nova, K. Shamseddine, Inner product on B*-algebras of operators on a Free Banach space over the Levi-Civita field, submitted.
[3] N.L. Alling, Foundations of Analysis over Surreal Number Fields, North Holland, 1987.
[4] M. Berz, Analysis on a nonarchimedean extension of the real numbers, Lecture Notes, 1992 and 1995 Mathematics Summer Graduate Schools of the German National Merit Foundation, MSUCL-933, Department of Physics, Michigan State University, 1994.
[5] M. Berz, Calculus and numerics on Levi-Civita fields, in: M. Berz, C. Bischof, G. Corliss, A. Griewank (Eds.), Computational Differentiation: Techniques, Applications, and Tools, SIAM, Philadelphia, 1996, pp. 19–35.
[6] M. Berz, Nonarchimedean analysis and rigorous computation, Int. J. Appl. Math. 2 (2000) 889–930.
[7] M. Berz, Cauchy theory on Levi-Civita fields, Contemp. Math. 319 (2003) 39–52.
[8] M. Berz, C. Bischof, A. Griewank, G. Corliss (Eds.), Computational Differentiation: Techniques, Applications, and Tools, SIAM, Philadelphia, 1996.
[9] M. Berz, K. Makino, K. Shamseddine, G. Hoffstätter, W. Wan, COSY INFINITY and its applications to nonlinear dynamics, in: M. Berz, C. Bischof, G. Corliss, A. Griewank (Eds.), Computational Differentiation: Techniques, Applications, and Tools, SIAM, Philadelphia, 1996, pp. 363–367.
[10] Anthony V. Fiacco, Garth P. McCormick, Nonlinear Programming; Sequential Unconstrained Minimization Techniques, SIAM, Philadelphia, 1990.
[11] W. Krull, Allgemeine Bewertungstheorie, J. Reine Angew. Math. 167 (1932) 160–196.
[12] D. Laugwitz, Tullio Levi-Civita's work on nonarchimedean structures (with an appendix: properties of Levi-Civita fields), in: Atti Dei Convegni Lincei 8: Convegno Internazionale Celebrativo Del Centenario Della Nascita De Tullio Levi-Civita, Academia Nazionale dei Lincei, Roma, 1975.
[13] T. Levi-Civita, Sugli infiniti ed infinitesimi attuali quali elementi analitici, Atti Ist. Veneto di Sc., Lett. ed Art., 7a 4 (1765) (1892).
[14] T. Levi-Civita, Sui numeri transfiniti, Rend. Acc. Lincei, 5a 7 (91) (1898) 113.
[15] David G. Luenberger, Linear and Nonlinear Programming, second ed., Addison-Wesley, Reading, Massachusetts, 1984.
[16] K. Makino, M. Berz, COSY INFINITY version 9, Nucl. Instrum. Methods A558 (2005) 346–350.
[17] L. Neder, Modell einer Leibnizschen Differentialrechnung mit aktual unendlich kleinen Größen, Math. Annalen 118 (1941–1943) 718–732.
[18] S. Priess-Crampe, Angeordnete Strukturen: Gruppen, Körper, projektive Ebenen, Springer, Berlin, 1983.
[19] P. Ribenboim, Fields: algebraically closed and others, Manuscripta Math. 75 (1992) 115–150.
[20] W.H. Schikhof, Ultrametric Calculus: An Introduction to p-Adic Analysis, Cambridge University Press, 1985.
[21] Ya. D. Sergeyev, Numerical computations and mathematical modelling with infinite and infinitesimal numbers, J. Appl. Math. Comput. 29 (2009) 177–195.

[22] Ya.D. Sergeyev, Numerical point of view on calculus for functions assuming finite, infinite, and infinitesimal values over finite, infinite, and infinitesimal domains, Nonlinear Anal. Ser. A: Theory Methods Appl. 71 (12) (2009) 1688–1707.

[23] Ya. D. Sergeyev, Lagrange lecture: methodology of numerical computations with infinities and infinitesimals, Rendiconti del Seminario Matematico dell'Università e del Politecnico di Torino 68 (2) (2010) 95–113.

[24] Ya.D. Sergeyev, Higher order numerical differentiation on the infinity computer, Optim. Lett. 5 (4) (2011) 575–585.

[25] Ya. D. Sergeyev, Solving ordinary differential equations by working with infinitesimals numerically on the infinity computer, Appl. Math. Comput. 219 (22) (2013) 10668–10681.

[26] K. Shamseddine, New elements of analysis on the Levi-Civita field (Ph.D. thesis), Michigan State University, East Lansing, Michigan, USA, also Michigan State University report MSUCL-1147, 1999.

[27] K. Shamseddine, On the topological structure of the Levi-Civita field, J. Math. Anal. Appl. 368 (2010) 281–292.

[28] K. Shamseddine, Absolute and relative extrema, the mean value theorem and the inverse function theorem for analytic functions on a Levi-Civita field,, Contemp. Math. 551 (2011) 257–268.

[29] K. Shamseddine, Nontrivial order preserving automorphisms of non-Archimedean fields, Contemp. Math. 547 (2011) 217–225.

[30] K. Shamseddine, New results on integration on the Levi-Civita field, Indag. Math. (N.S.) 24 (1) (2013) 199–211.

[31] K. Shamseddine, One-variable and multi-variable calculus on a non-Archimedean field extension of the real numbers, p-Adic Numbers Ultrametric Anal. Appl. 5 (2) (2013) 160–175.

[32] K. Shamseddine, M. Berz, Exception handling in derivative computation with non-Archimedean calculus, in: M. Berz, C. Bischof, G. Corliss, A. Griewank (Eds.), Computational Differentiation: Techniques, Applications, and Tools, SIAM, Philadelphia, 1996, pp. 37–51.

[33] K. Shamseddine, M. Berz, Convergence on the Levi-Civita field and study of power series, in: Proc. Sixth International Conference on p-adic Functional Analysis, Marcel Dekker, New York, NY, 2000, pp. 283–299.

[34] K. Shamseddine, M. Berz, The differential algebraic structure of the Levi-Civita field and applications, Int. J. Appl. Math. 3 (2000) 449–465.

[35] K. Shamseddine, M. Berz, Measure theory and integration on the Levi-Civita field, Contemp. Math. 319 (2003) 369–387.

[36] K. Shamseddine, M. Berz, Analytical properties of power series on Levi-Civita fields, Ann. Math. Blaise Pascal 12 (2) (2005) 309–329.

[37] K. Shamseddine, M. Berz, Intermediate value theorem for analytic functions on a Levi-Civita field, Bull. Belg. Math. Soc. Simon Stevin 14 (2007) 1001–1015.

[38] K. Shamseddine, T. Rempel, T. Sierens, The implicit function theorem in a non-Archimedean setting, Indag. Math. (N.S.) 20 (4) (2009) 603–617.

[39] K. Shamseddine, T. Sierens, On locally uniformly differentiable functions on a complete non-Archimedean ordered field extension of the real numbers, ISRN Math. Anal. 2012 (2012) 20.

[40] K. Shamseddine, V. Zeidan, One-dimensional optimization on non-Archimedean fields, J. Nonlinear Convex Anal. 2 (2001) 351–361.

[41] K. Shamseddine, V. Zeidan, Constrained second order optimization on non-Archimedean fields, Indag. Math. (N.S.) 14 (2003) 81–101.