Michigan State University

National Superconducting Cyclotron Laboratory

ANALYSIS ON A NONARCHIMEDEAN EXTENSION OF THE REAL NUMBERS

LECTURE NOTES, STUDIENSTIFTUNG SUMMER SCHOOL, BUDAPEST, SEPTEMBER 1992

M. BERZ

MSUCL-933

MAY 1994

Analysis on a Nonarchimedean, Extension of the Real Numbers

Lecture Notes

Studienstiftung Summer School
Budapest, September 1992

Martin Berz
Department of Physics and Astronomy and
National Superconducting Cyclotron Laboratory
Michigan State University
East Lansing, MI 48824

MSUCL-933

Analysis on a Nonarchimedean Extension of the Real Numbers

Lecture Notes

Studienstiftung Summer School
Budapest, September 1992

Martin Berz
Department of Physics and Astronomy and
National Superconducting Cyclotron Laboratory
Michigan State University
East Lansing, MI 48824

MSUCL-933

Abstract

A field extension $\mathcal{R}$ of the real numbers is presented. It has similar algebraic properties as R; for example, all roots of positive numbers **exist,** and the structure C obtained by adjoining the imaginary unit is algebraically complete. The set can be totally ordered and contains infinitely small and infinitely large quantities.

Under the topology induced by the ordering, the set becomes Cauchy complete; but different from R, there is a second natural way of introducing a topology. It is shown that $\mathcal{R}$ is the smallest totally ordered algebraically complete extension of R.

Power series have identical convergence properties as in R, and thus important transcendental functions exist and behave as in R. Furthermore, there is a natural way to extend any other real function under preservation of its smoothness properties. In addition to these common functions, delta functions can be introduced directly.

A calculus involving continuity, differentiability and integrability is developed. Central concepts like the intermediate value theorem, mean value theorem, and Taylor's theorem with remainder hold under slightly stronger conditions.

It is shown that, up to infinitely small errors, derivatives are differential quotients, i.e. slopes of infinitely small secants. While justifying the intuitive concept of derivatives of the fathers of analysis, it also offers a practical way of calculating exact derivatives numerically.

The existence of infinitely small and large numbers allows an introduction of delta functions in a natural way, and the important theorems about delta functions can be shown using the calculus on $\mathcal{R}$.

Contents

1 Introduction

The real numbers owe their fundamental role in mathematics and natural 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. Further yet, they allow the introduction of certain transcendental functions like exp, which are important in the sciences and arise from the concepts of power series. Furthermore, they offer an analysis involving differentiation and integration, a requirement needed for the expression of even very simple laws of nature.

While the first two properties are readily satisfied by the rational numbers, the geometric requirements demand to use 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. Being at such a basic level of our scientific language, hardly any thought is spent on the fundamental question whether there may be other useful number systems having the required properties.

This question is perhaps even more interesting in the 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. Firstly, 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. Secondly, 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", that is slopes of secants with infinitely small abscissa and ordinate differences, cannot be formulated rigorously within the real numbers.

The problems mentioned in the preceding paragraphs could 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 nonarchimedean. Since any archimedean Cauchy complete field is isomorphic to R, it is indeed this property 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 rather 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 nonarchimedean extensions of the real numbers (see for example the books of Rudin [1], Hewitt and Stromberg [2], or Stromberg [3], or at a deeper level the works of Fuchs [4], Ebbinghaus et al.[5] or Lightstone and Robinson [6]), but such extensions usually quickly fail to fulfill one or several of the above criteria of a "useful" field, usually even regarding the existence of roots.

A very important idea for the problem of the infinite came from Schmieden and Laugwitz [7] which was applied to Delta Functions[8] [9] and Distributions[10]. Certain equivalence classes of sequences of real numbers become the new number set, and perhaps most interestingly, 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 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 evolving structure, while very large, is not a field. The ring contains zero divisors and is not totally ordered. Robinson [11] then recognized that the intuitive method can be generalized [12] by a nonconstructive process based on model theory to obtain a totally ordered field, and initiated the branch of nonstandard analysis. Some of the standard works describing this field are from Robinson[13], Stroyan and Luxemburg[14], and Davies[15]. In this discipline, the transfer of theorems about real numbers is extremely simple, however at the expense of a non-constructive process invoking the axiom of choice, leading to an exceedingly large structure of numbers and theorems. The non-constructiveness makes practical use difficult and leads to several oddities, for example 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" [16]. A humerous yet 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" [17]. (We wonder about the applicability of the method to the problem of socially disadvantaged children, and also happily follow the author in using his fabulous TEX typesetting system). Other important accounts on surreal numbers are by Alling[18] and Gonshor.[19] .

In this paper, analysis on a different nonarchimedean extension of the real numbers is discussed. The numbers $\mathcal{R}$ were first discovered by the brilliant young Levi-Civita [20] ,[21], who succeeded to show 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. The subject appeared again in the work by Ostrowski[22], Neder[23], and later in the work of Laugwitz[24]. Two modern accounts of this work can be found in the book by Lightstone and Robinson[6], which ends with the proof of Cauchy completeness, as well as in Laugwitz' account on Levi-Civita's work [25], which also contains a summary of properties of Levi-Civita fields.

In this paper, we extend the previous work and attempt to formulate the basis of a workable analysis on the Levi-Civita field $\mathcal{R}$. Section 2 discusses questions about the structure of the field. We show that $\mathcal{R}$ admits nth roots of positive elements and that the field obtained by adjoining the imaginary unit is algebraically closed. We also introduce a new topology, complementing the order topology. In section 3, we apply these to the study of sequences and series; in particular, we show that any power series with complex coefficients converges within the conventional radius of convergence; this allows for the direct use of a large class of functions. In section 4, we develop a differential calculus on $\mathcal{R}$, and we prove certain fundamental tools like the intermediate value and mean value theorems, which hold under slightly stronger conditions.

2 The Nonarchimedean Fields $\mathcal{R}$ and $\mathcal{C}$

2.1 Algebraic Structure

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 will be 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 *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; so 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 $M = \bigcup_n M_n$. So 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. So we conclude that the sequence is divergent. ∎

Lemma 3 *Let $M, N \in \mathcal{F}$. Then we have*

a) $X \subset M \Rightarrow X \in \mathcal{F}$

b) $M \cup N \in \mathcal{F}$

c) $M \cap N \in \mathcal{F}$

d) $M + N = \{x + y \mid x \in M, y \in N\} \in \mathcal{F}$

e) For every $x \in M + N$, there are only finitely many pairs $(a, b) \in M \times N$ such that $x = a + b$.

Proof:
Statements a) - c) follow directly from the definition.

For d), 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

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

and set

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

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 $(a,b) \in M \times N$ which satisfy $x = a+b$ lie in the finite set $M^u \times N^u$. ∎

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

Definition 4 (The Sets $\mathcal{R}$ and $\mathcal{C}$) *We define*

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

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

So 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, i.e. they have left-finite support.

Obviously, we have $\mathcal{R} \subset \mathcal{C}$. In the following, we will denote elements of $\mathcal{R}$ and $\mathcal{C}$ by x, y, etc. and identify their values at $q \in Q$ with brackets like $x[q]$. 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 utilize the properties of left-finite sets (2) for their description:

Definition 5 (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 will refer to the pair of sequences $((q_n),(x[q_n]))$ as the table of x.*

For the further discussion, it is convenient to introduce the following terminology:

Definition 6 (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 which, among other things, is useful for the concept of differentiation on $\mathcal{R}$.

Lemma 7 *The relations $\sim$, $\approx$ and $=_r$ are equivalence relations. They satisfy*

$$x \approx y \Rightarrow x \sim y$$

$$\text{If } a, b \in Q, \quad a > b, \text{ then } x =_a y \Rightarrow x =_b y$$

Furthermore, we have

$$\lambda(\partial x) \leq \lambda(x); \text{ and if } \lambda(x) \neq 0, \infty, \text{ even } \lambda(\partial x) = \lambda(x) - 1$$

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

Definition 8 (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 $\operatorname{supp}(x+y) \subseteq \operatorname{supp}(x) \cup \operatorname{supp}(y)$, so by Lemma (3), with x and y having left-finite support, so does $x+y$. Lemma (3) 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 $\operatorname{supp}(x \cdot y) \subseteq \operatorname{supp}(x) + \operatorname{supp}(y)$, application of Lemma (3) shows that $\operatorname{supp}(x \cdot y) \in \mathcal{F}$.

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 9 *$(\mathcal{R},+,\cdot)$ and $(\mathcal{C},+,\cdot)$ are commutative rings with units.*

Proof:
$\mathcal{R}$ and $\mathcal{C}$ form abelian groups with respect to addition, where the neutral element is the function that vanishes on all Q, and the additive inverse is the function that is the pointwise inverse; obviously, the inverse is itself an element of $\mathcal{R}$ or $\mathcal{C}$ respectively.

The unit of multiplication is the function that vanishes everywhere except at 0, where it takes the value 1.

Multiplication is commutative since for all $q \in Q$, and for all $x, y \in \mathcal{C}$

$$(x \cdot y)[q] = \sum_{q_x+q_y=q} x[q_x] \cdot y[q_y] = \sum_{q_y+q_x=q} y[q_y] \cdot x[q_x] = (y \cdot x)[q]$$

For the proof of the associativity of multiplication, consider $x, y, z \in \mathcal{C}$. Then for all $q \in Q$ we have:

$$((x \cdot y) \cdot z)[q] = \sum_{q_{x \cdot y}+q_z=q} (x \cdot y)[q_{x \cdot y}] \cdot z[q_z] = \sum_{q_{x \cdot y}+q_z=q} \left(\sum_{q_x+q_y=q_{x \cdot y}} x[q_x] \cdot y[q_y] \right) \cdot z[q_z]$$

$$= \sum_{q_x+q_y+q_z=q} x[q_x] \cdot y[q_y] \cdot z[q_z] = \sum_{q_x+q_{y \cdot z}=q} x[q_x] \cdot \left(\sum_{q_y+q_z=q_{y \cdot z}} y[q_y] \cdot z[q_z] \right)$$

$$= \sum_{q_x+q_{y \cdot z}=q} x[q_x] \cdot (y \cdot z)[q_{y \cdot z}] = (x \cdot (y \cdot z))[q]$$

We note again that all sums arising are actually finite. In a similar manner, we prove the distributive law:

$$((x+y) \cdot z)[q] = \sum_{q_{x+y}+q_z=q} (x+y)[q_{x+y}] \cdot z[q_z]$$

$$= \sum_{q_x+q_z=q} x[q_x] \cdot z[q_z] + \sum_{q_y+q_z=q} y[q_y] \cdot z[q_z] = (x \cdot z + y \cdot z)[q],$$

and this concludes the proof. ∎

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

Theorem 10 (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 11 *In the following, we identify an element $x \in C$ with its image $\Pi(x) \in \mathcal{C}$ under the embedding. We remind 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 12 *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 $z_1 = z_2$.*

So the restrictions of the relations $\sim$ and $\approx$ to R and C do not produce anything new. Besides presenting themselves as ring extensions of R and C, because of the embeddings of R and C, the new sets also become linear spaces:

Theorem 13 (Differential Algebraic Structure) *The sets $\mathcal{R}$ and $\mathcal{C}$ form infinite dimensional vector spaces over R and C, respectively. Via the multiplication on $\mathcal{R}$ and $\mathcal{C}$, they are also algebras. The operation ∂ is a derivation, i.e.*

$$\partial(a+b) = \partial a + \partial b \text{ and } \partial(a \cdot b) = (\partial a) \cdot b + a \cdot (\partial b) \text{ for all } a, b \in \mathcal{C},$$

and so $\mathcal{R}$ and $\mathcal{C}$ form differential algebras.

The proof is obvious. It is also worth noting that the quantity λ is actually a valuation:

Theorem 14 (Valuation Structure) *The operation λ has the following properties:*

$$\lambda(x \cdot y) = \lambda(x) + \lambda(y) \text{ and } \lambda(x+y) \geq \min(\lambda(x), \lambda(y))$$

So it is a valuation of $\mathcal{C}$.

Definition 15 (The Number d) *Define the element $d \in \mathcal{R}$ as*

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

Lemma 16 (Algebraic Properties of d) *The number d has an inverse and admits n-th roots in $\mathcal{R}$*

Proof:
Obviously the numbers denoted d^{-1} and $d^{1/n}$, where

$$d^{-1}[q] = \begin{cases} 1 & \text{if} \quad q = -1 \\ 0 & \text{else} \end{cases}$$

$$d^{1/n}[q] = \begin{cases} 1 & \text{if} \quad q = 1/n \\ 0 & \text{else} \end{cases}$$

satisfy the requirements. ∎

Note that now, rational powers d^q of d are defined.

We have shown that $\mathcal{R}$ and $\mathcal{C}$ contain the real and complex numbers respectively, but in addition contain many more elements. The next theorem shows that, from the point of view of set theory, $\mathcal{C}$ is not larger than R.

Theorem 17 *The sets $\mathcal{C}$ and R have the same cardinality.*

Proof:
Since we constructed an injective mapping $\Pi : R \to \mathcal{C}$, we have card$(R) = c \leq$ card(C). On the other hand, every element of $\mathcal{C}$ is uniquely determined by a sequence of support points and two sequences of function values (for the real and imaginary parts respectively). So $\mathcal{C}$ can be mapped injectively to a subset of the set of functions $N \to R^3$ (where we agree to append triplets of zeroes if the set of support points is finite). Thus by the laws for cardinal number arithmetic, it follows that

$$\text{card(C)} \leq (c^3)^{\aleph_0} = c^{\aleph_0} = c = \text{card}(R),$$

and altogether we obtain card(R)=card(C). ∎

The only nontrivial step towards 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 central theorem that will be of key importance for a variety of proofs.

Lemma 18 (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 19 *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 want to 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 gives 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 get $f(x) =_{q+k} f(a_i)$, which in turn gives

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

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 $y = x$. ∎

Remark 20 *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. Among others, this is of significant importance for practical purposes.*

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

Theorem 21 *$(\mathcal{R},+,\cdot)$ 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 $f(x) = -y \cdot x - y$, the problem is thus reduced 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 get $y \cdot x_1 =_{q+k} y \cdot x_2$, and hence

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

thus f satisfies the hypothesis of the fixed point theorem (18), and consequently a fixed point of f exists. ∎

Now we examine the existence of roots in $\mathcal{R}$ and $\mathcal{C}$ and find that, regarding this important property, the new fields behave just like R and C respectively:

Theorem 22 *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 n-th roots in $\mathcal{R}$. If n is even and $z[q]$ is negative, z has no n-th roots in $\mathcal{R}$. If n is odd, z has a unique n-th root in $\mathcal{R}$.*

Let $z \in \mathcal{C}$ be nonzero. Then z has n distinct n-th 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 n-th 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 n-th 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 get $x_1^2 =_{q+k_y} x_2^2$. By induction on m, we get $x_1^m =_{q+k_y} x_2^m$ for all $m \in N$, $m > 1$.

In particular, this gives $x_1^2 \cdot P(x_1) =_{q+k_y} x_2^2 \cdot P(x_2)$ and finally $f(x_1) =_{q+k_y} f(x_2)$. So f and M satisfy the hypothesis of the fixed point theorem (18) 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; so we could not have chosen anything smaller than Q as the domain of the functions which are the elements of our new fields.

We will end the section on the algebraic properties of $\mathcal{R}$ and $\mathcal{C}$ by showing 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 23 (Fundamental Theorem of Algebra for $\mathcal{C}$) *Every polynomial of positive degree with coefficients in $\mathcal{C}$ has a root in $\mathcal{C}$.*

Proof:
Let the polynomial $P(x) = \sum_{\nu=0}^{n} a_\nu \cdot x^\nu$ of degree $n > 0$ be given. We can assume $a_0 \neq 0$ because otherwise 0 is a root of $P(x)$. Furthermore, multiplying the polynomial by a nonzero factor does not change the roots. Hence we can assume $\lambda(a_0) = 0$. Set $r = \min_{\nu \geq 1}(\lambda(a_\nu)/\nu)$, and let j be an index where this minimum is assumed. Replacing x by $d^{-r} \cdot x$ changes the coefficient a_ν into $\bar{a}_\nu = a_\nu \cdot d^{-r \cdot \nu}$, and, by the choice of r, we have $\lambda(\bar{a}_\nu) \geq 0$ and $\lambda(\bar{a}_j) = 0$. Since it suffices to find a root of the new polynomial, we can assume $\lambda(a_\nu) \geq 0$, $\lambda(a_0) = \lambda(a_j) = 0$ for the coefficients of the original polynomial.

Write $a_\nu = b_\nu + c_\nu$, where $b_\nu = a_\nu[0] \in C$, and form the polynomial $P_C(x) = \sum_{\nu=0}^{n} b_\nu \cdot x^\nu$ over C. Because $b_j \neq 0$, P_C is nonconstant. Hence by the fundamental theorem of algebra, P_C has a root $X_C \in C$.

Set $q = \min(\lambda(c_\nu))$, $s_C = P(X_C) = \sum_{\nu=0}^{n} c_\nu \cdot X_C^\nu$. If $s_C = 0$, we have found a root of $P(x)$. Otherwise, we have $\lambda(s_C) \geq q$. If $\lambda(s_C) = q$, we simply set $X = X_C$ and obtain $\lambda(P(X)) = q$. If $\lambda(s_C) > q$, we define X in a different way. Since P_C is nonconstant, not all of the formal derivatives of P_C can vanish at X_C. Let k denote the smallest number such that $P_C^{(k)}(X_C) \neq 0$. We define $X = X_C + d^{q/k}$. Taylor's theorem for polynomials then gives

$$P(X) = P(X_C) + d^{q/k} \cdot P'(X_C) + d^{2q/k} \cdot \frac{P''(X_C)}{2} + ... + d^{nq/k} \cdot \frac{P^{(n)}(X_C)}{n!}.$$

We have $\lambda(P(X_C)) > q$ and $\lambda(d^{iq/k} \cdot P^{(i)}(X_C)) \geq iq/k + q > q$ for $1 \leq i < k$, by definition of k and q. Furthermore, we have $\lambda(d^{iq/k} \cdot P^{(i)}(X_C) \geq iq/k > q$ for $i > k$. Hence we conclude that $\lambda(P(X)) = q = \lambda(d^{kq/k} P^{(k)}(X_C))$.

Now we set $s = P(X)$ and obtain $\lambda(s) = q$.

After these preparations, we can apply the fixed point theorem. Like we did in search for multiplicative inverses and n-th roots, we look for roots of the polynomial near the classical solution X_C, or to be precise, near X. We are looking for x such that $P(X + x) = 0$. For this we write

$$P(X + x) = P(X) + x \cdot P'(X) + x^2 \cdot \frac{P''(X)}{2} + ... + x^n \cdot \frac{P^{(n)}(X)}{n!},$$

according to Taylor's theorem for polynomials (which is a merely algebraic consequence). Since the complex parts of $P(X)$ and $P(X_C)$ agree, we still have $P^{(k)}(X)[0] \neq 0$; and k is the smallest number with this property.

Now let t denote a k-th root of $-k! \cdot s/P^{(k)}(X)$, which exists by theorem (22). Set $u(x) = u_1(x) + u_2(x)$, where

$$u_1(x) = \frac{1}{s}[\frac{P'(X)}{1!} + x \cdot \frac{P''(X)}{2!} + ... + x^{k-2} \cdot \frac{P^{(k-1)}(X)}{(k-1)!}]$$

$$u_2(x) = \frac{x^k}{s} \cdot [\frac{P^{(k+1)}(X)}{(k+1)!} + x \cdot \frac{P^{(k+2)}(X)}{(k+2)!} + ... + x^{(n-k-1)} \cdot \frac{P^{(n)}(X)}{n!}]$$

Then it suffices to find a solution of the fixed point equation

$$x = f(x) = t \cdot \sqrt[k]{1 + x \cdot u(x)},$$

which can be seen by raising both sides of the equation to the k-th power and performing a few calculations. Here the k-th root is to be interpreted as the unique k-th root with complex part equal to 1.

We have $\lambda(t) = \lambda(s)/k > 0$. Set $M = \{x \in \mathcal{C} \mid \lambda(x) \geq \lambda(t)\}$.

Since $\lambda(P^{(k)}(X)) = 0$ and $\lambda(P^{(\nu)}(X)) \geq q = \lambda(s)$ for all $\nu < k$, we get $\lambda(u_1(x)) \geq 0$ for $x \in M$. Since furthermore, $\lambda(x^k) \geq \lambda(t^k) = \lambda(s)$ for $x \in M$, we conclude that $\lambda(x^k/s) \geq 0$ holds. Thus we obtain $\lambda(u_2(x)) \geq 0$ and $\lambda(u(x)) \geq 0$, and hence $\lambda(f(x)) \geq \lambda(t)$ for all $x \in M$, giving $f(M) \subset M$ as needed. Now let $x_1, x_2 \in M$ satisfying $x_1 =_p x_2$ be given. Then we have $1 + x_1 \cdot u(x_1) =_p 1 + x_2 \cdot u(x_2)$, and we show that $\sqrt[k]{1 + x_1 \cdot u(x_1)} =_p \sqrt[k]{1 + x_2 \cdot u(x_2)}$ holds. For let $y = \sqrt[k]{1 + x_2 \cdot u(x_2)} - \sqrt[k]{1 + x_1 \cdot u(x_1)}$; since $\sqrt[k]{1 + x_1 u(x_1)} \approx 1 \approx \sqrt[k]{1 + x_2 u(x_2)}$, it follows first that $\lambda(y) > 0$. We write $\sqrt[k]{1 + x_2 u(x_2)} = \sqrt[k]{1 + x_1 u(x_1)} + y$; raising this equation to the k-th power, we obtain $1 + x_2 \cdot u(x_2) = 1 + x_1 \cdot u(x_1) + k \cdot y \cdot \sqrt[k]{1 + x_1 \cdot u(x_1)}^{(k-1)} + Q(y)$. Here Q is a polynomial in y, with neither constant nor linear parts, the coefficients of which are at most finite. Hence it follows that $\lambda(Q(y)) \geq 2\lambda(y) > \lambda(y)$. We write $y = k^{-1}(1 + x_1 \cdot u(x_1))^{(1-k)/k}[x_2 u(x_2) - x_1 u(x_1) - Q(y)]$. Then $\lambda(x_2 u(x_2) - x_1 u(x_1)) > p$ and $\lambda(k^{-1}(1 + x_1 u(x_1))^{(1-k)/k}) = 0$ imply $\lambda(y) > p$, so we conclude that $\sqrt[k]{1 + x_1 \cdot u(x_1)} =_p \sqrt[k]{1 + x_2 \cdot u(x_2)}$ holds.

But this means $t \cdot \sqrt[k]{1 + x_1 \cdot u(x_1)} =_{p+\lambda(t)} t \cdot \sqrt[k]{1 + x_2 \cdot u(x_2)}$. Since $\lambda(t) > 0$, the requirements of the fixed point theorem are satisfied, assuring the existence of a fixed point and in turn a root of the polynomial $P(x)$. ∎

As in the case with the determination of inverses and n-th roots, the roots of polynomials can be calculated to a given depth in finitely many steps.

Since the roots of polynomials over $\mathcal{R}$ which lie in $\mathcal{C}$ always appear in conjugate pairs, we obtain:

Corollary 24 ($\mathcal{R}$ is Real Closed) *Every Polynomial over $\mathcal{R}$ with odd degree has a root in $\mathcal{R}$.*

Alternative proof of the fundamental theorem of algebra for $\mathcal{C}$:

Theorem 25 *The field $\mathcal{C}$ is algebraically closed.*

Proof:
First we introduce some terminology which will be used only in this proof: a polynomial $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ over $\mathcal{C}$ is called normalized if $\lambda(a_\nu) \geq 0$ for all $\nu = 0, \ldots, n$, if $\lambda(a_0) = 0$ and there exists $j \in \{1, \ldots, n\}$

such that $\lambda(a_j) = 0$. A number $c \in C$ is called a quasi-root of a normalized polynomial $P(x)$ if $\lambda(P(c)) > 0$. If furthermore there exists $j \in \{1, \ldots, n-1\}$ such that $\lambda(P^{(j)}(c)) = 0$, c will be called a good quasi-root of $P(x)$ and j will be called an index of c.

We note that every polynomial $P(x)$ of positive degree with $a_0 \neq 0$ can be normalized without changing its degree nor the number of its roots: set $r = \min\{(\lambda(a_\nu) - \lambda(a_0))/\nu \mid \nu = 1, \ldots, n\}$ and consider $\bar{P}(x) = d^{-\lambda(a_0)} \cdot P(d^{-r}x)$ instead of $P(x)$. Also note that every normalized polynomial $P(x)$ of positive degree possesses a quasi-root: form the polynomial $P_C(x) = a_n[0]x^n + \cdots + a_1[0]x + a_0[0]$ which is a nonconstant polynomial over C. Hence by the fundamental theorem of algebra, $P_C(x)$ has a root in C which is a quasi-root of $P(x)$. Note that quasi-roots of normalized polynomials are nonzero because $\lambda(a_0) = 0$.

Now let a polynomial $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ with degree $n > 0$ be given. $P(x)$ has a root in $\mathcal{C}$ if $n = 1$. From this we can assume $n \geq 2$ and do induction on the degree of $P(x)$. If $a_0 = 0$, then zero is a root. Otherwise, by the above remark we can assume $P(x)$ to be normalized and we distinguish two cases:

Case 1: $P(x)$ has a good quasi-root x_0 with index j. We inductively construct a sequence (x_m) such that

$$\lambda(P(x_m)) \geq \left(\frac{n}{n-1}\right)^m \cdot q \quad \text{and}$$

$$\lambda(x_{m+1} - x_m) \geq \left(\frac{n}{n-1}\right)^m \cdot \frac{q}{n-1}$$

where $q = \lambda(P(x_0)) > 0$. Note that this implies $x_m =_0 x_0$.

Assume that x_m is already constructed and has the above properties. Form the polynomial

$$Q(x) = P(x_m) + P'(x_m)x + \cdots + \frac{P^{(j)}(x_m)}{j!}x^j.$$

Since we have $x_m =_0 x_0$, we get $\lambda(P^{(j)}(x_m)) = \lambda(P^{(j)}(x_0)) = 0$. So $Q(x)$ is a nonconstant polynomial of degree j less than n. hence by induction hypothesis, $Q(x)$ has j roots $t_1, \ldots, t_j$ (not necessarily distinct) in $\mathcal{C}$. Since $Q(x) = \frac{P^{(j)}(x_m)}{j!} \prod_{i=1}^{j}(x - t_i)$, we get $\frac{P(x_m) \cdot j!}{P^{(j)}(x_m)} = (-1)^j \prod_{i=1}^{j} t_i$. So there exists

$i \in \{1, \ldots, j\}$ such that

$$\lambda(t_i) \geq \frac{1}{j}\lambda\left(\frac{P(x_m)\cdot j!}{P^{(j)}(x_m)}\right) = \frac{\lambda(P(x_m))}{j}.$$

We set $x_{m+1} = x_m + t_i$ and obtain

$$\lambda(x_{m+1} - x_m) = \lambda(t_i) \geq \frac{\lambda(P(x_m))}{n-1} \geq \left(\frac{n}{n-1}\right)^m \frac{q}{n-1}.$$

Also we have

$$P(x_{m+1}) = P(x_m + t_i) = Q(t_i) + \frac{P^{(j+1)}(x_m)}{(j+1)!}t_i^{j+1} + \cdots + \frac{P^{(n)}(x_m)}{n!}t_i^n.$$

Since $Q(t_i) = 0$ and $\lambda(P^{(\nu)}(x_m)) \geq 0$ for all $\nu = j+1, \ldots, n$, we thus get

$$\lambda(P(x_{m+1})) \geq (j+1)\lambda(t_i) \geq \frac{j+1}{j}\lambda(P(x_m)) \geq \left(\frac{n}{n-1}\right)^{m+1}\cdot q.$$

Similar as in the fixed point theorem, the constructed sequence (x_m) defines a number that turns out to be a root of $P(x)$: for $p \in Q$ we choose $m \in N$ large enough to satisfy $(\frac{n}{n-1})^{m+1}\cdot\frac{q}{n-1} > p$ and put $X[p] := x_m[p]$, which is independent of the choice of m. Since below every given rational number the function X agrees with an element of $\mathcal{C}$, X itself has left-finite support and is thus an element of $\mathcal{C}$. For every $m \in N$ we have $\lambda(X - x_m) \geq (\frac{n}{n-1})^m \cdot \frac{q}{n-1}$ by definition of X. Since

$$P(X) = P(x_m) + P'(x_m)(X - x_m) + \cdots + \frac{P^{(n)}(x_m)}{n!}(X - x_m)^n,$$

we obtain $\lambda(P(X)) \geq \min(\lambda(P(x_m)), \lambda(X - x_m)) \geq (\frac{n}{n-1})^m \cdot \frac{q}{n-1}$ for every $m \in N$. This implies $\lambda(P(X)) = +\infty$. Thus we have found a root of $P(x)$.

Case 2: None of the quasi-roots of $P(x)$ is good. Nevertheless $P(x)$ has a quasi-root $c_0 \in C^*$. Let k be the largest index such that $\lambda(a_k) = 0$. Then $\lambda(P^{(k)}(c_0)) = 0$, so we conclude that $k = n$ because otherwise c_0 would be good. But this means that $\lambda(a_n) = 0$.

Now we repeatedly perform a certain process until we either find a root of $P(x)$ or we eventually construct sequences (c_m) of elements in C^* and (r_m)

of elements of Q with certain properties. Setting $r_0 = 0$ and $x_m = \sum_{i=0}^{m} c_i d^{r_i}$, these properties are

$$\lambda(P(x_{m+1})) > \lambda(P(x_m)), \quad r_{m+1} = \lambda(P(x_m))/n$$

$$\text{and} \quad \lambda(P^{(n-1)}(x_m)) > r_m$$

Note that this implies that the sequence (r_m) is strictly increasing.

Assume that $c_0, \ldots, c_{m-1}$ and $r_0, \ldots .r_{m-1}$ have already been constructed without revealing a root of $P(x)$. We consider the polynomial

$$\bar{P}(x) := P(x + x_{m-1}) = P(x_{m-1}) + P'(x_{m-1})x + \cdots + \frac{P^{(n)}(x_{m-1})}{n!}x^n.$$

If $P(x_{m-1}) = 0$, we have found a root; otherwise we normalize $\bar{P}(x)$ and obtain

$$Q(x) := d^{-\lambda(P(x_{m-1}))}P(d^{r_m}x + x_{m-1})$$

where $r_m = -\min\{(\lambda(P^{(\nu)}(x_{m-1})) - \lambda(P(x_{m-1})))/\nu \mid \nu = 1, \ldots, n\}$. $Q(x)$ has a quasi-root $c_m \in C^*$. If $Q(x)$ possesses a good quasi-root case 1 gives us a root of $Q(x)$ which in turn leads to a root of $P(x)$. So we assume that $Q(x)$ has no good quasi-roots either. As was noted at the beginning of this case, this implies that the leading coefficient of $Q(x)$ has lambda equal to 0. On the other hand the lambda of that leading coefficient is equal to $n \cdot r_m + \lambda(a_n) - \lambda(P(x_{m-1}))$; thus we obtain $r_m = \lambda(P(x_{m-1}))/n$ because $\lambda(a_n) = 0$.

We have $P(x_m) = P(d^{r_m}c_m + x_{m-1}) = d^{\lambda(P(x_{m-1}))}Q(c_m)$, which implies

$$\lambda(P(x_m)) = \lambda(P(x_{m-1})) + \lambda(Q(c_m)) > \lambda(P(x_{m-1})).$$

Since c_m is not a good quasi-root of $Q(x)$, we have $\lambda(Q^{(n-1)}(c_m)) > 0$. But

$$Q^{(n-1)}(c_m) = d^{(n-1)r_m - \lambda(P(x_{m-1}))} \cdot P^{(n-1)}(d^{r_m}c_m + x_{m-1}) = d^{-r_m}P^{(n-1)}(x_m).$$

So we also obtain

$$\lambda(P^{(n-1)}(x_m)) = r_m + \lambda(Q^{(n-1)}(c_m)) > r_m.$$

If the sequences (r_m) and (x_m) are constructed as above and we have not yet found a root, we now show that $t := -\frac{a_{n-1}}{n \cdot a_n}$ is a root of $P(x)$. For all $m \in N$ we have

$$\lambda(n! \cdot a_n x_m + (n-1)! \cdot a_{n-1}) = \lambda(P^{(n-1)}(x_m)) > r_m,$$

which means $t =_{r_m} x_m$. Since $x_m = \sum_{i=0}^{m} c_i d^{r_i}$ the set of support points of t contains all the r_m. But the sequence (r_m) is strictly increasing and the support points of t form a left-finite set. Hence the sequence (r_m) can not be bounded.

We have

$$P(t) = P(x_m) + P'(x_m)(t - x_m) + \cdots + \frac{P^{(n)}(x_m)}{n!}(t - x_m)^n$$

and since $\lambda(t - x_m) > r_m$ as well as $\lambda(P(x_m)) = n \cdot r_{m+1} > r_m$ we get $\lambda(P(t)) > r_m$ for all $m \in N$. Since (r_m) is unbounded we conclude that $\lambda(P(t)) = +\infty$ and we detected a root of $P(x)$. ∎

2.2 Order Structure

In the last section we have shown 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 will discuss the ordering.

The simplest way of introducing an order is to define a set of 'positive' numbers:

Definition 26 (The Set $\mathcal{R}^+$) *Let $\mathcal{R}^+$ be the set of all non-vanishing elements x of $\mathcal{R}$ which satisfy $x[\lambda(x)] > 0$:*

Lemma 27 (Properties of $\mathcal{R}^+$) *The set $\mathcal{R}^+$ has the following properties:*

$\mathcal{R}^+ \cap (-\mathcal{R}^+) = \emptyset$, $\mathcal{R}^+ \cap \{0\} = \emptyset$

$\mathcal{R}^+ \cup \{0\} \cup (-\mathcal{R}^+) = \mathcal{R}$

$x, y \in \mathcal{R}^+ \Rightarrow x + y \in \mathcal{R}^+$

$x, y \in \mathcal{R}^+ \Rightarrow x \cdot y \in \mathcal{R}^+$

The proofs follow rather directly from the respective definitions.

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

Definition 28 (Ordering in $\mathcal{R}$). *Let x, y be elements of $\mathcal{R}$. We say $x > y$, if and only if $x - y \in \mathcal{R}^+$. Furthermore, we say $x < y$, iff $y > x$.*

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

Theorem 29 (Properties of the Order)

With the order relation defined in (28), $(\mathcal{R},+,\cdot)$ becomes a totally ordered field, i.e.

For any x, y, exactly one of the following holds: $x < y$, $x = y$, $x > y$.

For any x, y, z with $x > y$, $y > z$, we have $x > z$.

Furthermore, the order is compatible with the algebraic structure of $\mathcal{R}$, i.e.

For any x, y, z, we have: $x > y \Rightarrow x + z > y + z$.

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

Proof:

The first statement directly follows from theorem (27).

To prove the second statement we write $x-z = (x-y)+(y-z)$. According to the requirements, we have $x-y \in \mathcal{R}^+$, $y-z \in \mathcal{R}^+$, and thus the statement follows by theorem (27).

The third statement is obtained because $(x+z)-(y+z) = x-y \in \mathcal{R}^+$.

The fourth statement follows analogously to the second because $(x \cdot z) - (y \cdot z) = z \cdot (x - y)$. Since, according to the requirements, $z \in \mathcal{R}^+$ and $(x-y) \in \mathcal{R}^+$, the proof follows from theorem (27). ∎

We immediately obtain:

Corollary 30 *The embedding Π from theorem (10) is compatible with the order, i.e. $x < y \Rightarrow \Pi(x) < \Pi(y)$.*

Remark 31 *Note that $\mathcal{C}$ cannot be ordered since it contains C, which cannot be ordered.*

Thus $\mathcal{R}$, like C, is a proper field extension of R. Note that this is not a contradiction to the well-known uniqueness of C as a field extension of R. The respective theorem of Frobenius asserts only the non-existence 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 32 **($\ll, \gg$)** *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 33 *For all $a, b, c \in \mathcal{R}$, we have*

$a \ll b \;\Rightarrow a < b$

$a \ll b,\; b \ll c \;\Rightarrow a \ll c.$

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

Corollary 34 *The field $\mathcal{R}$ is nonarchimedian, i.e there are elements which are not exceeded by any natural number.*

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

One of the consequences of $\mathcal{R}$ being nonarchimedian is that the idea of Dedekind cuts and the existence of suprema are no longer valid:

Example 35 **(Dedekind Cuts and Suprema)** *Let M_u, M_o be defined as follows:*

$M_u = \{x \in \mathcal{R} \;|x < 0$ *and* $|x|$ *is not infinitely small*$\}$

$M_o = \mathcal{R} \setminus M_u$

Then we obviously have $M_u \cup M_o = \mathcal{R}$, $M_u < M_o$ and $M_u \neq \emptyset \neq M_o$. Nevertheless, there is no cut s satisfying $M_u \leq s \leq M_o$: Assume s was such a cut.

If s is positive or zero we have $s \in M_o$, but $-d$ also is an element of M_o and it is smaller than s. Thus $s \not\leq M_o$.

If s is negative and $|s|$ not infinitely small, we have $s \in M_u$. But then $s/2$ also is negative and $|s/2|$ not infinitely small, and therefore $s/2$ is an element of M_u. From $s/2 > s$, we infer $M_u \not\leq s$.

Finally, if s is negative and $|s|$ infinitely small, we have $s \in M_o$. But $2 \cdot s$ also is an element of M_o and $2 \cdot s < s$. Thus, $s \not\leq M_o$.

Hence such a cut s cannot exist. Furthermore, M_u also does not have a supremum: M_u is bounded above by any element of M_o, but it is impossible to select one least of these upper bounds.

Remark 36 *It is apparent that the nonexistence of suprema is a consequence of the nonarchimedicity and is not specific to $\mathcal{R}$. Obviously, the same argument holds for any nonarchimedian totally ordered field if d is chosen to be any positive infinitely small quantity.*

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

Example 37 (Calculation of Derivatives with Differentials) *Let us consider the function $f(x) = x^2 - 2x$. f is differentiable on R, and we have $f'(x) = 2x - 2$. As we know, we can get 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 $f : R \to R$, then this is given exactly by the 'real part' of the difference quotient.

This observation is of great fundamental and practical importance. It enables us to replace differentiation by algebraic operations.

As we will show later, all algebraic operations on $\mathcal{R}$ can be implemented directly on a computer. Thus we are now able to determine exact derivatives numerically. This is a drastic improvement compared to all numerical methods operating with differences.

2.3 Topological Structure

In this section we will examine the topological structures of $\mathcal{R}$ and the related sets. We will see that on $\mathcal{R}$, in contrast to R, several different non-trivial 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 38 (Absolute Value on $\mathcal{R}$) *Let $x \in \mathcal{R}$. We define the absolute value of x as follows:*

If $x \geq 0$, we say $|x| = x$.

If $x < 0$, we say $|x| = -x$.

Lemma 39 (Properties of the Absolute Value)

The mapping "$|\ |$" $: \mathcal{R} \to \mathcal{R}$ has the following properties:

$|x| = 0$ *iff* $x = 0$

$|x| = |-x|$

$|x \cdot y| = |x| \cdot |y|$

$|x + y| \leq |x| + |y|$

$||x| - |y|| \leq |x - y|$

Proof:

The first two properties are obvious. For the third one, it suffices to consider the different cases which are given by the signs of x and y respectively.

To show the triangle inequality, we first note that, for any x, we have $-|x| \leq x \leq |x|$. Adding this inequality applied to x and y, we obtain

$-(|x| + |y|) \leq x + y \leq (|x| + |y|)$, that is $|x + y| \leq |x| + |y|$.

The last property directly follows from the triangle inequality, because

$|x| \leq |x - y| + |y|$, that is $|x| - |y| \leq |x - y|$; analogously, we obtain $|x - y| = |y - x| \geq |y| - |x| = -(|x| - |y|)$, and the statement is shown. ∎

Definition 40 (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 (22).

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

Definition 41 (Order Topology) *We call a subset M of $\mathcal{R}$, $\mathcal{C}$ or $\mathcal{R}^n$ open iff for any $x_0 \in M$ 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:

Lemma 42 (Properties of the Order Topology) *With the above topology, $\mathcal{R}$, $\mathcal{C}$ and $\mathcal{R}^n$ become non connected 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.*

Proof:
We first observe that the balls $O(x_0, \epsilon)$ and the whole space are open. Furthermore, all unions and finite intersections of open sets are obviously open. Since $M_1 = \{x \leq 0 \text{ or } (x > 0, x \ll 1)\}$ and $M_2 = \{x > 0 \text{ and } x \not\ll 1\}$ are disjoint and open, but $\mathcal{R} = M_1 \cup M_2$, $\mathcal{R}$ is non connected. Let x, y be different elements. Then $O(x, |x - y|/2)$ and $O(y, |x - y|/2)$ are disjoint and open and contain x and y respectively. Thus $\mathcal{R}$, $\mathcal{C}$ and $\mathcal{R}^n$ are Hausdorff.

There cannot be any countable basis because the uncountably many open sets $M_X = O(X, d)$, $X \in R, C$ or R^n are disjoint. Obviously, the open sets induced on R, C or R^n respectively by the sets M_X are just the single points. Thus, in the induced topology, all sets are open and the topology is therefore discrete.

To prove that the space is not locally compact, consider $x \in \mathcal{R}$ and let U be a neighbourhood of x. Let $\epsilon \in \mathcal{R}$ be such that $O(x, \epsilon) \subset U$. We have to show that the closure U^- of U is not compact. Let the sets M_i be defined as follows:

$M_{-1} = \{y | y - x \gg d \cdot \epsilon\} \cup \{y | y < x\}$

$M_i = (x + (i - 1) \cdot d \cdot \epsilon, x + (i + 1) \cdot d \cdot \epsilon)$ for $i = 0, 1, 2, \ldots$

Then the sets M_i cover $\mathcal{R}$, and, in particular, the closure U^- of any neighbourhood of x:

If $y < x$, we have $y \in M_{-1}$.

$y = x \in M_0$

If $y > x$ and $y - x \ll d \cdot \epsilon$ we obtain $y \in M_1$
If $y > x$ and $y - x \gg d \cdot \epsilon$ we obtain $y \in M_{-1}$
Otherwise, y is contained in one of the M_i, $i = 1, 2, ...$

Furthermore, the sets are open: the M_i, $i = 0, 1, 2, ...$ are open intervals. The set $\{y | y - x \gg d \cdot \epsilon\}$ is open because, with any y, it also contains $O(y, d \cdot \epsilon)$. Obviously, $\{y | y < x\}$ is also open. Thus M_{-1} is a union of open sets and hence itself open.

But it is impossible to select finitely many sets of the M_i which cover U^-, because each of the infinitely many numbers $x + i \cdot d\epsilon \in U$ is contained only in the set M_i. ∎

Remark 43 *A detailed study of the proof reveals that it can be executed in the same way on any other nonarchimedean structure, and thus the above unusual properties are not specific to $\mathcal{R}$.*

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

Definition 44 (Semi Norm on $\mathcal{C}$) *We introduce the semi norm "$|| \ ||_r$" as a function from $\mathcal{C}$ into R as follows:*

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

The supremum is finite and it is even a maximum since for any r, only finitely many of the $x[q]$ considered do not vanish. Thus the semi-norm is similar to the supremum norm for continuous functions. Its properties also are quite similar.

Lemma 45 (Properties of the Semi Norm) *For an arbitrary r, the mapping "$|| \ ||_r$" : $\mathcal{R} \to R$ satisfies the following:*

$||0||_r = 0$
$||x||_r \geq 0$
$||x||_r = ||-x||_r$
$||x + y||_r \leq ||x||_r + ||y||_r$
$| \, ||x||_r - ||y||_r \, | \leq ||x - y||_r$

Using the family of these semi norms, we can now define another topology:

Definition 46 (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.

Lemma 47 (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:
We can easily check that the balls $S(x_0, \epsilon)$ are open: If $x \in S(x_0, \epsilon)$, we also have $S(x, \epsilon - ||x - x_0||_{1/\epsilon}) \subset S(x_0, \epsilon)$. Furthermore, the whole space is open, and unions as well as finite intersections of open sets are also open. The balls $S(r, q)$ with r, q rational form a basis of the topology. We obtain a Hausdorff space: Let $x \neq y$ be given; let $r = \lambda(x - y)$. We define $\epsilon = \min(|(x - y)[r]|/2, 1/2|r|)$. Then $S(x, \epsilon)$ and $S(y, \epsilon)$ are disjoint and open, and contain x and y respectively.

Considering elements of R, their supports can only consist of zero. Therefore, the open subsets of $\mathcal{R}$ correspond to the open subsets of R. ∎

In addition to the topologies discussed, there is another topology which takes into account that, in any practical scenario, it will not be possible to detect infinitely small errors, nor will it be possible to measure infinitely large quantities. We obtain this topology by a suitable continuation of the order topology on R.

Definition 48 (Measure Topology) *Given any open subset of R, C or R^n, we form a subset of $\mathcal{R}$, $\mathcal{C}$ or $\mathcal{R}^n$ containing the elements of the original set as well as all the elements infinitely close to them. To the family of sets we obtain in this way, we add one more set, namely the one containing every element with infinitely large absolute values.*

Thus a basis of this topology would be all ϵ-balls with real ϵ and the set of numbers with infinitely large absolute value.

Lemma 49 (Properties of the Measure Topology) *With the above topology, $\mathcal{R}$, $\mathcal{C}$ and $\mathcal{R}^n$ are connected topological spaces. They are not Hausdorff. The topology is locally compact and induces the usual order topology on R, C and R^n respectively.*

Proof:
We can directly show that the whole space is open and that unions and finite intersections of open sets are also open. Obviously, elements with an infinitely small difference cannot be separated; they always are simultaneously inside or outside of any open set. The rest directly follows by transferring the properties of the order topology on R.

Lemma 50 (Comparison of the Topologies) *The order topology is a refinement of both the semi norm topology and the measure topology.*

3 Sequences and Series

3.1 Convergence and Completeness

In this section, we will discuss convergence with respect to the topologies introduced in the last chapter. We begin by introducing a special property of sequences:

Definition 51 (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, i.e. iff $\cup_{i=0}^{\infty}\text{supp}(a_i) \in \mathcal{F}$.*

This property is not automatically assured, as it 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 52 (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 $A = \cup_{i=0}^{\infty}\text{supp}(a_i)$, $B = \cup_{i=0}^{\infty}\text{supp}(b_i)$ 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 53 (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, i.e. 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 $\mathcal{C}$:

Theorem 54 (Expansion in Powers of Differentials) *Let $((q_i), (x[q_i]))$ be the table of $x \in \mathcal{R}$ or $\mathcal{C}$ (cf. 5). 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), 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. ∎

As it turns out, there is a very clear criterion describing the sequences and series that converge strongly:

Theorem 55 (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.

Proof:
The first property is obvious. The limit is an element of $\mathcal{R}$ or $\mathcal{C}$ since its support below any bound is equal to the support of a certain member of the sequence.

To prove the second statement, assume the series converges strongly. According to the first property, this means that, for any $r \in Q$, there exists $n \in N$ such that $0 =_r (\sum_{i=0}^{j+1} a_i - \sum_{i=0}^{j} a_i) = a_j$ for all $j \geq n$. But then, again according to the first statement, the sequence is a null sequence. The other direction follows analogously. ∎

Lemma 56 *Every strongly convergent sequence is regular.*

Proof:
Let $r \in Q$ be given. Use the convergence criterion (55) 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 $\cup_{i=0}^{\infty}\mathrm{supp}(a_i)$ smaller than r do already occur in $\cup_{i=0}^{n}\mathrm{supp}(a_i)$. This finite union, however, is contained in $\mathcal{F}$; and thus there are only finitely many elements of $\cup_{i=0}^{\infty}\mathrm{supp}(a_i)$ below r. ∎

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

Theorem 57 (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 (55) for series.

The other direction is proved analogously as in R: Let (a_n) converge strongly to the limit a. Let $\epsilon > 0$ be given. Choose $n \in N$ such that $|a_\nu - a| < \frac{\epsilon}{2} \, \forall \, \nu > n$. Let now $n_1, n_2 > n$ be given. Then we have $|a_{n_1} - a_{n_2}| \leq |a_{n_1} - a| + |a_{n_2} - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. The proof for C is analogous. ∎

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 we have to introduce another kind of convergence:

Definition 58 (Weak Convergence) *We call the sequence (a_i) weakly convergent if there is an $a \in C$ such that (a_i) converges to a with respect to the semi norm topology, i.e. 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 59 (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.*

Let on the other hand (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 $A \doteq \cup_i \text{supp}(a_i) \in \mathcal{F}$, the limit function a is an element of 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. So 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 60 (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 (i.e. the semi norm topology) and locally converges to the function which assumes the value $1/n$ at $-n \in Z^-$ and vanishes elsewhere. But this limit function is not an element of $\mathcal{C}$.*

Example 61 (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:

Theorem 62 *Strong convergence implies weak convergence to the same limit.*

Proof:
Let (a_i) converge strongly to the limit a. According to lemma (56), the sequence is regular. Furthermore, according to (55), for any $r \in Q$, there exists $n \in N$ such that $a_i[q]$ does not differ any more from $a[q]$ for all $q < r$ and for all $i > n$. In particular, this implies that $a_i[q] \to a[q]$ for all q. According to the convergence criterion for weak convergence (59), we obtain weak convergence to the limit a. ∎

The theorem also follows directly from the fact that the order topology is a refinement of the semi norm topology. The converse is not true, as exemplified by the unbounded null sequence above.

It is worthwhile to study sequences of purely complex numbers in the light of the two concepts of convergence in $\mathcal{C}$:

Theorem 63 *Let (a_i) be a purely complex sequence in C converging to the limit a. Then, regarded as a sequence in $\mathcal{C}$, (a_i) converges weakly to the same limit. On the other hand, let (a_i) be a sequence in $\mathcal{C}$ with purely complex members converging weakly to the limit a. Then a is purely complex, and the sequence (a_i) converges to a in the complex sense.*

Proof:
Note that for a_i, $a \in C$, we have that $|a_i - a|$ (as real absolute value) is equal to $||a_i - a||_r$ for every $r \geq 0$; $r \in Q$. ∎

The statement is not valid for strong convergence: Obviously complex sequences converge strongly if and only if they are ultimately constant. Thus the idea of weak convergence provides a relation between the "natural" convergence on $\mathcal{C}$, i.e. the strong convergence, and the usual convergence on C. Moreover, it is the most useful tool for the examination of power series as we will see in the next paragraph.

To finish this section about the convergence of sequences, we will show that the field $\mathcal{R}$ is indeed the smallest nonarchimedean extension of R satisfying the basic requirements demanded in the beginning, which gives it a unique position among all other field extensions.

Theorem 64 (Uniqueness of $\mathcal{R}$) *The field $\mathcal{R}$ is the smallest totally ordered non archimedian field extension of R that is complete with respect to the order topology, in which every positive number has an n-th 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.*

Proof:
Obviously, $\mathcal{R}$ satisfies the mentioned conditions. We now show that $\mathcal{R}$ can be embedded in any other field extension of R that is equipped with the above properties. So let $\mathcal{S}$ be such a field.

Let $\delta \in \mathcal{S}$ be positive and infinitely small such that (δ^n) is a null sequence. Let $\delta^{1/n}$ be an n-th root of δ. Such a root exists according to the requirements. Now observe that $(\delta^{1/n})^m = (\delta^{1/n \cdot p})^{m \cdot p}\ \forall\ p \in N$. So let $q = \frac{m}{n} \in Q$, and let $\delta^q = (\delta^{1/n})^m$. This element is unique. Furthermore, δ^q is still infinitely small for $q > 0$. Let $q_1 < q_2$. Then we clearly have $\delta^{q_1} > \delta^{q_2}$. Now let $a \in R$. Since $\mathcal{S}$ is an extension of R, we also have $a \in \mathcal{S}$, and thus $a \cdot \delta^q \in \mathcal{S}$.

Now let $((q_i), (x[q_i]))$ be the table of an element x of $\mathcal{R}$. Consider the sequence

$$s_n = \sum_{i=1}^{n} x[q_i]\delta^{q_i}.$$

Then in fact this sequence converges in $\mathcal{S}$: Let $\epsilon > 0$ be given. Since, according to the requirements, (δ^n) converges to zero, there exists $n \in N$ such that $|\delta^\nu| < \epsilon\ \forall\ \nu \geq n$. Since the sequence (q_i) strictly diverges, there is $m \in N$ such that $q_\mu > n+1 \; \forall \; \mu > m$. But then we have for arbitrary $\mu_1 > \mu_2 > m$:

$$|s_{\mu_1} - s_{\mu_2}| = |\sum_{i=\mu_2+1}^{\mu_1} x[q_i]\delta^{q_i}| \leq \sum_{i=\mu_2+1}^{\mu_1} |x[q_i]|\delta^{q_i}$$

$$\leq \left(\sum_{i=\mu_2+1}^{\mu_1} |x[q_i]|\right) \delta^{q_{\mu_2+1}} \leq \left(\sum_{i=\mu_2+1}^{\mu_1} |x[q_i]|\right) \delta^{n+1}$$

$$< \delta^n < \epsilon,$$

and thus the sequence converges because of the Cauchy completeness of $\mathcal{S}$. We now assign to every element $\sum_{i=1}^{\infty} x[q_i] \cdot d^{q_i}$ of $\mathcal{R}$ the element $\sum_{i=1}^{\infty} x[q_i] \cdot \delta^{q_i}$ of $\mathcal{S}$. This mapping is injective. Furthermore, we immediately verify that it is compatible with the algebraic operations and the order on $\mathcal{R}$. ∎

Remark 65 *We note that a field with the properties of $\mathcal{R}$ could also be obtained by successively extending a simpler non archimedian field, e.g. the well known field of rational functions. To do this, we first would have to Cauchy complete the field. After that, the algebraic closure had to be done, for example by the method of Kronecker-Steinitz. This method, however, is non-constructive, whereas the direct path followed here is entirely constructive.*

Remark 66 *In the proof of the uniqueness, we noted that δ was only required to be positive and infinitely small and such that (δ^n) is a null sequence. But besides that, its actual magnitude was irrelevant. Thus, none of the infinitely small quantities is significantly different from the others. In particular, there is an isomorphism of $\mathcal{R}$ onto itself given by the mapping $x \mapsto x'$, where $x'[q] = b^q x[a \cdot q]$; $a \in Q$, $b \in R$, $a, b > 0$ fixed. This remarkable property has no analogy in R.*

3.2 Power Series

In this section we discuss a very important class of sequences, namely that of power series. Transcendental functions are very important, especially for application, and one of the nice properties of the structures at hand here is that power series can be introduced in much the same way as in R and C.

Furthermore, power series will prove important for the understanding of other topics of analysis on $\mathcal{R}$, especially for the problem of continuation of arbitrary real functions.

We start our discussion of power series with an observation

Lemma 67 *Let $M \in \mathcal{F}$, i. e. a left finite set. For M define*

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

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, i.e. M_Σ contains infinitely many elements smaller than zero and is therefore not left finite.

Let on the other hand $\min(M) \geq 0$. 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. Let now $r \in Q$; 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 $r < g$. Let now $r \geq g$, 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 68 *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 $M = \text{supp}(x)$, and is thus 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 (52) which asserts that the product of regular sequences is regular. ∎

Theorem 69 (Power Series with Purely Complex Coefficients) *Let $\sum_{n=0}^{\infty} a_n z^n$, $a_n \in C$ be a power series with radius of convergence equal to 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 (68), as 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 gives:

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

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

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

$$\leq \left(\sum_{j=0}^{m} \frac{|x^j[r]||X|^{m-j}}{j!}\right) \cdot \left(\sum_{n=\nu_1}^{\nu_2} |a_n| \cdot n^m \cdot |X|^{n-m}\right)$$

Note that the right hand sum contains only real terms. As $|X|$ is within the radius of convergence, the series converges (note that the additional factor n^m does not influence this since $\lim_{n\to\infty} \sqrt[n]{n^m} = 1$). As the left hand term does not depend on ν, we therefore 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 70 (Pointformula 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 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 gives:

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

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 $a_i = (f(z_0 + d))[i]$. ∎

We will see that power series on C find a useful application in discussion of so called formal power series. As we will see in the following theorem, any power series with purely complex coefficients converges for infinitely small arguments; furthermore, multiplication can be done term by term in the usual formal power series sense, and convergence is always assured. Therefore, formal power series form a natural part of the theory of power series.

Theorem 71 (Formal Power Series) *Any Power series with purely complex coefficients converges strongly on any infinitely small ball, even if the classical radius of convergence is zero. Furthermore, on any infinitely small ball we have, again independently of the radius of convergence, that*

$$(\sum_{n=0}^{\infty} a_n x^n) \cdot (\sum_{n=0}^{\infty} b_n x^n) = (\sum_{n=0}^{\infty} c_n x^n),$$

where $c_n = \sum_{j=0}^{n} a_j \cdot b_{n-j}$

Proof:
Note that for infinitely small x and any $r \in Q$, we find an m with $x^i[r] = 0$ for any $i > m$. Hence for a fixed r, the above summation includes only finitely many terms, which may be resorted according to the distributive law. ∎

In the following, it will prove extremely useful that any power series with purely complex coefficients converges for infinitely small arguments since it will allow us to find continuations of real functions in a natural way.

Above we discussed power series with purely complex coefficients only. This allows us to define functions on $\mathcal{R}$ as continuations of important transcendental functions on R like the sine and exponential functions. We will finish this section with a discussion of power series in which the coefficients are not restricted to C.

Theorem 72 (General Convergence Criterion for Power Series) *Define* $a_i = \sum_{j=0}^{\infty} a_{i,j} \cdot d^{r_{i,j}}$ *as a sequence in* $\mathcal{C}$. *Let*

$$r = -\liminf(\frac{r_{i,0}}{i})$$

Let $x \in C$ *with* $x \approx x_0 \cdot d^{r_x}$. *Then the power series* $\sum_{i=1}^{\infty} a_i x^i$ *converges strongly for* $r_x > r$ *and diverges for* $r_x < r$.

For $r = r_x$, *we obtain:*

If $-r_{i,0}/i < r$ *for infinitely many* i, *the series diverges.*

If $-r_{i,0}/i \le r$ *for only finitely many* i, *the series converges.*

If $-r_{i,0}/i < r$ *for only finitely many* i, *and* $-r_{i,0}/i = r$ *for infinitely many* i, *the series diverges if* (a_i) *is not regular; if* (a_i) *is regular, let* $s = 1/\limsup_{\{i|r_{i,0}/i=r\}}(a_{i,0})$ *(convention* $1/0 = \infty, 1/\infty = 0$*). Then the sequence converges if* $|x_{i,0}| < s$ *and diverges if* $|x_{i,0}| > s$.

4 Calculus on $\mathcal{R}$

4.1 Continuity and Differentiability

We will introduce the concepts of continuity and differentiability on $\mathcal{R}$ and $\mathcal{C}$ in this section. This is done as in R via the $\epsilon - \delta$- method. Unlike in R, ϵ and δ may be of a completely different order of magnitude.

Definition 73 (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.

We note that the stronger condition of equicontinuity is automatically satisfied in R, since there we always have $\epsilon \sim \delta$. Besides the concept of equicontinuity, we also introduce a connected concept.

Definition 74 (Equicontinuity of Order k) *The Function $f : D \subset \mathcal{R} \to \mathcal{R}$ is called order k equicontinuous at the point $x_0 \in D$, if for any positive $\epsilon \in \mathcal{R}$, there is a positive $\delta \in \mathcal{R}$ satisfying $\delta \sim \epsilon \cdot d^k$ such that*

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

Remark 75 *While technically more general, conceptually the order k equicontinuity is not very different from plain equicontinuity since by a mere scaling of function values by a factor of d^k, an order k equicontinuous function can be transformed to an equicontinuous function. This allows to almost directly generalize statements involving equicontinuity to order k equicontinuity. For the sake of convenience of notation, in the following we restrict ourselves mostly to the study of equicontinuity.*

Theorem 76 (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.*

Proof:
The proof is analogous to the case of R. ∎

Definition 77 (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 $\delta \sim \epsilon$.

Analogously, we define differentiability on $\mathcal{C}$ using absolute values. Similar to the case of continuity, the concept of equidifferentiability can be generalized formally without significant conceptual consequences.

Definition 78 (Order k Equidifferentiability) *The function $f : D \subset \mathcal{R} \to \mathcal{R}$ is called order k equidifferentiable with derivative g at the point $x_0 \in D$, if for any positive $\epsilon \in \mathcal{R}$, we can find a positive $\delta \in \mathcal{R}$ satisfying $\delta \sim \epsilon \cdot d^k$ such that*

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

Theorem 79 (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 $(h \circ f)'(x) = h'(f(x)) \cdot f'(x)$.*

Proof:
The proofs are done as in the case of R. For equidifferentiability we also get $\epsilon \sim \delta$. ∎

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. In particular, we obtain

Corollary 80 (Differentiability of Rational Functions) *A rational function (with purely complex coefficients) is (equi)differentiable at any (finite) point where the denominator does not vanish (is ~ 1) .*

One of the most important concepts of conventional analysis is that of power series. As we will show in our next theorem, even power series have analogous properties of smoothness as in conventional analysis; they are infinitely often differentiable.

Theorem 81 (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 C with real radius of convergence $\eta > 0$. Then 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 $|z - z_0| \not\approx \eta$. 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:
Observing that $\lim_{n\to\infty} \sqrt[n]{n} = 1$ and using 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 get $g_1(z_c) =_0 g_1(z)$ and $f(z_c) =_0 f(z)$. As $h_c \neq 0$, we obtain

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

Let on the other hand 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 as the corresponding powers of h are much smaller than d^s in absolute value. Therefore we get

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

Let now $\epsilon > 0$ in $\mathcal{R}$ be given. First consider the case of $\epsilon \sim 1$. As 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, i.e. $h_c =_0 h, h_c \in C$, and $|h_c| < 2\delta$. Then we get using (i):

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

In the second subcase, we consider $|h| \ll 1$; 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 get 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 $\delta = \epsilon / |(\sum_{i=2}^{\infty} \frac{i \cdot (i-1)}{2} a_i (z - z_0)^{i-2})[0]|$ if the sum does not vanish, $\delta = \epsilon$ otherwise. Obviously, we reach $\delta \sim \epsilon$ 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 from (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 thus $|(f(z+h) - f(z))/h - g_1(z)| < \epsilon$. 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 get

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

and the proof is completed. ∎

We complete this section with a theorem that in a sense reduces the calculation of derivatives to arithmetic operations and is therefore of importance for practical purposes.

Theorem 82 (Derivatives are Differential Quotients) *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 $|h| < \delta$. ∎

This is a central theorem, because it allows the calculation of derivatives of functions on R by simple arithmetic on $\mathcal{R}$, as we mentioned before and saw in a special example in (37).

The following consequence is often important for practical purposes.

Corollary 83 (Remainder Formula) *Let f be a function equidifferentiable at x, let $|h| \ll 1$. Then we obtain:*

$$f(x+h) = f(x) + h \cdot f'(x) + r(x,h) \cdot h^2,$$

with an at most finite remainder $r(x,h)$.

Proof:
Let $q = \lambda(h)$. Then we have by the above theorem

$$f'(x) =_r \frac{f(x+h) - f(x)}{h} \;\forall\, r < q,$$

from which we get by multiplication with h and rearrangement of terms

$$f(x+h) =_{r+q} f(x) + f'(x) \cdot h \;\forall\, r < q.$$

Let D be the difference between the left and the right hand side. Clearly $D[r] = 0 \;\forall\, r < 2q$. Let $r(x,h) = D/h^2$. Then we have $r(x,h)[r] = 0 \;\; \forall\, r < 0$, and therefore the expected result

$$f(x+h) = f(x) + f'(x) \cdot h + r(x,h) \cdot h^2,$$

as claimed. ∎

For functions that are extensions of real functions, i.e. assume real values at real points, the concept of derivatives as differential quotients can also be expressed in a rather concise form using the differential algebraic structure of $\mathcal{R}$:

Theorem 84 (Differential Algebraic Computation of Derivatives) *Let f be a function that is equidifferentiable at the real point x and assumes real value and derivative there. Then,*

$$f'(x) =_0 \partial\left[f(x+d)\right].$$

Proof:
Use the remainder formula of the last theorem and observe that $\partial(f(x)) = \partial(f'(x)) = 0$, $\partial d = 1$.

4.2 Continuation of Real and Complex Functions

In this section we will discuss under what circumstances an arbitrary real function can be extended, or continued, from R to $\mathcal{R}$. For two important classes of functions, rational functions and power series, we have already found such continuations via corollary (80) and theorem (81). In both of these cases, this continuation could be done in a rather natural way, as both algebraic operations and the calculations of limits transfer directly to our new field.

However, for functions that cannot be expressed only in terms of algebraic operations and limits, this method is not applicable, and other methods to define continuations are needed. In particular, we are interested in preserving as many of the original smoothnesss properties as possible. It turns out that this is possible in a rather general fashion, and thus allows to increase the pool of functions on the new set drastically.

Definition 85 (Normal Continuation on $\mathcal{R}$) *Let f be a real function on the real interval $[a_r, b_r] \subset R$, ($a = -\infty$ or $b = +\infty$ permitted), and let f be n times differentiable there, ($n = 0$ or $n = \infty$ permitted). Let $a, b \in \mathcal{R}$ be infinitely close to a_r, b_r. To the function f, we then define the order n continuation $\bar{f}_n$ on $[a, b] \subset \mathcal{R}$ as follows: Let $\bar{x} \in [a, b] \subset \mathcal{R}$. Write $\bar{x} = X + x$, with $X \in R$ and $|x|$ at most infinitely small, and set $\bar{f}_n(\bar{x})$ as:*

$$\bar{f}_n(\bar{x}) = \sum_{i=0}^{n} f^{(i)}(X) \cdot \frac{x^i}{i!}$$

A function on $\mathcal{R}$ is called an order n normal function if it is the order n normal continuation of a real function.

Note that for $n = \infty$, the above sum is strongly convergent independent of the size of the derivatives $f^{(i)}$ according to theorem (71) and thus well defined.

Clearly, the restriction of $\bar{f}_n$ to R is just f. Furthermore, in any infinitely small neighborhood of a real number, the function is given by its Taylor series. Analogously, functions defined on regions of C may be continued. Of particular interest is the case of analytic functions, which automatically posses order ∞ continuations.

Definition 86 (Analytic Continuation on $\mathcal{C}$) *Let f be an analytic function on the region $D \in C$. To the function f, we construct an analytic continuation $\bar{f}_\infty$ on $D \subset \mathcal{C}$ as follows: Let $\bar{x} \in (a,b) \subset \mathcal{C}$. Write $\bar{x} = X + x$, with $X \in C$, $|x|$ at most infinitely small, and define $\bar{f}_\infty(\bar{x})$ as:*

$$\bar{f}_\infty(\bar{x}) = \sum_{i=0}^{\infty} f^{(i)}(X) \cdot \frac{x^i}{i!}$$

Theorem 87 (Uniqueness of Continuation) *Let f_1 and f_2 be two order n continuations that agree on all real or complex points of their domain. Then $f_1 = f_2$.*

Proof:
The condition implies that the underlying real or complex functions agree, and so do their derivatives, entailing that their continuations agree. ∎

We observe that the normal continuation of a function has the same properties of smoothness as the original function.

Theorem 88 (Continuation of Continuous Functions) *Let f be a continuous function on $[a_r, b_r] \subset R$, and at least n times differentiable. Then the order n continuation $\bar{f}_n$ on $[a,b] \subset \mathcal{R}$ is equicontinuous.*

Proof:
Let $x \in [a,b]$, $\epsilon > 0$ in $\mathcal{R}$ be given. In case ϵ is finite, choose δ such that in R, $|f(x+h) - f(x)| < \epsilon/2$ for all real h with $|h| < 2\delta$, which is possible because of the continuity of f. But since for continued functions, function values of infinitely close points differ by at most infinitely small quantities, for $h \in \mathcal{R}$ we have $|f(x+h) - f(x)| < \epsilon$ for all h with $|h| < \delta$ as needed. On the other hand, for infinitely small ϵ, as δ has to be chosen with $\delta \sim \epsilon$, it is sufficient to study only the points that are infinitely close to x, in which region the function is given by a power series, which is known to be equicontinuous.

Theorem 89 (Continuation of Differentiable Functions) *Let f be a function on $[a_r, b_r] \subset R$, n times differentiable ($n = 0$ or $n = \infty$ permitted). Then the continued function $\bar{f}_n$ on $[a,b] \subset \mathcal{R}$ is n times equidifferentiable on $[a,b] \subset \mathcal{R}$, and for real points in $[a,b]$, the derivatives of f and $\bar{f}_n$ agree.*

Proof:
Let $x \in [a, b]$. We will first consider the case of finite ϵ. We choose a δ such that for all real h with $|h| < 2\delta$, the difference quotient $(f(Re(x) + h) - f(Re(x)))/h$ does not differ from the derivative by more than $\epsilon/2$. Let now $h \in \mathcal{R}$ be positive with $|h| < \delta$, and let h_c be its real part. For $h_c = 0$, the difference between the derivative and the difference quotient is infinitely small, and therefore certainly smaller than the finite ϵ. Otherwise, since $|h_c| < 2\delta$, we infer that the difference quotient does not disagree with the derivative by more than ϵ.

On the other hand, for $\epsilon \ll 1$, observe that since δ has to be chosen with $\delta \sim \epsilon$, it is sufficient to study only the points that are infinitely near to x; but for those points, the function $\bar{f}_n$ is given by a power series, which is differentiable to the advertised values. ∎

As mentioned before, functions defined by algebraic operations and limits, especially rational functions and power series, can also be continued directly by virtue of their algebraic and convergence properties. However, in this case the same result is obtained.

Theorem 90 (Continuation of Rational Functions and Power Series) *The order ∞ continuations of a rational function or a power series agree with the results obtained from the algebraic and limiting procedures, respectively.*

To close, let us study order 0 continuations of real functions:

Example 91 (Non-Constant Functions with Vanishing Derivative) *An interesting case is the order 0 continuation of real functions. According to the definition of the continuation, such functions are constant in every interval of infinitely small width. By choosing $\delta = d$, we immediately verify that they are differentiable for any inner point within their domain, and their derivative vanishes.*

4.3 Improper Functions

Clearly the class of normal functions, which are built as continuations of real functions, is rather small compared to the class of all possible smooth functions on $\mathcal{R}$ or $\mathcal{C}$. Furthermore, we are also interested in certain functions that cannot be obtained by continuation from R or C, like Delta Functions. Finally, when developing a theory of smooth functions, we want the class of functions to be as large as possible. So it is desirable to discuss functions that go beyond what can be obtained by continuations of functions on R. We begin by extending the concept of normal functions.

Definition 92 (Scaled Normal Functions) *Let f be a function on D in $\mathcal{R}$ or $\mathcal{C}$. Then we will call f a scaled normal function if f can be written as*

$$f = l_1 \circ f_n \circ l_2,$$

where $l_1(x) = a_1 + b_1 \cdot x$ and $l_2(x) = a_2 + b_2 \cdot x$ are linear functions with coefficients from $\mathcal{R}$ or $\mathcal{C}$ and where f_n is a normal function of order n.

We will see that while enhancing our pool of interesting functions substantially, the above introduced scaled normal functions behave very similarly to the normal functions.

Another interesting class of improper functions are the delta functions:

Definition 93 (Delta Functions) *Let $\bar{f}_d$:$R \to R$ be continuous, n times differentiable with $\int_{-\infty}^{\infty} \bar{f}_d(x)\,dx = 1$. Let f_d be the order n continuation of $\bar{f}_d$, and let $c \gg 1$. Then the function f with*

$$f(x) = \begin{cases} 0 & \text{for } |x| \gg 1/c \\ c \cdot f_d(c \cdot x) & \text{else} \end{cases}$$

is called a delta function.

Lemma 94 *Delta Functions vanish for all arguments with finite or infinitely large absolute value, and there are points infinitely close to the origin where they assume infinitely large values.*

So apparently the definition of delta functions just follows the intuitive concept. We will later see that they can be integrated, and we will also prove the famous integral projection property.

Example 95 (Some Delta Functions) *The following functions are delta functions:*

$$
\begin{aligned}
\delta_1(x) &= \begin{cases} 1/d & \textit{for } x \in [-d/2, d/2] \\ 0 & \textit{else} \end{cases} \\
\delta_2(x) &= \begin{cases} (1-|x|/d)/d & \textit{for } x \in [-d, d] \\ 0 & \textit{else} \end{cases} \\
\delta_3(x) &= \begin{cases} (1-x^2/2d^2)/2d & \textit{for } x \in [-d, d] \\ (|x|-2d)^2/4d^3 & \textit{for } d < |x| \leq 2d \\ 0 & \textit{else} \end{cases} \\
\delta_4(x) &= \begin{cases} \exp[-x^2/d^2]/\sqrt{2\pi}d & \textit{for } |x|/d \textit{ not infinite} \\ 0 & \textit{else} \end{cases}
\end{aligned}
$$

The second example is continuous on $\mathcal{R}$, the third and fourth even differentiable on $\mathcal{R}$.

4.4 Intermediate Values and Extrema

In this section we will discuss certain fundamental and important concepts of analysis, namely those of intermediate values and of extrema of functions. In the case of real functions, continuity is sufficient for the function to assume intermediate values and extrema. However, in $\mathcal{R}$, somewhat stronger conditions are required. We begin by demonstrating that in $\mathcal{R}$, continuity is not enough to guarantee that intermediate values be assumed.

Example 96 (Continuous Functions and Intermediate Values) *Let us consider two functions, defined on the interval* $[-1, +1]$*:*

$$f_1(x) = \begin{cases} -1 & \textit{if } x \leq 0 \textit{ or } (x > 0 \textit{ and } x \ll 1) \\ 1 & \textit{else} \end{cases}$$

$$f_2(x) = \mathcal{R}e(x)$$

we refer to f_2 as the Micro Gauss bracket, as it determines the (unique) real part of x.

Both f_1 and f_2 are continuous; for any ϵ just choose $\delta = d$ and utilize that both functions are constant on the d neighborhood around x for any $x \in [-1, +1] \subset \mathcal{R}$. The function f_2 is even equicontinuous: for any $\epsilon > 0$ in $\mathcal{R}$, choose $\delta = \epsilon/2$.

But the function f_1 does not assume the value 0 which certainly lies between $f_1(-1)$ and $f_1(+1)$.The values of the function f_2 are purely real, which implies that d will not be assumed, while it is obviously an intermediate value. On the other hand, f_2 at least comes infinitely close to any intermediate value.

The next theorem will show that intermediate values are assumed whenever the function is equidifferentiable and its derivative does not vanish.

Theorem 97 (Intermediate Value Theorem) *Let f be a function defined on the finite interval $[a, b]$, and let f be equidifferentiable there. Furthermore, assume $f(x)$ is finite, $f'(x) \sim 1$ in $[a, b]$. Then f assumes every intermediate value between $f(a)$ and $f(b)$.*

Proof:
Let S be an intermediate value between $f(a)$ and $f(b)$. We begin by determining an $X \in [a, b]$ such that $|S - f(X)|$ is infinitely small.

In case S lies infinitely near $f(a)$, choose $X = a$; otherwise, if S lies infinitely near $f(b)$, choose $X = b$. Otherwise, let S_R, a_R, b_R be the real parts of S, a, b, respectively. Define a real function $f_R : [a_R, b_R] \to R$ as follows:

$$f_R(r) = \begin{cases} Re(f(r)) & \text{if } r \in (a_R, b_R) \\ Re(f(a)) & \text{if } r = a_R \\ Re(f(b)) & \text{if } r = b_R \end{cases},$$

where "Re" denotes the real part. Then as a real function, f_R is continuous on $[a_R, b_R]$. Since S is not infinitely near $f(a)$ or $f(b)$, we infer that S_R lies between $f_R(a_R)$ and $f_R(b_R)$, and hence there is a real $X \in (a_R, b_R)$ such that $f_R(X) = S_R$. Because $a_R < X < b_R$ and all three numbers are real, we have $X \in [a, b]$. Furthermore, $|S - f(X)| \leq |S - S_R| + |f_R(X) - f(X)|$ is infinitely small as desired.

Now let $s = S - f(X)$. We try to find an infinitely small x such that $X + x \in [a, b]$ and $S = f(X + x)$. Because of equidifferentiability of f, we get according to (83):

$$S = f(X + x) = f(X) + f'(X) \cdot x + r(X, x) \cdot x^2,$$

where $r(X, x)$ is at most finite, and by assumption $f'(X)$ is finite as well. Transforming the condition on x to a fixed point problem, we obtain

$$x = \frac{s}{f'(X)} - \frac{r(X,x)}{f'(X)} \cdot x^2 = F(x).$$

Choose now $M = \{x|\lambda(x) \geq \lambda(s), X + x \in [a,b]\}$. Then $r(X,x)$ and hence F are defined on M. And we have $F(M) \subset M$: Clearly on M, $\lambda(F(x)) = \lambda(s)$. Furthermore, if $X = a$, s has the same sign as $f'(X)$, and hence x is positive, entailing $X + x \in [a,b]$; if $X = b$, s and f' have opposite signs, and hence x is negative, entailing $X + x \in [a,b]$; and otherwise, X is finitely far away from both a and b, entailing $X + x \in [a,b]$. Thus there is a fixed point of F, and the intermediate value is assumed. ∎

Remark 98 *The proof shows that, at the expense of clarity, the requirements of the theorem can be reduced to asking that the derivative not vanish at the real intermediate value. For the most important application of the intermediate value theorem in practice, namely the construction of inverse functions, this however does not represent a major restriction, since inverses are usually needed over extended ranges.*

As pointed out before, mere continuity of the function is not sufficient to assert the existence of intermediate values. It turns out that for the more special class of normal functions, an intermediate value theorem can also be obtained for vanishing first derivative as long as at each point one of the higher derivatives does not vanish.

Theorem 99 (Intermediate Value Theorem for Normal Functions) *Let f be an order n normal function defined on the interval $[a,b]$, and let at every point of the interval at least one of the derivatives not vanish. Then f assumes every intermediate value between $f(a)$ and $f(b)$.*

Proof:
Let $s \in \mathcal{R}$ be between $f(a)$ and $f(b)$ and let S_R be the real part of S and f_R the underlying real function. Then S_R is between $f_R(Re(a))$ and $f_R(Re(b))$, and since f_R is continuous on $[a,b]$, there exists a real $X \in [a,b]$ such that $f_R(X) = S_R$. Let $s = S - f_R(X) = S - f(X) = S - S_R$; then s is infinitely small. Since f is normal, we have

$$f(X + x) = f_R(X) + \sum_{i=1}^{n} f_R^{(i)}(X)\frac{x^i}{i!};$$

let now k be the index of the first nonvanishing derivative, which exists by assumption. Then

$$f(X+x) = f_R(X) + \frac{f_R^{(k)}(X)}{k!} \cdot x^k + r(X,x) \cdot x^{k+1},$$

where $\lambda(r(X,x)) \geq 0$, i.e. $r(X,x)$ is at most finite. Note that for every infinitely small x, we have $\lambda(x^{k+1}) > \lambda(x^k)$, and thus

$$f(X+x) - f_R(X) \approx \frac{f_R^{(k)}(X)}{k!} \cdot x^k.$$

In particular, if k is even, we infer that $f(X+x) - f_R(X) > 0$ if and only if $f_R^{(k)}(X) > 0$ (*). Now we try to find an infinitely small $x \in \mathcal{R}$ such that $S = f(X+x)$, which is equivalent to

$$s = \frac{f_R^{(k)}(X)}{k!} \cdot x^k + r(X,x) \cdot x^{k+1}.$$

Since according to (*), if k is even then $sk!/f^{(k)}(X) > 0$. Therefore,

$$g(x) = \left(\frac{k!s}{f^{(k)}(X)} - \frac{k!r(X,x)}{f^{(k)}(X)} \cdot x^{k+1} \right)^{1/k}$$

exists in $\mathcal{R}$ whether k is odd or even. The proof is now reduced to finding a solution of the fixed point problem $x = g(x)$. Let $M = \{x \in \mathcal{R} | \lambda(x) \geq \lambda(s)/k\}$, and let $x \in M$ be given; then $\lambda(x^{k+1}) > \lambda(x^k) > \lambda(s)$. Hence $\lambda(g(x)) = \lambda(s)/k$, and thus $g(x) \in M$. Let x_1, x_2 in M be given such that $x_1 =_q x_2$. Then $x_1^2 =_{q+\lambda(s)/k} x_2^2$, and by induction on m we obtain

$$x_1^m =_{q+(m-1)\lambda(s)/k} x_2^m \text{ for all } m \geq 1.$$

In particular, $x_1^{k+1} =_{q+\lambda(s)} x_2^{k+1}$. Hence $[g(x_1)]^k =_{q+\lambda(s)} [g(x_2)]^k$. Since $g(x_1) \approx \left(\frac{sk!}{f_R^{(k)}(X)}\right)^{1/k} \approx g(x_2)$, we have $\lambda[g(x_1) - g(x_2)] > \lambda(s)/k$. Let $y = g(x_2) - g(x_1)$; then

$$[g(x_2)]^k - [g(x_1)]^k = k \cdot y \cdot [g(x_1)]^{k-1} + Q(y),$$

where $Q(y)$ is a polynomial in y with neither constant nor linear terms. Since $\lambda(y) > \lambda(s)/k = \lambda(g(x_1))$, we obtain $\lambda(Q(y)) > \lambda[y \cdot (g(x_1))^{k-1}]$. Therefore,

$$\lambda(y \cdot (g(x_1))^{k-1}) = \lambda\left((g(x_2))^k - (g(x_1))^k\right) > q + \lambda(s).$$

Hence, $\lambda(y) > q + \lambda(s) - (k-1) \cdot \lambda(g(x_1)) = q + \lambda(s) - (k-1) \cdot \lambda(s)/k$. So $\lambda(y) > q + \lambda(s)/k$, which implies

$$g(x_1) =_{q+\lambda(s)/k} g(x_2), \text{ where } \lambda(s)/k > 0.$$

Therefore, g and M satisfy the requirements of the fixed point theorem. This provides a solution of the fixed point problem and hence completes the proof. ∎

While the restrictions of the intermediate value theorem regarding finiteness of functions and derivatives may appear somewhat stringent, it is obvious that the theorem can be utilized in a much more general way by subjecting the function under consideration to suitable coordinate transformations that bring it into the proper form. In particular, using linear transformations, we obtain the intermediate value theorem for scaled normal functions:

Corollary 100 (Intermediate Value Theorem for Scaled Normal Functions) *Let f be a scaled normal function, and let the underlying normal function satisfy the requirements of the intermediate value theorem for normal functions. Then f takes any intermediate value between $f(a)$ and $f(b)$.*

Example 101 (Intermediate Values of Delta Functions) *The delta function δ defined as*

$$\delta(x) = \begin{cases} \exp[-x^2/d^2]/\sqrt{2\pi}d & \text{for } |x|/d \text{ not infinite} \\ 0 & \text{else} \end{cases}$$

assumes every positive real number in an infinitely small neighborhood of the origin.

The proof is obvious.

Remark 102 *We note that the existence of roots can now be shown similarly to the real case by use of the intermediate value theorem.*

Similar to the question of intermediate values, we find that continuity is also not sufficient for the existence of extrema:

Example 103 (Continuous and Differentiable Functions and Extrema) *We define f_1 and f_2 on $[-1,+1]$ as follows:*

$$\begin{aligned} f_1(x) &= x - Re(x) \\ f_2(x) &= (x - Re(x))^2 \end{aligned}$$

where "$Re(x)$" denotes the real part of x. Then f_1 is equicontinuous on $[-1,1]$, as the choice $\delta = \epsilon$ reveals. The function f_2 is even equidifferentiable on $[-1,1]$ with $f_2' = 2f_1$, as the choice $\delta = \epsilon$ reveals. But neither of the functions assumes a maximum: all positive infinitely small numbers are exceeded, while no positive finite number is reached.

Theorem 104 (Maximum Theorem for Normal Functions) *Let f be a continuous order n normal function on the interval $[a,b] \subset \mathcal{R}$, and let a, b be real. Then f assumes a maximum inside the interval.*

Proof:
Like in the case of the intermediate value theorem, consider first the real function f_R obtained by restricting f to R. Since this function is continuous, it assumes a maximum m; let M be the set of points where this happens. We will show that m is also a maximum for f. Apparently it is assumed for all points in M, and we will see that each point at which the maximum is assumed is infinitely close to an element in M.

First let x be not infinitely close to a point in M; then $X = Re(x) \notin M$, and therefore $f_R(X) < m$. But since $|f_R(X) - f(x)| \leq |f_R(X) - f(X)| + |f(X)) - f(x)|$ is infinitely small and both $f_R(X)$ and m are real, we even have $f(x) < m$.

On the other hand, let x be infinitely close to an element of M, i.e. $X = Re(x) \in M$. If all n derivatives of f vanish at X, the continued function is constant on any infinitely small neighborhood of X, and hence $f(x) = m$. Otherwise, let $j \leq n$ be the number of the first non-vanishing derivative. Then according to the Taylor formula with remainder for R, j must be even because otherwise X could not yield a local maximum of f_R in R; furthermore, $f_R^{(j)}(X)$ is negative. But since in the infinitely small neighborhood of X, the dominating term of the continuation is $f_R^{(j)}(X) \cdot (x - X)^j$, we infer that $f(x) < m$. ∎

4.5 Mean Value Theorem and Taylor Theorem

Like in conventional calculus, we obtain the general mean value theorem from a special case of it, the theorem of Rolle. Similar to before, slightly stronger smoothness conditions than in R are required, and we present two versions of Rolle's theorem.

Theorem 105 (Rolle's Theorem) *Let f be a function on the finite interval $[a,b]$. Let f be equidifferentiable twice, and let $f'' \sim 1$ on $[a,b]$. Then there exists $\xi \in [a,b]$ with $f'(\xi) = 0$.*

Proof:
Consider the function f' and apply the intermediate value theorem. ∎

Theorem 106 (Rolle's Theorem for Normal Functions) *Let f be an order n normal function on $[a,b]$ with $n \geq 1$, and let $f(a) = f(b) = 0$. Then there is a $\xi \in [a,b]$ with $f'(\xi) = 0$.*

Proof:
Let x be a point in $[a,b]$ where f assumes a maximum, and let $\xi = Re(x)$. Then according to the last theorem, the real restriction f_R of f assumes a maximum at ξ, and thus $f_R'(\xi) = 0$. But since for real points, the derivatives of f and f_R agree, we even have $f'(\xi) = 0$, as desired. ∎

As mentioned before, Rolle's theorem conveniently allows to prove the mean value theorem. Because for the case of twice equidifferentiable functions, the conditions on the second derivative are somewhat cumbersome to phrase, we restrict ourselves here to the case of normal functions.

Theorem 107 (Mean Value Theorem for Normal Functions) *Let f, g be order n normal functions on the interval $[a,b]$ with $n \geq 1$, and let $g(b) \neq g(a)$, g' nonzero on $[a,b]$. Then there is a $\xi \in [a,b]$ such that*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = f'(\xi)/g'(\xi)$$

Proof:
Define the function h on $[a,b]$ as follows:

$$h(x) = f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} \cdot (g(x) - g(a))$$

Then clearly h is order n normal on $[a, b]$ with $n \geq 1$; also $h(a) = h(b) = 0$, and therefore there is a $\xi \in (a, b)$ with $h'(\xi) = 0$. Differentiating h and dividing by $g'(\xi)$ give the desired result. ∎

Again as in conventional analysis, we obtain the Taylor theorem from the mean value theorem:

Theorem 108 (Taylor Theorem) *Let f be an order k normal function on the interval $[x_0, x_0 + h]$, and let $n \leq k - 1$. Let g be an order $l \geq 1$ normal function in the same interval with nowhere vanishing derivative. Then we obtain:*

$$f(x_0 + h) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x_0)}{\nu!} \cdot h^\nu + \frac{g(x_0 + h) - g(x_0)}{g'(x_0 + \theta h)} \cdot \frac{(1 - \theta)^n}{n!} h^n f^{(n+1)}(x_0 + \theta h)$$

Proof:
Use the mean value theorem and substitute $x_1 - x_0 = h$ and $\xi = x_0 + \theta h$ to obtain:

$$f(x_0 + h) = f(x_0) + \frac{g(x_0 + h) - g(x_0)}{g'(x_0 + \theta h)} \cdot f'(x_0 + \theta h)$$

Use this formula on the function

$$F(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x)}{\nu!}(x_0 + h - x)^\nu$$

Then we obtain for $x = x_0$ and $x = x_0 + h$:

$$F(x_0) = \sum_{\nu=0}^{n} \frac{f^{(\nu)}(x)}{\nu!} h^\nu$$

$$F(x_0 + h) = f(x_0 + h).$$

By differentiation, we obtain

$$F'(x) = \sum_{\nu=0}^{n} \frac{f^{(\nu+1)}(x)}{\nu!}(x_0 + h - x)^\nu - \sum_{\nu=1}^{n} \frac{f^{(\nu)}(x)}{(\nu - 1)!}(x_0 + h - x)^{\nu-1}$$

$$= \frac{f^{(n+1)}(x)}{n!} \cdot (x_0 + h - x)^n$$

and finally

$$f(x_0+h)=F(x_0+h)=F(x_0)+\frac{g(x_0+h)-g(x_0)}{g'(x_0+\theta h)}F'(x_0+\theta h)$$

$$=\sum_{\nu=0}^{n}\frac{f^{(\nu)}(x_0)}{\nu!}h^\nu+\frac{g(x_0+h)-g(x_0)}{g'(x_0+\theta h)}\cdot\frac{(1-\theta)^n}{n!}h^n f^{(n+1)}(x_0+\theta h)$$

For different choices of $g(x)$, one obtains various forms of the remainder, similar to the situation in R ∎

4.6 Integration on $\mathcal{R}$

In this section we will define an integral extending the concept of the Riemann integral on R. In the theory of integration in R, in particular in connection with the fundamental theorem, it proved important that primitives to functions are unique up to constants. This is connected to the fact that if the derivative of a differentiable function vanishes everywhere, the function must be constant. In R, this is a direct consequence of the mean value theorem, and here we will proceed in the same way. In a previous example, we showed that continuity is not enough for this condition; but even equidifferentiability is not sufficient, as the next example shows.

Example 109 (Non-Constant Equidifferentiable Function with Vanishing Derivative) *Let $x\in[-1,1]$. Write $x=a_0+\sum_{\nu=1}^{\infty}a_\nu\cdot d^{r_\nu}$, and define $f:[-1,1]\to\mathcal{R}$ via*

$$f(x)=\sum_{\nu=1}^{\infty}a_\nu\cdot d^{3r_\nu}$$

Then f is equidifferentiable on $[-1,1]$ with $f'(x)=0$ there. To see this, note first that for all $a,b\in[-1,1]$ with $a+b\in[-1,1]$ we have that $f(a+b)=f(a)+f(b)$. Let now $h=a_0+\sum_{\nu=1}^{\infty}a_\nu\cdot d^{r_\nu}\neq 0$ in $[-1,1]$ be given. Then for any $x\in[-1,1]$, we have

$$\left|\frac{f(x+h)-f(x)}{h}\right|=\left|\frac{f(h)}{h}\right|=\left|\frac{\sum_{\nu=1}^{\infty}a_\nu\cdot d^{3r_\nu}}{a_{0+}\sum_{\nu=1}^{\infty}a_\nu\cdot d^{r_\nu}}\right|.$$

Let now $\epsilon>0$ be given, and choose $\delta=\epsilon$. Let $|h|<\delta$ be nonzero. If ϵ is finite, observe that the difference quotient is always infinitely small and *hence less than ϵ; if ϵ is infinitely small, observe that the difference quotient is of the same magnitude as h^2, which is infinitely much smaller than ϵ. So in both cases, the difference quotient does not differ from 0 by more than ϵ, implying $f'(x) = 0$.*

Remark 110 *We note that the situation is not specific to $\mathcal{R}$, but holds similarly in other nonarchimedean ordered fields because of the existence of nontrivial field automorphisms on all such fields.*

Definition 111 (Primitive for Order n Normal Functions) *Let f be piecewise an order n normal function on the finite interval $[a,b] \in \mathcal{R}$. We say the function F is a primitive to f on $[a,b]$ if F is piecewise order $(n+1)$ normal and satisfies*

$$F'(x) = f(x) \text{ for all } x \in [a,b]$$

Theorem 112 (Existence and Uniqueness of Primitives) *Let f be piecewise continuous order n normal function on the interval $[a,b] \in \mathcal{R}$. Then f has a primitive F on $[a,b]$. Furthermore, if F_1 and F_2 are two primitives to f on $[a,b]$, then*

$$F_1 - F_2 = const. \text{ on } [a,b].$$

Proof:
Let f be as stated, f_R a real function having f as piecewise continuation. Then f_R is piecewise continuous and n times differentiable. Define $F_R(x) = \int_{Re(a)}^{x} f_R(x')dx'$; then F_R is piecewise $(n+1)$ times differentiable with derivative f_R. Let F be its piecewise order $(n+1)$ continuation. Then on all real points $x \in [a,b]$, $F'(x) = f(x)$. Because of the uniqueness theorem for continuations (87), f and F' agree on $[a,b]$.

On the other hand, let $F = F_1 - F_2$ on $[a,b]$. Let F_R be the restriction of $\mathcal{R}](\{)$ to R. Then on R, we have $F_R' = 0$ on $[Re(a), Re(b)]$, and thus F_R is constant there. But then also $F_R^{(i)} = 0$, implying that its order n continuation F is constant on $[a,b]$. ∎

After these preparations, we are ready to introduce an integral for the class of piecewise continuous normal functions:

Definition 113 (Integral for Piecewise Continuous Normal Functions) *Let f be a piecewise continuous order n normal function on the finite* *interval* $[a, b]$. *Let* F *be a primitive of* f *on* $[a, b]$. *We define the integral of* f *over the interval* $[a, b]$ *as follows:*

$$\int_a^b f \, dx = F(b) - F(a)$$

We also say $\int_b^a f \, dx = -\int_a^b f \, dx$.

We note that the definition is unique, independent of the particular choice of the primitive according to the uniqueness theorem for primitives.

Besides integrals over finite ranges, we also define those over infinite ranges similar to how it is done in R:

Definition 114 (Infinite Integrals) *Let* f *be a piecewise continuous order* n *normal function on the interval* $[a, \infty]$. *Let* F *be a primitive, and for real* x, *let* $\bar{L} = \lim_{x \to \infty} F(x)$ *exist. Let* B *be positive and infinitely large in magnitude. Then we define the two integrals*

$$\int_a^\infty f \, dx = \int_a^B f \, dx = \bar{L} - F(a).$$

Similarly, let f *be a piecewise continuous order* n *normal function on the interval* $[-\infty, b]$, *let* F *be its primitive, and for real* x, *let* $\underline{L} = \lim_{x \to -\infty} F(x)$ *exist; let* A *be negative with infinitely large magnitude, and define*

$$\int_{-\infty}^b f \, dx = \int_A^b f \, dx = F(b) - \underline{L}.$$

Furthermore, if both of the above conditions are satisfied, define

$$\int_{-\infty}^\infty f \, dx = \int_A^B f \, dx = \bar{L} - \underline{L}.$$

We obtain the following simple properties of the integral of a normal function:

Theorem 115 (Properties of the Integral) *Let* f, g *be piecewise continuous order* n *normal functions on the interval* $[a, b] \in \mathcal{R}$, *let* $c \in [a, b]$, *let* $x_1, x_2, k \in \mathcal{R}$. *Then*

$$\int_a^b (k_1 \cdot f + k_2 \cdot g)\ dx = k_1 \cdot \int_a^b f\ dx\ +\ k_2 \cdot \int_a^b g\ dx$$

$$\int_a^b f\ dx = \int_a^c f\ dx\ + \int_c^b f\ dx$$

$$\int_a^b k\ dx = k \cdot (b - a)\ \textit{(area of rectangle)}$$

$$\textit{If } f(x) \leq g(x)\ \textit{ on } [a, b], \int_a^b f(x)dx \leq \int_a^b g(x)dx$$

The proof follows directly from the definition of the integral. Similar to the situation in R, also here the integral as a function of the right boundary is a primitive:

Theorem 116 (Fundamental Theorem for Normal Functions) *Let f be a piecewise continuous order n normal function on the finite interval $[a, b] \in \mathcal{R}$, $c \in \mathcal{R}$ with $a \leq c \leq b$. For $x \in [a, b]$ we define a function g as*

$$g(x) = \int_c^x f(x)\ dx.$$

Then g is a primitive of f.

The integral can readily be extended to scaled normal functions:

Definition 117 (Integral for Scaled Normal Functions) *Let $f = l_1 \circ f_n \circ l_2$ be a scaled piecewise normal function on the interval $[a, b]$ with linear transformations $l_1(x) = a_1 + b_1 \cdot x$ and $l_2(x) = a_2 + b_2 \cdot x$ and a piecewise normal function f_n as in the definition (92). Then we define the integral of the function f as :*

$$\int_a^b f(x)\ dx\ \ =\ \ (b - a) \cdot a_1 + \frac{b_1}{b_2} \cdot \int_{l_2(a)}^{l_2(b)} f_n\ dx$$

Apparently the integral for scaled normal functions is particularly useful for studying delta functions and other improper functions. One obtains:

Theorem 118 (Integral of Delta Functions) *Let δ be a delta function. Then for any at least finite a, we have*

$$\int_{-a}^{a} \delta(x)dx = 1.$$

Proof:
Since δ is a delta function, there is an order n normal function δ_n such that $\delta(x) = c\delta_n(cx)$ if $|cx|$ is not infinitely large, zero else. Using the rules about integration of scaled normal functions and infinite integrals, we have

$$\int_{-a}^{a} \delta(x)\ dx = \int_{-a}^{a} c\delta_n(cx)\ dx$$

$$= \frac{c}{c}\int_{-ca}^{ca} \delta_n(x)\ dx = \int_{-\infty}^{\infty} \delta_n(x)\ dx = 1$$

In a similar way, we obtain the fundamental theorem of delta functions:

Theorem 119 (Fundamental Theorem of Delta Functions) *Let δ be an order n delta function and f a continuous order n normal function on $[-a, a]$ with $a \in R$. Then if the integral exists, we have*

$$\int_{-a}^{a} f(x)\cdot\delta(x)dx =_0 f(0),$$

i.e. the integral agrees with $f(0)$ up to at most infinitely small error. Furthermore, if $n = 0$, the integral always exists, and exactly equals $f(0)$.

Proof:
Since f is a continuous order n normal function, for any infinitely small x, we have $f(x) = f(0) + \sum_{i=1}^{n} f^{(i)}(0)\cdot x^i/i!$. Since δ is a delta function, there is an order n normal function δ_n such that $\delta(x) = c\delta_n(cx)$ if $|cx|$ is not infinitely large, zero else. Using the rules about integration of scaled normal functions and infinite integrals, we obtain

$$\int_{-a}^{a} f(x)\cdot\delta(x)\ dx = \int_{-a}^{a} f(x)\cdot c\delta_n(cx)\ dx$$

$$= \frac{c}{c}\int_{-ca}^{ca} f(\frac{x}{c})\cdot\delta_n(x)\ dx = \int_{-\infty}^{\infty} f(\frac{x}{c})\cdot\delta_n(x)\ dx$$

$$= \int_{-\infty}^{\infty} \left[f(0) + \sum_{i=1}^{n} f^{(i)} \cdot \frac{x^i}{c^i i!} \right] \cdot \delta_n(x) \, dx.$$

Since all the $f^{(i)}$ are finite, we have $f(0) + \sum_{i=1}^{n} f^{(i)} \cdot \frac{x^i}{c^i i!} =_0 f(0)$, and the statement follows. In the special case of $n = 0$, the integral exactly equals $f(0)$. ∎

Acknowledgements

In the beginning stages, I benefitted from several comments and suggestions by Detlev Laugwitz, who pointed out Levi-Civita's early work and his own contributions to Levi-Civita fields in connection with a seminar given at his group in Darmstadt. Also, many comments from Norman Alling about his work on surreal numbers and several other questions of nonarchimedean number systems are appreciated.

For enjoyable discussions about his "Numbers and Games" and several other topics on the occasion of Kansas State University colloquia of him and me on the same day, I would like to thank John Conway. For making several useful comments about this work, and more generally for fostering my appreciation of the beauty of mathematics, I would like to thank my father Edgar Berz as well as Robert B. Burckel.

Several students helped with the translation of the German original [26], including Khodr Shamseddine, Volker Gebhardt, Stefan Schwede, and Achim Kasper. For financial support, I am grateful to the Alfred P. Sloan Foundation and the Studienstiftung des Deutschen Volkes.

References

[1] W. Rudin. *Real and Complex Analysis.* McGraw Hill, 1987.

[2] E. Hewitt and K. Stromberg. *Real and Abstract Analysis.* Springer, 1969.

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

[4] L. Fuchs. *Partially Ordered Algebraic Systems.* Pergamon Press, Addison Wesley, 1963.

[5] H.-D. Ebbinghaus et al. *Zahlen.* Springer, 1992.

[6] A. H. Lightstone and A. Robinson. *Nonarchimedean Fields and Asymptotic Expansions.* North Holland, New York, 1975.

[7] C. Schmieden and D. Laugwitz. Eine Erweiterung der Infinitesimalrechnung. *Mathematische Zeitschrift*, 69:1–39, 1958.

[8] D. Laugwitz. Eine Einführung der Delta-Funktionen. *Sitzungsberichte der Bayerischen Akademie der Wissenschaften*, 4:41, 1959.

[9] D. Laugwitz. Anwendungen unendlich kleiner Zahlen II. *Journal für die reine und angewandte Mathematik*, 208:22–34, 1961.

[10] D. Laugwitz. Anwendungen unendlich kleiner Zahlen I. *Journal für die reine und angewandte Mathematik*, 207:53–60, 1961.

[11] A. Robinson. Non-standard analysis. In *Proceedings Royal Academy Amsterdam, Series A*, volume 64, page 432, 1961.

[12] D. Laugwitz. Ein Weg zur Nonstandard-Analysis. *Jahresberichte der Deutschen Mathematischen Vereinigung*, 75:66–93, 1973.

[13] A. Robinson. *Non-Standard Analysis.* North-Holland, 1974.

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

[15] M. Davies. *Applied Nonstandard Analysis.* John Wiley and Sons, 1977.

[16] J. H. Conway. *On Numbers and Games.* North Holland, 1976.

[17] D. E. Knuth. *Surreal Numbers: How two ex-students turned on to pure mathematics and found total happiness.* Addison-Wesley, 1974.

[18] N. L. Alling. *Foundations of Analysis over Surreal Number Fields.* North Holland, 1987.

[19] H. Gonshor. *An Introduction to the Theory of Surreal Numbers.* Cambrindge University Press, 1986.

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

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

[22] A. Ostrowski. Untersuchungen zur arithmetischen Theorie der Körper. *Mathematische Zeitschrift*, 39:269–404, 1935.

[23] L. Neder. Modell einer Leibnizschen Differentialrechnung mit aktual unendlich kleinen Größen. *Mathematische Annalen*, 118:718–732, 1941-1943.

[24] D. Laugwitz. Eine nichtarchimedische Erweiterung angeordneter Körper. *Mathematische Nachrichten*, 37:225–236, 1968.

[25] 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.

[26] M. Berz. Analysis auf einer nichtarchimedischen Erweiterung der reellen Zahlen. Report (in German) MSUCL-753, Department of Physics, Michigan State University, 1990.

Index