A Consistent and Categorical Axiomatization of Differentiation Arithmetic Applicable to First and Higher Order Derivatives

Hend Dawood
Department of Mathematics, Faculty of Science,
Cairo University, Giza 12613, Egypt
Email: hend.dawood@sci.cu.edu.eg

Nefertiti Megahed
Department of Mathematics, Faculty of Science,
Cairo University, Giza 12613, Egypt
Email: nefertiti@sci.cu.edu.eg

Abstract. Differentiation arithmetic is a principal and accurate technique for the computational evaluation of derivatives of first and higher order. This article aims at recasting real differentiation arithmetic in a formalized theory of dyadic real differentiation numbers that provides a foundation for first and higher order automatic derivatives. After we set the stage by putting on a systematic basis certain fundamental notions of the algebra of differentiation numbers, we begin by setting up an axiomatic theory of real differentiation arithmetic, as a many-sorted extension of the theory of a continuously ordered field, and then establish the proofs for its consistency and categoricity. Next, we carefully construct the algebraic system of real differentiation arithmetic, deduce its fundamental properties, and prove that it constitutes a commutative unital ring. Furthermore, we describe briefly the extensionality of the system to an interval differentiation arithmetic and to an algebraically closed commutative ring of complex differentiation arithmetic. Finally, a word is said on machine realization of real differentiation arithmetic and its correctness, with an addendum on how to compute automatic derivatives of first and higher order.

1. INTRODUCTION

Many applications and algorithms in mathematics and scientific computing require the value of the derivative of a given function at some point. There are two main classes of differentiation algorithms that are used today to compute the derivative at a point: *symbolic* methods and *numerical* methods. The usual *symbolic* method requires the expression of the function from which the expression of its derivative is derived, using some known rules, and then the value of the derivative is evaluated at the given point by substituting in this derived expression. The downside of this method is that it is difficult to be computationally implemented and usually inefficient, especially when dealing with complex expressions and higher order derivatives. On the other hand, most *numerical* methods, which can be easily performed by a computer, depend on approximating the derivative using *finite differences*. That is, for a differentiable function f,

$$\text{first derivative} \cong \frac{f(x+h) - f(x)}{h},$$

for a *small* nonzero value of h. As h approaches zero, the derivative is better approximated. The drawback of this method is the difficulty in choosing the values for h. *Small* values could enlarge the rounding errors on the computer and *large* values will lead to a bad approximation of the derivative. This method could be improved by using relatively small values of h and fixing the round-off errors by using *interval enclosures* of the function f instead (see, e.g., [14], [18], [20], and [39]). However, with multivariate functions and higher derivatives, the complexity and round-off errors may inevitably increase. One approach that proved to be promising in coping with these challenges is *automatic differentiation*. Automatic differentiation (also called "algorithmic differentiation", "computational differentiation", or "differentiation arithmetic") is a principal and reliable technique for the *concurrent* computation of the values of a function and its derivative without any need for the *symbolic* expression of the derivative, only the expression or the algorithm of the function is needed. Automatic differentiation is neither numeric nor symbolic: In comparison to the ordinary numerical method of finite differences, automatic differentiation is 'theoretically' exact, and in contrast to symbolic differentiation, it is computationally cheap. The literature on automatic differentiation is very extensive and diverse. For further reading, see, e.g., [12], [16], [38], [41], [35], and [26]. Nowadays, for its accuracy and efficiency in the evaluation of derivatives of first and higher order, automatic differentiation is becoming a mainstream with numerous applications in mathematics and scientific computing. And, not surprisingly, there are many computational implementations of automatic differentiation. As instances, we may mention INTLAB, Fortran 95 AD Compiler, ADOL-C, Sollya, Tapenade, and InCLosure (see, e.g., [49], [40], [56], [8], [27], and [17]).

The problem of automatically computing the derivative of a given function is not new. The concept has its roots back to the dawning of the twentieth century. In 1916, Armin Elmendorf, an instructor in the department of mechanic, university of Wisconsin, designed a *mechanical differentiating machine* for drawing the differential or rate curve of any given curve, whether the latter be a curve plotted between two variables connected by an algebraic equation or an empirical curve obtained from experimental data [22]. Following the steps of Elmendorf, many mechanical differentiating machines were built based on the geometric concept of the derivative, using the proper levers, linkages, and gears. Perhaps the most

notable among these is the *mechanical differentiator* designed and built by C. P. Atkinson in 1951 [2]. Later, the idea seemed to be rediscovered in the works of Beda, Wengert, and Moore (see, e.g., [3], [57], and [37]). Modern developments of the subject appeared in [44] and [12]. Moore in [38] introduced a generalized notion under the title "*recursive evaluation of derivatives*" and used it to evaluate Taylor's coefficients (More extensive bibliographies[1] can be found in, e.g., [11] and [26]). Since then, applications and algorithms that employ automatic differentiation have been proliferating into many scientific disciplines.

Throughout this article, we shall understand by "differentiation arithmetic" (or "differentiation algebra") the fundamental mathematical structure underlying automatic differentiation as it is now implemented and practised. Although differentiation arithmetic is of great importance in both fundamental research and practical applications, no attempt has been made to put on a systematic basis its theory. This article thus aims at presenting a concrete and categorical account of a theory of dyadic real differentiation numbers that provides a foundation for first and higher order automatic derivatives. The importance of categoricity is that if a characterization of differentiation arithmetic is categorical, then it correctly describes, up to isomorphism, *every structure* of differentiation arithmetic. The role of categoricity in mathematics is best described by Corcoran in [10] and reworded by Shapiro in [50] as : "*a categorical axiomatization is the best one can do*". In this sense, the principal goal of this article is to present this "best" axiomatization. In order for this goal to be met, it is imperative to reformalize a number of fundamental algebraic and analytic notions in the symbolism of the theory to be presented, in a way that makes it possible to establish the *metatheoretic* statements about the theory. This is mainly done in section 2 in which we give an axiomatization of the theory $\mathsf{Th}_{d\mathcal{K}}$ of a differential continuously ordered field. In section 3, we set up a formalized theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of real differentiation arithmetic as a many-sorted extension of the theory $\mathsf{Th}_{d\mathcal{K}}$. We axiomatize the basic operations and relations of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$, deduce their fundamental properties, then we establish two important model-theoretic assertions concerning the categoricity and consistency of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$. In order for the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ to handle higher and partial derivatives, in section 4, we introduce the notion of differentiation-extensionality of a real function, characterize a differentiability predicate for real differentiation numbers, and establish the differentiability criterion thereof. In section 5, we carefully construct the algebraic system of real differentiation arithmetic, deduce its fundamental properties, and finally prove it constitutes a commutative unital ring. Finally, in section 6, we say a word on machine realization of real differentiation arithmetic and provide the proofs for its algorithmic correctness. The computational algorithms of section 6 are implemented in Lisp as a part of version 1.0 of InCLosure[2] [17]. The InCLosure commands to compute the results of the examples are described with comparison to the results obtained using Wolfram Mathematica [58]. In addition, an InCLosure input file and its corresponding output containing, respectively, the code and results of the examples are also available as a supplementary material to this article (see Supplementary Materials).

---

[1]The interested reader may also refer to the excellent web site *http://www.autodiff.org* which is a community portal to automatic differentiation.

[2]InCLosure (Interval enCLosure) is a language and environment for reliable scientific computing, which is coded entirely in Lisp. Latest version of InCLosure is available for free download via *https://doi.org/10.5281/zenodo.2702404*.

In practice, there are two modes of automatic differentiation, a *forward-mode* and a *reverse-mode*. Both modes, or combinations of them, are widely used and have numerous applications including sensitivity analysis, non-linear optimization, machine learning, robotics, computer graphics, automated theorem proving, and computer vision (see, e.g., [1], [16], [23], [34], [51], and [54]). Being an axiomatic extension of the theory of a continuously ordered field, our theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of real differentiation algebra provides a rigorous and unified mathematical foundation for the various approaches of automatic differentiation as it is now practised. To the best of the authors' knowledge, in almost all automatic differentiation literature, *Clifford's dual numbers* and *Grassmann numbers*[3] are usually 'borrowed' and 'reinvented' under different names as proposed algebraic foundations respectively for automatic first and higher-order derivatives. In this connection, we would like to remark that our theory differs in that it provides a foundation for higher and partial derivatives without the need for defining Clifford's or Grassmann algebras of higher dimensions. Noteworthy, moreover, is that with some basic alterations, the categorical system presented in this article can be extended analogously to handle automatic derivatives of *interval-valued* functions, and also can be seamlessly carried over to an *algebraically closed commutative ring* of *complex differentiation arithmetic*.

2. A Differential Continuously Ordered Field

In order to set up our formalized theory of real differentiation arithmetic in section 3, it is imperative to give in this section an axiomatization of the theory $\mathsf{Th}_{d\mathcal{K}}$ of a differential continuously ordered field. The intended interpretation of the system $\mathsf{Th}_{d\mathcal{K}}$ is the differential ordered field $\langle \mathbb{R}; +_{\mathbb{R}}, \times_{\mathbb{R}}; 0_{\mathbb{R}}, 1_{\mathbb{R}}; d \rangle$ of real numbers, where d is the differential operator for unary real-valued functions.

For the consistency and categoricity results to be provable, we need to characterize a differential operator and a differentiability criterion thereof in a *merely syntactical manner* (without any reference to real analysis or any interpretation). Thus, before turning to the axioms of $\mathsf{Th}_{d\mathcal{K}}$, it is imperative to reformalize a number of fundamental algebraic and analytic notions in the symbolism of the theory to be presented (For other formal approaches to these notions, see, e.g., [6], [15], [36], [48], and [52]).

Let f be an n-ary function (or, generally, a finitary relation). We shall denote by $\mathrm{dom}\,(f)$ and $\mathrm{ran}\,(f)$, respectively, the *domain* and *range* of f. Without loss of generality, in the remaining of this section and the succeeding sections, we consider only *unary* functions. The theory $\mathsf{Th}_{\mathcal{K}}$ of a continuously ordered field can be characterized as follows.

Definition 2.1 (Theory of Continuously Ordered Fields)**.** *The theory* $\mathsf{Th}_{\mathcal{K}}$ *of a continuously*[4] *ordered field (or, in short, a* co*-field) is the theory of a totally ordered field* $\mathfrak{K} = \langle \mathcal{K}; +_{\mathcal{K}}, \times_{\mathcal{K}}; 0_{\mathcal{K}}, 1_{\mathcal{K}}; \leq_{\mathcal{K}} \rangle$ *with the following axiom of continuity*

---

[3]A *Grassmann algebra* (*exterior algebra*, or a *superalgebra*) [25] was introduced by and named after the German linguist and mathematician Hermann Gunther Grassmann (1809–1877). A *Clifford's algebra*, named after the English mathematician and philosopher William Kingdon Clifford (1845–1879), is the special case of a one-dimensional Grassmann algebra [9]. Both algebras have widespread applications (see, e.g., [55]).

[4]A *continuously* ordered field is usually called a *complete* ordered field. Following Tarski (see, e.g.,[53]), we shall adopt the word "*continuously*" instead of "*complete*", as we use the word "*complete*" in a different logical sense.

$$\text{(ACO)} \ (\forall S \subseteq \mathcal{K})\,(\forall T \subseteq \mathcal{K}) \left(\begin{array}{c} (\forall x \in S)(\forall y \in T)(x <_\mathcal{K} y) \Rightarrow \\ (\exists z \in \mathcal{K})(\forall x \in S)(\forall y \in T)(x \neq z \wedge y \neq z \\ \Rightarrow x <_\mathcal{K} z \wedge z <_\mathcal{K} y) \end{array} \right).$$

Since the axiom of continuity is a *second-order* axiom, it follows that the theory of a co-field is a second-order theory and henceforth it is assumed that our formalism is higher-order. Let $\geq_\mathcal{K}$ denote the converse relation of $\leq_\mathcal{K}$, and let "$-_\mathcal{K}$" and "$^{-1}\mathcal{K}$" denote the unary $\mathcal{K}$-operations of *negation* and *reciprocal*, respectively. The binary $\mathcal{K}$-operations of *subtraction* and *division* are defined as usual. Hereafter, if confusion is unlikely, the subscript "$\mathcal{K}$" may be suppressed. Also, we shall denote by $\mathcal{K}_{\langle x \rangle}$ the set of all $\mathcal{K}$-valued unary functions. From now on, the letters f, g, and h (with or without subscripts) shall be employed as variable symbols to denote elements of the set $\mathcal{K}_{\langle x \rangle}$.

Next we extend the theory $\mathsf{Th}_\mathcal{K}$ of a co-field by introducing two $\mathcal{K}$-operators, namely *limit* ("lim") and *differential* ("d"), and one $\mathcal{K}$-predicate, namely the *differentiability* predicate ("diff"). Let $f \in \mathcal{K}_{\langle x \rangle}$, and let x and l be, respectively, a $\mathcal{K}$-variable symbol and a $\mathcal{K}$-constant symbol. The limit operator of $f(x)$ with respect to l, denoted $\lim_{x \to l} f(x)$, is defined as follows.

$$\lim_{x \to l} f(x) = L \Leftrightarrow (\forall \epsilon > 0)\,(\exists \delta > 0)\,(\forall x \in \mathrm{dom}\,(f))\,(0 < |x - l| < \delta \Rightarrow |f(x) - L| < \epsilon).$$

where the one-place operation symbol $|\cdot|$, called a $\mathcal{K}$-modulus (or absolute value), is defined as

$$(\forall y \in \mathcal{K})\,(\exists z \in \mathcal{K})\,(|y| = z \Leftrightarrow (0 \leq y \wedge z = y) \vee (\neg 0 \leq y \wedge z = -y)).$$

If there is no such $L \in \mathcal{K}$, we say that the limit of f at l does not exist in $\mathcal{K}$.

Definition 2.2 (Differential $\mathcal{K}$-Operator)**.** *Let $f \in \mathcal{K}_{\langle x \rangle}$, and let x and h be $\mathcal{K}$-variable symbols. For a nonnegative integer n, the n-differential operator of $f(x)$, denoted $d^n f(x)$, is defined recursively as follows.*

(i) $d^0 f(x) = f(x)$,

(ii) $d^1 f(x) = \lim\limits_{h \to 0_\mathcal{K}} \dfrac{f(x + h) - f(x)}{h} = d^1 d^0 f(x)$,

(iii) $n \geq 1 \Rightarrow d^n f(x) = d^1 d^{n-1} f(x)$.

It is obvious that if the limit in (ii) of the preceding definition exists, then the differential $d^n f(x)$ of f is in turn a $\mathcal{K}$-valued unary function. From now on, we shall usually write $d^n f$ and df for $d^n f(x)$ and $d^1 f(x)$, respectively.

The *differentiability* predicate is characterized in the following definition.

Definition 2.3 (Differentiability $\mathcal{K}$-Predicate)**.** *Let $f \in \mathcal{K}_{\langle x \rangle}$, and let $x_0 \in \mathrm{dom}\,(f)$ be a $\mathcal{K}$-constant symbol. For a nonnegative integer n, the n-differentiability predicate is a ternary predicate, $\mathrm{diff}^n(f, x_0)$, defined by*

$$\mathrm{diff}^n(f, x_0) \Leftrightarrow d^n f(x_0) \in \mathcal{K}.$$

If $\mathrm{diff}^n(f, x_0)$ is true, *we say that f is n-differentiable at x_0.*

Since for $x_0 \in \mathrm{dom}\,(f)$, $d^0 f(x_0) = f(x_0) \in \mathcal{K}$, it follows that $\mathrm{diff}^0(f, x_0)$ is always true and accordingly every $f \in \mathcal{K}_{\langle x \rangle}$ is 0-differentiable at $x_0 \in \mathrm{dom}\,(f)$. Obviously, if $\mathrm{diff}^n(f, x_0)$ is true, then $d^0 f(x_0), d^1 f(x_0), ..., d^n f(x_0)$ are in $\mathcal{K}$.

On the basis of the notions introduced so far, the theory $\mathsf{Th}_{d\mathcal{K}}$ of a differential co-field can then be axiomatized.

Definition 2.4 (Theory of a Differential co-Field). *Take $\sigma = \{+, \times; -, ^{-1}; 0, 1, \leq\}$ as a set of non-logical constants, and let $\mathsf{Th}_{\mathcal{K}}$ be the theory of a co-field $\mathfrak{K} = \langle \mathcal{K}; \sigma^{\mathfrak{K}} \rangle$. The theory $\mathsf{Th}_{d\mathcal{K}}$ of a differential co-field $\mathfrak{K}_d = \langle \mathcal{K}; \sigma^{\mathfrak{K}}; d \rangle$ is the deductive closure of $\mathsf{Th}_{\mathcal{K}}$ and the following two sentences.*

(i) $\left(\forall f, g \in \mathcal{K}_{\langle x \rangle}\right)\left(d\left(f + g\right) = df + dg\right),$

(ii) $\left(\forall f, g \in \mathcal{K}_{\langle x \rangle}\right)\left(d\left(f \times g\right) = f \times dg + g \times df\right).$

Consider the constant functions $f(x) = 0_{\mathcal{K}}$ and $g(x) = 1_{\mathcal{K}}$. By means of definition 2.2, it is clear that $d(0_{\mathcal{K}}) = d(1_{\mathcal{K}}) = 0_{\mathcal{K}}$. In general, for any $\mathcal{K}$-constant symbol c, $d(c) = 0_{\mathcal{K}}$ and $d(cx) = c$. In accordance with this, the set $\mathcal{K}$ can be defined to be $\mathcal{K} = \{f \in \mathcal{K}_{\langle x \rangle} | df = 0_{\mathcal{K}}\}$. By virtue of definition 2.2 and the sentences (i) and (ii) of definition 2.4, further properties of real differentiation are derivable analogously.

3. A Categorical Axiomatization of Real Differentiation Arithmetic

Having axiomatized the theory $\mathsf{Th}_{d\mathcal{K}}$ of a differential co-field in the preceding section, in this section we set up a formalized theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of real differentiation arithmetic as a many-sorted[5] extension of the theory $\mathsf{Th}_{d\mathcal{K}}$. Having at our disposal the notions formalized in section 2, we axiomatize the basic operations and relations of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ and deduce their fundamental properties. Furthermore, we establish two important model-theoretic assertions concerning the categoricity and consistency of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$.

First we define what a finitary *differentiation tuple* is.

Definition 3.1 (Differentiation Tuples). *Let $\mathfrak{K}_d = \langle \mathcal{K}; \sigma^{\mathfrak{K}}; d \rangle$ be a differential co-field, and for a nonnegative[6] integer n, let $\mathcal{K}^n$ be the n-th Cartesian power of $\mathcal{K}$. The set of all n-ary differentiation tuples over $\mathcal{K}$, with respect to a constant $x_0 \in \mathcal{K}$, is defined to be*

$$^n\mathbf{D}_{\mathcal{K}} = \left\{ \mathsf{f} \in \mathcal{K}^{n+1} | \left(\exists f \in \mathcal{K}_{\langle x \rangle}\right) \left(\begin{array}{c} \mathsf{f} = \left(d^0 f\left(x_0\right), d^1 f\left(x_0\right), ..., d^n f\left(x_0\right)\right) \\ \wedge\ x_0 \in \mathrm{dom}\left(f\right) \wedge \mathrm{diff}^n\left(f, x_0\right) \end{array} \right) \right\}.$$

That is, a differentiation tuple is an ordered tuple of $\mathcal{K}$-constants. For brevity henceforth, in differentiation tuples, we shall usually write f, $f^{(1)}$, and so forth to $f^{(n)}$ in place of, respectively, $d^0 f(x_0)$, $d^1 f(x_0)$, and so forth to $d^n f(x_0)$. In this work, we are concerned with *dyadic* differentiation tuples, that is tuples with $n = 1$. We shall use the name "*differentiation numbers*", or simply "D-*numbers*", for dyadic differentiation tuples. The set of differentiation numbers at some point x_0 shall be denoted by $\mathbf{D}_{\mathcal{K}}$. The letters f, g, and h, or equivalently $\left(f, f^{(1)}\right)_{x_0}$, $\left(g, g^{(1)}\right)_{x_0}$, and $\left(h, h^{(1)}\right)_{x_0}$, shall be employed as variable symbols to denote elements of $\mathbf{D}_{\mathcal{K}}$. Also, the letters a, b, and c, or equivalently $(a, 0_{\mathcal{K}})_{x_0}$, $(b, 0_{\mathcal{K}})_{x_0}$, and $(c, 0_{\mathcal{K}})_{x_0}$, shall be used to denote constants of $\mathbf{D}_{\mathcal{K}}$. In particular, we shall use $1_{\mathbf{D}}$ to denote the differentiation number $(1_{\mathcal{K}}, 0_{\mathcal{K}})_{x_0}$ and $0_{\mathbf{D}}$ to denote the differentiation

---

[5]Loosely speaking, a *many-sorted* structure is a structure with an arbitrary number of universe sets.

[6]For any set (empty or not) $\mathcal{S}$, the *zeroth* Cartesian power $\mathcal{S}^0$, of $\mathcal{S}$, is the singleton set $\{\varnothing\}$.

number $(0_\mathcal{K}, 0_\mathcal{K})_{x_0}$. Moreover, it is convenient for our purpose to define a proper subset of $\mathbf{D}_\mathcal{K}$ as

$$\mathbf{D}_{(\mathcal{K},0)} = \left\{ \mathsf{f} \in \mathbf{D}_\mathcal{K} | \mathsf{f} = (f, 0_\mathcal{K})_{x_0} \right\}.$$

We are now ready to pass to our formal characterization of the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of a differentiation algebra over a co-field.

Definition 3.2 (Theory of Differentiation Algebra). *Take $\sigma = \{+, \times; -, ^{-1}; 0, 1\}$ as a set of non-logical constants, let $\mathsf{Th}_{d\mathcal{K}}$ be the theory of a differential co-field $\mathfrak{K}_d = \langle \mathcal{K}; \sigma^\mathfrak{K}; d \rangle$, and let $\left(f, f^{(1)}\right)_{x_0}$, $\left(g, g^{(1)}\right)_{x_0}$, and $\left(h, h^{(1)}\right)_{x_0}$ be in $\mathbf{D}_\mathcal{K}$. A differentiation algebra (or, in short, a D-algebra) over $\mathfrak{K}_d$ is a two-sorted structure $\mathfrak{D}_\mathcal{K} = \langle \mathbf{D}_\mathcal{K}; \mathcal{K}; \sigma^{\mathfrak{D}\mathcal{K}} \rangle$. The theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of $\mathfrak{D}_\mathcal{K}$ is the deductive closure of $\mathsf{Th}_{d\mathcal{K}}$ and the following three sentences.*

(DA1) $\left(f, f^{(1)}\right)_{x_0} =_{\mathbf{D}} \left(g, g^{(1)}\right)_{x_0} \Leftrightarrow f(x_0) =_\mathcal{K} g(x_0) \wedge f^{(1)}(x_0) =_\mathcal{K} g^{(1)}(x_0),$

(DA2) $\circ \in \{+, \times\} \Rightarrow \left(f, f^{(1)}\right)_{x_0} \circ_{\mathbf{D}} \left(g, g^{(1)}\right)_{x_0} = \left(f \circ_\mathcal{K} g, (f \circ_\mathcal{K} g)^{(1)}\right)_{x_0},$

(DA3) $\diamond \in \{-\} \vee \left(\diamond \in \{^{-1}\} \wedge f(x_0) \neq 0_\mathcal{K}\right) \Rightarrow \diamond_{\mathbf{D}} \left(f, f^{(1)}\right)_{x_0} = \left(\diamond_\mathcal{K} f, (\diamond_\mathcal{K} f)^{(1)}\right)_{x_0}.$

The sentence (DA1) of the preceding definition is a definitional axiom characterizing the equality relation on $\mathbf{D}_\mathcal{K}$. The sentences (DA2) and (DA3) are finite axiom schemata prescribing, respectively, two binary operations, namely *addition* ("$+$") and *multiplication* ("$\times$"), and two unary operations, namely *negation* ("$-$") and *reciprocal* ("$^{-1}$"). The intended interpretation of the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ corresponds the classes "$\mathcal{K}$" and "$\mathbf{D}_\mathcal{K}$" to the classes "$\mathbb{R}$" and "$\mathbf{D}_\mathbb{R}$" (of *real numbers* and *real differentiation numbers*), respectively, and the symbols "$\circ_\mathcal{K}$", and "$\diamond_\mathcal{K}$" to the ordinary binary and unary operations for the reals. For simplicity of the language, hereafter, where no confusion is likely, the subscripts "$\mathbf{D}$", "$\mathcal{K}$", and "x_0" will be omitted. Also, we shall usually write the structure $\mathfrak{D}_\mathcal{K}$ as $\langle \mathbf{D}_\mathcal{K}; +_{\mathbf{D}}, \times_{\mathbf{D}}; 0_{\mathbf{D}}, 1_{\mathbf{D}} \rangle$, omitting the universe set $\mathcal{K}$.

By *Leibniz's rules* of differential sum and product, prescribed by (i) and (ii) in definition 2.4, definition 3.2 implies the following theorem.

Theorem 3.1 (Algebraic Operations of Differentiation Numbers). *Let $\left(f, f^{(1)}\right)$ and $\left(g, g^{(1)}\right)$ be two differentiation numbers. Then, the binary and unary differentiation operations are formulated as follows.*

(i) $\left(f, f^{(1)}\right) + \left(g, g^{(1)}\right) = \left(f + g, f^{(1)} + g^{(1)}\right),$

(ii) $\left(f, f^{(1)}\right) \times \left(g, g^{(1)}\right) = \left(f \times g, f^{(1)} \times g + f \times g^{(1)}\right),$

(iii) $-\left(f, f^{(1)}\right) = \left(-f, -(f)^{(1)}\right),$

(iv) $f(x_0) \neq 0_\mathbb{R} \Rightarrow \left(f, f^{(1)}\right)^{-1} = \left(f^{-1}, -f^{-2} \times f^{(1)}\right).$

By analogy with the ordinary language of arithmetic, differentiation subtraction and division are defined respectively as $\mathsf{f} - \mathsf{g} = \mathsf{f} + (-\mathsf{g})$ and $\mathsf{f} \div \mathsf{g} = \mathsf{f} \times \left(\mathsf{g}^{-1}\right)$. For the sake of perspicuity, and if confusion is unlikely, we may use the name "D-*operation*" ("D-*addition*", "D-*multiplication*", and so forth), or equivalently "*differentiation operation*" ("*differentiation*

addition", "*differentiation multiplication*", and so forth) to mean algebraic operations for real differentiation numbers (D-numbers).

We are now in a position to establish some important model-theoretic assertions about the theory $\mathbf{D}_{\mathcal{K}}$ of differentiation numbers. Assertions of isomorphism, categoricity, and consistency are the main questions of the metamathematical[7] investigation conducted below. Before discussing these questions, we deal first with some semantical preliminaries of particular importance for our purpose. An interpretation (a structure) $\mathfrak{M}$ is a *model* of a theory $\mathfrak{T}$, in symbols $\mathfrak{M} \models \mathfrak{T}$, iff every formula of $\mathfrak{T}$ is true for $\mathfrak{M}$. A theory $\mathfrak{T}$ is *categorical*[8], in symbols $\mathrm{Cat}\,(\mathfrak{T})$, iff any two models of $\mathfrak{T}$ are isomorphic. Otherwise $\mathfrak{T}$ is uncategorical (or disjunctive), in symbols $\mathrm{Uncat}\,(\mathfrak{T})$. A theory $\mathfrak{T}$ is a model-theoretically *consistent* (*m-consistent*) theory, in symbols $\mathrm{Con}\,(\mathfrak{T})$, iff $\mathfrak{T}$ has a model. Otherwise $\mathfrak{T}$ is m-inconsistent, in symbols $\mathrm{Inc}\,(\mathfrak{T})$.

On the basis of the above notions, we next prove two metatheorems[9] about the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$, concerning the *existence* and *uniqueness* of a differentiation algebra.

Theorem 3.2 (Existence of a Differentiation Algebra). *There exists at least one differentiation algebra.*

Proof. Since the theory $\mathsf{Th}_{\mathcal{K}}$ of a continuously ordered field has the model $\langle \mathbb{R}; +_{\mathbb{R}}, \times_{\mathbb{R}}; 0_{\mathbb{R}}, 1_{\mathbb{R}}; \leq_{\mathbb{R}} \rangle$ of real numbers, it follows that the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ has a model $\langle \mathbf{D}_{\mathbb{R}}; \mathbb{R}; +_{\mathbf{D}}, \times_{\mathbf{D}}; 0_{\mathbf{D}}, 1_{\mathbf{D}} \rangle$, and accordingly the existence assertion is established. $\square$

Theorem 3.3 (Categoricity of the Differentiation Number Theory). *The theory* $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ *of differentiation numbers is categorical.*

Proof. Take $\sigma = \{+, \times; -, {}^{-1}; 0, 1\}$ as a set of non-logical constants, and let $\mathfrak{D}_1 = \langle \mathcal{D}_1; \mathcal{K}_1; \sigma^{\mathfrak{D}1} \rangle$ and $\mathfrak{D}_2 = \langle \mathcal{D}_2; \mathcal{K}_2; \sigma^{\mathfrak{D}2} \rangle$ be two structures such that $\mathfrak{D}_1 \models \mathsf{Th}_{\mathbf{D}\mathcal{K}} \wedge \mathfrak{D}_2 \models \mathsf{Th}_{\mathbf{D}\mathcal{K}}$. Accordingly, $\langle \mathcal{K}_1; \sigma^{\mathcal{K}1} \rangle$ and $\langle \mathcal{K}_2; \sigma^{\mathcal{K}2} \rangle$ are two continuously ordered fields. A theory of continuously ordered fields is categorical, that is, there is one and up to isomorphism only one continuously ordered field[10]. $\langle \mathbb{R}; +_{\mathbb{R}}, \times_{\mathbb{R}}; 0_{\mathbb{R}}, 1_{\mathbb{R}}; \leq_{\mathbb{R}} \rangle$ is characterized, up to isomorphism, as the only continuously ordered field. Let $j : \mathcal{K}_1 \hookrightarrow \mathcal{K}_2$ be the isomorphism from $\mathcal{K}_1$ onto $\mathcal{K}_2$. We can then define $J : \mathcal{D}_1 \hookrightarrow \mathcal{D}_2$ by

$$J\,(\mathsf{f}) = J\left(\left(f, f^{(1)} \right) \right) = \left(j\,(f), j\left(f^{(1)} \right) \right),$$

---

[7]Metamathematics ("*metatheory*", "*epitheory*", or "*methodology of the deductive sciences*") is the theory concerned with reasoning *about* formalized languages and theories, and their interpretations. Thus, metamathematics takes formalized deductive theories as its objects of study (see, e.g., [13], [28], [33], [47], and [48]).

[8]The notion of categoricity originated in 1904 with Oswald Veblen. The terms "*categorical*" and "*disjunctive*" were suggested to Veblen by the American pragmatic philosopher John Dewey (see [29]).

[9]It must be noted that the symbols $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$, $\mathfrak{T}$, $\mathfrak{M}$, $\mathrm{Cat}\,(\mathfrak{T})$, and $\mathrm{Con}\,(\mathfrak{T})$ are not symbols of the object language, and they are not employed in constructing the sentences of our theory of differentiation arithmetic; rather, they are "*metalinguistic symbols*" (or "*metasymbols*") of an associated *metalanguage*. For each *object language* $\mathfrak{L}$ of a formalized theory $\mathfrak{T}$, there are *metalanguages* in which the metamathematics of the object theory can be symbolized. Our metalanguage is English equipped with some special symbols. A *metatheorem* is a true metalinguistic assertion *about* an object $\mathfrak{L}$-theory $\mathfrak{T}$.

[10]The proof of the existence assertion ("*there is a continuously ordered field*") is due to Dedekind, and the proof of the uniqueness assertion ("*a continuously ordered field is unique*") is due to Cantor (see, [21] and [5]). Dedekind and Cantor did not work with a first-order axiomatization of the linear continuum. Instead, they *informally* worked with a *second-order* formalization.

for all $\mathsf{f} = \left(f, f^{(1)}\right)$ in $\mathcal{D}_1$ where $f, f^{(1)} \in \mathcal{K}_1$. By means of definition 3.2 and that j is an isomorphism, we can show that J is an isomorphism from $\mathcal{D}_1$ onto $\mathcal{D}_2$ as follows.

- J is a bijection from $\mathcal{D}_1$ onto $\mathcal{D}_2$ since the range of J is $\mathcal{D}_2$ and

$$\left(\forall \mathsf{f} \in \mathcal{D}_1\right)\left(\forall \mathsf{g} \in \mathcal{D}_1\right)\left(J\left(\mathsf{f}\right) = J\left(\mathsf{g}\right) \Rightarrow \left(j\left(f\right), j\left(f^{(1)}\right)\right) = \left(j\left(g\right), j\left(g^{(1)}\right)\right) \Rightarrow \mathsf{f} = \mathsf{g}\right).$$

- J is function-preserving for "+" since we have

$$\begin{aligned}
J\left(\mathsf{f} + \mathsf{g}\right) &= \left(j\left(f + g\right), j\left(f^{(1)} + g^{(1)}\right)\right) \\
&= \left(j\left(f\right) + j\left(g\right), j\left(f^{(1)}\right) + j\left(g^{(1)}\right)\right) \\
&= \left(j\left(f\right), j\left(f^{(1)}\right)\right) + \left(j\left(g\right), j\left(g^{(1)}\right)\right) \\
&= J\left(\mathsf{f}\right) + J\left(\mathsf{g}\right).
\end{aligned}$$

- J is function-preserving for "$\times$" since we have

$$\begin{aligned}
J\left(\mathsf{f} \times \mathsf{g}\right) &= \left(j\left(f \times g\right), j\left(f^{(1)} \times g + f \times g^{(1)}\right)\right) \\
&= \left(j\left(f\right) \times j\left(g\right), j\left(f^{(1)}\right) \times j\left(g\right) + j\left(f\right) \times j\left(g^{(1)}\right)\right) \\
&= \left(j\left(f\right), j\left(f^{(1)}\right)\right) \times \left(j\left(g\right), j\left(g^{(1)}\right)\right) \\
&= J\left(\mathsf{f}\right) \times J\left(\mathsf{g}\right).
\end{aligned}$$

This establishes the categoricity of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ and completes the proof. $\qquad\square$

The theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$, thus, uniquely defines the algebra of differentiation numbers, and the structure $\langle \mathbf{D}_{\mathbb{R}}; \mathbb{R}; +_{\mathbf{D}}, \times_{\mathbf{D}}; 0_{\mathbf{D}}, 1_{\mathbf{D}}\rangle$ is, up to isomorphism, the only possible model of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$. In other words, according to theorem 3.3, the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ is the "best" axiomatization of real differentiation numbers, in the sense that it correctly describes, up to isomorphism, *every structure* of real differentiation arithmetic[11].

Now, the consistency of $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ can be easily established.

Theorem 3.4 (Consistency of the Differentiation Number Theory). *The theory* $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ *of differentiation numbers is consistent.*

Proof. By theorem 3.2, the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ is satisfiable by the model $\langle \mathbf{D}_{\mathbb{R}}; \mathbb{R}; +_{\mathbf{D}}, \times_{\mathbf{D}}; 0_{\mathbf{D}}, 1_{\mathbf{D}}\rangle$. $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ thus is *m-consistent*. $\qquad\square$

By virtue of the categoricity theorem for $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$, the properties of real numbers are naturally assumed priori. Thus, from now on, we shall speak of the structure $\langle \mathbf{D}_{\mathbb{R}}; \mathbb{R}; +_{\mathbf{D}}, \times_{\mathbf{D}}; 0_{\mathbf{D}}, 1_{\mathbf{D}}\rangle$ of *real* differentiation arithmetic and, accordingly, we shall write $\mathbb{R}$, $\mathbf{D}_{\mathbb{R}}$, and $\mathbb{R}_{\langle x\rangle}$ in place of $\mathcal{K}$, $\mathbf{D}_{\mathcal{K}}$, and $\mathcal{K}_{\langle x\rangle}$, respectively.

---

[11]The role of categoricity is of great importance in mathematics. Corcoran in [10] described this role by saying: "*the best possible characterization of an interpretation would be a characterization up to isomorphism*". This is reworded by Shapiro in [50] as: "*It is clear that if an axiomatization correctly describes a structure, then it also correctly describes any isomorphic structure. Thus, for the purpose of description, a categorical axiomatization is the best one can do*".

Let $\|\mathcal{S}\|$ denote the cardinality of a set $\mathcal{S}$. We close this section by establishing two theorems concerning the cardinality of the set $\mathbf{D}_{\mathbb{R}}$ of real differentiation numbers.

Theorem 3.5. *The sets $\mathbf{D}_{\mathbb{R}}$ and $\mathbb{R}^2$ are equal.*

Proof. Our proof reduces to showing that $\mathbf{D}_{\mathbb{R}} \subseteq \mathbb{R}^2$ and $\mathbb{R}^2 \subseteq \mathbf{D}_{\mathbb{R}}$. First, we establish that $\mathbf{D}_{\mathbb{R}} \subseteq \mathbb{R}^2$. Let $\mathsf{f} \in \mathbf{D}_{\mathbb{R}}$. Then we have $\mathsf{f} = \left(f, f^{(1)}\right)_{x_0} = \left(f(x_0), f^{(1)}(x_0)\right) \in \mathbb{R}^2$. Hence $\mathbf{D}_{\mathbb{R}} \subseteq \mathbb{R}^2$. It remains to show that $\mathbb{R}^2 \subseteq \mathbf{D}_{\mathbb{R}}$. Now let $(a, b) \in \mathbb{R}^2$. Then there exist $x_0, \beta \in \mathbb{R}$ such that $a = bx_0 + \beta$. We want to find $\left(f, f^{(1)}\right)_{x_0} \in \mathbf{D}_{\mathbb{R}}$ such that $a = f(x_0)$ and $b = f^{(1)}(x_0)$. Let $f(x) = bx + \beta$, then $f(x_0) = bx_0 + \beta = a$ and $f^{(1)}(x_0) = b$, that is $\left(f, f^{(1)}\right)_{x_0} = (a, b)$. Hence we deduce that for any $(a, b) \in \mathbb{R}^2$ we have $(a, b) \in \mathbf{D}_{\mathbb{R}}$. That is, $\mathbb{R}^2 \subseteq \mathbf{D}_{\mathbb{R}}$, which completes the proof. □

What is noteworthy here is that the constant x_0 is completely *arbitrary* in the definition of $\mathbf{D}_{\mathbb{R}}$, and thus replacing x_0 with another real constant, say x_1, will not make a difference in the proof of theorem 3.5. Accordingly, we will always have the same set $\mathbf{D}_{\mathbb{R}} = \mathbb{R}^2$.

Theorem 3.6 (Cardinality of Differentiation Numbers)**.** *The sets $\mathbf{D}_{\mathbb{R}}$ and $\mathbb{R}$ are equicardinal. In symbols, $\|\mathbf{D}_{\mathbb{R}}\| = \|\mathbb{R}\| = \mathfrak{c} = 2^{\aleph_0}$.*

Proof. By theorem 3.5, $\mathbf{D}_{\mathbb{R}} = \mathbb{R}^2$. Accordingly, $\|\mathbf{D}_{\mathbb{R}}\| = \|\mathbb{R}^2\|$. Since $\|\mathbb{R}^2\| = \|\mathbb{R}\|$, it follows that $\mathbf{D}_{\mathbb{R}}$ has the cardinality $\mathfrak{c} = 2^{\aleph_0}$ of the continuum. □

4. Differentiation-Extensionality of Real Functions: Higher and Partial Auto-Derivatives

In this section, we are to extend the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ in order that higher and partial derivatives can be manipulated without the need for defining *Grassmann algebras* of higher dimensions[12]. Toward doing this, we introduce the notion of *differentiation-extensionality* of a real function, characterize a *differentiability predicate* for real differentiation numbers, and establish the *differentiability criterion* thereof. In the first place, we want to define functions on real differentiation numbers beyond the rational ones. In order to be able to do this, we do need a *substitution rule* (or an *extension principle*). In other words, we need to extend functions of real numbers to functions of real D-numbers. On grounds of our characterization of real differentiation numbers, we have the following equivalence that gives a new reformulation of the differentiability predicate.

$$\mathrm{diff}^1\left(f, x_0\right) \Leftrightarrow d^1 f(x_0) \in \mathbb{R} \Leftrightarrow \left(f, f^{(1)}\right)_{x_0} \in \mathbf{D}_{\mathbb{R}}.$$

---

[12]Kalman in [32] considered tackling this point, using a very different approach, but his system turns out to be a *recursive* treatment of *n-tuples*. At first, he introduces the notation $f^{[n,m]}$, where n is the number of variables and m is the order of the derivative. This is followed on page 6 by the definition $f^{[1,3]} = \left(f(x), f^{(1)}(x), f^{(2)}(x), f^{(3)}(x)\right)$, which is clearly a *quadruple*, and the representation $[1, 3]$ is merely a shorthand. On page 17, he defines

$$f^{[n,m]}(x) = \begin{cases} f(x) & n = 0 \text{ or } m = 0, \\ \left(f^{[n-1,m]}(x), (\partial_n f)^{[n,m-1]}(x)\right) & \text{otherwise.} \end{cases}$$

The first component of the last pair is an $(m+1)$-tuple and the other component is an m-tuple. Accordingly, the pair is a shorthand for a $(2m+1)$-tuple, not a dyadic differentiation number. For a comparison of higher-order automatic differentiation methods, see, e.g., [42].

In accordance with the previous equivalence, we make this *extension principle* precise in the following definition.

Definition 4.1 (Differentiation Extension of a Real Function). *For $k \in \{1, ..., n\}$, let g_k be real functions differentiable at some $x_0 \in \mathrm{dom}\,(g_k)$, that is for each g_k there is $\mathbf{g}_k = \left(g_k\,(x_0)\,, g_k^{(1)}\,(x_0)\right) \in \mathbf{D}_{\mathbb{R}}$. Let $\mathcal{F}_{\mathbb{R}}\,(g_1, ..., g_n)$ be an n-place real-valued function of $g_1, ..., g_n$ which is differentiable at x_0. A differentiation extension of $\mathcal{F}_{\mathbb{R}}$ is an n-place $\mathbf{D}_{\mathbb{R}}$-valued function $\mathcal{F}_{\mathbf{D}_{\mathbb{R}}}\,(\mathbf{g}_1, ..., \mathbf{g}_n)$ defined to be*

$$\mathcal{F}_{\mathbf{D}_{\mathbb{R}}}\,(\mathbf{g}_1, ..., \mathbf{g}_n) = \left(\mathcal{F}_{\mathbb{R}}\,(g_1, ..., g_n)\,, (\mathcal{F}_{\mathbb{R}}\,(g_1, ..., g_n))^{(1)}\right),$$

and obtained from $\mathcal{F}_{\mathbb{R}}$ by replacing each real function symbol g_k, whenever it occurs in $\mathcal{F}_{\mathbb{R}}$, by the corresponding differentiation variable symbol $\mathbf{g}_k$.

Clearly, since $\mathcal{F}_{\mathbb{R}}$ is differentiable at x_0, its differentiation extension is in $\mathbf{D}_{\mathbb{R}}$. We thus understand by $\mathcal{F}_{\mathbb{R}}$ and $\mathcal{F}_{\mathbf{D}}$ two functions defined by the same rule but with different arguments; the former is a real-valued function and the latter is a $\mathbf{D}_{\mathbb{R}}$-valued function (*differentiation function*, or D-*function*). Analogously to rational real-valued functions, a *rational differentiation function* is a (multivariate) function obtained by means of a finite number of the differentiation arithmetic operations $\circ_{\mathbf{D}} \in \{+, \times\}$ and $\diamond_{\mathbf{D}} \in \{-, ^{-1}\}$. Hereafter, if the type of function is clear from its arguments, and if confusion is unlikely, we shall usually drop the subscripts $\mathbb{R}$ and $\mathbf{D}$. Thus, we may, for instance, write $\mathcal{F}\,(g_1, ..., g_n)$ and $\mathcal{F}\,(\mathbf{g}_1, ..., \mathbf{g}_n)$ for, respectively, a real-valued function and its differentiation extension.

To further illustrate, let, for example, $g_1\,(x) = \sin x$ and $g_2\,(x) = x^2$ be both differentiable real functions at some x_0, and let $\mathcal{F}_{\mathbb{R}}\,(g_1, g_2)$ be differentiable at x_0 such that

$$\mathcal{F}_{\mathbb{R}}\,(g_1, g_2) = g_1\,(x) + g_2\,(x) = \sin x + x^2.$$

The differentiation extension of $\mathcal{F}_{\mathbb{R}}$ is then

$$\begin{aligned}
\mathcal{F}_{\mathbf{D}_{\mathbb{R}}}\,(\mathbf{g}_1, \mathbf{g}_2) &= \left(g_1, g_1^{(1)}\right)_{x_0} + \left(g_2, g_2^{(1)}\right)_{x_0} \\
&= \left(\sin x, (\sin x)^{(1)}\right)_{x_0} + \left(x^2, \left(x^2\right)^{(1)}\right)_{x_0} \\
&= \left(\sin x + x^2, \left(\sin x + x^2\right)^{(1)}\right)_{x_0} \\
&= \left(\mathcal{F}_{\mathbb{R}}\,(g_1, g_2)\,, (\mathcal{F}_{\mathbb{R}}\,(g_1, g_2))^{(1)}\right)_{x_0}.
\end{aligned}$$

Further numerical examples, along with a description of the computational techniques, are presented in section 6.

By dint of the *extension principle* characterized in definition 4.1, one can now define fundamental functions of real differentiation numbers. As an instance, replacing $\mathcal{F}$ by the "sin" function, we obtain $\sin\left(g, g^{(1)}\right)_{x_0} = \left(\sin\,(g)\,, (\sin\,(g))^{(1)}\right)_{x_0}$.

Moreover, *higher order derivatives* can be handled based on the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of differentiation numbers. Again, on grounds of the *extension principle* formulated in definition 4.1, we can introduce the *differential operator* and the *differentiability predicate* for differentiation numbers.

Definition 4.2 (Differential Operator for Real D-Numbers). *Let* $\mathsf{g} = \left(g, g^{(1)}\right)$ *be in* $\mathbf{D}_{\mathbb{R}}$. *For a nonnegative integer* n, *the* n-*differential operator of* g, *denoted* $d^n\mathsf{g}$, *is defined recursively as follows.*

(i) $d^0\mathsf{g} = \mathsf{g}$,

(ii) $d^1\mathsf{g} = \left(dg, dg^{(1)}\right) = \left(g^{(1)}, g^{(2)}\right) = d^1 d^0\mathsf{g}$,

(iii) $n \geq 1 \Rightarrow d^n\mathsf{g} = \left(g^{(n)}, g^{(n+1)}\right) = d^1 d^{n-1}\mathsf{g}$.

Definition 4.3 (Differentiability Predicate for Real D-Numbers). *For a differentiation number* $\mathsf{g} = \left(g, g^{(1)}\right)_{x_0}$ *and a nonnegative integer* n, *the* n-*differentiability predicate,* $\mathrm{diff}^n(\mathsf{g})$, *is defined by*

$$\mathrm{diff}^n(\mathsf{g}) \Leftrightarrow d^n\mathsf{g} \in \mathbf{D}_{\mathbb{R}}.$$

In accordance with this definition, we have then the following theorem.

Theorem 4.1 (Differentiability Criterion for Real D-Numbers). *Let* $\mathsf{g} = \left(g, g^{(1)}\right)_{x_0}$ *be a differentiation number. Then, for* $n \geq 0$, $\mathrm{diff}^n(\mathsf{g}) \Leftrightarrow \mathrm{diff}^{n+1}(g, x_0)$.

Proof. It is clear that if the real function g is $(n+1)$-differentiable at x_0, then, for $n \geq 0$, $\left(g^{(n)}, g^{(n+1)}\right)_{x_0}$ is in turn an element of $\mathbf{D}_{\mathbb{R}}$, which, by definition 4.3, establishes the theorem. □

On the strength of the results obtained so far, we have the nice consequence that, within the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$, we can do differentiation arithmetic on the pairs $\left(g, g^{(1)}\right)$, $\left(g^{(1)}, g^{(2)}\right)$, ..., $\left(g^{(n)}, g^{(n+1)}\right)$ and we can implement automatic differentiations for higher order derivatives without the need for defining an arithmetic for n-tuples of the form $\left(g, g^{(1)}, ..., g^{(n)}\right)$, provided that we have included "seeds" for higher order derivatives in our machine implementation. In fact, for $n \geq 1$, the n-tuple construction of the theory of higher-order differentiation arithmetic can be exploited from the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ via writing, for example, $\left(g, d\left(g, g^{(1)}\right)\right)$ for $\left(g, g^{(1)}, g^{(2)}\right)$. We will discuss this further, along with some numerical examples, in section 6.

Finally, before we close this section, let us stress that restricting our functions to be *unary* is not a loss of generality. The reason for this is that an n-place function can be considered a family of n unary functions. To illustrate, consider the 2-place function $f(x, y) = x^2 + xy$. To *partially* differentiate $f(x, y)$ with respect to x at a point $(a, b) \in \mathbb{R}^2$, we need to differentiate the unary function $f_b(x) = x^2 + bx$, at the point a, and the corresponding differentiation number will be $\left(f_b, f_b^{(1)}\right)_a$.

Moreover, for a real closed interval X, let $f(x)$ be a differentiable real-valued function with $x \in X$. By employing *interval arithmetic* and utilizing the interval extensions $f(X)$ and $f^{(1)}(X)$ of, respectively, $f(x)$ and its derivative, the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of real differentiation arithmetic can be carried over to a theory of *interval differentiation arithmetic*, to automatically compute derivatives of interval-valued functions (For further details about interval arithmetic, see, e.g., [15], [19], [20], [30], [31], and [39]). Analogously, the theory is extensible to a differentiation arithmetic of *fuzzy-valued functions* (For more details on fuzzy calculus, see, e.g., [24] and [43]).

5. The Algebraic System of Real Differentiation Arithmetic

In this section, we further investigate the algebraic system of real differentiation numbers. We shall now make use of the part of the theory developed in sections 3 and 4 to delve deep into the algebraic properties of real differentiation arithmetic. In the sequel, by virtue of our definition of a differentiation number, the properties of real numbers are naturally assumed in advance.

The properties of addition and multiplication in $\mathbf{D}_\mathbb{R}$ are figured in the first two theorems of this section.

Theorem 5.1 (Additive Properties of Differentiation Numbers). *Differentiation addition satisfies the following properties.*

(i) *Identity for $+$.* $(\forall \mathsf{f} \in \mathbf{D}_\mathbb{R})\,(0_\mathbf{D} + \mathsf{f} = \mathsf{f} + 0_\mathbf{D} = \mathsf{f})$,

(ii) *Additive Inverses .* $(\forall \mathsf{f} \in \mathbf{D}_\mathbb{R})\,(\mathsf{f} + (-\mathsf{f}) = 0_\mathbf{D})$,

(iii) *Cancellativity for $+$.* $(\forall \mathsf{f}, \mathsf{g}, \mathsf{h} \in \mathbf{D}_\mathbb{R})\,(\mathsf{f} + \mathsf{h} = \mathsf{g} + \mathsf{h} \Rightarrow \mathsf{f} = \mathsf{g})$,

(iv) *Commutativity for $+$.* $(\forall \mathsf{f}, \mathsf{g} \in \mathbf{D}_\mathbb{R})\,(\mathsf{f} + \mathsf{g} = \mathsf{g} + \mathsf{f})$,

(v) *Associativity for $+$.* $(\forall \mathsf{f}, \mathsf{g}, \mathsf{h} \in \mathbf{D}_\mathbb{R})\,(\mathsf{f} + (\mathsf{g} + \mathsf{h}) = (\mathsf{f} + \mathsf{g}) + \mathsf{h})$.

Proof. The proof for (i) and (ii) follows directly from theorem 3.1. (iii) is immediate by the fact that every invertible element is cancellable. By commutativity and associativity of real addition, (iv) and (v) are easily derivable from theorem 3.1. $\qquad\square$

Theorem 5.2 (Multiplicative Properties of Differentiation Numbers). *Differentiation multiplication satisfies the following properties.*

(i) *Absorbing Element.* $(\forall \mathsf{f} \in \mathbf{D}_\mathbb{R})\,(0_\mathbf{D} \times \mathsf{f} = \mathsf{f} \times 0_\mathbf{D} = 0_\mathbf{D})$,

(ii) *Identity for $\times$.* $(\forall \mathsf{f} \in \mathbf{D}_\mathbb{R})\,(1_\mathbf{D} \times \mathsf{f} = \mathsf{f} \times 1_\mathbf{D} = \mathsf{f})$,

(iii) *Multiplicative Inverses.* $(\forall \mathsf{f} \in \mathbf{D}_\mathbb{R})\,\left(f(x_0) \neq 0 \Leftrightarrow \mathsf{f} \times (\mathsf{f}^{-1}) = 1_\mathbf{D}\right)$,

(iv) *Cancellativity for $\times$.* $(\forall \mathsf{f}, \mathsf{g} \in \mathbf{D}_\mathbb{R})\,((\mathsf{f} \times \mathsf{h} = \mathsf{g} \times \mathsf{h} \Rightarrow \mathsf{f} = \mathsf{g}) \Leftrightarrow h(x_0) \neq 0)$,

(v) *Commutativity for $\times$.* $(\forall \mathsf{f}, \mathsf{g} \in \mathbf{D}_\mathbb{R})\,(\mathsf{f} \times \mathsf{g} = \mathsf{g} \times \mathsf{f})$,

(vi) *Associativity for $\times$.* $(\forall \mathsf{f}, \mathsf{g}, \mathsf{h} \in \mathbf{D}_\mathbb{R})\,(\mathsf{f} \times (\mathsf{g} \times \mathsf{h}) = (\mathsf{f} \times \mathsf{g}) \times \mathsf{h})$.

Proof. The proof for (i) and (ii) is immediate from theorem 3.1. For (iii), let $\mathsf{f} \in \mathbf{D}_\mathbb{R}$ such that $f(x_0) \neq 0$. By theorem 3.1, and assuming the properties of real multiplication, $\mathsf{f} \times (\mathsf{f}^{-1}) = (f, f^{(1)}) \times (f^{-1}, -f^{-2} \times f^{(1)}) = (1, 0) = 1_\mathbf{D}$. The converse direction is derivable by assuming $\mathsf{f} \times (\mathsf{f}^{-1}) = 1_\mathbf{D}$. Then $\mathsf{f}^{-1} \in \mathbf{D}_\mathbb{R}$, and hence $f(x_0) \neq 0$. (iv) is entailed by the fact that an element is not cancellable for multiplication iff it is a zero divisor. By commutativity and associativity of real multiplication, (v) and (vi) are easily derivable from theorem 3.1. $\qquad\square$

Distributivity of differentiation arithmetic is established in the next theorem.

Theorem 5.3 (Distributivity in Differentiation Numbers). *Multiplication distributes over addition in differentiation arithmetic, that is*

$$(\forall \mathsf{f}, \mathsf{g}, \mathsf{h} \in \mathbf{D}_{\mathbb{R}}) \, (\mathsf{h} \times (\mathsf{f} + \mathsf{g}) = \mathsf{h} \times \mathsf{f} + \mathsf{h} \times \mathsf{g}) \, .$$

Proof. Let f, g, and h be any three differentiation numbers. According to theorem 3.1, we have

$$
\begin{aligned}
\mathsf{h} \times (\mathsf{f} + \mathsf{g}) &= \left(h \times (f + g) \, , h^{(1)} \times (f + g) + h \times \left(f^{(1)} + g^{(1)} \right) \right) \\
&= \left(h \times f + h \times g, \left(h^{(1)} \times f + h \times f^{(1)} \right) + \left(h^{(1)} \times g + h \times g^{(1)} \right) \right) \\
&= \left(h \times f, \left(h^{(1)} \times f + h \times f^{(1)} \right) \right) + \left(h \times g, \left(h^{(1)} \times g + h \times g^{(1)} \right) \right) \\
&= \mathsf{h} \times \mathsf{f} + \mathsf{h} \times \mathsf{g},
\end{aligned}
$$

and therefore multiplication distributes over addition in $\mathbf{D}_{\mathbb{R}}$. $\qquad \square$

With the preceding theorems at our disposal, we now pass to our main question concerning the algebraic system of differentiation arithmetic[13]. The following theorem clarifies an answer.

Theorem 5.4 (Commutative Ring of Differentiation Numbers). *The structure $\langle \mathbf{D}_{\mathbb{R}}; +_{\mathbf{D}}, \times_{\mathbf{D}}; 0_{\mathbf{D}}, 1_{\mathbf{D}} \rangle$ is a commutative unital ring with every element whose first component is not zero has a multiplicative inverse.*

Proof. By theorem 5.1, the additive structure $\langle \mathbf{D}_{\mathbb{R}}; +_{\mathbf{D}}; 0_{\mathbf{D}} \rangle$ is an abelian group. By theorem 5.2, the multiplicative structure $\langle \mathbf{D}_{\mathbb{R}}; \times_{\mathbf{D}}; 1_{\mathbf{D}} \rangle$ is a noncancellative abelian monoid. According to theorem 5.3, $\times_{\mathbf{D}}$ distributes over $+_{\mathbf{D}}$. Hence, $\mathbf{D}_{\mathbb{R}}$ forms a commutative unital ring. By theorem 5.2, every element whose first component is not zero has a multiplicative inverse. The proof of the theorem is therefore complete. $\qquad \square$

The structure of real differentiation arithmetic is not, though, an integral domain since $(0, \alpha) \times (0, \beta) = 0_{\mathbf{D}}$, and accordingly there are nonzero *zero divisors*.

Our last result of this section, concerning the *isomorphism theorem* for differentiation arithmetic, is figured in the following theorem.

Theorem 5.5 (Isomorphism Theorem for Differentiation Numbers). *The structure $\langle \mathbf{D}_{(\mathbb{R}, 0)}; +_{\mathbf{D}}, \times_{\mathbf{D}} \rangle$ is isomorphically equivalent to the field $\langle \mathbb{R}; +_{\mathbb{R}}, \times_{\mathbb{R}} \rangle$ of real numbers.*

Proof. Let $\iota : \mathbb{R} \hookrightarrow \mathbf{D}_{(\mathbb{R}, 0)}$ be the mapping from $\mathbb{R}$ to $\mathbf{D}_{(\mathbb{R}, 0)}$ given by $\iota (\alpha) = (\alpha, 0)$. It is easy to show that ι is an isomorphism. $\qquad \square$

That is, up to isomorphism, the sets $\mathbb{R}$ and $\mathbf{D}_{(\mathbb{R}, 0)}$ are equivalent, and accordingly the subalgebra $\langle \mathbf{D}_{(\mathbb{R}, 0)}; +_{\mathbf{D}}, \times_{\mathbf{D}} \rangle$ is a field.

---

[13]Rall, on page 9 of [45], stated without a proof that "There are *no divisors of zero*. A mathematical system with these properties is called an *integral domain*. [...] It is important that [the structure] is an integral domain because this means that the same results will be obtained independently of the order in which equivalent sequences of arithmetic operations are performed", which is an incorrect statement. Later in his article [46], he provided a correct statement about his structure and gave numerical instances, but again without a proof. In the present article we prove generalized statements for *every structure* of differentiation arithmetic in a *categorical* sense.

It is noteworthy at this point to mention that extensionality of the theory to *complex numbers* is readily possible with nice consequences. To illustrate this, we introduce a definition: a commutative ring $\mathfrak{R}$ is said to be "algebraically closed" if every finite system of polynomial equations in one or more variables with coefficients in $\mathfrak{R}$ which has a solution in some (commutative) extension of $\mathfrak{R}$ already has a solution in $\mathfrak{R}$ [7]. In fact, going a little further by starting with a categorical characterization of the field of complex numbers[14], in place of the continuously ordered field of the reals, we get a categorical theory of an *algebraically closed commutative ring* of *complex differentiation numbers*.

6. MACHINE REALIZATION OF REAL DIFFERENTIATION ARITHMETIC

This final section is devoted to discussing the fundamentals of machine implementation of automatic differentiation. The two modes of automatic differentiation are both realizable in the framework of our theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of differentiation algebra. We give here a *mathematical* flavor of the forward-mode. With some basic alterations, the reverse-mode can be realized as well. The computational algorithm of this section is implemented in Lisp as a part of version 1.0 of InCLosure [17] (see Supplementary Materials). After prescribing the algorithm and establishing its correctness, to offer insights of the theory we first give a simple example that can be worked by hand, then we introduce a more sophisticated example that will be computed to an arbitrary precision using InCLosure, and finally, we present a brief account of how to compute auto-derivatives of higher order using *dyadic* differentiation arithmetic. The InCLosure commands to compute the results of the examples are described with comparison to the results obtained using Wolfram Mathematica [58].

Toward computing the differentiation number of a differentiable real function at some point x_0, we start with a (minimal) set of symbolic expressions of differentiable real functions and their derivatives which works as *seeds* for performing the computation.

Definition 6.1 (Differentiation Seeds). *Let $f(x), f^{(1)}(x) \in \mathbb{R}_{\langle x \rangle}$ be respectively the symbolic expressions of a differentiable function and its derivative. The set $\mathbf{P}_{\langle x \rangle}$ of differentiation seeds is a finite set of ordered pairs $\left(f(x), f^{(1)}(x) \right)$. We define $\mathbf{P}_{\langle x \rangle}$ to be the union of the following sets.*

Real powers: $\left(ax^b, abx^{b-1} \right)$ *with a and b are real constants, for instance, $(a, 0)$, $(x, 1)$, $\left(x^2, 2x \right)$, $(\sqrt{x}, 1/(2\sqrt{x}))$, and so forth.*

Logarithms and exponentials: $(\ln(x), 1/x)$, (e^x, e^x), $(\log_a(x), 1/(x \ln(a)))$, *and* $(a^x, a^x \ln(a))$ *where $a \neq 1$ is a positive real constant.*

Trigonometrics and inverse trigonometrics: $(\sin(x), \cos(x))$, $(\cos(x), -\sin(x))$, $\left(\sin^{-1}(x), 1/\sqrt{1-x^2} \right)$, *and so forth.*

Absolute values: $(|x|, |x|/x)$.

We understand by function seeds a set $\mathbf{S}_{\langle x \rangle} = \left\{ f \in \mathbb{R}_{\langle x \rangle} | \left(f, f^{(1)} \right) \in \mathbf{P}_{\langle x \rangle} \right\}$.

It is imperative here to mention that differentiation seeds are not elements of $\mathbf{D}_{\mathbb{R}}$, as we do not have a numerical value, x_0, for the variable symbol x. Rather, $\mathbf{P}_{\langle x \rangle}$ is a finite set of

---

[14]A categorical axiomatization of complex numbers was presented by Bosch and Krajkiewicz in [4].

symbolic expressions of functions $f(x)$ and their derivatives $f^{(1)}(x)$. In practice, several variants of this definition are possible depending upon the choice of the *minimal* set $\mathbf{S}_{\langle x \rangle}$ of function seeds.

On the basis of definition 6.1, the following two theorems can be proved.

Theorem 6.1 (Compositions of Seeds). *Let $\mathbf{C}_{\langle x \rangle}$ be the set of all $f \in \mathbb{R}_{\langle x \rangle}$ such that f is a finite composition of seeds $g_1, ..., g_n$ from $\mathbf{S}_{\langle x \rangle}$. If an $f \in \mathbf{C}_{\langle x \rangle}$ is differentiable at some $x_0 \in \mathrm{dom}(f)$, then f is automatically differentiable at x_0; in other words, there are $\left(g_1, g_1^{(1)}\right), ..., \left(g_n, g_n^{(1)}\right) \in \mathbf{D}_{\mathbb{R}}$ from which $\left(f, f^{(1)}\right)_{x_0}$ is computable.*

Proof. The theorem is immediate by a finite application of the chain rule. $\square$

Obviously, if $g, h \in \mathbf{S}_{\langle x \rangle}$ and $f(x) = g \circ h(x) = g(h(x))$ is a composition of g and h, then

$$\left(f, f^{(1)}\right)_{x_0} = \left(g(h(x)), h^{(1)}(x) g^{(1)}(h(x))\right)_{x_0}$$
$$= \left(g(h(x_0)), h^{(1)}(x_0) g^{(1)}(h(x_0))\right).$$

For f to be differentiable at x_0, h and g should be respectively differentiable at x_0 and $h(x_0)$. Thus, there are $\left(h, h^{(1)}\right)_{x_0}, \left(g, g^{(1)}\right)_{h(x_0)} \in \mathbf{D}_{\mathbb{R}}$ from which $\left(f, f^{(1)}\right)_{x_0}$ is computable.

As examples of compositions, we can mention $\left(\sin(x^2), 2x\cos(x^2)\right)_{x_0}$ and $(e^{\sqrt{x}}, e^{\sqrt{x}}/2\sqrt{x})_{x_0}$.

Theorem 6.2 (Algebraic Combinations of Seeds). *Let $\mathbf{A}_{\langle x \rangle}$ be the set of all $f \in \mathbb{R}_{\langle x \rangle}$ such that f is a rational function of elements $g_1, ..., g_n$ from $\mathbf{C}_{\langle x \rangle}$. If an $f \in \mathbf{A}_{\langle x \rangle}$ is differentiable at some $x_0 \in \mathrm{dom}(f)$, then f is automatically differentiable at x_0.*

Proof. Let $f \in \mathbf{A}_{\langle x \rangle}$. By hypothesis, f is a rational real function $\mathcal{F}_{\mathbb{R}}(g_1, ..., g_i, ..., g_n)$, where $g_i \in \mathbf{C}_{\langle x \rangle}$. By theorem 6.1, each $g_i \in \mathbf{C}_{\langle x \rangle}$ is automatically differentiable, that is, $\mathbf{g}_i = \left(g_i, g_i^{(1)}\right) \in \mathbf{D}_{\mathbb{R}}$. Then, by definition 4.1, there is a rational differentiation function $\mathcal{F}_{\mathbf{D}}$ such that

$$\left(f, f^{(1)}\right) = \left(\mathcal{F}_{\mathbb{R}}(g_1, ..., g_n), \mathcal{F}_{\mathbb{R}}^{(1)}(g_1, ..., g_n)\right)$$
$$= \mathcal{F}_{\mathbf{D}}(\mathbf{g}_1, ..., \mathbf{g}_n) \in \mathbf{D}_{\mathbb{R}},$$

and the theorem follows. $\square$

To illustrate this, let $f \in \mathbf{A}_{\langle x \rangle}$ be defined by $f(x) = \ln(x^2) + \cos(x^2) = g_1(x) + g_2(x) = \mathcal{F}_{\mathbb{R}}(g_1, g_2)$, where $g_1, g_2 \in \mathbf{C}_{\langle x \rangle}$. Then f is automatically differentiable and

$$\left(f, f^{(1)}\right)_{x_0} = \left(\mathcal{F}_{\mathbb{R}}(g_1, g_2), \mathcal{F}_{\mathbb{R}}^{(1)}(g_1, g_2)\right)_{x_0}$$
$$= \mathcal{F}_{\mathbf{D}}(\mathbf{g}_1, \mathbf{g}_2)$$
$$= \mathbf{g}_1 + \mathbf{g}_2$$
$$= \left(\ln(x^2), 2/x\right)_{x_0} + \left(\cos(x^2), -2x\sin(x^2)\right)_{x_0}$$
$$= \left(\ln(x^2) + \cos(x^2), (2/x) - 2x\sin(x^2)\right)_{x_0} \in \mathbf{D}_{\mathbb{R}}.$$

In order to make full use of the results obtained so far, we would like to get the maximal set of *automatically differentiable functions* with respect to our definition of the set $\mathbf{S}_{\langle x \rangle}$ of seeds. We therefore make the following definition.

Definition 6.2 (Whole Combinations of Seeds)**.** *Let* $\mathbf{W}_{\langle x \rangle}$ *be the set defined by the following recursion scheme.*

(i) $\left(\forall f \in \mathbb{R}_{\langle x \rangle} \right) \left(f \in \mathbf{A}_{\langle x \rangle} \Rightarrow f \in \mathbf{W}_{\langle x \rangle} \right),$

(ii) $\left(\forall f \in \mathbb{R}_{\langle x \rangle} \right) \left(f \text{ is a finite composition of elements of } \mathbf{W}_{\langle x \rangle} \Rightarrow f \in \mathbf{W}_{\langle x \rangle} \right),$

(iii) $\left(\forall f \in \mathbb{R}_{\langle x \rangle} \right) \left(f \text{ is a rational function of elements of } \mathbf{W}_{\langle x \rangle} \Rightarrow f \in \mathbf{W}_{\langle x \rangle} \right).$

In consequence of this definition, by recursive use of theorems 6.1 and 6.2, the following theorem is easily proved.

Theorem 6.3 (Auto-Differentiable Functions)**.** *Let* f *be in* $\mathbf{W}_{\langle x \rangle}$*. If* f *is differentiable at some* $x_0 \in \mathrm{dom}\,(f)$*, then* f *is automatically differentiable at* x_0*.*

Since the hypotheses of theorems 6.1, 6.2, and 6.3 imply the inclusion

$$\mathbf{S}_{\langle x \rangle} \subset \mathbf{C}_{\langle x \rangle} \subset \mathbf{A}_{\langle x \rangle} \subset \mathbf{W}_{\langle x \rangle} \subset \mathbb{R}_{\langle x \rangle},$$

it is then clear that whether a function is an element of the set $\mathbf{W}_{\langle x \rangle}$ of automatically differentiable functions or not is dependent upon the choice of the finite set $\mathbf{S}_{\langle x \rangle}$ of seeds.

Now with the aid of the notions prescribed in this section and by virtue of the *extension principle* (definition 4.1), an algorithm for real automatic differentiation can be sketched as follows.

Algorithm 6.1 (Computing Real Differentiation Numbers)**.** *Given a real-valued function* $f\,(x)$ *and a real constant* x_0*, a real differentiation number* $\left(f, f^{(1)} \right)_{x_0}$*, if any, is computed through the following steps.*

Input: $f\,(x)$, x_0.

Step 1: $f \in \mathbf{S}_{\langle x \rangle}$.

   : *If* $\left(f, f^{(1)} \right)_{x_0} \in \mathbf{D}_{\mathbb{R}}$*, go to Output; Else go to Error.*

Step 2: $f \in \mathbf{W}_{\langle x \rangle}$.

   : *Decompose* f *to seeds* $g_1, ..., g_n$ *in* $\mathbf{S}_{\langle x \rangle}$,

   : *For each seed* g_k *in* f*, if* $\left(g_k, g_k^{(1)} \right) \in \mathbf{D}_{\mathbb{R}}$*, compute* $\left(f, f^{(1)} \right)_{x_0}$ *by applying the chain rule or performing differentiation arithmetic on the differentiation numbers* $\left(g_k, g_k^{(1)} \right)$*; Else go to Error.*

   : *If* $\left(f, f^{(1)} \right)_{x_0} \in \mathbf{D}_{\mathbb{R}}$*, go to Output; Else go to Error.*

Error: *Return Error and Exit.*

Output: *Return* $\left(f, f^{(1)} \right)_{x_0}$ *and Exit.*

The correctness of algorithm 6.1 is established by theorems 6.1, 6.2, and 6.3. The truth value of the differentiability predicate $\text{diff}^1\left(f, x_0\right)$ is obtained from the fact that

$$\text{diff}^1\left(f, x_0\right) \Leftrightarrow d^1 f\left(x_0\right) \in \mathbb{R} \Leftrightarrow \left(f, f^{(1)}\right)_{x_0} \in \mathbf{D}_{\mathbb{R}}.$$

To sum up, *first-order* automatic differentiation can be viewed as

$$f\left(x\right), x_0 \overset{\text{Input}}{\Longrightarrow} \begin{bmatrix} \text{Differentiation Seeds} \\ \text{Chain Rule} \\ \text{Real Differentiation Algebra} \end{bmatrix} \overset{\text{Output}}{\Longrightarrow} \left(f, f^{(1)}\right)_{x_0}.$$

To further illustrate, it is sufficient to give a simple example that can be worked by hand.

Example 6.1 (Differentiation Number for a Rational Function). *Consider the function*

$$f\left(x\right) = \frac{2\left(x+1\right)}{x+3} \text{ with } x \neq -3.$$

We want to compute the real differentiation number $\left(f, f^{(1)}\right)_3$. *Applying algorithm 6.1 we get*

$$\left(f, f^{(1)}\right)_3 = \frac{(2,0)_3 \times ((3,1)_3 + (1,0)_3)}{(3,1)_3 + (3,0)_3} = \frac{(2,0)_3 \times (4,1)_3}{(6,1)_3} = \frac{(8,2)_3}{(6,1)_3} = (4/3, 1/9)_3,$$

hence the value of f *at* 3 *is* $4/3$ *and the value of its first derivative* $f^{(1)}$ *at* 3 *is* $1/9$.

Now we turn to a more sophisticated example whose result will be computed to an *arbitrary precision* using InCLosure, and then compared to the result obtained using the ordinary symbolic methods of Wolfram Mathematica evaluated in floating-point machine numbers.

Example 6.2 (Real Automatic Differentiation in InCLosure). *Consider the real function*

$$f\left(x\right) = \sin\left(e^{\left(x\left(\sin\left(\cos\left(\tan\left(\sec\left(\csc\left(e^{\left(0.5^x + \frac{7e^{-x^2}}{8}\right)}\right)\right)\right)\right)\right)\right)}\right),$$

InCLosure provides an arbitrary precision with the default precision is 20 *significant digits. To compute the real differentiation number* $\left(f, f^{(1)}\right)_2$, *for the function* f *at the point* $x = 2$, *to the default precision, we write the following InCLosure command.*

```
ADReal "sin(e^(x*(sin(cos(tan(sec(csc(e^(0.5^x+e^-x^2/8*7)))))))))"
 "x=2" ; Default precision 20
```

This will result in $(0.25432188631704068941, 0.31246117673851504682)$. *To compute the result to a higher precision, say for example to* 50 *significant digits, we just add '50' as a last parameter as follows.*

```
ADReal "sin(e^(x*(sin(cos(tan(sec(csc(e^(0.5^x+e^-x^2/8*7)))))))))"
 "x=2" 50 ; With precision 50
```

which will result in

$$\left(\begin{array}{c} 0.25432188631704068940938563131188319748318849000912, \\ 0.3124611767385150468218334269699529824673549079415 \end{array} \right).$$

Now, let us compare these results to those obtained using Wolfram Mathematica. *For the values of the function and its derivative at $x = 2$, we write the following two free-form* Mathematica *commands*

```
sin(e^(x*(sin(cos(tan(sec(csc(e^(0.5^x+e^-x^2/8*7))))))))) at x=2

derivative sin(e^(x*(sin(cos(tan(sec(csc(e^(0.5^x+e^-x^2/8*7)))))))))
 at x=2
```

which will give the values 0.254322 and 0.312461 for $f(2)$ and $f^{(1)}(2)$, respectively.

Mathematica can certainly be set to use arbitrarily many digits, but noteworthy is that the symbolic capability of Mathematica gives accurate results only if the decimal inputs (e.g., '0.5' in the previous example) are written as rationals. This is a default feature of InCLosure. The philosophy of InCLosure is to be more to accuracy and guaranteed enclosures than to performance. Toward achieving this, InCLosure deploys the '*exact rationals*' property of Lisp in such a way that all computations on inputted and intermediate values are done by default in 'exact rationals' not in floating-point machine numbers. That is, more accurate results are always readily available and rounding is applied only on the final decimal result according to how many digits the user wants to display. We provide a supplementary text file containing InCLosure results to '10, 000' significant digits. Here, the result is abridged to 50 digits for the shortage of space.

Finally, we describe how *higher-order* automatic derivatives can be handled based on the theory $Th_{\mathbf{D}\mathcal{K}}$ of *dyadic* differentiation numbers. Extending the set $\mathbf{S}_{\langle x \rangle}$ of *function seeds* by including the symbolic expressions of $f^{(0)}, f^{(1)}, ..., f^{(n)}, f^{(n+1)}$, for an arbitrary n, we can do differentiation arithmetic on the pairs $\left(f, f^{(1)} \right), \left(f^{(1)}, f^{(2)} \right), ..., \left(f^{(n)}, f^{(n+1)} \right)$ and we can implement automatic differentiation for higher order derivatives without the need for neither defining an arithmetic for n-tuples of the form $\left(g, g^{(1)}, ..., g^{(n)} \right)$ nor using *Grassmann algebras* of higher dimensions.

As mentioned above, first-order automatic differentiation is "*differentiation arithmetic equipped with seeds and the chain rule*". To be able to compute automatic derivatives of higher order, we need to include also *Leibniz' product rule*. To illustrate, consider the real function

$$f(x) = \sin\left(x^2\right) + \ln(x).$$

We want to compute the differentiation numbers $\left(f^{(1)}, f^{(2)} \right)$ and $\left(f^{(2)}, f^{(3)} \right)$ at some real number x_0. As mentioned above**,** the set of symbolic *function seeds* should include the derivatives up to the third order. So, for $f_1(x) = \sin(x)$, $f_2(x) = x^2$, and $f_3(x) = \ln(x)$, we will have respectively the following differentiation seeds

$$(\sin(x), \cos(x)), (\cos(x), -\sin(x)), (-\sin(x), -\cos(x));$$
$$\left(x^2, 2x\right), (2x, 2), (2, 0);$$
$$\left(\ln(x), \frac{1}{x}\right), \left(\frac{1}{x}, -\frac{1}{x^2}\right), \left(-\frac{1}{x^2}, \frac{2}{x^3}\right).$$

Now to compute $d\left(f, f^{(1)}\right) = \left(f^{(1)}, f^{(2)}\right)$ at x_0, we have the following *dyadic* differentiation numbers for respectively $\sin\left(x^2\right)$ and $\ln\left(x\right)$

$$\left(f_1\left(f_2\right)^{(1)}, f_1\left(f_2\right)^{(2)}\right)_{x_0} = \begin{pmatrix} f_1^{(1)}\left(f_2\left(x_0\right)\right) f_2^{(1)}\left(x_0\right), \\ f_1^{(2)}\left(f_2\left(x_0\right)\right) \left(f_2^{(1)}\left(x_0\right)\right)^2 + f_1^{(1)}\left(f_2\left(x_0\right)\right) f_2^{(2)}\left(x_0\right) \end{pmatrix}$$
$$= \left(2x_0\cos\left(x_0^2\right), -4x_0^2\sin\left(x_0^2\right) + 2\cos\left(x_0^2\right)\right),$$
$$\left(f_3^{(1)}, f_3^{(2)}\right)_{x_0} = \left(\frac{1}{x_0}, -\frac{1}{x_0^2}\right),$$

where all the values in the above pairs are computed by direct evaluation of the seeds. Having now the required differentiation numbers for $\sin\left(x^2\right)$ and $\ln\left(x\right)$, we simply add the resultant pairs by differentiation addition to get $\left(f^{(1)}, f^{(2)}\right)$ at x_0.

Similarly, to compute $\left(f^{(2)}, f^{(3)}\right)$, we have the following dyadic differentiation numbers for respectively $\sin\left(x^2\right)$ and $\ln\left(x\right)$

$$\left(f_1\left(f_2\right)^{(2)}, f_1\left(f_2\right)^{(3)}\right)_{x_0} = \begin{pmatrix} f_1^{(2)}\left(f_2\left(x_0\right)\right) \left(f_2^{(1)}\left(x_0\right)\right)^2 + f_1^{(1)}\left(f_2\left(x_0\right)\right) f_2^{(2)}\left(x_0\right), \\ \left(f_1^{(3)}\left(f_2\left(x_0\right)\right) \left(f_2^{(1)}\left(x_0\right)\right)^3 + 2f_2^{(1)}\left(x_0\right) f_2^{(2)}\left(x_0\right) f_1^{(2)}\left(f_2\left(x_0\right)\right)\right) \\ + \left(f_1^{(2)}\left(f_2\left(x_0\right)\right) f_2^{(1)}\left(x_0\right) f_2^{(2)}\left(x_0\right) + f_1^{(1)}\left(f_2\left(x_0\right)\right) f_2^{(3)}\left(x_0\right)\right) \end{pmatrix}$$
$$= \left(-4x_0^2\sin\left(x_0^2\right) + 2\cos\left(x_0^2\right), -8x_0^3\cos\left(x_0^2\right) - 12x_0\sin\left(x_0^2\right)\right),$$
$$\left(f_3^{(2)}, f_3^{(3)}\right)_{x_0} = \left(-\frac{1}{x_0^2}, \frac{2}{x_0^3}\right),$$

and by differentiation addition of the resultant pairs we get $\left(f^{(2)}, f^{(3)}\right)$ at x_0.

We close this section by illustrating the case of multiplication. Consider the real function

$$h\left(x\right) = f\left(x\right) \times g\left(x\right) = \sin\left(x\right) \times \ln\left(x\right).$$

To compute the differentiation number $\left(h^{(2)}, h^{(3)}\right)$ at some real number x_0, we will have the following differentiation seeds, for $f\left(x\right) = \sin\left(x\right)$ and $g\left(x\right) = \ln\left(x\right)$, respectively

$$\left(\sin\left(x\right), \cos\left(x\right)\right), \left(\cos\left(x\right), -\sin\left(x\right)\right), \left(-\sin\left(x\right), -\cos\left(x\right)\right);$$
$$\left(\ln\left(x\right), \frac{1}{x}\right), \left(\frac{1}{x}, -\frac{1}{x^2}\right), \left(-\frac{1}{x^2}, \frac{2}{x^3}\right).$$

Now to compute $\left(h^{(2)}, h^{(3)}\right)$, we have the following six dyadic differentiation numbers

$$\left(f, f^{(1)}\right)_{x_0} = \left(\sin\left(x_0\right), \cos\left(x_0\right)\right), \left(f^{(1)}, f^{(2)}\right)_{x_0} = \left(\cos\left(x_0\right), -\sin\left(x_0\right)\right),$$
$$\left(f^{(2)}, f^{(3)}\right)_{x_0} = \left(-\sin\left(x_0\right), -\cos\left(x_0\right)\right), \left(g, g^{(1)}\right)_{x_0} = \left(\ln\left(x_0\right), \frac{1}{x_0}\right),$$
$$\left(g^{(1)}, g^{(2)}\right)_{x_0} = \left(\frac{1}{x_0}, -\frac{1}{x_0^2}\right), \left(g^{(2)}, g^{(3)}\right)_{x_0} = \left(-\frac{1}{x_0^2}, \frac{2}{x_0^3}\right),$$

and by differentiation multiplication and addition of the above pairs, according to Leibniz' formula, we get $\left(h^{(2)}, h^{(3)}\right)$ at x_0 as follows,

$$
\begin{aligned}
\left(h^{(2)}, h^{(3)}\right) &= \left(f^{(2)}, f^{(3)}\right)_{x_0} \times \left(g, g^{(1)}\right)_{x_0} + 2 \times \left(f^{(1)}, f^{(2)}\right)_{x_0} \times \left(g^{(1)}, g^{(2)}\right)_{x_0} \\
&\quad + \left(f, f^{(1)}\right)_{x_0} \times \left(g^{(2)}, g^{(3)}\right)_{x_0} \\
&= \left(\begin{array}{c} \frac{2\cos(x_0)}{x_0} - \sin\left(x_0\right)\left(\ln\left(x_0\right) + \frac{1}{x_0^2}\right), \\ \sin\left(x_0\right)\left(\frac{2}{x_0^3} - \frac{3}{x_0}\right) - \cos\left(x_0\right)\left(\ln\left(x_0\right) + \frac{3}{x_0^2}\right) \end{array} \right).
\end{aligned}
$$

The above examples illustrate that *automatic differentiation* as based on the theory $\mathsf{Th}_{\mathbf{D}\mathcal{K}}$ of *dyadic* differentiation numbers, equipped with "differentiation seeds", the "chain rule" and "Leibniz' product rule", is completely sufficient for computing automatic derivatives of first and higher orders.

7. CONCLUSION

Computers are now taking an increasingly paramount and highly efficient role in practising mathematics and in producing and verifying scientific knowledge. An important issue in the state of the art, which is of great importance, is to computationally evaluate the derivatives of a given function. Differentiation arithmetic is a principal and reliable technique that makes this very desirable task possible; and so, in this article, we recasted real differentiation arithmetic in a consistent and categorical axiomatic theory of dyadic differentiation numbers that provides a foundation for first and higher order automatic derivatives. We, next constructed the algebraic system of real differentiation arithmetic, deduced its fundamental properties, and proved that it constitutes a commutative unital ring. Finally, we presented a brief account of machine realization of the theory of dyadic differentiation numbers, gave numerical examples that showed how to compute automatic derivatives of first and higher order to an arbitrary precision using InCLosure, and thereupon compared our results to the results obtained using the symbolic differentiation capability of Wolfram Mathematica.

What is then the ultimate importance of an axiomatic theory of differentiation arithmetic? Not only are formal theories indispensable for pure mathematics, but they are also of great importance for applied and computationally-oriented mathematics. Being an axiomatic extension of the theory of a continuously ordered field, our theory of real differentiation algebra provides a rigorous and unified mathematical foundation for the various approaches of automatic differentiation as it is now practised. One main advantage of our axiomatization is that, although it considers only 'dyadic' differentiation numbers, the presented theory nevertheless underlies higher and partial automatic derivatives without the need for defining Clifford's or Grassmann algebras of higher dimensions. Moreover, it is a well-known logical fact that a "categorical" axiomatization of a theory is the "best" axiomatization possible. By virtue of being *categorical*, the axiomatic system presented in this article is "best" in the sense that it correctly describes, up to isomorphism, *every structure* of real differentiation numbers. A further novelty of this axiomatization is gaining the advantage of deducing the fundamental properties of differentiation numbers in a merely logical manner. Such a formalization, it is hoped, will have a substantial impact on both fundamental research

and practical applications of differentiation arithmetic. Noteworthy also is that with some basic alterations, the categorical system presented in this article is extensible seamlessly to an interval differentiation arithmetic, and to an algebraically closed commutative ring of complex differentiation arithmetic. To reiterate, the research conducted in this article is intended not just as a new logical foundation of differentiation arithmetic, but as a concrete systematic basis that leads to deeper understanding, and from which new investigations, techniques, and insights hopefully might accrue.

8. Supplementary Materials

To reproduce the results of the calculations in this article, latest version of InCLosure is available for free download via *https://doi.org/10.5281/zenodo.2702404* or from the first author's website at: *http://scholar.cu.edu.eg/henddawood/software/InCLosure*. An InCLosure input file and its corresponding output containing, respectively, the code and results of the examples are also available as a supplementary material to this article, via *https://doi.org/10.5281/zenodo.3352442*.

Acknowledgments

References

[1] M. Abdelhafez, D. Schuster, and J. Koch. *Gradient-Based Optimal Control of Open Systems Using Quantum Trajectories and Automatic Differentiation*. ArXiv e-prints, eprint arXiv:1901.05541, January 2019. URL *https://arxiv.org/abs/1901.05541*.

[2] C. P. Atkinson and A. S. Levens. *A New Differentiating Machine*. Mathematical Tables and Other Aids to Computation, 5(34):99–102, 1951.

[3] L. M. Beda, L. N. Korolev, N. V. Sukkikh, and T. S. Frolova. *Programs for Automatic Differentiation for the Machine BESM*. Technical report, Institute for Precise Mechanics and Computation Techniques, Academy of Science, Moscow, USSR, 1959. (In Russian).

[4] W. Bosch and P. Krajkiewicz. *A Categorical System of Axioms for the Complex Numbers*. Mathematics Magazine, 43(2):67–70, March 1970, doi:10.2307/2688967.

[5] G. Cantor. *Beitrage zur Begrundung der transfiniten Mengenlehre II*. Mathematische Annalen, 49:207–246, 1897. Translated with introduction and commentary by Philip E. B. Jourdain as "*Contributions to the Founding of the Theory of Transfinite Numbers*", Dover Publications, New York, 1955.

[6] R. L. Causey. *Logic, Sets, and Recursion*. Jones and Bartlett Publishers, Boston, 1994.

[7] G. L. Cherlin. *Algebraically Closed Commutative Rings*. The Journal of Symbolic Logic, 38(3):493–499, December 1973, doi:10.2307/2273048.

[8] S. Chevillard, M. Joldes, and C. Lauter. *Sollya: An Environment for the Development of Numerical Codes*. In K. Fukuda, J. van der Hoeven, M. Joswig, and N. Takayama, editors, *Mathematical Software - ICMS 2010*, volume 6327 of *Lecture Notes in Computer Science*, pp. 28–31, Springer, Heidelberg, Germany, September 2010.

[9] W. K. Clifford. *Preliminary Sketch of Biquaternions*. Proceedings of the London Mathematical Society, 4(1):381–395, November 1873.

[10] J. Corcoran. *Categoricity*. History and Philosophy of Logic, 1(1):187–207, 1980, doi:10.1080/01445348008837010. URL *https://doi.org/10.1080/01445348008837010*.

[11] G. F. Corliss. *Automatic Differentiation Bibliography*. In G. G. Corliss, editor, *SIAM Proceedings of Automatic Differentiation of Algorithms: Theory, Implementation and Application*, SIAM, 1992.

[12] G. F. Corliss and L. B. Rall. *An Introduction to Automatic Differentiation*. In M. Berz, C. H. Bischof, G. F. Corliss, and A. Griewank, editors, *Computational Differentiation: Techniques, Applications, and Tools*, pp. 1–17, SIAM, Philadelphia, PA, 1996.

[13] H. Curry. *Foundations of Mathematical Logic*. Dover Publications, New York, 1977.

[14] H. Dawood. *Theories of Interval Arithmetic: Mathematical Foundations and Applications*. LAP Lambert Academic Publishing, Saarbrücken, 2011, ISBN 978-3-8465-0154-2.

[15] H. Dawood. *Interval Mathematics: Foundations, Algebraic Structures, and Applications*. Master's thesis, Department of Mathematics, Faculty of Science, Cairo University, Giza, 2012, doi:10.13140/RG.2.2.24252. 13449. Doi:10.13140/RG.2.2.24252.13449, URL *http://dx.doi.org/10.13140/RG.2.2.24252.13449*.

[16] H. Dawood. *Interval Mathematics as a Potential Weapon against Uncertainty*. In S. Chakraverty, editor, *Mathematics of Uncertainty Modeling in the Analysis of Engineering and Science Problems*, chapter 1, pp. 1–38, IGI Global, Hershey, PA, January 2014, ISBN 978-1-4666-4991-0, doi:10.4018/978-1-4666-4991-0. ch001. URL *http://dx.doi.org/10.4018/978-1-4666-4991-0.ch001*.

[17] H. Dawood. *InCLosure (Interval enCLosure)–A Language and Environment for Reliable Scientific Computing*. Computer Software, Version 1.0, Department of Mathematics, Faculty of Science, Cairo University, Giza, Egypt, July 2018, doi:10.5281/zenodo.2702405. InCLosure Support: *http://scholar.cu.edu.eg/henddawood/ software/InCLosure*, URL *https://doi.org/10.5281/zenodo.2702405*.

[18] H. Dawood. *On Some Algebraic and Order-Theoretic Aspects of Machine Interval Arithmetic*. Online Mathematics Journal, 1(2):1–13, April 2019, doi:10.5281/zenodo.2656089. URL *http://doi.org/10.5281/ zenodo.2656089*.

[19] H. Dawood and Y. Dawood. *A Logical Formalization of the Notion of Interval Dependency: Towards Reliable Intervalizations of Quantifiable Uncertainties*. Online Mathematics Journal, 1(3), July 2019, doi: 10.5281/zenodo.3234184. URL *http://doi.org/10.5281/zenodo.3234184*.

[20] H. Dawood and Y. Dawood. *Universal Intervals: Towards a Dependency-Aware Interval Algebra*. In S. Chakraverty, editor, *Mathematical Methods in Interdisciplinary Sciences*, John Wiley & Sons, Hoboken, New Jersey, March 2020, ISBN 978-1-119-58550-3.

[21] R. Dedekind. *Essays on the Theory of Numbers: Continuity and Irrational Numbers*. Dover Publications, New York, 1963.

[22] A. Elmendorf. *A Differentiating Machine*. The American Mathematical Monthly, 23:292–295, 1916.

[23] C. P. Fries. *Stochastic Automatic Differentiation: Automatic Differentiation for Monte-Carlo Simulations*. Quantitative Finance, pp. 1–17, 2019, doi:10.1080/14697688.2018.1556398. URL *https://doi.org/10.1080/ 14697688.2018.1556398*.

[24] R. Goetschel and W. Voxman. *Elementary Fuzzy Calculus*. Fuzzy sets and systems, 18(1):31–43, 1986.

[25] H. Grassmann. *Die Lineale Ausdehnungslehre ein neuer Zweig der Mathematik: dargestellt und durch Anwendungen auf die übrigen Zweige der Mathematik, wie auch auf die Statik, Mechanik, die Lehre vom Magnetismus und die Krystallonomie erläutert*. Verlag von Otto Wigand, Leipzig, 1844.

[26] A. Griewank and A. Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. SIAM, Philadelphia, PA, second edition, 2008.

[27] L. Hascoet and V. Pascual. *The Tapenade Automatic Differentiation Tool: Principles, Model, and Specification*. ACM Transactions on Mathematical Software, 39(3):20:1–20:43, May 2013, doi:10.1145/2450153.2450158.

[28] G. Hunter. *Metalogic: An Introduction to the Metatheory of Standard First Order Logic*. Macmillan, New York, 1971.

[29] E. V. Huntington. *A Set of Postulates for Ordinary Complex Algebra*. Transactions of the American Mathematical Society, 6(2):209–229, 1905.

[30] IEEE 1788 Committee. *IEEE Standard for Interval Arithmetic*. IEEE Std 1788-2015, pp. 1–97, June 2015, doi:10.1109/IEEESTD.2015.7140721. URL *https://ieeexplore.ieee.org/document/7140721*.

[31] IEEE 1788 Committee. *IEEE Standard for Interval Arithmetic (Simplified)*. IEEE Std 1788.1-2017, pp. 1–38, January 2018, doi:10.1109/IEEESTD.2018.8277144. URL *https://ieeexplore.ieee.org/document/8277144*.

[32] D. Kalman. *Doubly Recursive Multivariate Automatic Differentiation*. Mathematics Magazine, 75(3):187–202, 2002, doi:10.1080/0025570X.2002.11953128.

[33] S. C. Kleene. *Introduction to Metamathematics*. North-Holland Publishing Company, Amsterdam, 1952.

[34] G. Melquiond. *Proving Bounds on Real-Valued Functions with Computations*. In A. Armando, P. Baumgartner, and G. Dowek, editors, *International Joint Conference on Automated Reasoning IJCAR*, volume 5195 of *Lecture Notes in Artificial Intelligence*, pp. 2–17, Springer-Verlag, August 2008, doi: 10.1007/978-3-540-71070-7_2.

[35] D. P. Mitchell. *Three Applications of Interval Analysis in Computer Graphics*. In *Frontiers in Rendering Course Notes, SIGGRAPH '91 Conference Proceedings*, pp. 1–13, Las Vegas, Nevada, 28 July–2 August 1991.

[36] R. Montague. *Formal Philosophy: Selected Papers*. Yale University Press, London, 1974. Edited with introduction by Richmond H. Thomason.

[37] R. E. Moore. *Interval Analysis*. Prentice Hall, 1966.

[38] R. E. Moore. *Methods and Applications of Interval Analysis*. Number 2 in SIAM studies in Applied Mathematics, SIAM, Philadelphia, 1979.

[39] R. E. Moore, R. B. Kearfott, and M. J. Cloud. *Introduction to Interval Analysis*. SIAM, 2009.

[40] U. Naumann and J. Riehme. *A Differentiation-Enabled Fortran 95 Compiler*. ACM Transactions on Mathematical Software, 31(4):458–474, December 2005, doi:10.1145/1114268.1114270.

[41] R. D. Neidinger. *Introduction to Automatic Differentiation and MATLAB Object-Oriented Programming*. SIAM Review, 52(3):545–563, 2010.

[42] R. D. Neidinger and B. Altman. *Comparing High-Order Multivariate AD Methods*. Optimization Methods and Software, 33(4-6):995–1009, May 2018, doi:10.1080/10556788.2018.1472256.

[43] M. L. Puri and D. A. Ralescu. *Differentials of Fuzzy Functions*. Journal of Mathematical Analysis and Applications, 91(2):552–558, 1983.

[44] L. B. Rall. *Automatic Differentiation: Techniques and Applications*, volume 120 of *Lecture Notes in Computer Science*. Springer Verlag, Berlin, 1981.

[45] L. B. Rall. *The Arithmetic of Differentiation*. MRC Technical Report 2688, Mathematics Research Center, University of Wisconsin-Madison, Madison, Wisconsin, 1984.

[46] L. B. Rall. *The Arithmetic of Differentiation*. Mathematics Magazine, 59(5):275–282, December 1986, doi:10.1080/0025570X.1986.11977261.

[47] H. Rasiowa and R. Sikorski. *The Mathematics of Metamathematics*. Panstwowe Wydawnictwo Naukowe, Warszawa, 1963.

[48] A. Robinson. *On the Metamathematics of Algebra*. North-Holland Publishing Company, Amsterdam, 1951.

[49] S. M. Rump. *INTLAB–INTerval LABoratory*. In T. Csendes, editor, *Developments in Reliable Computing*, pp. 77–104, Kluwer Academic Publishers, Dordrecht, 1999.

[50] S. Shapiro. *Second-Order Languages and Mathematical Practice*. The Journal of Symbolic Logic, 50(3):714–742, 1985.

[51] H. Sommer, C. Pradalier, and P. Furgale. *Automatic Differentiation on Differentiable Manifolds as a Tool for Robotics*. In M. Inaba and P. Corke, editors, *Robotics Research*, volume 114 of *Springer Tracts in Advanced Robotics*, pp. 505–520, Springer Verlag, 2016.

[52] A. Tarski. *Logic, Semantics, Metamathematics*. Oxford At the Clarendon Press, 1965. Translated from Polish by J. H. Woodger.

[53] A. Tarski. *Introduction to Logic and to the Methodology of the Deductive Sciences*. Oxford University Press, New York, fourth edition, 1994. Translated from Polish by Olaf Helmer.

[54] L. Tingelstad and O. Egeland. *Automatic Multivector Differentiation and Optimization*. Advances in Applied Clifford Algebras, 27(1):707–731, March 2017.

[55] M. A. S. Trindade, E. Pinto, and S. Floquet. *Clifford Algebras, Multipartite Systems and Gauge Theory Gravity*. Advances in Applied Clifford Algebras, 29(1), 2019, doi:10.1007/s00006-018-0917-0. URL *https://doi.org/10.1007/s00006-018-0917-0*.

[56] A. Walther. *Getting Started with ADOL-C*. In U. Naumann, O. Schenk, H. D. Simon, and S. Toledo, editors, *Combinatorial Scientific Computing*, number 09061 in Dagstuhl Seminar Proceedings, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009. URL *http://drops.dagstuhl.de/opus/volltexte/2009/2084*.

[57] R. Wengert. *A Simple Automatic Derivative Evaluation Program*. Communications of the ACM, 7(8):463–464, 1964.

[58] Wolfram Research, Inc.. *Mathematica, Version 11.3, 'Computer Software'*. Wolfram Research, Inc., Champaign, Illinois, 2018. URL *https://www.wolfram.com*.