2. Calculus and Numerics on Levi-Civita Fields*

(Chapter of "Computational Differentiation: Techniques, Applications, and Tools", Martin Berz, Christian Bischof, George Corliss, and Andreas Griewank, eds., SIAM, 1996.)

Martin Berz†

Abstract

The formal process of the evaluation of derivatives using some of the various modern methods of computational differentiation can be recognized as an example for the application of conventional "approximate" numerical techniques on a non-archimedean extension of the real numbers. In many cases, the application of "infinitely small" numbers instead of "small but finite" numbers allows the use of the old numerical algorithm, but now with an error that in a rigorous way can be shown to become infinitely small (and hence irrelevant).

While intuitive ideas in this direction have accompanied analysis from the early days of Newton and Leibniz, the first rigorous work goes back to Levi-Civita, who introduced a number field that in the past few years turned out to be particularly suitable for numerical problems. While Levi-Civita's field appears to be of fundamental importance and simplicity, efforts to introduce advanced concepts of calculus on it are only very new. In this paper, we address several of the basic questions providing a foundation for such a calculus. After addressing questions of algebra and convergence, we study questions of differentiability, in particular with an eye to usefulness for practical work.

Keywords: Levi-Civita, non-standard analysis, non-Archimedean analysis, analysis with infinitesimals, differentials, infinitesimals, derivatives as differential quotients, computer functions, differential quotients, computation of derivatives, R.

1 Introduction

The real numbers owe their fundamental role in mathematics and the sciences to certain special properties. To begin, like all fields, they allow arithmetic calculation. Furthermore, they allow measurement; any result of even the finest measurement can be expressed as a real number. Additionally, they allow expression of geometric concepts, which (for example because of Pythagoras) requires the existence of roots–a property that at the same time is beneficial for algebra. Furthermore, they allow the introduction of certain transcendental functions such as exp, which are important in the sciences and arise from the concept of power series. In addition, they allow the formulation of an analysis involving differentiation and integration, a requirement for the expression of even simple laws of nature.

While the first two properties are readily satisfied by the rational numbers, the geometric requirements demand using at least the set of algebraic numbers. Transcendental functions, being the result of limiting processes, require Cauchy completeness, and it is easily shown that the real numbers are the smallest ordered field having this property. Because it is at such a basic level of our scientific language, hardly any thought is spent on the

*Work supported in part by the US Department of Energy, Grant Number DE-FG02-95ER40931, and the Alfred P. Sloan Foundation.

†Department of Physics and Astronomy, Michigan State University, East Lansing, MI. 48824, USA, (berz@pilot.msu.edu).

fundamental question of whether there may be other useful number systems having the required properties.

This question is perhaps even more intriguing in light of the observation that, while the real numbers R and their algebraic completion C as well as the vector space R^n have certainly proven extremely successful for the expression and rigorous mathematical formulation of many physical concepts, they have two shortcomings in interpreting intuitive scientific concepts. First, they do not permit a direct representation of improper functions such as those used for the description of point charges; of course, within the framework of distributions, these concepts can be accounted for in a rigorous fashion, but at the expense of the intuitive interpretation. Second, another intuitive concept of the fathers of analysis, and for that matter quite a number of modern scientists sacrificing rigor for intuition, the idea of derivatives as "differential quotients" as slopes of secants with infinitely small abscissa and ordinate differences cannot be formulated rigorously within the real numbers. Especially for our purposes of computational differentiation, the concept of "derivatives are differential quotients" would of course be a remedy to many a problem, since it would replace any attempted limiting process involving the unavoidable cancellation of digits by computer-friendly algebra in a new number system.

The problems mentioned in the preceding paragraphs might be solved if, in addition to the real numbers, there were also "infinitely small" and "infinitely large" numbers; that is if the number system were non-archimedean. Since any archimedean Cauchy complete field is isomorphic to R, it is indeed the absence of such numbers that makes the real numbers unique. However, since the "fine structure" of the continuum is not observable by means of science, archimedicity is not required by nature, and leaving it behind would possibly allow the treatment of the above two concepts. So it appears on the one hand legitimate and on the other hand intriguing to study such number systems, as long as the above mentioned essential properties of the real numbers are preserved.

There are simple ways to construct non-archimedean extensions of the real numbers (see for example the books of Rudin [Rudin1987a], Hewitt and Stromberg [Hewitt1969a], or Stromberg [Stromberg1981a], or at a deeper level the works of Fuchs [Fuchs1963a], Ebbinghaus et al.[Ebbinghaus1992a] or Lightstone and Robinson [Lightstone1975a]), but such extensions usually quickly fail to satisfy one or several of the above criteria of a "useful" field, often already regarding the universal existence of roots.

An important idea for the problem of the infinite came from Schmieden and Laugwitz [Schmieden1958a], which was then quickly applied to delta functions [Laugwitz1959a] [Laugwitz1961b] and Distributions [Laugwitz1961a]. Certain equivalence classes of sequences of real numbers become the new number set, and, perhaps most interesting, logical statements are considered proved if they hold for "most" of the elements of the sequences. This approach lends itself to the introduction of a general scheme that allows the transfer of many properties of the real numbers to the new structure. This method supplies an elegant tool that, in particular, permits the determination of derivatives as differential quotients.

Unfortunately, the resulting structure has two shortcomings. On the one hand, while very large, it is not a field; there are zero divisors, and the ring is also not totally ordered. On the other hand, the structure is already so large that individual numbers can never be represented by only a finite amount of information and are thus out of reach for computational problems. Robinson [Robinson1961a] recognized that the intuitive method can be generalized [Laugwitz1973a] by a nonconstructive process based on model theory to obtain a totally ordered field, and initiated the branch of non-standard analysis. Some of the standard works (pun certainly intended!) describing this field are from Robinson

[Robinson1974a], Stroyan and Luxemburg [Stroyan1976a], and Davis [Davis1977a]. In this discipline, the transfer of theorems about real numbers is extremely simple, although at the expense of a nonconstructive process invoking the axiom of choice, leading to an exceedingly large structure of numbers and theorems. The nonconstructiveness makes practical use difficult and leads to several oddities, for example, the fact that the sign of certain elements, although assured to be either positive or negative, cannot be decided.

Another approach to a theory of infinitely small numbers originated in game theory, of all places, and was pioneered by John Conway in his marvel "On Numbers and Games" [Conway1976a]. A humorous and totally non-standard (pun again intended) yet at the same time very insightful account of these numbers can also be found in Donald Knuth's mathematical novelette "Surreal Numbers: How Two Ex-Students Turned to Pure Mathematics and Found Total Happiness" [Knuth1974a]. (We wonder about the applicability of the method to social problems of a larger scale, and as editors of these proceedings gladly acknowledge another work of Knuth, namely his fabulous TeX typesetting system.) Other important accounts on surreal numbers are by Alling [Alling1987a] and Gonshor [Gonshor1986a].

In this paper, analysis on a different non-archimedean extension of the real numbers is discussed. The numbers $\mathcal{R}$ were first discovered by the brilliant young Levi-Civita [Levi-Civita1892a], [Levi-Civita1898a], who succeeded in showing that they form a totally ordered field that is Cauchy complete. He concluded by showing that any power series with real or complex coefficients converges for infinitely small arguments and used this to extend real differentiable functions to the field. His number system has subsequently been rediscovered independently by a handful of people, including the author, and the subject appeared in the work of Ostrowski [Ostrowski1935a], Neder [Neder1943a], and later in the work of Laugwitz [Laugwitz1968a]. Two modern and rather complete accounts of Levi-Civita's work can be found in the book by Lightstone and Robinson [Lightstone1975a], which ends with the proof of Cauchy completeness, and in Laugwitz's account on Levi-Civita's work [Laugwitz1975a], which also contains a summary of properties of Levi-Civita fields.

In this paper, we extend the previous work and formulate the basis of a workable analysis on the Levi-Civita field $\mathcal{R}$. More extensive treatments of the matter can be found in [Berz1994a] and [Berz1990c], and a summary of some of the important concepts is contained in [Berz1992b]. We begin with questions about the structure of the field and show that $\mathcal{R}$ admits nth roots of positive elements; more so, the field obtained by adjoining the imaginary unit is algebraically closed. We also introduce a new topology, complementing the order topology, that is useful for a novel study of power series, which can be shown to converge not only for infinitely small arguments, but even within the conventional radius of convergence. This fact allows for the direct use of a large class of functions, in particular all the functional dependencies that can be formulated on a von Neuman computer. A differential calculus on $\mathcal{R}$ is developed, culminating in the proof of the theorem that derivatives are differential quotients after all, and that this offers a pretty way of doing computational differentiation.

2 Algebraic Properties of $\mathcal{R}$

We begin the discussion by introducing a specific family of sets.

Definition 1. **(The Family of Left-Finite Sets)** *A subset M of the rational numbers Q is called left-finite iff for every number $r \in Q$ there are only finitely many elements of M that are smaller than r. The set of all left-finite subsets of Q will be denoted by $\mathcal{F}$.*

The next lemma gives some insight into the structure of left-finite sets:

LEMMA 2.1. *Let $M \in \mathcal{F}$. If $M \neq \emptyset$, the elements of M can be arranged in ascending order, and there exists a minimum of M. If M is infinite, the resulting strictly monotonic sequence is divergent.*

Proof: A finite totally ordered set can always be arranged in ascending order; hence, we may assume that M is infinite.

For $n \in N$, set $M_n = \{x \in M \mid x \leq n\}$. Then M_n is finite by the definition of left-finiteness and we have $M = \bigcup_n M_n$.Hence, we first arrange the finitely many elements of M_0 in ascending order, append the finitely many elements of M_1 not in M_0 in ascending order, and continue inductively.

If the resulting strictly monotonic sequence were bounded, there would also be a rational bound below which there would be infinitely many elements of M, contrary to the assumption that M be left-finite. Therefore, we conclude that the sequence is divergent. $\square$

LEMMA 2.2. *Let $M, N \in \mathcal{F}$. Then we have $X \subset M \Rightarrow X \in \mathcal{F}$, $M \cup N \in \mathcal{F}$,and $M \cap N \in \mathcal{F}$. We also have $M + N = \{x + y \mid x \in M, y \in N\} \in \mathcal{F}$,and for every $x \in M + N$, there are only finitely many pairs $(a, b) \in M \times N$ such that $x = a + b$.*

Proof: The first three statements follow directly from the definition. For the proof of the fourth statement, let x_M, x_N denote the smallest elements in M, N respectively; these exist by the preceding lemma. Let r in Q be given. Set

$$M^u = \{x \in M | x < r - x_N\}, \quad N^u = \{x \in N | x < r - x_M\}$$

$$M^o = M \setminus M^u, \quad N^o = N \setminus N^u.$$

Then we have $M + N = (M^u \cup M^o) + (N^u \cup N^o) = (M^u + N^u) \cup (M^o + N^u) \cup (M^u + N^o) \cup (M^o + N^o) = (M^u + N^u) \cup (M^o + N) \cup (M + N^o)$. By definition of M^o and N^o, $(M^o + N)$ and $(M + N^o)$ do not contain any elements smaller than r. Thus all elements of $M + N$ that are smaller than r must actually be contained in $M^u + N^u$. Since both M^u and N^u are finite because of the left-finiteness of M and N, $M^u + N^u$ is also finite. Thus there are only finitely many elements in $M + N$ that are smaller than r.

To show the last statement, let $x \in M + N$ be given. Set $r = x + 1$ and define M^u, N^u as in the preceding paragraph. Then we have $x \notin (M^o + N)$, $x \notin (M + N^o)$. Hence all pairs $(a, b) \in M \times N$ that satisfy $x = a + b$ lie in the finite set $M^u \times N^u$. $\square$

Having discussed the family of left-finite sets, we introduce two sets of functions from the rational numbers into R and C.

DEFINITION 2. **(The Sets $\mathcal{R}$ and $\mathcal{C}$)** *We define*

$$\mathcal{R} = \{f : Q \to R \mid \{x | f(x) \neq 0\} \in \mathcal{F}\} \text{ and } \mathcal{C} = \{f : Q \to C \mid \{x | f(x) \neq 0\} \in \mathcal{F}\}.$$

Hence, the elements of $\mathcal{R}$ and $\mathcal{C}$ are those real or complex-valued functions on Q that are nonzero only on a left-finite set, that is, they have left-finite support.

Obviously, we have $\mathcal{R} \subset \mathcal{C}$. In the following, we denote elements of $\mathcal{R}$ and $\mathcal{C}$ by x, y, etc. and identify their values at $q \in Q$ with brackets, like in $x[q]$. This avoids confusion when we later consider functions on $\mathcal{R}$ and $\mathcal{C}$. Since the elements of $\mathcal{R}$ and $\mathcal{C}$ are functions with left-finite support, it is convenient to use the properties of left-finite sets (2.1) for their description.

DEFINITION 3. **(Notation for Elements of $\mathcal{R}$ and $\mathcal{C}$)** *An element x of $\mathcal{R}$ or $\mathcal{C}$ is uniquely characterized by an ascending (finite or infinite) sequence (q_n) of support points* *and a corresponding sequence* $(x[q_n])$ *of function values. We refer to the pair of sequences* $((q_n),(x[q_n]))$ *as the table of* x.

Already at this point it is worth noting that for questions of implementation, it is usually sufficient to store only the first few of the support point and remember carefully up to what "depth" a given number in $\mathcal{R}$ is known.

For subsequent discussion, it is convenient to introduce the following terminology.

DEFINITION 4. **(supp, λ, $\sim$, $\approx$, $=_r$,∂)** *For* $x, y \in \mathcal{C}$, *we define*

$\mathrm{supp}(x) = \{q \in Q \mid x[q] \neq 0\}$ *and call it the support of* x.

$\lambda(x) = \min(\mathrm{supp}(x))$ *for* $x \neq 0$ *(which exists because of left-finiteness) and* $\lambda(0) = +\infty$.

Comparing two elements, we say

$x \sim y$ *iff* $\lambda(x) = \lambda(y)$;

$x \approx y$ *iff* $\lambda(x) = \lambda(y)$ *and* $x[\lambda(x)] = y[\lambda(y)]$;

$x =_r y$ *iff* $x[q] = y[q]$ *for all* $q \leq r$;

Furthermore, we define an operation $\partial : \mathcal{C} \to \mathcal{C}$ *via* $(\partial x)[q] = (q+1) \cdot x[q+1]$.

At this point, these definitions may feel somewhat arbitrary; but after having introduced the concept of ordering on $\mathcal{R}$, we will see that λ describes "orders of infinite largeness or smallness," the relation "$\approx$ " corresponds to agreement up to infinitely small relative error, while "$\sim$ " corresponds to agreement of order of magnitude. The operation "∂" will prove to be a derivation that, among other things, is useful for the concept of differentiation on $\mathcal{R}$.

LEMMA 2.3. *The relations* $\sim$, $\approx$ *and* $=_r$ *are equivalence relations. They satisfy*

$$x \approx y \Rightarrow x \sim y; \text{ and if } a, b \in Q,\ a > b, \text{ then } x =_a y \Rightarrow x =_b y.$$

Furthermore, we have $\lambda(\partial x) \leq \lambda(x)$; *and if* $\lambda(x) \neq 0, \infty$, *even* $\lambda(\partial x) = \lambda(x) - 1$.

We now define arithmetic on $\mathcal{R}$ and $\mathcal{C}$:

DEFINITION 5. **(Addition and Multiplication on $\mathcal{R}$ and $\mathcal{C}$)** *We define addition on* $\mathcal{R}$ *and* $\mathcal{C}$ *componentwise:*

$$(x+y)[q] = x[q] + y[q].$$

Multiplication is defined as follows. For $q \in Q$ *we set*

$$(x \cdot y)[q] = \sum_{\substack{q_x, q_y \in Q, \\ q_x + q_y = q}} x[q_x] \cdot y[q_y].$$

We remark that $\mathcal{R}$ and $\mathcal{C}$ are closed under addition since $\mathrm{supp}(x+y) \subseteq \mathrm{supp}(x) \cup \mathrm{supp}(y)$, so by Lemma (2.2), with x and y having left-finite support, so does $x+y$. Lemma (2.2) also shows that only finitely many terms contribute to the sum in the definition of the product.

Furthermore, the product defined above is itself an element of $\mathcal{R}$ or $\mathcal{C}$, respectively, since the sets of support points satisfy $\mathrm{supp}(x \cdot y) \subseteq \mathrm{supp}(x) + \mathrm{supp}(y)$; application of Lemma (2.2) shows that $\mathrm{supp}(x \cdot y) \in \mathcal{F}$.

It turns out that the operations $+$ and $\cdot$ we just defined on $\mathcal{R}$ and $\mathcal{C}$ make ($\mathcal{R}$,+,$\cdot$) and ($\mathcal{C}$,+,$\cdot$) into fields. We begin by showing the ring structure. THEOREM 2.1. *($\mathcal{R}$,+,$\cdot$) and ($\mathcal{C}$,+,$\cdot$) are commutative rings with units.*

As it turns out, $\mathcal{R}$ and $\mathcal{C}$ can be viewed as extensions of R and C, respectively.

THEOREM 2.2. **(Embeddings of R into $\mathcal{R}$ and C into $\mathcal{C}$)** *R and C can be embedded into $\mathcal{R}$ and $\mathcal{C}$ under preservation of their arithmetic structures.*

Proof: Let $x \in C$. Define Π by

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

Then $\Pi(x) \in \mathcal{C}$, and if $x \in R$, $\Pi(x)$ is contained in $\mathcal{R}$. Π is injective, and direct calculation shows that $\Pi(x+y) = \Pi(x) + \Pi(y)$ and $\Pi(x \cdot y) = \Pi(x) \cdot \Pi(y)$. So R and C are embedded as subfields in the rings $\mathcal{R}$ and $\mathcal{C}$ respectively. However, the embedding is not surjective, since only elements with support $\{0\}$ are actually reached. □

REMARK 2.1. *In the following, we identify an element $x \in C$ with its image $\Pi(x) \in \mathcal{C}$ under the embedding. We recall that the sum of a complex number and an element of $\mathcal{C}$ has to be distinguished from the componentwise addition of a constant to a function.*

Furthermore, we note that every element in $\mathcal{C}$ has a unique representation as $a + b \cdot i$, where i denotes the imaginary unit in C and where $a, b \in \mathcal{R}$.

We also make the following observation. REMARK 2.2. *Let z_1 and z_2 be complex numbers. Then if both z_1 and z_2 are nonzero, we have $z_1 \sim z_2$. Furthermore, $z_1 \approx z_2$ is equivalent to $z_1 = z_2$.*

The only nontrivial step toward the proof that $\mathcal{R}$ and $\mathcal{C}$ are fields is the existence of multiplicative inverses of nonzero elements. For this purpose, we prove a new theorem that will be of key importance for a variety of proofs and applications.

LEMMA 2.4. **(Fixed Point Theorem)** *Let $q_M \in Q$ be given. Define $M \subset \mathcal{R}$ ($M \subset \mathcal{C}$) to be the set of all elements x of $\mathcal{R}$ ($\mathcal{C}$) such that $\lambda(x) \geq q_M$. Let $f : M \to \mathcal{C}$ satisfy $f(M) \subset M$. Suppose there exists $k \in Q$, $k > 0$ such that for all $x_1, x_2 \in M$ and all $q \in Q$, we have*

$$x_1 =_q x_2 \;\Rightarrow\; f(x_1) =_{q+k} f(x_2).$$

Then there is a unique solution $x \in M$ of the fixed point equation

$$x = f(x).$$

REMARK 2.3. *Without further knowledge about $\mathcal{R}$ and $\mathcal{C}$, the requirements and meaning of the fixed point theorem are not very intuitive. However, as we will see later, the assumption about f means that f is a contracting function with an infinitely small contraction factor. Furthermore, the sequence (a_i) that is constructed in the proof is indeed a Cauchy sequence, which is assured convergence because of the Cauchy completeness of our fields with respect to the order topology, as discussed below. However, while making the situation more transparent, the properties of ordering and Cauchy completeness are not required to formulate and prove the fixed point theorem, and so we refrain from invoking them here.*

Proof: We choose an arbitrary $a_0 \in M$ and define recursively

$$a_i = f(a_{i-1}), \quad i = 1, 2, \ldots.$$

Since f maps M into itself, this generates a sequence of elements of M. First we note that

$$a_i[p] = a_{i-1}[p] \text{ for all } p < (i-1) \cdot k + q_M \quad (*).$$

Since $a_0, a_1 \in M$, we have $a_1[p] = 0 = a_0[p]$ for all $p < q_M$. So $(*)$ holds for $i = 1$, and induction shows that it holds for all $i \geq 1$.

Next we define a function $x : Q \to C$ in the following way: for $q \in Q$ choose $i \in N$ such that $(i-1) \cdot k + q_M > q$. Set $x[q] := a_i[q]$; note that, by virtue of (*), this is independent of the choice of i.

Furthermore, we have $x =_q a_i$. So in particular x is an element of $\mathcal{R}$ or $\mathcal{C}$, respectively, since for every $q \in Q$, the set of its support points smaller than q agrees with the set of support points smaller than q of one of the $a_i \in M$. Also, since $x[p] = 0$ for all $p < q_M$, x is contained in M.

Now we show that x defined as above is a solution of the fixed point equation. For $q \in Q$ choose again $i \in N$ such that $(i-1) \cdot k + q_M > q$. Then it follows that $x =_q a_i =_q a_{i+1}$. By the contraction property of f, we thus obtain $f(x) =_{q+k} f(a_i)$, which in turn implies

$$x[q] = a_{i+1}[q] = f(a_i)[q] = f(x)[q].$$

Since this holds for all $q \in Q$, x is a fixed point of f.

It remains to show that x is a unique fixed point. Assume that $y \in M$ is a fixed point of f. The contraction property of f is equivalent to $\lambda(f(x_1) - f(x_2)) \geq \lambda(x_1 - x_2) + k$ for all $x_1, x_2 \in M$. This implies

$$\lambda(x - y) = \lambda(f(x) - f(y)) \geq \lambda(x - y) + k,$$

which is possible only if $y = x$. $\square$

REMARK 2.4. *It is worthwhile to point out that, in spite of the iterative character of the fixed point theorem, for every $q \in Q$, the value of the fixed point x at q can be calculated in finitely many steps. This is of significant importance especially for practical purposes.*

Using the fixed point theorem, we can now easily show the existence of multiplicative inverses.

THEOREM 2.3. *$(\mathcal{R},+,\cdot)$ and $(\mathcal{C},+,\cdot)$ are fields.*

Proof: We prove the theorem for $\mathcal{R}$; the proof for $\mathcal{C}$ is completely analogous. It remains to show the existence of multiplicative inverses of nonzero elements.

Let $z \in \mathcal{R} \setminus \{0\}$ be given. Set $q = \lambda(z)$, $a = z[q]$ and $z^* = 1/a \cdot d^{-q} \cdot z$. Then $\lambda(z^*) = 0$ and $z^*[0] = 1$. If an inverse of z^* exists, then $1/a \cdot d^{-q}(z^*)^{-1}$ is an inverse of z; so without loss of generality, we may assume $\lambda(z) = 0$ and $z[0] = 1$.

If $z = 1$, there exists an inverse. Otherwise, z is of the form $z = 1 + y$ with $0 < k = \lambda(y) < +\infty$. It suffices to find $x \in \mathcal{R}$ such that $(1 + x) \cdot (1 + y) = 1$. This is equivalent to

$$x = -y \cdot x - y$$

Setting $f(x) = -y \cdot x - y$ reduces the problem to finding a fixed point of f. Let $M = \{x \in \mathcal{R} \mid \lambda(x) \geq k\}$, then $f(M) \subset M$. Let $x_1, x_2 \in M$ satisfying $x_1 =_q x_2$ be given. Since the smallest support point of y is k, we obtain $y \cdot x_1 =_{q+k} y \cdot x_2$, and hence

$$-y \cdot x_1 - y =_{q+k} -y \cdot x_2 - y.$$

Thus f satisfies the hypothesis of the fixed point theorem (2.4), and consequently a fixed point of f exists. $\square$

Now we examine the existence of roots in $\mathcal{R}$ and $\mathcal{C}$. Using the fixed point theorem, we find the new result that, regarding this important property, the new fields behave just like R and C, respectively. THEOREM 2.4. *Let $z \in \mathcal{R}$ be nonzero, and set $q = \lambda(z)$. If $n \in N$ is even and $z[q]$ is positive, z has two nth roots in $\mathcal{R}$. If n is even and $z[q]$ is negative, z has no nth roots in $\mathcal{R}$. If n is odd, z has a unique nth root in $\mathcal{R}$.*

Let $z \in \mathcal{C}$ be nonzero. Then z has n distinct nth roots in $\mathcal{C}$.

Proof: Let z be a nonzero number and write $z = a \cdot d^q \cdot (1 + y)$, where $a \in C$, $q \in Q$, and $\lambda(y) > 0$. Assume that w is an nth root of z. Since $q = \lambda(z) = \lambda(w^n) = n \cdot \lambda(w)$, we can write $w = b \cdot d^{q/n} \cdot (1 + x)$, where $b \in C$, $\lambda(x) > 0$. Raising to the nth power, we see that $b^n = a$ and $(1 + x)^n = 1 + y$ have to hold simultaneously. The first of these equations has a solution if and only if the corresponding roots exist in R or C. So it suffices to show that the equation

$$(1 + x)^n = 1 + y$$

has a unique solution with $\lambda(x) > 0$. But this equation is equivalent to $nx + x^2 \cdot P(x) = y$, where $P(x)$ is a polynomial with integer coefficients. Because $\lambda(x) > 0$, also $\lambda(P(x)) \geq 0$, and hence $\lambda(x^2 \cdot P(x)) = 2\lambda(x) + \lambda(P(x)) > \lambda(x) > 0$; so finally we have $\lambda(x) = \lambda(y)$ for all such x. The equation can be rewritten as a fixed point problem $x = f(x)$, where

$$f(x) = \frac{y}{n} - x^2 \cdot \frac{P(x)}{n}.$$

Now let M be the set of all numbers in $\mathcal{C}$ (or in $\mathcal{R}$ if $z \in \mathcal{R}$) whose smallest support point does not lie below $k_y = \lambda(y)$. Then as we just concluded, any solution of the fixed point equation must lie in M. We further have $f(M) \subset M$; for if $x \in M$, then $\lambda(x^2 \cdot P(x)) \geq 2 \cdot k_y > k_y$. Hence it follows that $f(x) = y/n - x^2 \cdot P(x)/n$ has k_y as smallest support point, and thus $f(x) \in M$.

Let $x_1, x_2 \in M$ satisfying $x_1 =_q x_2$ be given. Then $\lambda(x_1) \geq k_y$, $\lambda(x_2) \geq k_y$, and the definition of multiplication shows that we obtain $x_1^2 =_{q+k_y} x_2^2$. By induction on m, we obtain $x_1^m =_{q+k_y} x_2^m$ for all $m \in N$, $m > 1$.

In particular, this implies $x_1^2 \cdot P(x_1) =_{q+k_y} x_2^2 \cdot P(x_2)$ and finally $f(x_1) =_{q+k_y} f(x_2)$. So f and M satisfy the hypothesis of the fixed point theorem (2.4) which provides a unique solution of $(1 + x)^n = 1 + y$ in M and hence in $\mathcal{R}$. □

We remark that a crucial point to the proof was the existence of roots of the numbers d^q; hence we could not have chosen anything smaller than Q as the domain of the functions that are the elements of our new fields.

We end the section on the algebraic properties of $\mathcal{R}$ and $\mathcal{C}$ by remarking that $\mathcal{C}$ is algebraically closed. Although a rather deep result, it is obtained with limited effort using the fixed point theorem as well as the algebraic completeness of C.

THEOREM 2.5. **(Fundamental Theorem of Algebra for $\mathcal{C}$)** *Every polynomial of positive degree with coefficients in $\mathcal{C}$ has a root in $\mathcal{C}$.*

The proof is omitted for reasons of space; for details, see [Berz1994a]

3 Order Structure

In the previous section we showed that $\mathcal{R}$ and $\mathcal{C}$ do not differ significantly from R and C, respectively, as far as their algebraic properties are concerned. In this section we discuss the ordering.

The simplest way of introducing an order is to define a set of "positive" numbers.

DEFINITION 6. **(The Set $\mathcal{R}^+$)** *Let $\mathcal{R}^+$ be the set of all nonvanishing elements x of $\mathcal{R}$ that satisfy $x[\lambda(x)] > 0$.*

LEMMA 3.1. **(Properties of $\mathcal{R}^+$)** *The set $\mathcal{R}^+$ has the following properties:*
$\mathcal{R}^+ \cap (-\mathcal{R}^+) = \emptyset$, $\mathcal{R}^+ \cap \{0\} = \emptyset$, *and* $\mathcal{R}^+ \cup \{0\} \cup (-\mathcal{R}^+) = \mathcal{R}$
$x, y \in \mathcal{R}^+ \;\Rightarrow\; x + y \in \mathcal{R}^+$ *and* $x, y \in \mathcal{R}^+ \;\Rightarrow\; x \cdot y \in \mathcal{R}^+$

The proofs follow rather directly from the respective definitions.

Having defined $\mathcal{R}^+$, we can now easily introduce an order in $\mathcal{R}$.

DEFINITION 7. **(Ordering in $\mathcal{R}$)**. *Let x, y be elements of $\mathcal{R}$. We say $x > y$ iff $x - y \in \mathcal{R}^+$. Furthermore, we say $x < y$ iff $y > x$.*

With this definition of the order relation, $\mathcal{R}$ is a totally ordered field.

THEOREM 3.1. **(Properties of the Order)**. *With the order relation defined in (7), $(\mathcal{R},+,\cdot)$ becomes a totally ordered field.*
Furthermore, the order is compatible with the algebraic structure of $\mathcal{R}$, that is

For any x, y, z, we have: $x > y \Rightarrow x+z > y+z$; and if $z > 0$, we have $x > y \Rightarrow x\cdot z > y\cdot z$.

Since the proof follows the same arguments as the corresponding one for R, the details are omitted here. We immediately obtain that the embedding Π is compatible with the ordering, that is $x < y \Rightarrow \Pi(x) < \Pi(y)$. Furthermore $\mathcal{C}$, like C, cannot be ordered.

Thus $\mathcal{R}$, like C, is a proper field extension of R. Note that this is not a contradiction of the well-known uniqueness of C as a field extension of R. The respective theorem of Frobenius asserts only the nonexistence of any (commutative) field on $R^n$for $n > 2$. However, regarded as an R-vector space, $\mathcal{R}$ is infinite dimensional.

Besides the usual order relations, some other notations are convenient.

DEFINITION 8. **(The Number** d**)** *We define the number $d \in \mathcal{R}$ as follows:*

$$d[q] = \begin{cases} 1 & x = 1 \\ 0 & else \end{cases} .$$

Apparently, the number d admits roots, and we have $d^{m/n}[q] = 1$ for $x = m/n$, zero otherwise. As we shall see, it plays the role of an infinitesimal and thus satisfies what Rall suspected about the number $(0, 1)$ in his arithmetic of differentiation [Rall1986a].

DEFINITION 9. **($\ll, \gg$)** *Let a, b be positive. We say a is infinitely smaller than b (and write $a \ll b$), iff $n \cdot a < b$ for all natural n; we say a is infinitely larger than b (and write $a \gg b$) iff $b \ll a$. If $a \ll 1$, we say a is infinitely small; if $1 \ll a$, we say a is infinitely large. Infinitely small numbers are also called infinitesimals or differentials. Infinitely large numbers are also called infinite. Numbers that are neither infinitely small nor infinitely large are also called finite.*

COROLLARY 3.1. *For all $a, b, c \in \mathcal{R}^+$, we have*
$a \ll b \ \Rightarrow a < b$ and $a \ll b,\ b \ll c \ \Rightarrow a \ll c$.

We observe $d^q \ll 1$ iff $q > 0$, $d^q \gg 1$ iff $q < 0$.

COROLLARY 3.2. *The field $\mathcal{R}$ is non-archimedean, that is, there are elements that are not exceeded by any natural number.*

Proof: For example, we have $n < d^{-1}\ \forall\, n \in N$. □

It is a crucial property of the field $\mathcal{R}$ that the differentials, especially the formerly defined number d, satisfy Leibniz's intuitive idea of derivatives as differential quotients. This will be discussed in great detail below; but already here we want to give a simple example.

EXAMPLE 1. **(Calculation of Derivatives with Differentials)** *Let us consider the function $f(x) = x^2 - 2x$. Obviously, f is differentiable on R, and we have $f'(x) = 2x - 2$. As we know, we can obtain certain approximations to the derivative at the position x by calculating the difference quotient*

$$\frac{f(x + \Delta x) - f(x)}{\Delta x}$$

at the position x. Roughly speaking, the accuracy increases if Δx gets smaller. In our enlarged field $\mathcal{R}$, infinitely small quantities are available, and thus it is natural to calculate *the difference quotient for such infinitely small numbers. For example, let* $\Delta x = d$*; we obtain*

$$\frac{f(x+d)-f(x)}{d} = \frac{(x^2+2xd+d^2-2x-2d)-(x^2-2x)}{d} = 2x-2+d.$$

We realize that the difference quotient differs from the exact value of the derivative by only an infinitely small error. If all we are interested in is the usual real derivative of the real function $f : R \to R$*, this is given exactly by the "real part" of the difference quotient.*

This phenomenon will be studied in a fully rigorous way below.

4 Topology, Convergence, and Cauchy-Completeness

In this section we examine the topological structures of $\mathcal{R}$ and the related sets. We will see that on $\mathcal{R}$, in contrast to R, several different nontrivial topologies can be defined, all of which have certain advantages.

We begin with the introduction of an absolute value; this is done as in any totally ordered field.

DEFINITION 10. **(Absolute Value on $\mathcal{R}$)** *Let* $x \in \mathcal{R}$*. We define the absolute value of* x *as follows:*

If $x \geq 0$*, we set* $|x| = x$*; and if* $x < 0$*, we set* $|x| = -x$*.*

LEMMA 4.1. **(Properties of the Absolute Value)** *The mapping "*$|\ |$*"* $: \mathcal{R} \to \mathcal{R}$ *has the following properties:*

$|x| = 0$ *iff* $x = 0$; $|x \cdot y| = |x| \cdot |y|$ *; and* $|x+y| \leq |x| + |y|$.

The proof follows the same lines as the counterpart in R.

DEFINITION 11. **(Absolute Value on $\mathcal{C}$ and $\mathcal{R}^n$)** *On* $\mathcal{C}$ *and* $\mathcal{R}^n$*, we define absolute values as follows: Any element* $z \in \mathcal{C}$ *can be written* $z = a + bi$ *with* $a, b \in \mathcal{R}$*, and this representation is unique. We then define* $|a+bi| = \sqrt{a^2+b^2}$*.*

Furthermore, for any $(x_1, ..., x_n) \in \mathcal{R}^n$*, we define* $|(x_1, ..., x_n)| = \sqrt{x_1^2 + ... + x_n^2}$

The roots exist according to theorem (2.4).

Just as in any totally ordered set, we can now introduce the so-called order topology:

DEFINITION 12. **(Order Topology)** *We call a subset* M *of* $\mathcal{R}$*,* $\mathcal{C}$ *or* $\mathcal{R}^n$ *open iff for any* $x_0 \in M$ *there exists an* $\epsilon > 0$ *;* $\epsilon \in \mathcal{R}$ *such that* $O(x_0, \epsilon)$*, the set of points* x *with* $|x - x_0| < \epsilon$*, is a subset of* M*.*

Thus all ϵ-balls form a basis of the topology. We obtain the following theorem.

THEOREM 4.1. **(Properties of the Order Topology)** *With the above topology,* $\mathcal{R}$*,* $\mathcal{C}$ *and* $\mathcal{R}^n$ *become nonconnected topological spaces. They are Hausdorff. There are no countable bases. The topology induced to* R *is the discrete topology. The topology is not locally compact.*

The proof is omitted for reasons of space. We remark that a detailed study of the properties reveals that they hold in an identical way on any other non-archimedean structure, and thus the above unusual properties are not specific to R.

Besides the absolute value, it is useful to introduce a semi-norm that is not based on the order. For this purpose, we regard $\mathcal{C}$ as a space of functions as in the beginning, and define the semi-norm as a mapping from $\mathcal{C}$ into R.

DEFINITION 13. **(Semi-Norm on $\mathcal{C}$)** *We introduce the semi-norm "*$||\ ||_r$*" as a function from* $\mathcal{C}$ *into* R *as follows:*

$$||x||_r = \sup_{q \leq r}\{|x[q]|\}.$$

Note that the supremum is finite and even a maximum, since for any r, only finitely many of the $x[q]$ under consideration do not vanish; thus the semi-norm has a certain similarity to the supremum norm for continuous functions. Its properties also are quite similar.

LEMMA 4.2. **(Properties of the Semi Norm)** *For arbitrary r, the map "$|| \;\; ||_r$" : $\mathcal{R} \to R$ satisfies the following:*

$||x||_r \geq 0$; $||a \cdot x||_r = |a| \cdot ||x||_r$ *for all* $a \in R$; *and* $||x+y||_r \leq ||x||_r + ||y||_r$.

Thus $|| \;\; ||_r$ is a semi-norm in the usual sense. Using the family of these semi-norms, we can now define another topology.

DEFINITION 14. **(Semi-Norm Topology)** *We call a subset M of $\mathcal{R}$, $\mathcal{C}$, or $\mathcal{R}^n$ open with respect to the semi-norm topology iff for any $x_0 \in M$ there is a real $\epsilon > 0$ such that $S(x_0, \epsilon) = \{x | \; ||x - x_0||_{1/\epsilon} < \epsilon\} \subset M$.*

We will see that the semi-norm topology is the most useful topology for considering convergence in general. Moreover, it is of great importance for the implementation of the calculus on $\mathcal{R}$ and $\mathcal{C}$ on computers.

THEOREM 4.2. **(Properties of the Semi-Norm Topology)** *With the above definition of the semi norm topology, $\mathcal{R}$, $\mathcal{C}$, and $\mathcal{R}^n$ are topological spaces. They are Hausdorff with countable bases. The topology induced on R by the semi norm topology is the usual order topology on R.*

Proof omitted.

We now discuss convergence with respect to the topologies just introduced. We begin by studying a special property of sequences.

DEFINITION 15. **(Regularity of a Sequence)** *A sequence (a_i) in $\mathcal{C}$ is called regular iff the union of the supports of all members of the sequence is a left-finite set, that is iff $\cup_{i=0}^{\infty} \mathrm{supp}(a_i) \in \mathcal{F}$.*

This property is not automatically assured, as becomes apparent from considering the sequence (d^{-i}). As the next theorem shows, the property of regularity is compatible with the common operations of sequences:

LEMMA 4.3. **(Properties of Regularity)** *Let (a_i), (b_i) be regular sequences. Then the sequence of the sums, the sequence of the products, any rearrangement, as well as any subsequence of one of the sequences, and the merged sequence $c_{2i} = a_i$, $c_{2i+1} = b_i$ are regular.*

Proof: Let $A = \cup_{i=0}^{\infty} \mathrm{supp}(a_i)$, $B = \cup_{i=0}^{\infty} \mathrm{supp}(b_i)$ be the unions of the support points of all members of the sequences. According to the requirements, we have $A \in \mathcal{F}$ and $B \in \mathcal{F}$.

Every support point of the sequence of the sums is a support point of either one of the a_i or one of the b_i and is thus contained in $(A \cup B) \in \mathcal{F}$. Every support point of the sequence of the products is contained in $(A + B) \in \mathcal{F}$.

The support points of any subsequence of (a_i) are contained in A , and the support points of the joined sequence (c_i) are contained in $A \cup B$. □

DEFINITION 16. **(Strong Convergence)** *We call the sequence (a_i) in $\mathcal{R}$ or $\mathcal{C}$ strongly convergent to the limit $a \in \mathcal{R}$ or $\mathcal{C}$ respectively iff it converges to a with respect to the order topology, that is, iff for every $\epsilon > 0$, $\epsilon \in \mathcal{R}$ there exists $n \in N$ such that $|a_i - a| < \epsilon \; \forall \; i > n$.*

Using the idea of strong convergence allows a simple representation of the elements of $\mathcal{R}$ and $\mathcal{C}$ that is indeed strongly reminiscent of the familiar expansion of real numbers in powers of ten, and that enjoys a similar usefulness for practical calculations.

THEOREM 4.3. **(Expansion in Powers of Differentials)** *Let $((q_i), (x[q_i]))$ be the table of $x \in \mathcal{R}$ or $\mathcal{C}$ (cf. 3). Then the sequence $x_n = \sum_{i=1}^{n} x[q_i] \cdot d^{q_i}$ converges strongly to* *the limit x. Hence we can write*

$$x = \sum_{i=1}^{\infty} x[q_i] \cdot d^{q_i}.$$

Proof: Without loss of generality, let the set $\{q_i\}$ be infinite. Let $\epsilon > 0$ in $\mathcal{R}$ be given. Choose $n \in N$ such that $d^n < \epsilon$. Since q_i diverges strictly according to Lemma 2.1, there is $m \in N$ such that $q_\nu > n \; \forall \; \nu > m$. Hence we have $(x_\nu - x)[i] = 0$ for all $i \leq n$ and for all $\nu > m$. Thus $|x_\nu - x| < \epsilon$ for all $\nu > m$. Therefore, (x_n) converges strongly to x. □

A convenient criterion describes the sequences and series that converge strongly.

THEOREM 4.4. **(Convergence Criterion for Strong Convergence)** *Let (a_i) be a sequence in $\mathcal{R}$ or $\mathcal{C}$. Then (a_i) converges strongly iff for all $r \in Q$ there exists $n \in N$ such that $a_{i_1} =_r a_{i_2}$ for all $i_1, i_2 > n$. The series $\sum_{i=0}^{\infty} a_i$ converges strongly iff the sequence (a_i) is a null sequence.*

The proof is straightforward.

LEMMA 4.4. *Every strongly convergent sequence is regular.*

Proof: Let $r \in Q$ be given. Use the convergence criterion (4.4) to choose $n \in N$ such that the values of the members of the sequence do not change any more below r. Then we have that all the elements of $\cup_{i=0}^{\infty}\text{supp}(a_i)$ smaller than r do already occur in $\cup_{i=0}^{n}\text{supp}(a_i)$. This finite union, however, is contained in $\mathcal{F}$; and thus there are only finitely many elements of $\cup_{i=0}^{\infty}\text{supp}(a_i)$ below r. □

We will now prove that $\mathcal{R}$ and $\mathcal{C}$ are complete with respect to strong convergence.

THEOREM 4.5. **(Cauchy Completeness of $\mathcal{R}$ and $\mathcal{C}$)** *(a_n) is a Cauchy sequence in $\mathcal{R}$ or $\mathcal{C}$ (for any positive $\epsilon \in \mathcal{R}$ exists $n \in N$ such that $|a_{n_1} - a_{n_2}| \leq \epsilon$ for all $n_1, n_2 \geq n$), if and only if (a_n) converges strongly (there is $a \in \mathcal{R}$ or $\mathcal{C}$, respectively, such that for any positive $\epsilon \in \mathcal{R} \; \exists \; n \in N : \; |a - a_\nu| < \epsilon \; \forall \; \nu > n$).*

Proof: Let (a_n) be a Cauchy sequence in $\mathcal{R}$. Write $b_n = a_{n+1} - a_n$. Then (b_n) is a null sequence. Since we have $a_n = a_0 + \sum_{i=0}^{n-1} b_i$, (a_n) converges strongly according to the convergence criterion (4.4) for series. The other direction is proved analogously as in R.

As we see, the concept of strong convergence provides very nice properties, and moreover strong convergence can be checked easily by virtue of the convergence criterion. However, for some applications it is not sufficient, and it is advantageous to study a new kind of convergence.

DEFINITION 17. **(Weak Convergence)** *We call the sequence (a_i) weakly convergent if there is an $a \in \mathcal{C}$ such that (a_i) converges to a with respect to the semi-norm topology, that is, for any $\epsilon > 0$; $\epsilon \in R$ there exists $n \in N$ such that $||a_i - a||_{1/\epsilon} < \epsilon \; \forall \; i > n$. In this case, we call a the weak limit of (a_i).*

THEOREM 4.6. **(Convergence Criterion for Weak Convergence)** *Let the sequence (a_i) converge weakly to the limit a. Then the sequence $(a_i[q])$ converges pointwise to $a[q]$, and the convergence is uniform on every subset of Q bounded above.*

On the other hand, let (a_i) be regular, and let the sequence $(a_i[q])$ converge pointwise to $a[q]$. Then (a_i) converges weakly to a.

Proof: Let (a_i) converge weakly to a. Let $r \in Q$ and $\epsilon > 0$; $\epsilon \in R$ be given. Choose $\epsilon_1 < \min(\epsilon, 1/(1 + |r|))$ such that, for all rational $q \leq r$, we have $q < 1/\epsilon_1$. Choose $n \in N$ such that $|(a_i - a)[q]| < \epsilon_1 \; \forall \; i > n, q < 1/\epsilon_1$. Then we obtain $|(a_i - a)[q]| < \epsilon \; \forall \; q < r$ and $\forall i > n$, and uniform convergence is proved.

Let on the other hand the sequence be regular and pointwise convergent. Since every support point of the limit function agrees at least with one support point of one member of the sequence, and therefore is contained in $A = \cup_i\text{supp}(a_i) \in \mathcal{F}$, the limit function a

is an element of $\mathcal{C}$. Let now $\epsilon > 0$; $\epsilon \in R$ be given. Let $r > 1/\epsilon$. We show first that the sequence of functions (a_i) converges uniformly on $\{q \in Q | q \leq r\}$: Any point at which the limit function a can differ from any a_i has to be in A. Thus there are only finitely many points to be studied below r. For any such q, find N_q such that $|a_i[q] - a[q]| < \epsilon$ for all $i > N_q$, and let $N = \max(N_q)$. Then we have $|a_i[q] - a[q]| < \epsilon$ for all $i > N$ and for all $q \leq r$. In particular, we obtain $||a_i - a||_{1/\epsilon} < \epsilon$ for all $i > N$. □

Whereas $\mathcal{R}$ is complete with respect to strong convergence, it is not with respect to weak convergence, as we see in the following example.

Example. **(Weak Convergence and Completeness)** Let $a_n = \sum_{i=1}^{n} d^{-i}/i$. Then the sequence (a_n) is Cauchy with respect to weak convergence (that is, the semi-norm topology) and locally converges to the function that assumes the value $1/n$ at -$n \in Z^-$ and vanishes elsewhere. But this limit function is not an element of $\mathcal{C}$.

Example. **(Unbounded Null Sequence)** Let $a_n = d^{-n}/n$. Then (a_n) is obviously unbounded, but converges weakly to zero.

The relationship between strong convergence and weak convergence is provided by the following theorem, which follows rather directly from the convergence criterion:

THEOREM 4.7. *Strong convergence implies weak convergence to the same limit.*

The proof is straightforward.

THEOREM 4.8. **(Uniqueness of $\mathcal{R}$)** *The field $\mathcal{R}$ is the smallest totally ordered non archimedean field extension of R that is complete with respect to the order topology, in which every positive number has an nth root, and in which there is a positive infinitely small element a such that (a^n) is a null sequence with respect to the order topology.*

We now discuss a very important class of sequences, namely, the power series. Once their convergence properties are established, they will allow the extension of many important real functions, and they will also provide the key for an exhaustive study of differentiability of all functions that can be represented on a computer [Shamseddine1996a]. We begin our discussion of power series with an observation.

LEMMA 4.5. *Let $M \in \mathcal{F}$, that is, a left finite set. For M define*

$$M_\Sigma = \{x_1 + ... + x_n | n \in N \ \text{ and } x_1, ..., x_n \in M\};$$

then M_Σ is left finite if and only if $\min(M) \geq 0$.

Proof: First let $\min(M) = g < 0$. Clearly, all multiples of g are in M_Σ. In other words, M_Σ contains infinitely many elements smaller than zero and is therefore not left finite.

On the other hand. let $\min(M) \geq 0$. For $\min(M) = 0$, we start the discussion by considering $\bar{M} = M \setminus \{0\}$, which has a minimum greater than zero. But since M differs from $\bar{M}$ only by containing zero, and since inclusion of zero does not change a sum, we obviously have $\bar{M}_\Sigma = M_\Sigma$. It therefore suffices to consider sets with a positive minimum. Now let ow $r \in Q$; we show that there are only finitely many elements in M_Σ that are smaller than r. Since all elements in M_Σ are greater than or equal to the minimum g, the property holds for $r < g$. Now let $r \geq g$, and let $n = [r/g]$ be the greatest integer less than or equal to r/g. Let $x < r$ in M_Σ. Then at most n terms can sum up to x, since any sum with more than n terms exceeds r and thus x. Furthermore, the sum can contain only finitely many different elements of M, namely those below r. But this means that there are only finitely many ways of forming sums, and thus only finitely many results of summations below r. □

COROLLARY 4.1. *A sequence $x_i = x^i$ is regular iff x is at most finite.*

A sequence $x_i = a_i \cdot x^i$ or $x_i = \sum_{j=0}^{i} a_j \cdot x^j$ is regular if x is at most finite and a_i is regular.

Proof: First observe that the set $\cup_{i=1}^{\infty}\text{supp}(x^i)$ is identical with the set M_Σ in the previous lemma if we set $M = \text{supp}(x)$. This it is left finite iff $\text{supp}(x)$ has a minimum greater than or equal to zero; this is the case iff x is at most finite.

To prove the second part, we employ Corollary 4.3, which asserts that the product of regular sequences is regular. ◻

THEOREM 4.9. **(Power Series with Purely Complex Coefficients)** *Let $\sum_{n=0}^{\infty} a_n z^n$, $a_n \in C$ be a complex power series with radius of convergence η. Let $z \in \mathcal{C}$, and let $A_n(z) = \sum_{i=0}^{n} a_i z^i \in \mathcal{C}$. Then, for $|z| < \eta$ and $|z| \not\approx \eta$, the sequence is weakly convergent, and for any $q \in Q$, the sequence $A_n(z)[q]$ converges absolutely. We define the limit to be the* **continuation** *of the power series on $\mathcal{C}$.*

Proof: First note that the sequence is regular for any at most finite z, which follows from Corollary 4.1. Since the sequence a_i has only purely complex terms and is therefore regular.

Now we have to show that the sequence $A_n(z)$ converges for any fixed z with $|z| < \eta$ and $|z| \not\approx \eta$. Write z as a sum of a purely complex X and an at most infinitely small x. For $x = 0$, we are done. Otherwise, let $r \in Q$ be given. Choose $m \in N$ with $m \cdot \lambda(x) > r$. Then $(X + x)^n$ evaluated at r yields

$$((X + x)^n)[r] = (\sum_{j=0}^{n} x^j \cdot \frac{n!}{(n-j)!j!} \cdot X^{n-j})[r] = \sum_{j=0}^{min(m,n)} x^j[r] \cdot \frac{n!}{(n-j)!j!} \cdot X^{n-j}.$$

For the last equality, we use that x^j vanishes at r for $j > m$. We obtain the following chain of inequalities for any $\nu_2 > \nu_1 > m$:

$$\begin{aligned}
\sum_{n=\nu_1}^{\nu_2} |a_n (X+x)^n [r]| &= \sum_{n=\nu_1}^{\nu_2} |a_n| \cdot |\sum_{j=0}^{\min(m,n)} x^j[r] \cdot \frac{n!}{(n-j)!j!} \cdot X^{n-j}| \\
&\leq \sum_{n=\nu_1}^{\nu_2} \sum_{j=0}^{m} |a_n||x^j[r]| \frac{n!}{(n-j)!j!} |X|^{n-j} \\
&\leq \left(\sum_{j=0}^{m} \frac{|x^j[r]||X|^{m-j}}{j!} \right) \cdot \left(\sum_{n=\nu_1}^{\nu_2} |a_n| \cdot n^m \cdot |X|^{n-m} \right). \qquad (1)
\end{aligned}$$

Note that the righthand sum contains only real terms. Since $|X|$ is within the radius of convergence, the series converges; the additional factor n^m does not influence this since $\lim_{n\to\infty} \sqrt[n]{n^m} = 1$. Since the lefthand term does not depend on ν, we obtain absolute convergence at r. ◻

A prominent result of the Cauchy theory of analytic functions is that an analytic function is completely determined by the values it takes on a closed path. Our theory guarantees the uniqueness of a function even from the knowledge of only its value at one suitable point, as the following theorem shows.

THEOREM 4.10. **(Point Formula à la Cauchy)** *Let $f(z) = \sum_{i=0}^{\infty} a_i (z - z_0)^i$ be the continuation of a complex power series on $\mathcal{C}$. Then the function is completely determined by its value at $z_0 + h$, where h is an arbitrary nonzero infinitely small number.*

Proof: Evaluating the power series yields

$$f(z_0 + h) = \sum_{i=0}^{\infty} a_i h^i.$$

Let $r = \lambda(h)$, $h_0 = h[\lambda(h)]$. Then we obtain

$a_0 = (f(z_0 + h))[0]$
$a_1 = (f(z_0 + h))[r]/h_0$,
$a_2 = (f(z_0 + h) - a_1 h)[2r]/h_0^2$,
$a_3 = (f(z_0 + h) - a_1 h - a_2 h^2)[3r]/h_0^3$,
....

Choosing $h = d$, we obtain the even simpler result $a_i = (f(z_0 + d))[i]$. □

5 Continuity and Differentiability

In this section we introduce the concepts of continuity and differentiability on $\mathcal{R}$ and $\mathcal{C}$; we do so, as in R, via the $\epsilon - \delta$- method. Unlike in R, however, ϵ and δ may be of a completely different order of magnitude.

DEFINITION 18. **(Continuity and Equicontinuity)** *The function $f : D \subset \mathcal{R} \to \mathcal{R}$ is called continuous at the point $x_0 \in D$, if for any positive $\epsilon \in \mathcal{R}$ there is a positive $\delta \in \mathcal{R}$ such that*

$$|f(x) - f(x_0)| < \epsilon \text{ for any } x \in D \text{ with } |x - x_0| < \delta.$$

The function is called equicontinuous at the point x_0 if for any ϵ it is possible to choose the δ in such a way that $\delta \sim \epsilon$.

Analogously, we define continuity on $\mathcal{C}$ or $\mathcal{R}^n$ by use of absolute values.

THEOREM 5.1. **(Rules about Continuity)** *Let $f, g : D \subset \mathcal{R} \to \mathcal{R}$ be (equi)continuous at the point $x \in D$ (and there ~ 1). Then $f + g$ and $f \cdot g$ are (equi)continuous at the point x. Let h be (equi)continuous at the point $f(x)$, then $h \circ f$ is (equi)continuous at the point x.*

The proof is analogous to the case of R.

DEFINITION 19. **(Differentiability, Equidifferentiability)** *The function $f : D \subset \mathcal{R} \to \mathcal{R}$ is called differentiable with derivative g at the point $x_0 \in D$ if, for any positive $\epsilon \in \mathcal{R}$, we can find a positive $\delta \in \mathcal{R}$ such that*

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - g \right| < \epsilon \text{ for any } x \in D \backslash \{x_0\} \text{ with } |x - x_0| < \delta.$$

If this is the case, we write $g = f'(x_0)$. The function is called equidifferentiable at the point x_0, if for any at most finite ϵ it is possible to choose δ such that $\delta \sim \epsilon$. The function is called k-equidifferentiable at the point x_0 if, for any at most finite ϵ, it is possible to choose $\epsilon \sim d^k \cdot \delta$.

Analogously, we define differentiability on $\mathcal{C}$ using absolute values.

THEOREM 5.2. **(Rules about Differentiability)** *Let $f, g : D \to \mathcal{R}$ be (equi)differentiable at the point $x \in D$ (and not infinitely large there). Then $f + g$ and $f \cdot g$ are (equi)differentiable at the point x, and the derivatives are given by $(f+g)'(x) = f'(x) + g'(x)$ and $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$. If $f(x) \neq 0$ ($f(x) \sim 1$), the function $1/f$ is (equi)differentiable at the point x with derivative $(1/f)'(x) = -f'(x)/f^2(x)$. Let h be differentiable at the point $f(x)$, then $h \circ f$ is differentiable at the point x, and the derivative is given by $(h \circ f)'(x) = h'(f(x)) \cdot f'(x)$.*

The proofs are analogous to the case of R. For equidifferentiability we also obtain $\epsilon \sim \delta$ as required.

Functions that are produced by a finite number of arithmetic operations from constants and the identity have therefore the same properties of smoothness as in R and C.

The following consequence is often important for practical purposes.

THEOREM 5.3. **(Equidifferentiability of Power Series)** *Let $f(z) = \sum_{i=0}^{\infty} a_i (z - z_0)^i$ be a power series with purely complex coefficients on $\mathcal{C}$ with radius of convergence $\eta > 0$. If η is finite, the series*

$$g_k(z) = \sum_{i=k}^{\infty} i \cdot (i-1) \cdot \ldots \cdot (i-k+1) a_i (z - z_0)^{i-k}$$

converges weakly for any $k \geq 1$ and for any z with $|z - z_0| < \eta$ and $|z - z_0| \not\approx \eta$; if $\eta = \infty$, the series converges for any $k \geq 1$ and for any z for which $|z - z_0|$ is finite. Furthermore, the function f is infinitely often equidifferentiable for such z, with derivatives $f^{(k)} = g_k$. In particular, for $i \geq 0$, we have $a_i = f^{(i)}(z_0)/i!$. For $z \in C$ the derivatives agree with the corresponding ones of the complex power series.

Proof: Observe that $\lim_{n \to \infty} \sqrt[n]{n} = 1$ and use induction on k, the first part is clear.

For the proof of the second part, let $|z - z_o| < \eta$, $|z - z_0| \not\approx \eta$. Let us first state two intermediate results concerning the term $|(f(z+h) - f(z))/h - g_1(z)|$. First let h be not infinitely small. Let $z_c \in C$ and $h_c \in C$ be the purely complex parts of z and h, therefore $z_c =_0 z$, $h_c =_0 h$. Evidently, we obtain $g_1(z_c) =_0 g_1(z)$ and $f(z_c) =_0 f(z)$. As $h_c \neq 0$, we obtain

$$\left| \frac{f(z+h) - f(z)}{h} - g_1(z) \right| =_0 \left| \frac{f(z_c + h_c) - f(z_c)}{h_c} - g_1(z_c) \right| . \quad (i)$$

On the other hand, let h be infinitely small. Write $h = h_0 \cdot d^r \cdot (1 + h_1)$ with $h_0 \in C$, $0 < r \in Q$, h_1 infinitely small. Then we obtain for any $s \leq 2r$

$$f(z+h)[s] = \sum_{i=0}^{\infty} a_i (z + h - z_0)^i [s] = \sum_{i=0}^{\infty} a_i \cdot \sum_{\nu=0}^{i} ((z - z_0)^{i-\nu} \frac{i!}{\nu!(i-\nu)!} h^\nu)[s]$$

$$= \sum_{i=0}^{\infty} a_i ((z - z_0)^i)[s] + \sum_{i=1}^{\infty} (h \cdot i \cdot a_i (z - z_0)^{i-1})[s] + \sum_{i=2}^{\infty} (h^2 \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2})[s].$$

Other terms are not relevant, since the corresponding powers of h are much smaller than d^s in absolute value. Therefore we obtain

$$\frac{f(z+h) - f(z)}{h} - g_1(z) =_r h_0 d^r \sum_{i=2}^{\infty} \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2} . \quad (ii)$$

Let now $\epsilon > 0$ in $\mathcal{R}$ be given. First consider the case of $\epsilon \sim 1$. Since f is differentiable in C, for any $z_c \in C$, we may choose a $\delta > 0$ in R such that $|(f(z_c + h_c) - f(z_c))/h_c - g_1(z_c)| < \epsilon/2$ for all nonzero $h_c \in C$ with $|h_c| < 2\delta$.

Let now $h \in \mathcal{C}$, $|h| < \delta$. As a first subcase, we consider $h \sim 1$; choose h_c as the purely complex part of h, namely, $h_c =_0 h, h_c \in C$, and $|h_c| < 2\delta$. Then, using (i), we obtain

$$\left| \frac{f(z+h) - f(z)}{h} - g_1(z) \right| < \left| \frac{f(z_c + h_c) - f(z_c)}{h_c} - g_1(z_c) \right| + \frac{\epsilon}{2} < \epsilon \ \forall h \text{ with } |h| < \delta.$$

In the second subcase, we consider $|h| \ll 1$. We write $h = h_0 \cdot d^r (1 + h_1)$, with h_0 purely complex, $r \in Q$ and positive, and h_1 infinitely small, to obtain from (ii)

$$\left| \frac{f(z+h) - f(z)}{h} - g_1(z) \right| < d^{r/2} < \epsilon.$$

For infinitely small ϵ, we write $\epsilon = \epsilon_0 \cdot d^{r_\epsilon}(1 + \epsilon_1)$, with $r_\epsilon \in Q$ positive, $\epsilon_0 \in R$, and ϵ_1 infinitely small. Choose now $\delta = \epsilon / 4 |(\sum_{i=2}^{\infty} \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2})[0]|$ if the sum does not vanish, $\delta = \epsilon$ otherwise. Obviously, $\delta \sim \epsilon$ in both cases. Consider now h with $|h| < \delta$, and write $h = h_0 \cdot d^{r_h}(1 + h_1)$ with $h_0 \in C$, $r_h \geq r_\epsilon$ in Q, and h_1 infinitely small. Then we obtain, again fro (ii),

$$\frac{f(z+h) - f(z)}{h} - g_1(z) =_{r_h} h_0 d^{r_h} \sum_{i=2}^{\infty} \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2}.$$

For $r_h > r_\epsilon$, we have $|(f(z+h) - f(z))/h - g_1(z)| =_{r_\epsilon} 0$, and hence it follows that $|(f(z+h) - f(z))/h - g_1(z)| < \epsilon$. Consider therefore $r_h = r_\epsilon = r$. For vanishing sum $\sum_{i=2}^{\infty} \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2}$, we have $(f(z+h) - f(z))/h - g_1(z) =_r 0$, and therefore less than ϵ in magnitude. Otherwise, we obtain

$$\left| \frac{f(z+h) - f(z)}{h} - g_1(z) \right| < 2|h_0| d^r | \sum_{i=2}^{\infty} \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2} | < \epsilon,$$

which concludes the proof. □

We complete our discussion with a key theorem that indeed reduces the calculation of derivatives to mere arithmetic operations and that represents the key for a rigorous computational study of differentiability, which is described in [Shamseddine1996a].

THEOREM 5.4. **(Derivatives Are Differential Quotients, After All!)** *Let $f : D \to \mathcal{R}$ be a function that is equidifferentiable at the point $x \in D$. Let $|h| \ll d^r$, and $x + h \in D$. Then the derivative of f satisfies*

$$f'(x) =_r \frac{f(x+h) - f(x)}{h}.$$

In particular, the real part of the derivative can be calculated exactly from the differential quotient for any infinitely small h.

Proof: Let h be as in the requirement, $h = h_0 \cdot d^{r_h}(1 + h_1)$, with $h_0 \in R$, h_1 as before, and therefore $r_h > r$. Choose now $\epsilon = d^{(r+r_h)/2}$. Since f is equidifferentiable, we can find a positive $\delta \sim \epsilon$ such that for any Δx with $|\Delta x| < \delta$, the differential quotient differs less than ϵ from the derivative, and hence $\left| \frac{f(x+\Delta x) - f(x)}{\Delta x} - f'(x) \right|$ is infinitely smaller than d^r. But apparently, the above h satisfies $|h| < \delta$. □

For reasons of space, our treatment of calculus on $\mathcal{R}$ concludes with this central theorem, which is the key to performing computational differentiation with $\mathcal{R}$. A detailed study of the practical issues, including methods to rigorously decide differentiability of computer functions, is given in [Shamseddine1996a]. There is a wealth of other calculus on $\mathcal{R}$, including versions of intermediate value theorem, Rolle's theorem, and Taylor's theorem, and a theory of integration that allows the treatment of delta functions in the expected manner; for details, the reader is referred to [Berz1994a].

Acknowledgments

I am grateful to Louis Rall for the first thought that "(0, 1) is something like a differential." For discussions about Levi-Civita fields, surreal numbers and non-standard Analysis, I would like to thank Detlef Laugwitz and John H. Conway. For a careful reading of the manuscript, I thank Khodr Shamseddine and Silke Rolles.

References

[Alling1987a] N. L. ALLING, *Foundations of Analysis over Surreal Number Fields*, North Holland, Amsterdam, 1987.

[Berz1990c] M. BERZ, *Analysis auf einer nichtarchimedischen Erweiterung der reellen Zahlen*, Tech. Report MSUCL – 753, Department of Physics, Michigan State University, East Lansing, Mich., 1990.

[Berz1992b] ——, *Automatic differentiation as nonarchimedean analysis*, in Computer Arithmetic and Enclosure Methods, Amsterdam, 1992, Elsevier Science Publishers.

[Berz1994a] ——, *Analysis on a nonarchimedean extension of the real numbers*, Lecture Notes, Studienstiftung Summer School, September 1992 MSUCL–933, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, Mich., 1994.

[Conway1976a] J. H. CONWAY, *Elementary Numerical Analysis*, North Holland, Amsterdam, 1978.

[Davis1977a] M. DAVIS, *Applied Nonstandard Analysis*, John Wiley, New York, 1977.

[Ebbinghaus1992a] H. E. ET AL., *Zahlen*, Springer-Verlag, Berlin, 1992.

[Fuchs1963a] L. FUCHS, *Partially Ordered Algebraic Systems*, Pergamon Press, Addison-Wesley, Reading, Mass., 1963.

[Gonshor1986a] H. GONSHOR, *An Introduction to the Theory of Surreal Numbers*, Cambridge University Press, 1986.

[Hewitt1969a] E. HEWITT AND K. STROMBERG, *Real and Abstract Analysis*, Springer-Verlag, Berlin, 1969.

[Knuth1974a] D. E. KNUTH, *Surreal Numbers: How Two Ex-Students Turned on to Pure Mathematics and Found Total Happiness*, Addison-Wesley, Reading, Mass., 1974.

[Laugwitz1959a] D. LAUGWITZ, *Eine Einführung der Delta-Funktionen*, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, 4 (1959), p. 41.

[Laugwitz1961a] ——, *Anwendungen unendlich kleiner Zahlen I*, Journal für die reine und angewandte Mathematik, 207 (1961), pp. 53–60.

[Laugwitz1961b] ——, *Anwendungen unendlich kleiner Zahlen II*, Journal für die reine und angewandte Mathematik, 208 (1961), pp. 22–34.

[Laugwitz1968a] ——, *Eine nichtarchimedische Erweiterung angeordneter Körper*, Mathematische Nachrichten, 37 (1968), pp. 225–236.

[Laugwitz1973a] ——, *Ein Weg zur Nonstandard-Analysis*, Jahresberichte der Deutschen Mathematischen Vereinigung, 75 (1973), pp. 66–93.

[Laugwitz1975a] ——, *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, Roma, 1975, Academia Nazionale dei Lincei.

[Levi-Civita1892a] T. LEVI-CIVITA, *Sugli infiniti ed infinitesimi attuali quali elementi analitici*, Atti Ist. Veneto di Sc., Lett. ed Art., 7a (1892), p. 1765.

[Levi-Civita1898a] ——, *Sui numeri transfiniti*, Rend. Acc. Lincei, 5a (1898), pp. 91, 113.

[Lightstone1975a] A. H. LIGHTSTONE AND A. ROBINSON, *Nonarchimedean Fields and Asymptotic Expansions*, North Holland, New York, 1975.

[Neder1943a] L. NEDER, *Modell einer Leibnizschen Differentialrechnung mit aktual unendlich kleinen Größen*, Mathematische Annalen, 118 (1941–43), pp. 718–732.

[Ostrowski1935a] A. OSTROWSKI, *Untersuchungen zur arithmetischen Theorie der Körper*, Mathematische Zeitschrift, 39 (1935), pp. 269–404.

[Rall1986a] L. B. RALL, *The arithmetic of differentiation*, Mathematics Magazine, 59 (1986), p. 275.

[Robinson1961a] A. ROBINSON, *Non-standard analysis*, Proceedings Royal Academy Amsterdam, Series A, 64 (1961), p. 432.

[Robinson1974a] ——, *Non-Standard Analysis*, North-Holland, 1974.

[Rudin1987a] W. RUDIN, *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

[Schmieden1958a] C. SCHMIEDEN AND D. LAUGWITZ, *Eine Erweiterung der Infinitesimalrechnung*, Mathematische Zeitschrift, 69 (1958), pp. 1–39.

[Shamseddine1996a] K. Shamseddine and M. Berz, *Exception handling in derivative computation with nonarchimedean calculus*, in Computational Differentiation: Techniques, Applications, and Tools, M. Berz, C. H. Bischof, G. F. Corliss, and A. Griewank, eds., SIAM, Philadelphia, Penn., 1996, pp. 37–51.

[Stromberg1981a] K. Stromberg, *An Introduction to Classical Real Analysis*, Wadsworth, 1981.

[Stroyan1976a] K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the Theory of Infinitesimals*, Academic Press, New York, 1976.