ISSN 1061-9208, Russian Journal of Mathematical Physics, Vol. 17, No. 2, 2010, pp. 159–191. © Pleiades Publishing, Ltd., 2010.

Infinitesimals without Logic

P. Giordano
Università della Svizzera Italiana, Academy of Architecture
Largo Bernasconi 2 CH-6850 Mendrisio, Switzerland
E-mail: paolo.giordano@usi.ch

Received September 22, 2009

Abstract. We introduce a ring of the so-called Fermat reals, which is an extension of the real field containing nilpotent infinitesimals. The construction is inspired by Smooth Infinitesimal Analysis (SIA) and provides a powerful theory of actual infinitesimals without any background in mathematical logic. In particular, in contrast to SIA, which admits models in intuitionistic logic only, the theory of Fermat reals is consistent with the classical logic. We face the problem of deciding whether or not a product of powers of nilpotent infinitesimals vanishes, study the identity principle for polynomials, and discuss the definition and properties of the total order relation. The construction is highly constructive, and every Fermat real admits a clear and order-preserving geometrical representation. Using nilpotent infinitesimals, every smooth function becomes a polynomial because the remainder in Taylor's formulas is now zero. Finally, we present several applications to informal classical calculations used in physics, and all these calculations now become rigorous, and at the same time, formally equal to the informal ones. In particular, an interesting rigorous deduction of the wave equation is given, which clarifies how to formalize the approximations tied with Hooke's law using the language of nilpotent infinitesimals.

DOI: 10.1134/S1061920810020032

1. INTRODUCTION AND GENERAL PROBLEM

Frequently, in works of physicists, it is possible to find informal calculations of the form

$$1/\sqrt{1-v^2/c^2} = 1 + v^2/(2c^2), \qquad \sqrt{1-h_{44}(x)} = 1 - \tfrac{1}{2}h_{44}(x), \tag{1}$$

with an explicit use of infinitesimals $v/c \ll 1$ or $h_{44}(x) \ll 1$ such that, e.g., $h_{44}(x)^2 = 0$. For example, Einstein [13] wrote the formula (using the equality sign rather than the approximate equality sign $\simeq$)

$$f(x, t+\tau) = f(x,t) + \tau \cdot \partial f/\partial t(x,t), \tag{2}$$

justifying it with the words *"since τ is very small;"* formulas (1) are a particular case of the general formula (2). Dirac [10] wrote an analogous equality when studying the Newtonian approximation in general relativity.

Using this type of infinitesimals, we can write out an *equality*, in some infinitesimal neighborhood, between a smooth function and its tangent straight line, or, in other words, a Taylor formula without remainder.

Obviously, there are many possibilities to formalize this kind of intuitive reasonings, obtaining a more or less good dialectic between informal and formal thinking, and indeed there are several theories of actual infinitesimals (from now on, for simplicity, we will say "infinitesimals" instead of "actual infinitesimals" as opposed to "potential infinitesimals"). Starting from these theories, we can distinguish between two types of definitions of infinitesimals. First, there can be (at least) a ring R containing the real field $\mathbb{R}$, and infinitesimals are elements $\varepsilon \in R$ such that $-r < \varepsilon < r$ for every positive standard real $r \in \mathbb{R}_{>0}$. The second type of infinitesimals is defined by using some algebraic property of nilpotency, i.e., $\varepsilon^n = 0$ for some natural number $n \in \mathbb{N}$. For some ring R, these definitions can coincide; however, anyway, they certainly lead only to the trivial infinitesimal $\varepsilon = 0$ if $R = \mathbb{R}$.

However, these definitions of infinitesimals correspond to theories which are completely different in nature and in the underlying ideas. Indeed, these theories can be seen in a more interesting way to belong to two different classes. To the first one, we can refer the theories that need a certain amount of nontrivial results of mathematical logic, whereas in the other class, we have attempts to define sufficiently strong theories of infinitesimals without using nontrivial results of mathematical logic. In the first class, we have the Non-Standard Analysis (NSA) and the Synthetic Differential Geometry (SDG, which is also referred to as the Smooth Infinitesimal Analysis, see, e.g., Bell [3], Kock [20], Lavendhomme [22], and Moerdijk and Reyes [23]), whereas in the second one, we have, e.g., the Weil functors (see Kriegl and Michor [21]), Levi-Civita fields (see Shamseddine [25] and Berz [7]), the surreal numbers (see Conway [9] and Ehresmann [12]), and geometries over rings containing infinitesimals (see Bertram [6]). More precisely, we can say that, to work in NSA and SDG, one needs a formal control deeply stronger than the one used in "standard mathematics." Indeed, to use NSA, one has to be able to formally write the sentences one needs to use the transfer theorem. Moreover, SDG admits no models in classical logic, but in intuitionistic logic only, and hence we must be sure that our proofs make no use of the law of the excluded middle, or, e.g., of the classical part of De Morgan's law, of some form of the axiom of choice, of the implication of double negation toward affirmation, or any other logical principle which is not valid in intuitionistic logic. Physicists, engineers, and also the majority of mathematicians are not used to have this strong formal control in their work, and, for this reason, there are attempts to present both NSA and SDG reducing the necessary formal control as much as possible, even if this is technically impossible at some level (see, e.g., Henson [19], and Benci and Di Nasso [4, 5] for NSA and Bell [3] and Lavendhomme [22] for SDG, where, using an axiomatic approach, the authors try to postpone a very difficult construction of an intuitionistic model of a whole set theory using topos).

On the other hand, NSA is essentially the only theory of infinitesimals with discrete diffusion and sufficiently great community of working mathematicians and published results in several areas of mathematics and its applications, see, e.g., [1], and SDG is the only theory of infinitesimals with nontrivial, new, and published results in differential geometry concerning infinite-dimensional spaces like the space of all diffeomorphisms of a generic (e.g., noncompact) smooth manifold. In NSA, we have only few results concerning differential geometry. Other theories of infinitesimals, at least up to now, have less formal strength than NSA or SDG or even less potentiality to be applied in several areas of mathematics.

Our main aim, for which the present work represents a first step, is to find a theory of infinitesimals within the "standard mathematics" (in the precise sense explained above of a formal control more "standard" and not so strong as the one needed, e.g., in NSA or SDG) with results comparable with those of SDG, without forcing the reader to learn a strong formal control of the mathematics he/she is doing. Because it has to be considered inside "standard mathematics," our theory of infinitesimals must be compatible with classical logic.

Concretely, the idea of the present work is to by-pass the impossibility theorem about the incompatibility of SDG with the classical logic that forces SDG to find models within intuitionistic logic.

Another point of view concerning current theories of infinitesimals is that, despite the fact that thay are frequently presented using opposed motivations, they lack the intuitive interpretation of what the powerful formalism permits to do. For a concrete example in this direction, see Giordano [16]. Another aim of the present work is to construct a new theory of infinitesimals always preserving a very good dialectic between formal properties and intuitive interpretation.

More technically, we want to show that the real field can be extended by adding nilpotent infinitesimals, arriving at an enlarged real line ${}^\bullet\mathbb{R}$, by means of a very simple construction completely within "standard mathematics." Indeed, to define the extension ${}^\bullet\mathbb{R} \supset \mathbb{R}$, we use elementary analysis only. To avoid any misunderstanding, is it important to clarify that the purpose of the present work is to obtain a theory of nilpotent infinitesimals as a first step for the foundation of a smooth ($\mathcal{C}^\infty$) differential geometry rather than to give an alternative foundation of differential and integral calculus (like NSA). For some preliminary results in this direction, see Giordano [16].

2. MOTIVATIONS FOR THE TITLE "FERMAT REALS"

As is well known, historically, two possible reductionist constructions of the real field starting from the rationals have been made. The first is Dedekind's order completion using sections of rationals, and the other one is Cauchy's metric space completion. Certainly, there are no historical reasons to attribute our extension ${}^\bullet\mathbb{R} \supset \mathbb{R}$ of the real field (to be described below) to Fermat, but there are motivations to say that the underlying spirit and some properties of our theory could possibly please him. Here are some arguments.

(1) A formalization of Fermat's infinitesimal method to derive functions is provable in our theory. We recall that Fermat's idea was, roughly speaking and not on the basis of an accurate historical analysis which goes beyond the scope of the present work (see, e.g., Edwards [11] and Eves [14]), to suppose first that $h \neq 0$, to construct the incremental ratio $(f(x+h) - f(x))/h$ and, after suitable simplifications (sometimes using infinitesimal properties), to set $h = 0$ in the final result.

(2) Fermat's method to find the maximum or minimum of a given function $f(x)$ at $x = a$ was to take e to be extremely small so that the value of $f(x+h)$ was approximately equal to that of $f(x)$. In modern, algebraic language, it can be said that $f(x+h) = f(x)$ only if $h^2 = 0$, i.e., if e is a first-order infinitesimal. Fermat was aware that this is not a "true" equality but some kind of approximation (ibidem). We follow a similar idea to define ${}^{\bullet}\mathbb{R}$ by introducing a suitable equivalence relation to represent the above equality.

(3) Fermat has been described by Bell [2] as "the king of amateurs" of mathematics, and hence we can suppose that in its mathematical work the informal/intuitive part was stronger with respect to the formal one. For this reason, we can think that he could be possibly pleased by our idea to obtain a theory of infinitesimals by preserving the intuitive meaning and without forcing the working mathematician to be much too formal.

For these reason, we chose the title "Fermat reals" for our ring ${}^{\bullet}\mathbb{R}$ (note that the possessive case is not used, to stress that we are not attributing our construction of ${}^{\bullet}\mathbb{R}$ to Fermat).

3. DEFINITION AND ALGEBRAIC PROPERTIES OF FERMAT REALS: THE BASIC IDEA

We start from the idea that a smooth ($\mathcal{C}^\infty$) function $f\colon {}^{\bullet}\mathbb{R} \longrightarrow {}^{\bullet}\mathbb{R}$ is actually *equal* to its tangent straight line in the first-order neighborhood, e.g., of the point $x = 0$, i.e.,

$$\forall h \in D: \; f(h) = f(0) + h \cdot f'(0), \tag{3}$$

where D is a subset of ${}^{\bullet}\mathbb{R}$ defining the above neighborhood of $x = 0$. Relation (3) can be viewed as a first-order Taylor's formula without remainder, because, intuitively, we think that $h^2 = 0$ for any $h \in D$ (indeed, the property $h^2 = 0$ defines the first-order neighborhood of $x = 0$ in ${}^{\bullet}\mathbb{R}$). These almost trivial considerations help us to understand many things. First, ${}^{\bullet}\mathbb{R}$ must necessarily be a ring rather than a field, because, in a field, the equation $h^2 = 0$ implies $h = 0$; moreover, we will surely have some limitation in the extension of some function from $\mathbb{R}$ to ${}^{\bullet}\mathbb{R}$, e.g., for the square root, because, when using this function with the standard properties, once again the equation $h^2 = 0$ would imply $|h| = 0$. On the other hand, we are also forced to ask whether or not formula (3) uniquely determines the derivative $f'(0)$ (because, even if it is true that we cannot simplify by h, we know that the polynomial coefficients of a Taylor's formula are unique in classical analysis). In fact, we shall prove that

$$\exists!\, m \in \mathbb{R} \quad \forall h \in D: \; f(h) = f(0) + h \cdot m, \tag{4}$$

i.e., the slope of the tangent is uniquely determined if it is an ordinary real number. We refer to formulas of the form (4) as *derivation formulas*.

When trying to construct a model for (3), a natural idea is to regard our new numbers in ${}^{\bullet}\mathbb{R}$ as equivalence classes $[h]$ of ordinary functions $h\colon \mathbb{R} \longrightarrow \mathbb{R}$. In this way, we may hope both to include the real field by using classes generated by constant functions and to have the class generated by the function $h(t) = t$ as a first-order infinitesimal number. To understand how to define this equivalence relation, we are to treat (3.1) as follows:

$$f(h(t)) \sim f(0) + h(t) \cdot f'(0), \tag{5}$$

where the idea is that we are going to define $\sim$. If we assume that $h(t)$ is "sufficiently similar to t," then we can define $\sim$ in such a way that (5) is equivalent to

$$\lim_{t \to 0^+} (f(h(t)) - f(0) - h(t) \cdot f'(0))/t = 0,$$

i.e.,

$$x \sim y \quad :\Longleftrightarrow \quad \lim_{t \to 0^+} (x_t - y_t)/t = 0. \tag{6}$$

In this way, formula (5) is very near to the definition of differentiability of f at 0.
It is important to note that, because of de L'Hôpital's theorem, we have an isomorphism

$$\mathcal{C}^1(\mathbb{R},\mathbb{R})/\!\sim\,\simeq \mathbb{R}[x]/(x^2),$$

where the left-hand side is (isomorphic to) the usual tangent bundle of $\mathbb{R}$, and thus we obtain nothing new. It is not easy to understand what set of functions we have to choose for x, y in (6) to obtain a nontrivial structure. The first idea is to take continuous functions at $t=0$, instead of more regular ones like $\mathcal{C}^1$-functions, in such a way that, e.g., $h_k(t)=|t|^{1/k}$ becomes a kth order nilpotent infinitesimal ($h^{k+1}\sim 0$); indeed, for almost all results presented in this paper, continuous functions at $t=0$ work well. However, only when proving the nontrivial property

$$(\forall x\in {}^\bullet\mathbb{R}:\; x\cdot f(x)=0) \quad\Longrightarrow\quad \forall x\in {}^\bullet\mathbb{R}:\; f(x)=0, \tag{7}$$

we can see that it is insufficient to take continuous functions at $t=0$. To prove (7), the following objects turn out to be useful.

Definition 1. If $x\colon \mathbb{R}_{\geqslant 0}\longrightarrow\mathbb{R}$, then we say that x is *nilpotent* if and only if $|x(t)-x(0)|^k=o(t)$ as $t\to 0^+$ for some $k\in\mathbb{N}$. Denote by $\mathcal{N}$ the set of all nilpotent functions.

For example, any Hölder function ($|x(t)-x(s)|\leqslant c\cdot|t-s|^\alpha$ for some constant $\alpha>0$) is nilpotent. The choice of nilpotent functions instead of more regular ones establishes a great difference of our approach from the classical definition of jets (see, e.g., Bröcker [8], Golubitsky and Guillemin [17]) which can be recalled by (6).

Another problem necessarily related to the basic idea (3) is that the use of nilpotent infinitesimals frequently leads to the consideration of terms like $h_1^{i_1}\cdots h_n^{i_n}$. For this type of products, the first problem is to know whether or not $h_1^{i_1}\cdots h_n^{i_n}\neq 0$ and what is the order k of this new infinitesimal, i.e., what is a k for which $(h_1^{i_1}\cdots h_n^{i_n})^k\neq 0$ and $(h_1^{i_1}\cdots h_n^{i_n})^{k+1}=0$. We shall have a good frame if we shall be able to solve these problems starting from the order of each infinitesimal h_j and from the values of the powers $i_j\in\mathbb{N}$. On the other hand, almost all examples of nilpotent infinitesimals are of the form $h(t)=t^\alpha$, with $0<\alpha<1$, and their sums; these functions also have important properties in the treatment of products of powers. For these reasons, we shall focus our attention on the following family of functions $x\colon\mathbb{R}_{\geqslant 0}\longrightarrow\mathbb{R}$ in the definition (6) of $\sim$.

Definition 2. We say that x *is a little-oh polynomial* and write $x\in\mathbb{R}_o[t]$ if and only if $x\colon\mathbb{R}_{\geqslant 0}\longrightarrow\mathbb{R}$ and

$$x_t=r+\sum_{i=1}^k \alpha_i\cdot t^{a_i}+o(t) \quad\text{as}\quad t\to 0^+$$

for suitable $k\in\mathbb{N}$, $r,\alpha_1,\dots,\alpha_k\in\mathbb{R}$, and $a_1,\dots,a_k\in\mathbb{R}_{\geqslant 0}$.

Hence a little-oh polynomial $x\in\mathbb{R}_o[t]$ is a polynomial function with real coefficients in the real variable $t\geqslant 0$ with generic positive powers of t and up to a little-oh function as $t\to 0^+$.

Remark 3. Below, when writing $x_t=y_t+o(t)$ as $t\to 0^+$, we mean that $\lim_{t\to 0^+}(x_t-y_t)/t=0$ and $x_0=y_0$. In other words, every little-oh function is treated as a continuous function as $t\to 0^+$.

Example. Simple examples of little-oh polynomials are (1) $x_t=1+t+t^{1/2}+t^{1/3}+o(t)$ and (2) $x_t=r\quad\forall t$. Note that, in Definition 2, we can take $k=0$, and hence α and a are void sequences of reals, i.e., $\alpha=a\colon\varnothing\longrightarrow\mathbb{R}$ if we think of an n-tuple x of reals as a function $x\colon\{1,\dots,n\}\longrightarrow\mathbb{R}$. Another example is (3) $x_t=r+o(t)$.

4. FIRST PROPERTIES OF LITTLE-OH POLYNOMIALS

Little-Oh Polynomials Are Nilpotent

The first properties of little-oh polynomials are as follows: if $x_t = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + o_1(t)$ and $y_t = s + \sum_{j=1}^N \beta_j \cdot t^{b_j} + o_2(t)$ as $t \to 0^+$, then

$$(x + y) = r + s + \sum_{i=1}^{k} \alpha_i \cdot t^{a_i} + \sum_{j=1}^{N} \beta_j \cdot t^{b_j} + o_3(t)$$

and

$$(x \cdot y)_t = rs + \sum_{i=1}^{k} s\alpha_i \cdot t^{a_i} + \sum_{j=1}^{N} r\beta_j \cdot t^{b_i} + \sum_{i=1}^{k}\sum_{j=1}^{N} \alpha_i\beta_j \cdot t^{a_i}t^{b_j} + o_4(t),$$

and hence the set of little-oh polynomials is closed with respect to pointwise sum and product. Moreover, little-oh polynomials are nilpotent functions (see Definition 1); to prove this fact, we firstly prove that the set of nilpotent functions $\mathcal{N}$ is a subalgebra of the algebra $\mathbb{R}^{\mathbb{R}}$ of real-valued functions. Indeed, let x and y be two nilpotent functions such that we have $|x - x(0)|^k = o_1(t)$ and $|y - y(0)|^N = o_2(t)$; then we can write $x \cdot y - x(0) \cdot y(0) = x \cdot [y - y(0)] + y(0) \cdot [x - x(0)]$, and thus we can consider $|x \cdot [y - y(0)]|^k = |x|^k \cdot |y - y(0)|^k = |x|^k \cdot o_1(t)$ and $|x|^k \cdot o_1(t)/t \to 0$ as $t \to 0^+$ because $|x|^k \to |x(0)|^k$, and hence $x \cdot [y - y(0)] \in \mathcal{N}$. Analogously, $y(0) \cdot [x - x(0)] \in \mathcal{N}$, and therefore, the closeness of $\mathcal{N}$ with respect to the product follows from the closeness with respect to the sum. The case of sum results from the following relations (using the formulas $x_t := x(t)$, $u := x - x_0$, $v := y - y_0$, $|u_t|^k = o_1(t)$, and $|v_t|^N = o_2(t)$ and the assumption $k \geqslant N$):

$$u^k = o_1(t), \quad v^k = o_2(t), \quad (u + v)^k = \sum_{i=0}^{k} \binom{k}{i} u^i \cdot v^{k-i},$$

and

$$\forall i = 0, \ldots, k: \ \frac{u_t^i \cdot v_t^{k-i}}{t} = \frac{\left(u_t^k\right)^{\frac{i}{k}} \cdot \left(v_t^k\right)^{\frac{k-i}{k}}}{t^{\frac{i}{k}} \cdot t^{\frac{k-i}{k}}} = \left(\frac{u_t^k}{t}\right)^{\frac{i}{k}} \cdot \left(\frac{v_t^k}{t}\right)^{\frac{k-i}{k}}.$$

We can now prove that $\mathbb{R}_o[t]$ is a subalgebra of $\mathcal{N}$. Indeed, every constant $r \in \mathbb{R}$ and every power t^{a_i} are elements of $\mathcal{N}$, and hence,

$$r + \sum_{i=1}^{k} \alpha_i \cdot t^{a_i} \in \mathcal{N},$$

and thus, it remains to prove that, if $y \in \mathcal{N}$ and $w = o(t)$, then $y + w \in \mathcal{N}$. However, this holds because every little-oh function is trivially nilpotent and follows from the closeness of $\mathcal{N}$ with respect to the sum.

Closeness of Little-Oh Polynomials with respect to Smooth Functions

We claim that the class of little-oh polynomials is kept by the smooth functions, i.e., if $x \in \mathbb{R}_o[t]$ and if $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ is smooth, then $f \circ x \in \mathbb{R}_o[t]$. Write

$$x_t = r + \sum_{i=1}^{k} \alpha_i \cdot t^{a_i} + w(t) \quad \text{with} \quad w(t) = o(t), \qquad h(t) := x(t) - x(0) \quad \forall t \in \mathbb{R}_{\geqslant 0};$$

hence, $x_t = x(0) + h_t = r + h_t$. The function $t \mapsto h(t) = \sum_{i=1}^k \alpha_i \cdot t^{a_i} + w(t)$ belongs to $\mathbb{R}_o[t] \subseteq \mathcal{N}$, and thus, we can write $|h|^N = o(t)$ for some $N \in \mathbb{N}$ as $t \to 0^+$. It follows from Taylor's formula that

$$f(x_t) = f(r + h_t) = f(r) + \sum_{i=1}^{N} (f^{(i)}(r)/i!) \cdot h_t^i + o(h_t^N). \tag{8}$$

However, $|o(h_t^N)|/|t| = (|o(h_t^N)|/|h_t^N|) \cdot (|h_t^N|/|t|) \to 0$, and hence,

$$o(h_t^N) = o(t) \in \mathbb{R}_o[t]. \tag{9}$$

Relation (9), formula (8), the property $h \in \mathbb{R}_o[t]$, and the closeness of the little-oh polynomials with respect to the ring operations imply that $f \circ x \in \mathbb{R}_o[t]$.

5. EQUALITY AND DECOMPOSITION OF FERMAT REALS

Definition 4. Let x, $y \in \mathbb{R}_o[t]$. We say that $x \sim y$ or that $x = y$ in ${}^\bullet\mathbb{R}$ if and only if $x(t) = y(t) + o(t)$ as $t \to 0^+$. It is easy to prove that $\sim$ is an equivalence relation. Thus, we can set ${}^\bullet\mathbb{R} := \mathbb{R}_o[t]/\sim$, i.e., ${}^\bullet\mathbb{R}$ is the quotient set of $\mathbb{R}_o[t]$ with respect to the equivalence relation $\sim$.

The equivalence relation $\sim$ is a congruence with respect to pointwise operations, and hence ${}^\bullet\mathbb{R}$ is a commutative ring. To simplify the notation, we sometimes write "$x = y$ in ${}^\bullet\mathbb{R}$" instead of $x \sim y$, and we speak of the elements of $\mathbb{R}_o[t]$ directly (instead of their equivalence classes); for example, we can say that $x = y$ in ${}^\bullet\mathbb{R}$ and $z = w$ in ${}^\bullet\mathbb{R}$ imply $x + z = y + w$ in ${}^\bullet\mathbb{R}$.

The immersion of $\mathbb{R}$ in ${}^\bullet\mathbb{R}$ is $r \longmapsto \hat{r}$ defined by $\hat{r}(t) := r$. Below we always identify $\hat{\mathbb{R}}$ with $\mathbb{R}$, which is thus a subring of ${}^\bullet\mathbb{R}$. Conversely, if $x \in {}^\bullet\mathbb{R}$, then the mapping ${}^\circ(-) : x \in {}^\bullet\mathbb{R} \mapsto {}^\circ x = x(0) \in \mathbb{R}$, which evaluates each extended real in 0, is well defined. We refer to ${}^\circ(-)$ as the *standard part mapping*. We also note that $\dim_{\mathbb{R}} {}^\bullet\mathbb{R} = \infty$ (as a vector space over the field $\mathbb{R}$), and this stresses how different our approach is from the classical definition of jets. Instead, our idea is similar to NSA, where standard sets can be extended by adding new infinitesimal points, and this differs from the point of view of jet theory.

The following theorem introduces a decomposition of a Fermat real $x \in {}^\bullet\mathbb{R}$, which chooses a unique notation for its standard part and all its infinitesimal parts.

Theorem 5. *If $x \in {}^\bullet\mathbb{R}$, then there is one and only one sequence $(k, r, \alpha_1, \dots, \alpha_k, a_1, \dots, a_k)$ such that $k \in \mathbb{N}, r, \alpha_1, \dots, \alpha_k, a_1, \dots, a_k \in \mathbb{R}$ and*

(1) $x = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i}$ *in* ${}^\bullet\mathbb{R}$,
(2) $0 < a_1 < a_2 < \dots < a_k \leqslant 1$,
(3) $\alpha_i \neq 0 \quad \forall i = 1, \dots, k$.

In this statement, we have also to include the void case $k = 0$ and $\alpha = a\colon \varnothing \longrightarrow \mathbb{R}$. Obviously, as usual, we use the definition $\sum_{i=1}^0 b_i = 0$ for the sum of an empty set of numbers. As we shall see, this is the case if x is a standard real, i.e., if $x \in \mathbb{R}$.

In the following, we use the notation $t^a := \mathrm{d}t_{1/a} := [t \in \mathbb{R}_{\geqslant 0} \mapsto t^a \in \mathbb{R}]_\sim \in {}^\bullet\mathbb{R}$, and thus, e.g., $\mathrm{d}t_2 = t^{1/2}$ is a second-order infinitesimal. In general, as we shall see from the definition of order for a generic infinitesimal, $\mathrm{d}t_a$ is an infinitesimal of order a. In other words, these two predicates for the same object enable us to stress the difference between an actual infinitesimal $\mathrm{d}t_a$ and a potential infinitesimal $t^{1/a}$, namely, an actual infinitesimal of order $a \geqslant 1$ corresponds to a potential infinitesimal of order $\frac{1}{a} \leqslant 1$ (with respect to the classical notion of the order of an infinitesimal function in calculus, see, e.g., Prodi [24] and Silov [26]).

Remark 6. Note that $\mathrm{d}t_a \cdot \mathrm{d}t_b = \mathrm{d}t_{\frac{ab}{a+b}}$. Moreover, $\mathrm{d}t_a^\alpha := (\mathrm{d}t_a)^\alpha = \mathrm{d}t_{\frac{a}{\alpha}}$ for every $\alpha \geqslant 1$ and, finally, $\mathrm{d}t_a = 0$ for every $a < 1$. For example, $\mathrm{d}t_a^{[a]+1} = 0$ for every $a \in \mathbb{R}_{>0}$, where $[a] \in \mathbb{N}$ is the integer part of a, i.e., $[a] \leqslant a < [a] + 1$.

Existence. Since $x \in \mathbb{R}_o[t]$, we can write

$$x_t = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i} + o(t) \quad \text{as} \quad t \to 0^+, \quad \text{where} \quad r, \alpha_i \in \mathbb{R},\, a_i \in \mathbb{R}_{\geqslant 0},\, k \in \mathbb{N}.$$

Hence $x = r + \sum_{i=1}^k \alpha_i \cdot t^{a_i}$ in ${}^\bullet\mathbb{R}$, and our purpose is to pass from this representation of x to another representation satisfying conditions (1), (2), and (3) of the statement. If $a_i > 1$, then

$\alpha_i \cdot t^{a_i} = 0$ in ${}^\bullet\mathbb{R}$. Therefore, we can assume that $a_i \leqslant 1$ for every $i = 1, \dots, k$. Moreover, we can also assume that $a_i > 0$ for every i (because otherwise, if $a_i = 0$, we can replace $r \in \mathbb{R}$ by $r + \sum\{\alpha_i \mid a_i = 0,\, i = 1, \dots, k\}$). Summing the terms t^{a_i} with the same a_i, we can consider the sums

$$\bar{\alpha}_i := \sum\{\alpha_j \mid a_j = a_i\,,\, j = 1, \dots, k\}$$

as the coefficients in

$$x = r + \sum_{i \in I} \bar{\alpha}_i \cdot t^{a_i}$$

in ${}^\bullet\mathbb{R}$, where $I \subseteq \{1, \dots, k\}$, $\{a_i \mid i \in I\} = \{a_{\min I}, \dots, a_{\max I}\}$, and $a_i \neq a_j$ for any $i,\, j \in I$ with $i \neq j$. Neglecting $\bar{\alpha}_i$ if $\bar{\alpha}_i = 0$ and renumbering a_i for $i \in I$ in such a way that $a_i < a_j$ if $i,\, j \in I$ with $i < j$, we prove the existence. Note that, if $x = r \in \mathbb{R}$, then $I = \varnothing$ in the final step of this proof. $\square$

Uniqueness. Suppose that

$$x = r + \sum_{i=1}^{k} \alpha_i \cdot t^{a_i} = s + \sum_{j=1}^{N} \beta_j \cdot t^{b_j} \tag{10}$$

in ${}^\bullet\mathbb{R}$, where α_i, β_j, a_i, and b_j satisfy the assumptions in Theorem 5. First of all, ${}^\circ x = x(0) = r = s$ because $a_i,\, b_j > 0$. Hence $\alpha_1 t^{a_1} - \beta_1 t^{b_1} + \sum_i \alpha_i \cdot t^{a_i} - \sum_j \beta_j \cdot t^{b_j} = o(t)$. By contradiction, if the inequality $a_1 < b_1$ were valid, then, collecting the terms t^{a_1}, we would have

$$\alpha_1 - \beta_1 t^{b_1 - a_1} + \sum_i \alpha_i \cdot t^{a_i - a_1} - \sum_j \beta_j \cdot t^{b_j - a_1} = \frac{o(t)}{t} \cdot t^{1 - a_1}. \tag{11}$$

In (11), we have $\beta_1 t^{b_1 - a_1} \to 0$ as $t \to 0^+$ because $a_1 < b_1$ by assumption; $\sum_i \alpha_i \cdot t^{a_i - a_1} \to 0$ because $a_1 < a_i$ for $i = 2, \dots, k$; $\sum_j \beta_j \cdot t^{b_j - a_1} \to 0$ because $a_1 < b_1 < b_j$ for $j = 2, \dots, N$, and, finally t^{1-a_1} is bounded because $a_1 \leqslant 1$. Hence, for $t \to 0^+$, we obtain $\alpha_1 = 0$, which contradicts condition (3) of Theorem 5. We can argue in a similar way for $b_1 < a_1$, which gives $a_1 = b_1$. This together with equation (11) yield

$$\alpha_1 - \beta_1 + \sum_i \alpha_i \cdot t^{a_i - a_1} - \sum_j \beta_j \cdot t^{b_j - a_1} = \frac{o(t)}{t} \cdot t^{1 - a_1} \tag{12}$$

and hence $\alpha_1 = \beta_1$ for $t \to 0^+$. We can now restart from (12) to prove that $a_2 = b_2$, $\alpha_2 = \beta_2$, etc., in the same way. At the end, we must have $k = N$, because otherwise, if, say, $k < N$ at the end of the above recursion process, then we would have $\sum_{j=k+1}^{N} \beta_j \cdot t^{b_j} = o(t)$. Collecting the terms containing $t^{b_{k+1}}$, we obtain

$$t^{b_{k+1} - 1} \cdot [\beta_{k+1} + \beta_{k+2} \cdot t^{b_{k+2} - b_{k+1}} + \dots + \beta_N \cdot t^{\beta_N - \beta_{k+1}}] \to 0. \tag{13}$$

In this sum, $\beta_{k+j} \cdot t^{b_{k+j} - b_{k+1}} \to 0$ as $t \to 0^+$, because $b_{k+1} < b_{k+j}$ for $j > 1$, and hence

$$\beta_{k+1} + \beta_{k+2} \cdot t^{b_{k+2} - b_{k+1}} + \dots + \beta_N \cdot t^{\beta_N - \beta_{k+1}} \to \beta_{k+1} \neq 0.$$

Thus, it follows from (13) that $t^{b_{k+1} - 1} \to 0$, i.e., $b_{k+1} > 1$, which contradicts the uniqueness assumption $b_{k+1} \leqslant 1$.

Let us note explicitly that the uniqueness proof enables us also to claim that the decomposition is well defined in ${}^\bullet\mathbb{R}$, i.e., if $x = y$ in ${}^\bullet\mathbb{R}$, then the decompositions of x and y are equal. $\square$

Using this theorem, we introduce two symbols. The first one stresses the potential nature of an infinitesimal $x \in {}^\bullet\mathbb{R}$, and the other its actual nature.

Definition 7. For $x \in {}^\bullet\mathbb{R}$, we say that

$$x = r + \sum_{i=1}^{k} \alpha_i \cdot t^{a_i} \text{ is the potential decomposition } (\text{of } x) \tag{14}$$

if and only if conditions (1), (2), and (3) of Theorem 5 hold. Certainly, it is implicitly assumed that equality in (14) is the equality in ${}^\bullet\mathbb{R}$.

For example, $x = 1 + t^{1/3} + t^{1/2} + t$ is a decomposition, because we face increasing powers of t. The only decomposition of a standard real $r \in \mathbb{R}$ is the void one, i.e., that with $k = 0$ and $\alpha = a \colon \varnothing \longrightarrow \mathbb{R}$; to see that this is the case, it suffices to go along the existence proof again with this case $x = r \in \mathbb{R}$ (or to prove it directly, e.g., by contradiction).

Definition 8. Considering $t^{a_i} = \mathrm{d}t_{1/a_i}$, we can also use the following notation, stressing the fact that $x \in {}^\bullet\mathbb{R}$ is an actual infinitesimal:

$$x = {}^\circ x + \sum_{i=1}^{k} {}^\circ x_i \cdot \mathrm{d}t_{b_i}, \tag{15}$$

where the notation ${}^\circ x_i := \alpha_i$ and $b_i := 1/a_i$ is used; thus, the condition uniquely identifying all b_i is $b_1 > b_2 > \cdots > b_k \geqslant 1$. We refer to (15) as the *actual decomposition* of x or simply the *decomposition* of x. We also use the notation $\mathrm{d}^i x := {}^\circ x_i \cdot \mathrm{d}t_{b_i}$ (and simply $\mathrm{d}x := \mathrm{d}^1 x$) and refer to ${}^\circ x_i$ as the ith *standard part of* x and to $\mathrm{d}^i x$ as the ith *infinitesimal part of* x or the ith *differential* of x. We can also write $x = {}^\circ x + \sum_i \mathrm{d}^i x$; in this notation, the summands are uniquely determined (and the number of summands as well). Finally, if $k \geqslant 1$ (i.e., if $x \in {}^\bullet\mathbb{R} \setminus \mathbb{R}$), then we set $\omega(x) := b_1$ and $\omega_i(x) := b_i$. The real number $\omega(x) = b_1$ is the greatest order in the actual decomposition (15) corresponding to the smallest order in the potential decomposition (14), and it is called the *order* of the Fermat real $x \in {}^\bullet\mathbb{R}$. The number $\omega_i(x) = b_i$ is referred to as the ith order of x. If $x \in \mathbb{R}$, we set $\omega(x) := 0$ and $\mathrm{d}^i x := 0$. Note that $\omega(x) = \omega(\mathrm{d}x)$, $\mathrm{d}(\mathrm{d}x) = \mathrm{d}x$ in general and, using the notation of (14), we have $\omega(x) = 1/a_1$.

Example. If $x = 1 + t^{1/3} + t^{1/2} + t$, then ${}^\circ x = 1$ and $\mathrm{d}x = \mathrm{d}t_3$, and hence x is a third-order infinitesimal, i.e., $\omega(x) = 3$, $\mathrm{d}^2 x = \mathrm{d}t_2$, and $\mathrm{d}^3 x = \mathrm{d}t$; finally, the standard parts are ${}^\circ x_i = 1$.

6. THE IDEALS D_k

In this section, we introduce the set of nilpotent infinitesimals corresponding to a kth-order neighborhood of 0. The restriction of every smooth function to this neighborhood is a polynomial of order k given by its kth-order Taylor formula (without any remainder). We begin with a theorem characterizing infinitesimals whose order is less than k.

Theorem 9. *If $x \in {}^\bullet\mathbb{R}$ and $k \in \mathbb{N}_{>1}$, then $x^k = 0$ in ${}^\bullet\mathbb{R}$ if and only if ${}^\circ x = 0$ and $\omega(x) < k$.*

Proof. If $x^k = 0$, then applying the standard part mapping gives ${}^\circ(x^k) = ({}^\circ x)^k = 0$, and hence ${}^\circ x = 0$. Moreover, $x^k = 0$ yields $x_t^k = o(t)$, and hence $\left(x_t/t^{1/k}\right)^k \to 0$ and $x_t/t^{1/k} \to 0$. Representing this condition by using the potential decomposition

$$x = \sum_{i=1}^{k} \alpha_i \cdot t^{a_i}$$

of x (this yields $\omega(x) = 1/a_1$) gives

$$\lim_{t \to 0^+} \sum_i \alpha_i \cdot t^{a_i - \frac{1}{k}} = 0 = \lim_{t \to 0^+} t^{a_1 - \frac{1}{k}} \cdot \left[\alpha_1 + \alpha_2 \cdot t^{a_2 - a_1} + \cdots + \alpha_k \cdot t^{a_k - a_1}\right].$$

However,

$$\alpha_1 + \alpha_2 \cdot t^{a_2 - a_1} + \cdots + \alpha_k \cdot t^{a_k - a_1} \to \alpha_1 \neq 0,$$

and hence $t^{a_1 - \frac{1}{k}} \to 0$ and $a_1 > \frac{1}{k}$, i.e., $\omega(x) < k$.
Vice versa, if $^\circ x = 0$ and $\omega(x) < k$, then

$$x = \sum_{i=1}^{k} \alpha_i \cdot t^{a_i} + o(t) \quad \text{and} \quad \lim_{t\to 0^+} x_t / t^{1/k} = \lim_{t \to 0^+} \sum_i \alpha_i \cdot t^{a_i - 1/k} + \lim_{t\to 0^+} o(t)/t \cdot t^{1-1/k}.$$

On the other hand, $t^{1-1/k} \to 0$ because $k > 1$ and $t^{a_i - 1/k} \to 0^+$ because $1/a_i \leqslant 1/a_1 = \omega(x) < k$, and hence, $x^k = 0$ in $^\bullet\mathbb{R}$. $\square$

If we want a smooth function to be equal to its kth Taylor formula in a kth-order infinitesimal neighborhood, we are to use infinitesimals which can delete the remainder, i.e., such that $h^{k+1} = 0$. The previous theorem enables us to extend the definition of the ideal D_k to real-number subscripts rather than positive integers k only.

Definition 10. If $a \in \mathbb{R}_{>0} \cup \{\infty\}$, then $D_a := \{x \in {}^\bullet\mathbb{R} \mid {}^\circ x = 0,\ \omega(x) < a + 1\}$. Moreover, we simply denote D_1 by D.

(1) If $x = \mathrm{d}t_3$, then $\omega(x) = 3$ and $x \in D_3$. In general, $\mathrm{d}t_k \in D_a$ if and only if $\omega(\mathrm{d}t_k) = k < a+1$. For example, $\mathrm{d}t_k \in D$ if and only if $1 \leqslant k < 2$.
(2) $D_\infty = \bigcup_a D_a = \{x \in {}^\bullet\mathbb{R} \mid {}^\circ x = 0\}$ is the set of all infinitesimals of $^\bullet\mathbb{R}$.
(3) $D_0 = \{0\}$ because the only infinitesimal whose order is strictly less than 1 is $x = 0$ by the definition of order (see Definition 8).

The following theorem gathers several expected properties of the sets D_a and of the order of an infinitesimal $\omega(x)$.

Theorem 11. *Let $a, b \in \mathbb{R}_{>0}$ and $x, y \in D_\infty$. Then the following assertions hold.*
(1) $a \leqslant b \implies D_a \subseteq D_b$.
(2) $x \in D_{\omega(x)}$.
(3) $a \in \mathbb{N} \implies D_a = \{x \in {}^\bullet\mathbb{R} \mid x^{a+1} = 0\}$.
(4) $x \in D_a \implies x^{\lceil a \rceil + 1} = 0$.
(5) $x \in D_\infty \setminus \{0\}$ *and* $k = [\omega(x)] \implies x \in D_k \setminus D_{k-1}$.
(6) $d(x \cdot y) = dx \cdot dy$.
(7) $x \cdot y \neq 0 \implies 1/\omega(x \cdot y) = 1/\omega(x) + 1/\omega(y)$.
(8) $x + y \neq 0 \implies \omega(x + y) = \omega(x) \vee \omega(y)$.
(9) D_a *is an ideal.*

In this statement, if $r \in \mathbb{R}$, then $\lceil r \rceil$ is the *ceiling* of the real r, i.e., the unique integer $\lceil r \rceil \in \mathbb{Z}$ such that $\lceil r \rceil - 1 < r \leqslant \lceil r \rceil$. Moreover, if $r, s \in \mathbb{R}$, then $r \vee s := \max(r, s)$.

Property (4) in Theorem 11 cannot be proved by substituting the ceiling $\lceil a \rceil$ with the integer part $[a]$. In fact, if $a = 1.2$ and $x = \mathrm{d}t_{2.1}$, then $\omega(x) = 2.1$ and $[a] + 1 = 2$, and therefore, $x^{[a]+1} = x^2 = \mathrm{d}t_{\frac{2.1}{2}} \neq 0$ in $^\bullet\mathbb{R}$, whereas $\lceil a \rceil + 1 = 3$ and $x^3 = \mathrm{d}t_{\frac{2.1}{3}} = 0$.

Finally, note the following increasing sequence of ideals/neighborhoods of zero:

$$\{0\} = D_0 \subset D = D_1 \subset D_2 \subset \cdots \subset D_k \subset \cdots \subset D_\infty. \tag{16}$$

By (16) and by the property $\mathrm{d}t_a = 0$ for $a < 1$, $\mathrm{d}t$ is the smallest infinitesimal and $\mathrm{d}t_2$, $\mathrm{d}t_3$, etc., are greater infinitesimals. We shall see that this agrees with order properties of these infinitesimals.

7. PRODUCTS OF POWERS OF NILPOTENT INFINITESIMALS

In this section, we introduce instruments used to decide whether or not a product of the form $h_1^{i_1} \cdots h_n^{i_n}$, $h_k \in D_\infty \setminus \{0\}$, vanishes or belongs to some D_k. Generally speaking, this problem is nontrivial in a ring (e.g., in SDG, there is no effective procedure to solve this problem; see, e.g., Lavendhomme [22]), and its solution is very useful in proofs of infinitesimal Taylor formulas.

Theorem 12. *Let $h_1,\dots,h_n \in D_\infty \setminus \{0\}$ and $i_1,\dots,i_n \in \mathbb{N}$. Then the following assertions hold.*

(1) $h_1^{i_1}\cdots h_n^{i_n}=0 \iff \sum_{k=1}^n i_k/\omega(h_k)>1$.

(2) $h_1^{i_1}\cdots h_n^{i_n}\neq 0 \implies 1/\omega(h_1^{i_1}\cdots h_n^{i_n})=\sum_{k=1}^n i_k/\omega(h_k)$.

Proof. Let

$$h_k=\sum_{r=1}^{N_k}\alpha_{kr}t^{a_{kr}} \tag{17}$$

be the potential decomposition of h_k for $k=1,\dots,n$. Then, by Definitions 7 (of potential decomposition) and 8 (of order), $0<a_{k1}<a_{k2}<\cdots<a_{kN_k}\leqslant 1$ and $j_k:=\omega(h_k)=1/a_{k1}$. Hence $1/j_k\leqslant a_{kr}$ for every $r=1,\dots,N_k$. Therefore, it follows from (17) by collecting the terms containing t^{1/j_k} that

$$h_k=t^{1/j_k}\cdot\left(\alpha_{k1}+\alpha_{k2}t^{a_{k2}-1/j_k}+\cdots+\alpha_{kN_k}t^{a_{kN_k}-1/j_k}\right),$$

and hence

$$\begin{aligned}h_1^{i_1}\cdots h_n^{i_n}=t^{i_1/j_1+\cdots+i_n/j_n}\cdot\left(\alpha_{11}+\alpha_{12}t^{a_{12}-1/j_1}+\cdots+\alpha_{1N_1}t^{a_{1N_1}-1/j_1}\right)^{i_1}\\ \cdots\left(\alpha_{n1}+\alpha_{n2}t^{a_{n2}-1/j_n}+\cdots+\alpha_{nN_n}t^{a_{nN_n}-1/j_n}\right)^{i_n}.\end{aligned} \tag{18}$$

Hence, if $\sum_k i_k/j_k>1$, then $t^{\frac{i_1}{j_1}+\cdots+\frac{i_n}{j_n}}=0$ in ${}^\bullet\mathbb{R}$, and thus $h_1^{i_1}\cdots h_n^{i_n}=0$ as well. Vice versa, if $h_1^{i_1}\cdots h_n^{i_n}=0$, then the right-hand side of (18) is an $o(t)$ as $t\to 0^+$, i.e.,

$$\begin{aligned}t^{\frac{i_1}{j_1}+\cdots+\frac{i_n}{j_n}-1}\cdot\left(\alpha_{11}+\alpha_{12}t^{a_{12}-\frac{1}{j_1}}+\cdots+\alpha_{1N_1}t^{a_{1N_1}-\frac{1}{j_1}}\right)^{i_1}\\ \cdots\left(\alpha_{n1}+\alpha_{n2}t^{a_{n2}-\frac{1}{j_n}}+\cdots+\alpha_{nN_n}t^{a_{nN_n}-\frac{1}{j_n}}\right)^{i_n}\to 0.\end{aligned}$$

However,

$$\left(\alpha_{k1}+\alpha_{k2}t^{a_{k2}-1/j_k}+\cdots+\alpha_{kN_k}t^{a_{kN_k}-1/j_k}\right)^{i_k}\to\alpha_k^{i_k}\neq 0,$$

and thus we must have $i_1/j_1+\cdots+i_n/j_n-1>0$. This completes the proof of part 1.
To prove part 2, it suffices to apply recursively property 7 of Theorem 11. $\square$

Example 13. *The following equality holds:*

$$\omega(\,dt_{a_1}^{i_1}\cdots\,dt_{a_n}^{i_n})^{-1}=\sum_k i_k/\omega(\,dt_{a_k})=\sum_k i_k/a_k$$

and $dt_{a_1}^{i_1}\cdots\,dt_{a_n}^{i_n}=0$ *if and only if* $\sum_k i_k/a_k>1$, *and thus, e.g.,* $dt\cdot h=0$ *for every* $h\in D_\infty$.

The following corollary gives a necessary and sufficient condition for $h_1^{i_1}\cdots h_n^{i_n}\in D_p\setminus\{0\}$.

Corollary 14. *Under the assumptions of Theorem 12, suppose that $p\in\mathbb{R}_{>0}$. Then*

$$h_1^{i_1}\cdots h_n^{i_n}\in D_p\setminus\{0\}\quad\iff\quad 1/(p+1)<\sum_{k=1}^n i_k/\omega(h_k)\leqslant 1.$$

Let h, $k\in D$; in this case, $\sum_k i_k/(j_k+1)=1/2+1/2=1$, and thus,

$$h\cdot k=0. \tag{19}$$

This is a great conceptual difference between Fermat reals and the ring of SDG, where the product of two first-order infinitesimal is not necessarily zero. The consequences of this property of

Fermat reals arrive very deeply in the development of the theory of Fermat reals, forcing us, e.g., to develop several new concepts if we want to generalize the derivation formula (4) to functions defined on infinitesimal domains, like $f\colon D \longrightarrow {}^\bullet\mathbb{R}$ (see Giordano [16]). We only mention here that, looking at the simple Definition 4, formula (19) has an intuitively clear meaning, and, to preserve this intuition, we keep this relation instead of changing the theory completely toward a less intuitive one.

Let us note explicitly that the possibility to prove these results about products of powers of nilpotent infinitesimals is essentially tied with the choice of little-oh polynomials in the definition of the equivalence relation $\sim$ in Definition 2. Equally effective and useful results cannot be proved for a more general family of nilpotent functions (see, e.g., Giordano [15]).

8. IDENTITY PRINCIPLE FOR POLYNOMIALS AND INVERTIBLE FERMAT REALS

In this section, we prove that, if a polynomial $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ of ${}^\bullet\mathbb{R}$ is identically zero, then $a_k = 0$ for all $k = 0, \ldots, n$. To prove this conclusion, it suffices to mean "identically zero" as "equal to zero for every x belonging to an extension of an open subset of $\mathbb{R}$." Therefore, we first define the extension.

Definition 15. If U is an open subset of $\mathbb{R}^n$, then ${}^\bullet U := \{x \in {}^\bullet\mathbb{R}^n \mid {}^\circ x \in U\}$. Here the symbol ${}^\bullet\mathbb{R}^n$ stands for ${}^\bullet\mathbb{R}^n := {}^\bullet\mathbb{R} \times \overset{n}{\ldots\ldots} \times {}^\bullet\mathbb{R}$.

The identity principle for polynomials can now be stated as follows (and proved in the standard way by using Vandermonde matrices).

Theorem 16. *Let $a_0, \ldots, a_n \in {}^\bullet\mathbb{R}$ and U be an open neighborhood of 0 in $\mathbb{R}$ such that*

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0 \quad \text{in } {}^\bullet\mathbb{R} \quad \forall x \in {}^\bullet U. \tag{20}$$

Then $a_0 = a_1 = \cdots = a_n = 0$ in ${}^\bullet\mathbb{R}$.

Let us now see more formally that, to prove (3), we must embed the reals $\mathbb{R}$ in a ring containing nilpotent elements rather than in a field. In fact, applying (3) to the function $f(h) = h^2$ for $h \in D$, where $D \subseteq {}^\bullet\mathbb{R}$ is a given subset of ${}^\bullet\mathbb{R}$, we obtain $f(h) = h^2 = f(0) + h \cdot f'(0) = 0$ for any $h \in D$. It is assumed here that the relation $f'(0) = 0$ is preserved when passing from $\mathbb{R}$ to ${}^\bullet\mathbb{R}$. In other words, if D and $f(h) = h^2$ verify (3), then each element $h \in D$ of this kind is a type of a number whose square is zero.

Since property (3) cannot thus hold for a field, we need a sufficiently good family of cancellation laws as substitutes. The simplest law of this kind is also useful to prove the uniqueness of (4).

Theorem 17. *If $x \in {}^\bullet\mathbb{R}$ is a Fermat real and $r, s \in \mathbb{R}$ are standard real numbers, then $x \cdot r = x \cdot s$ in ${}^\bullet\mathbb{R}$ and $x \neq 0$ imply $r = s$.*

Proof. It follows from Definition 4 of the equality relation in ${}^\bullet\mathbb{R}$ and from the assumption $x \cdot r = x \cdot s$ that $\lim_{t \to 0^+} x_t \cdot (r - s)/t = 0$. However, for $r \neq s$, this would imply that $\lim_{t \to 0^+} x_t/t = 0$, i.e., $x = 0$ in ${}^\bullet\mathbb{R}$, and this contradicts the assumption $x \neq 0$. $\square$

The last result of this section takes its ideas from similar situations of formal power series and also gives a formula for the inverse of an invertible Fermat real.

Theorem 18. *Let $x = {}^\circ x + \sum_{i=1}^n {}^\circ x_i \cdot \mathrm{d}t_{a_i}$ be the decomposition of a Fermat real $x \in {}^\bullet\mathbb{R}$. Then x is invertible if and only if ${}^\circ x \neq 0$. In this case,*

$$1/x = 1/{}^\circ x \cdot \sum_{j=0}^{+\infty} (-1)^j \cdot \Big(\sum_{i=1}^n {}^\circ x_i / {}^\circ x \cdot \mathrm{d}t_{a_i} \Big)^j. \tag{21}$$

In formula (21), the series is actually a finite sum, because any $\mathrm{d}t_{a_i}$ is nilpotent, for instance, we have $(1 + \mathrm{d}t_2)^{-1} = 1 - \mathrm{d}t_2 + \mathrm{d}t_2^2 - \mathrm{d}t_2^3 + \cdots = 1 - \mathrm{d}t_2 + \mathrm{d}t$ because $\mathrm{d}t_2^3 = 0$.

Proof. If $x \cdot y = 1$ for some $y \in {}^\bullet\mathbb{R}$, then, taking the standard parts of each side, we have ${}^\circ x \cdot {}^\circ y = 1$, and hence ${}^\circ x \neq 0$. Vice versa, let

$$y := {}^\circ x^{-1} \cdot \sum_{j=0}^{+\infty} (-1)^j \cdot \Big(\sum_i {}^\circ x_i / {}^\circ x \, \mathrm{d}t_{a_i} \Big)^j \quad \text{and} \quad h := x - {}^\circ x = \sum_i {}^\circ x_i \, \mathrm{d}t_{a_i} \in D_\infty.$$

Then we can also write

$$y = {}^\circ x^{-1} \cdot \sum_{j=0}^{+\infty} (-1)^j \cdot h^j / {}^\circ x^j .$$

However, since $h \in {}^\bullet\mathbb{R}$ is a little-oh polynomial with $h(0) = 0$, it is also continuous, and hence, for any $t \in (-\delta, \delta)$, we have $|h_t / {}^\circ x| < 1$ for a sufficiently small $\delta > 0$. Therefore,

$$\forall t \in (-\delta, \delta): \quad y_t = \frac{1}{{}^\circ x} \cdot \left(1 + \frac{h_t}{{}^\circ x} \right)^{-1} = \frac{1}{{}^\circ x + h_t} = \frac{1}{x_t}.$$

This relation and Definition 4 yield $x \cdot y = 1$ in ${}^\bullet\mathbb{R}$. $\square$

9. DERIVATION FORMULA

In this section, we give a proof of (4), which was the principal motivation for the construction of the ring of Fermat reals ${}^\bullet\mathbb{R}$. In any case, before proving the derivation formula, we must extend a given smooth function $f \colon \mathbb{R} \longrightarrow \mathbb{R}$ to a certain function ${}^\bullet f \colon {}^\bullet\mathbb{R} \longrightarrow {}^\bullet\mathbb{R}$.

Definition 19. Let A be an open subset of $\mathbb{R}^n$, let $f \colon A \longrightarrow \mathbb{R}$ be a smooth function, and let $x \in {}^\bullet A$. Then we write ${}^\bullet f(x) := f \circ x$.

This definition is correct, because little-oh polynomials are preserved by smooth functions and f is locally Lipschitzian. Therefore,

$$\left| \frac{f(x_t) - f(y_t)}{t} \right| \leqslant K \cdot \left| \frac{x_t - y_t}{t} \right| \quad \forall t \in (-\delta, \delta)$$

for a sufficiently small δ and for some constant K, and hence, if $x = y$ in ${}^\bullet\mathbb{R}$, then also ${}^\bullet f(x) = {}^\bullet f(y)$ in ${}^\bullet\mathbb{R}$.

The function ${}^\bullet f$ is an extension of f, i.e., ${}^\bullet f(r) = f(r)$ in ${}^\bullet\mathbb{R}$ for any $r \in \mathbb{R}$, which follows directly from the definition of equality in ${}^\bullet\mathbb{R}$ (in Definition 4). Thus, we can still use the symbol $f(x)$ both for $x \in {}^\bullet\mathbb{R}$ and $x \in \mathbb{R}$ without confusion. After the introduction of the extension of smooth functions, we can also state the following useful *elementary transfer theorem* for equalities, whose proof follows directly from the above definitions.

Theorem 20. *Let A be an open subset of $\mathbb{R}^n$, and τ, $\sigma : A \longrightarrow \mathbb{R}$ be smooth functions. Then* $:$ ${}^\bullet\tau(x) = {}^\bullet\sigma(x)$ *for any* $x \in {}^\bullet A$ *if and only if*$:$ $\tau(r) = \sigma(r)$ *for any* $r \in A$.

Let us now prove the derivation formula (4).

Theorem 21. *Let A be an open set in $\mathbb{R}$, $x \in A$ and $f : A \longrightarrow \mathbb{R}$ a smooth function, then*

$$\exists!\, m \in \mathbb{R} \;\; \forall h \in D: \;\; f(x+h) = f(x) + h \cdot m. \tag{22}$$

In this case, $m = f'(x)$, where $f'(x)$ stands for the ordinary derivative of f at x.

Proof. The *uniqueness* follows from the previous cancellation law theorem, Theorem 17. Indeed, if $m_1 \in \mathbb{R}$ and $m_2 \in \mathbb{R}$ verify (22), then $h \cdot m_1 = h \cdot m_2$ for every $h \in D$. However, there is a nonzero first-order infinitesimal, e.g., $\mathrm{d}t \in D$, and thus, Theorem 17 implies that $m_1 = m_2$.

To prove the *existence* part, take $h \in D$. Then $h^2 = 0$ in ${}^\bullet\mathbb{R}$, i.e., $h_t^2 = o(t)$ as $t \to 0^+$. However, f is smooth, and hence it follows from the second-order Taylor's formula that

$$f(x + h_t) = f(x) + h_t \cdot f'(x) + (h_t^2/2) \cdot f''(x) + o(h_t^2).$$

Moreover,

$$\frac{o(h_t^2)}{t} = (o(h_t^2)/h_t^2) \cdot (h_t^2/t) \to 0 \quad \text{as} \quad t \to 0^+,$$

and thus $(h_t^2/2) \cdot f''(x) + o(h_t^2) = o_1(t)$ as $t \to 0^+$, which gives $f(x+h_t) = f(x) + h_t \cdot f'(x) + o_1(t)$ as $t \to 0^+$, i.e., $f(x+h) = f(x) + h \cdot f'(x)$ in ${}^\bullet\mathbb{R}$. This proves the existence part, because $f'(x) \in \mathbb{R}$. □

For example, $e^h = 1 + h$, $\sin(h) = h$, and $\cos(h) = 1$ for every $h \in D$.

Analogously, we can prove the following infinitesimal Taylor's formula.

Lemma 22. *Let A be an open set in $\mathbb{R}^d$, $x \in A$, $n \in \mathbb{N}_{>0}$, and $f\colon A \longrightarrow \mathbb{R}$ a smooth function. Then $\forall h \in D_n^d : \; f(x + h) = \sum_{\substack{j \in \mathbb{N}^d \\ |j| \leqslant n}} (h^j/j!) \cdot \partial^{|j|} f / \partial x^j(x)$.*

For example, $\sin(h) = h - h^3/6$ if $h \in D_3$, and thus, $h^4 = 0$.

It is possible to generalize several results of the present work to functions of class $\mathcal{C}^n$ only, instead of smooth ones. However, it is an explicit purpose of this work to simplify statements of results, definitions, and notations, even if, as a result of this searching for simplicity, its applicability holds only for a more restricted class of functions. Some more general results stated for $\mathcal{C}^n$ functions (but less simple) can be found in Giordano [15].

Note that $m = f'(x) \in \mathbb{R}$, i.e., the slope is a standard real number, and we can use the previous formula with standard real numbers x only rather than with a generic $x \in {}^\bullet\mathbb{R}$. We shall remove this limitation in subsequent works (see also Giordano [16]).

Applying this theorem to the smooth function $p(r) := \int_x^{x+r} f(t)\,\mathrm{d}t$ (where f is assumed to be smooth), we immediately obtain the following result, which is frequently used in informal calculations.

Corollary 23. *Let A be open in $\mathbb{R}$, let $x \in A$, and let $f\colon A \longrightarrow \mathbb{R}$ be smooth. Then*

$$\forall h \in D : \int_x^{x+h} f(t)\, dt = h \cdot f(x).$$

Moreover, $f(x) \in \mathbb{R}$ is uniquely determined by this relation.

10. NILPOTENT INFINITESIMALS AND ORDER PROPERTIES

In mathematics, like in other disciplines, the layout of a work reflects the personal philosophical ideas of the authors. In particular, the present work is based on the idea that a good mathematical theory is able to construct a good dialectic between formal properties proved in the theory and their informal interpretations. The dialectic has to be, as far as possible, in both directions: theorems proved in the theory should have a clear and useful intuitive interpretation and, on the other hand, the intuition corresponding to the theory has to be able to suggest true sentences, i.e., conjectures or sketches of proofs that can then be converted into rigorous proofs.

In a theory of new numbers, like the present one (concerning Fermat reals), the introduction of an order relation can be a hard test of the excellence of this dialectic between formal properties and their informal interpretations. Indeed, if we introduce a new ring of numbers (like ${}^\bullet\mathbb{R}$) by extending the real field $\mathbb{R}$, we want the new order relation, defined on the new ring, to extend the standard one on $\mathbb{R}$. This extension naturally leads to the desire to find a geometrical representation of the new numbers, according to the above principle of having a good formal/informal dialectic.

We begin this section by showing that, in our setting, there is a strong relationship between order properties and algebraic properties. In particular, we claim that it is impossible to have simultaneously good order properties and uniqueness without limitations in the derivation formula. In the following theorem, we see that the property $h \cdot k = 0$ is a general consequence of the assumption that there is a total order on D.

Theorem 24. *Let $(R, \leqslant)$ be a generic ordered ring, and let $D \subseteq R$ be such that*

(1) $0 \in D$,
(2) $\forall h \in D: \ h^2 = 0$ *and* $-h \in D$,
(3) $(D, \leqslant)$ *is a total order,*

then $h \cdot k = 0$ *for every* $h, k \in D$.

This theorem implies that, if a total order in our theory of infinitesimal numbers is desired and if $D = \{h \mid h^2 = 0\}$, then we must accept that the product of any two elements of D vanishes. For example, if we think that a geometric representation of infinitesimals is impossible without the trichotomy law, then the product of two first-order infinitesimals in this theory must be zero.

Proof. Let $h, k \in D$ be two elements of D. By assumption, we have $0, -h, -k \in D$, and hence all these elements are comparable with respect to the order relation $\leqslant$, because, by assumption, this relation is total on D. For example, $h \leqslant k$ or $k \leqslant h$. Consider the case $h \leqslant k$ only, because the case $k \leqslant h$ can be studied in a similar way by transposing h with k.

First sub-case: $k \geqslant 0$. Multiplying the relation $h \leqslant k$ by $k \geqslant 0$, we obtain

$$hk \leqslant k^2. \tag{23}$$

If $h \geqslant 0$, then multiplying by $k \geqslant 0$ gives $0 \leqslant hk$, and thus it follows from (23) that $0 \leqslant hk \leqslant k^2 = 0$, and hence $hk = 0$.

If $h \leqslant 0$, then multiplying by $k \geqslant 0$ gives

$$hk \leqslant 0, \tag{24}$$

Furthermore, if $h \geqslant -k$, then multiplying by $k \geqslant 0$ gives $hk \geqslant -k^2$ and $0 \geqslant hk \geqslant -k^2 = 0$ by (24), and hence $hk = 0$.

Otherwise, if $h \leqslant -k$, then multiplying by $-h \geqslant 0$ gives $-h^2 = 0 \leqslant hk \leqslant 0$ by (24), and hence $hk = 0$. This completes the discussion of the case $k \geqslant 0$.

Second sub-case: $k \leqslant 0$. In this case, $h \leqslant k \leqslant 0$. Multiplying by $h \leqslant 0$ gives $h^2 = 0 \geqslant hk \geqslant 0$, and hence $hk = 0$. $\square$

Thus, the trichotomy law and uniqueness in a possible derivation formula of the form

$$\exists!\, m \in R: \ \forall h \in D: \ f(h) = f(0) + h \cdot m \tag{25}$$

framed in the ring R of Theorem 24 are incompatible. In fact, if $a, b \in D$ are elements of $D \subseteq R$, then both a and b can play the role of $m \in R$ in (25) for the linear function $f : h \in D \mapsto h \cdot a = 0 \in R$. Thus, if the derivation formula (25) is applied to linear functions (or even to constant functions), then the uniqueness property for this formula cannot hold in the ring R.

In the next section, we introduce a natural and meaningful total order relation on ${}^\bullet\mathbb{R}$. Therefore, the previous theorem, Theorem 24, strongly motivates the rule that the product of two first-order infinitesimals must be zero for the ring of Fermat reals ${}^\bullet\mathbb{R}$, and hence, for the derivation formula in ${}^\bullet\mathbb{R}$, the uniqueness cannot hold in its strongest form. Since we shall also see that the order relation enables us to have a geometric representation of Fermat reals, we can summarize the conclusions of this section by saying that the uniqueness in the derivation formula is incompatible with a natural geometric interpretation of Fermat reals, and hence, with a good dialectic between formal properties and informal interpretations of this theory.

11. ORDER RELATION

By the above sections, one can draw the conclusion that the ring of Fermat reals ${}^\bullet\mathbb{R}$ is essentially "the little-oh" calculus. On the other hand, the Fermat reals give us more flexibility than this calculus. Namely, when working with ${}^\bullet\mathbb{R}$, we need no remainders made of "little-oh," and we can neglect them and use the powerful algebraic calculus with nilpotent infinitesimals. However, thinking of the elements of ${}^\bullet\mathbb{R}$ as new numbers (rather than simply as "little-oh functions") permits us to treat them in a different and new way, for example, to define an order relation on these numbers with a clear geometrical interpretation.

First of all, introduce the useful notation $\forall^0 t \geqslant 0: \ \mathcal{P}(t)$. Let us read the quantifier $\forall^0 t \geqslant 0$ by saying "*for every* $t \geqslant 0$ (*sufficiently*) *small*," to indicate that the property $\mathcal{P}(t)$ is true for all t in some right neighborhood of $t=0$ (recall that, by Definition 2, our little-oh polynomials are always defined on $\mathbb{R}_{\geqslant 0}$), i.e., $\exists \delta > 0: \ \forall t \in [0, \delta): \ \mathcal{P}(t)$.

The first heuristic idea to define an order relation is $x \leqslant y \iff x - y \leqslant 0 \iff \exists z: \ z = 0$ in ${}^\bullet\mathbb{R}$ and $x - y \leqslant z$. More formally,

Definition 25. Let x, $y \in {}^\bullet\mathbb{R}$. Then we say that $x \leqslant y$ if and only if there is a $z \in {}^\bullet\mathbb{R}$ such that $z = 0$ in ${}^\bullet\mathbb{R}$ and $\forall^0 t \geqslant 0: \ x_t \leqslant y_t + z_t$.

Recall that the condition $z = 0$ in ${}^\bullet\mathbb{R}$ is equivalent to the condition $z_t = o(t)$ as $t \to 0^+$. It is immediate that, equivalently, $x \leqslant y$ if and only if there are $x' = x$ and $y' = y$ in ${}^\bullet\mathbb{R}$ such that $x_t \leqslant y_t$ for every t sufficiently small. This also implies that the relation $\leqslant$ is well defined on ${}^\bullet\mathbb{R}$, i.e., if $x' = x$ and $y' = y$ in ${}^\bullet\mathbb{R}$ and $x \leqslant y$, then $x' \leqslant y'$ (recall that, to simplify the notation, we use little-oh polynomials directly as elements of ${}^\bullet\mathbb{R}$ rather than equivalence classes). As usual, we use the notation $x < y$ for $x \leqslant y$ and $x \neq y$.

Theorem 26. *The relation $\leqslant$ is an order, i.e., it is reflexive, transitive, and antisymmetric*; *it extends the order relation of $\mathbb{R}$, and $({}^\bullet\mathbb{R}, \leqslant)$ is an ordered ring. Finally, the following assertions are equivalent*:

(1) *$h \in D_\infty$, i.e., h is an infinitesimal,*
(2) $\forall r \in \mathbb{R}_{>0}: \ -r < h < r$.

Hence an infinitesimal can be thought of as a number with zero standard part or as a number smaller than every standard positive real number and greater than every standard negative real number.

Proof. Let us prove that $x \leqslant y$ and $w \geqslant 0$ imply $x \cdot w \leqslant y \cdot w$ only (the other ones are simple consequences of Definition 25). Suppose that

$$x_t \leqslant y_t + z_t \quad \forall^0 t \geqslant 0, \qquad w_t \geqslant z'_t \quad \forall^0 t \geqslant 0; \tag{26}$$

then $w_t - z'_t \geqslant 0$ for every t small, and hence $x_t \cdot (w_t - z'_t) \leqslant y_t \cdot (w_t - z'_t) + z_t \cdot (w_t - z'_t) \ \forall^0 t \geqslant 0$ by (26), which yields $x_t \cdot w_t \leqslant y_t \cdot w_t + (-x_t z'_t - y_t z'_t + z_t w_t - z_t z'_t) \ \forall^0 t \geqslant 0$. However, $-xz' - yz' + zw - zz' = 0$ in ${}^\bullet\mathbb{R}$, because $z = 0$ and $z' = 0$, and hence the conclusion follows. $\square$

Example. We have, for example, $\mathrm{d}t > 0$ and $\mathrm{d}t_2 - 3\,\mathrm{d}t > 0$, because $t^{1/2} > 3t$ for $t \geqslant 0$ sufficiently small, and hence $t^{1/2} - 3t > 0 \ \forall^0 t \geqslant 0$. Examples of this kind suggest the idea that our little-oh polynomials are always locally comparable with respect to the pointwise order relation, and this is the first step to prove that the trichotomy law holds for our order relation. In the following statement, we use the notation $\forall^0 t > 0 : \mathcal{P}(t)$, which naturally means that $\forall^0 t \geqslant 0: \ t \neq 0 \implies \mathcal{P}(t)$, where $\mathcal{P}(t)$ is a generic property depending on t.

Lemma 27. *Let x, $y \in {}^\bullet\mathbb{R}$. In this case, the following assertions hold.*

(1) $^\circ x < {}^\circ y \implies \forall^0 t \geqslant 0: \ x_t < y_t$.
(2) *If $^\circ x = {}^\circ y$, then* $\left(\forall^0 t > 0: \ x_t < y_t\right)$ *or* $\left(\forall^0 t > 0: \ x_t > y_t\right)$ *or* $(x = y$ *in* ${}^\bullet\mathbb{R})$.

Proof. Suppose that $^\circ x < {}^\circ y$. In this case, the continuous function $t \geqslant 0 \mapsto y_t - x_t \in \mathbb{R}$ takes the value $y_0 - x_0 > 0$. Hence, it is locally positive, i.e., $\forall^0 t \geqslant 0: \ x_t < y_t$. Now suppose that $^\circ x = {}^\circ y$ and introduce a notation for the potential decompositions of x and y (see Definition 7). By the definition of equality in ${}^\bullet\mathbb{R}$, we can always write

$$x_t = {}^\circ x + \sum_{i=1}^{N} \alpha_i \cdot t^{a_i} + z_t \quad \text{and} \quad y_t = {}^\circ y + \sum_{j=1}^{M} \beta_j \cdot t^{b_j} + w_t \qquad \forall t \geqslant 0,$$

where $x = {}^\circ x + \sum_{i=1}^N \alpha_i \cdot t^{a_i}$ and $y = {}^\circ y + \sum_{j=1}^M \beta_j \cdot t^{b_j}$ are the potential decompositions of x and y (hence $0 < \alpha_i < \alpha_{i+1} \leqslant 1$ and $0 < \beta_j < \beta_{j+1} \leqslant 1$), whereas w and z are little-oh polynomials such that $z_t = o(t)$ and $w_t = o(t)$ as $t \to 0^+$.

Case $a_1 < b_1$. In this case, the least power in the two decompositions is $\alpha_1 \cdot t^{a_1}$, and hence, we expect that the second possibility in the assertion holds if $\alpha_1 > 0$, otherwise, the first possibility holds if $\alpha_1 < 0$ (recall that we always have $\alpha_i \neq 0$ in the decomposition). Indeed, consider the condition $x_t < y_t$ for $t > 0$ and list some equivalent formulas:

$$\sum_{i=1}^N \alpha_i \cdot t^{a_i} < \sum_{j=1}^N \beta_j \cdot t^{b_j} + w_t - z_t,$$

$$t^{a_1} \cdot \Big[\alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1}\Big] < t^{a_1} \cdot \Big[\sum_{j=1}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1}\Big],$$

$$\alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} < \sum_{j=1}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1}.$$

Therefore, consider the function

$$f(t) := \sum_{j=1}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1} - \alpha_1 - \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} \quad \forall t \geqslant 0.$$

Write $(w_t - z_t) \cdot t^{-a_1} = ((w_t - z_t)/t) \cdot t^{1-a_1}$; we have $(w_t - z_t)/t \to 0$ as $t \to 0^+$, because $w_t = o(t)$ and $z_t = o(t)$. Further, $a_1 \leqslant 1$, and hence t^{1-a_1} is bounded in a right neighborhood of $t = 0$. Therefore, $(w_t - z_t) \cdot t^{-a_1} \to 0$, and the function f is continuous at $t = 0$ as well, because $a_1 < a_i$ and $a_1 < b_1 < b_j$. By continuity, the function f is locally strictly positive if and only if $f(0) = -\alpha_1 > 0$, and thus $\big(\forall^0 t > 0 :\ x_t < y_t\big) \iff \alpha_1 < 0$ and $\big(\forall^0 t > 0 :\ x_t > y_t\big) \iff \alpha_1 > 0$.

Case $a_1 > b_1$. We can argue in a similar way using b_1 and β_1 instead of a_1 and α_1.

Case $a_1 = b_1$. We shall exploit the above idea. Let us study the condition $x_t < y_t$. The relations

$$t^{a_1} \cdot \Big[\alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1}\Big] < t^{a_1} \cdot \Big[\beta_1 + \sum_{j=2}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1}\Big],$$

$$\alpha_1 + \sum_{i=2}^N \alpha_i \cdot t^{a_i - a_1} < \beta_1 + \sum_{j=2}^N \beta_j \cdot t^{b_j - a_1} + (w_t - z_t) \cdot t^{-a_1}$$

are equivalent ways to express this condition. Hence, exactly as was proved above, we can claim that $\alpha_1 < \beta_1$ implies $\forall^0 t > 0 :\ x_t < y_t$ and $\alpha_1 > \beta_1$ implies $\forall^0 t > 0 :\ x_t > y_t$.

Otherwise $\alpha_1 = \beta_1$, and we can restart the same reasoning by using a_2, b_2, α_2, β_2, etc. If $N = M$ (the number of summands in the decompositions), then, using this procedure, we can prove that $\forall t \geqslant 0 :\ x_t = y_t + w_t - z_t$, i.e., $x = y$ in ${}^\bullet\mathbb{R}$.

It remains to consider the case, e.g., $N < M$. Under this assumption, using the above procedure, we arrive at the following consequences of the condition $x_t < y_t$:

$$0 < \sum_{j>N} \beta_j \cdot t^{b_j} + w_t - z_t,$$

$$0 < t^{b_{N+1}} \cdot \Big[\beta_{N+1} + \sum_{j>N+1} \beta_j \cdot t^{b_j - b_{N+1}} + (w_t - z_t) \cdot t^{-b_{N+1}}\Big],$$

$$0 < \beta_{N+1} + \sum_{j>N+1} \beta_j \cdot t^{b_j - b_{N+1}} + (w_t - z_t) \cdot t^{-b_{N+1}}.$$

Hence $\beta_{N+1} > 0 \implies \forall^0 t > 0 :\ x_t < y_t$, and $\beta_{N+1} < 0 \implies \forall^0 t > 0 :\ x_t > y_t$.

This lemma can be used to find an equivalent formulation of the order relation.

Theorem 28. *Let $x,\, y \in {}^\bullet\mathbb{R}$, then*

(1) $x \leqslant y \iff \left(\forall^0 t > 0: \ x_t < y_t\right)$ *or* $\quad (x = y \text{ in } {}^\bullet\mathbb{R})$,

(2) $x < y \iff \left(\forall^0 t > 0: \ x_t < y_t\right)$ *and* $(x \neq y \text{ in } {}^\bullet\mathbb{R})$.

Proof. $\Rightarrow$ If ${}^\circ x < {}^\circ y$, then by Lemma 27, we can conclude that the first alternative is true. If ${}^\circ x = {}^\circ y$, then from Lemma 27, we have

$$\left(\forall^0 t > 0: \ x_t < y_t\right) \quad \text{or} \quad (x = y \text{ in } {}^\bullet\mathbb{R}) \quad \text{or} \quad \left(\forall^0 t > 0: \ x_t > y_t\right). \tag{27}$$

The assertion follows in the first two cases. In the third case, it follows from $x \leqslant y$ that

$$\forall^0 t \geqslant 0: \ x_t \leqslant y_t + z_t \tag{28}$$

with $z_t = o(t)$. Hence, by the third possibility in (27), $0 < x_t - y_t \leqslant z_t \ \forall^0 t > 0$, and hence $\lim_{t\to 0^+} \frac{x_t - y_t}{t} = 0$, i.e., $x = y$ in ${}^\bullet\mathbb{R}$.

$\Leftarrow$ This follows immediately from the reflexive property of $\leqslant$ or from the Definition 25.

$\Rightarrow$ It follows from $x < y$ that $x \leqslant y$ and $x \neq y$, and thus, the conclusion follows from the previous part.

$\Leftarrow$ It follows from $\forall^0 t > 0 : x_t < y_t$ and from assertion 1 that $x \leqslant y$ and hence $x < y$ by the assumption $x \neq y$. $\square$

We can now prove that our order is total.

Corollary 29. *Let $x,\, y \in {}^\bullet\mathbb{R}$,. Then in ${}^\bullet\mathbb{R}$ we have*

(1) $x \leqslant y$ *or* $y \leqslant x$ *or* $x = y$,

(2) $x < y$ *or* $y < x$ *or* $x = y$.

Proof. If ${}^\circ x < {}^\circ y$, then it follows from Lemma 27 that $x_t < y_t$ for $t \geqslant 0$ sufficiently small. Hence, $x \leqslant y$ by Theorem 28. We can argue in the same way if ${}^\circ x > {}^\circ y$. The case ${}^\circ x = {}^\circ y$ can be handled in the same way by using assertion (1) of Lemma 27.

The other part is a general consequence of the previous one. $\square$

From the proof of Lemma 27 and from Theorem 28, we can derive the following assertion.

Theorem 30. *Let $x,\, y \in {}^\bullet\mathbb{R}$. If ${}^\circ x \neq {}^\circ y$, then $x < y \iff {}^\circ x < {}^\circ y$. Otherwise, if ${}^\circ x = {}^\circ y$, then*

(1) *if $\omega(x) > \omega(y)$, then $x > y$ if and only if ${}^\circ x_1 > 0$;*

(2) *if $\omega(x) = \omega(y)$, then ${}^\circ x_1 > {}^\circ y_1 \implies x > y$ and ${}^\circ x_1 < {}^\circ y_1 \implies x < y$.*

Example. The above theorem gives an effective criterion to decide whether or not $x < y$. Indeed, if the two standard parts are different, then the order relation can be decided on the basis of these standard parts only; e.g., $2 + \mathrm{d}t_2 > 3\,\mathrm{d}t$ and $1 + \mathrm{d}t_2 < 3 + \mathrm{d}t$.

Otherwise, if the standard parts are equal, we first have to look at the order and at the first standard parts, i.e., ${}^\circ x_1$ and ${}^\circ y_1$, which are the coefficients of the biggest infinitesimals in the decompositions of x and y. For example, $3\,\mathrm{d}t_2 > 5\,\mathrm{d}t$ and $\mathrm{d}t_2 > a\,\mathrm{d}t$ for every $a \in \mathbb{R}$, and $\mathrm{d}t < \mathrm{d}t_2 < \mathrm{d}t_3 < \cdots < \mathrm{d}t_k$ for every $k > 3$, where $\mathrm{d}t_k > 0$.

If the orders are equal, we must compare the first standard parts, e.g., $3\,\mathrm{d}t_5 > 2\,\mathrm{d}t_5$.

The other cases fall within the previous ones, because of the properties of the ordered ring ${}^\bullet\mathbb{R}$. For example, $\mathrm{d}t_5 - 2\,\mathrm{d}t_3 + 3\,\mathrm{d}t < \mathrm{d}t_5 - 2\,\mathrm{d}t_3 + \mathrm{d}t_{3/2}$ if and only if $3\,\mathrm{d}t < \mathrm{d}t_{3/2}$, which is true because $\omega(\mathrm{d}t) = 1 < \omega(\mathrm{d}t_{3/2}) = \frac{3}{2}$. Finally $\mathrm{d}t_5 - 2\,\mathrm{d}t_3 + 3\,\mathrm{d}t > \mathrm{d}t_5 - 2\,\mathrm{d}t_3 - \mathrm{d}t$ because $3\,\mathrm{d}t > -\mathrm{d}t$.

12. ABSOLUTE VALUE, POWERS AND LOGARITHMS

Having a total order, we can define the absolute value in the usual way and, exactly as in $\mathbb{R}$, we can prove the standard properties of the absolute value. Moreover, the following cancellation law can be proved.

Theorem 31. *Let $h \in {}^\bullet\mathbb{R} \setminus \{0\}$ and r, $s \in \mathbb{R}$, then $|h| \cdot r \leqslant |h| \cdot s \implies r \leqslant s$.*

Proof. In fact, if $|h| \cdot r \leqslant |h| \cdot s$, then from Theorem 28, we obtain that either

$$\forall^0 t > 0 : \ |h_t| \cdot r \leqslant |h_t| \cdot s \tag{29}$$

or $|h| \cdot r = |h| \cdot s$. Since $h \neq 0$, we have $\left(\forall^0 t > 0 : \ h_t > 0\right)$ or $\left(\forall^0 t > 0 : \ h_t < 0\right)$, and hence, we can always find a $\bar{t} > 0$ such that $|h_{\bar{t}}| \neq 0$ to which (29) is applicable. Therefore, we must have $r \leqslant s$ in the first case. In the other one, we have $|h| \cdot r = |h| \cdot s$ and $h \neq 0$. Hence $|h| \neq 0$, and the conclusion follows from Theorem 17. $\square$

Due to the presence of nilpotent elements in ${}^\bullet\mathbb{R}$, we cannot define powers x^y and logarithms $\log_x y$ without any limitation. For example, we cannot define the square root having the usual properties

$$x \in {}^\bullet\mathbb{R} \implies \sqrt{x} \in {}^\bullet\mathbb{R}, \tag{30}$$

$$x = y \text{ in } {}^\bullet\mathbb{R} \implies \sqrt{x} = \sqrt{y} \text{ in } {}^\bullet\mathbb{R}, \tag{31}$$

and $\sqrt{x^2} = |x|$, because these are incompatible with the existence of $h \in D$ such that $h^2 = 0$ and $h \neq 0$. Indeed, the general property stated in Section 4 permits one to obtain a property of the form (30) (i.e., the closure of ${}^\bullet\mathbb{R}$ with respect to a given operation) for smooth functions only. Moreover, Definition 19 states that, to obtain a well-defined operation, we need a locally Lipschitzian function. For these reasons, we limit x^y to $x > 0$ and x invertible only, and $\log_x y$ to x, $y > 0$ and to the case in which both x and y are invertible.

Definition 32. *Let x, $y \in {}^\bullet\mathbb{R}$, with x strictly positive and invertible. Then*

(1) $x^y := [t \geqslant 0 \mapsto x_t^{y_t}]_{= \text{ in } {}^\bullet\mathbb{R}}$;
(2) *if $y > 0$ and y is invertible, then* $\log_x y := [t \geqslant 0 \mapsto \log_{x_t} y_t]_{= \text{ in } {}^\bullet\mathbb{R}}$.

By Theorem 28, it follows from $x > 0$ that $\forall^0 t > 0 : \ x_t > 0$, and thus, exactly as in Section 4 and in Definition 19, the above operations are well defined on ${}^\bullet\mathbb{R}$, because ${}^\circ x \neq 0 \neq {}^\circ y$. The elementary transfer theorem, Theorem 20, ensures the usual properties. To prove the ordinary monotonicity properties, it suffices to use Theorem 28.

Finally, it can be useful to state here the *elementary transfer theorem for inequalities* whose proof follows immediately from the definition of $\leqslant$ and from Theorem 28.

Theorem 33. *Let A be an open subset of $\mathbb{R}^n$ and τ, $\sigma \colon A \longrightarrow \mathbb{R}$ be smooth functions. In this case, $\forall x \in {}^\bullet A : \ {}^\bullet\tau(x) \leqslant {}^\bullet\sigma(x)$ if and only if $\forall r \in A : \ \tau(r) \leqslant \sigma(r)$.*

13. GEOMETRICAL REPRESENTATION OF FERMAT REALS

At the beginning of this article, we argued that one of the conducting idea in the construction of Fermat reals is to always maintain a clear intuitive meaning. More precisely, we always tried, and we shall always try, to keep a good dialectic between provable formal properties and their intuitive meaning. In this direction, we can see the possibility of finding a geometrical representation of Fermat reals.

The idea is that, to any Fermat real $x \in {}^\bullet\mathbb{R}$, we can associate the function

$$t \in \mathbb{R}_{\geqslant 0} \mapsto {}^\circ x + \sum_{i=1}^{N} {}^\circ x_i \cdot t^{1/\omega_i(x)} \in \mathbb{R}, \tag{32}$$

where N is, of course, the number of summands in the decomposition of x. Therefore, a geometric representation of this function is also a geometric representation of the number x, because different Fermat reals have different decompositions, see Theorem 5. Finally, we can guess that, because the notion of equality in ${}^\bullet\mathbb{R}$ depends only on the germ generated by each little-oh polynomial (see Definition 4), we can represent each $x \in {}^\bullet\mathbb{R}$ with only the first small part of the function (32).

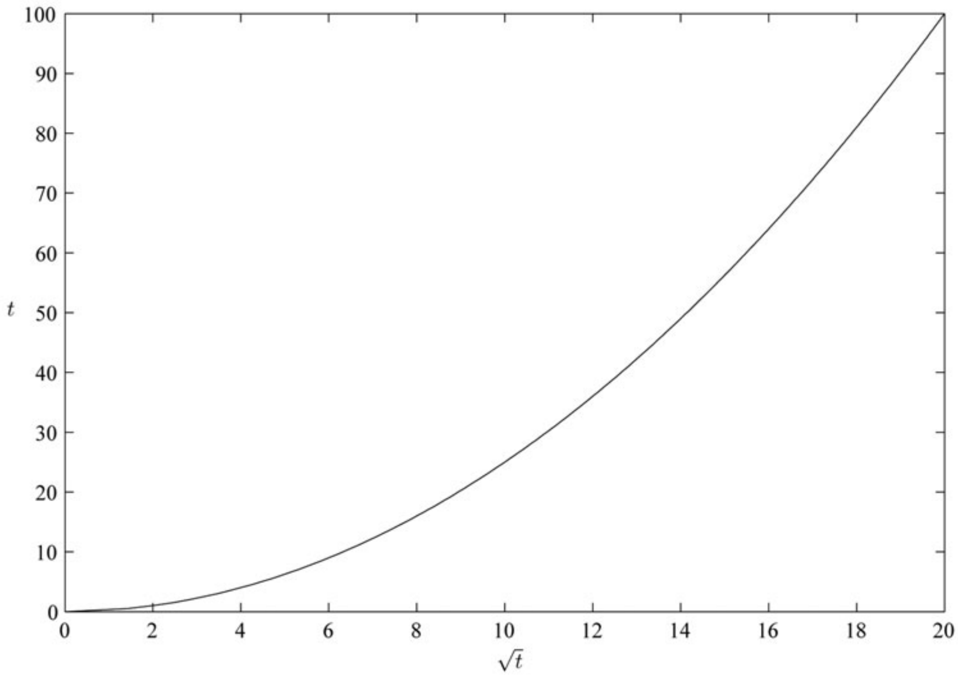

Fig. 1. The function representing the Fermat real $\mathrm{d}t_2 \in D_3$

Definition 34. For $x \in {}^\bullet\mathbb{R}$ and $\delta \in \mathbb{R}_{>0}$, set

$$\operatorname{graph}_\delta(x) := \left\{ \left({}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)}, t \right) \mid 0 \leqslant t < \delta \right\},$$

where N stands for the number of summands in the decomposition of x.

Note that the values of the function are placed in the abscissa position, and thus the correct representation of $\operatorname{graph}_\delta(x)$ is given by Fig. 1. This inversion of abscissa and ordinate in the $\operatorname{graph}_\delta(x)$ permits to represent this graph as a line tangent to the classical straight line $\mathbb{R}$ and hence, to have a better graphical picture. Finally, note that if $x \in \mathbb{R}$ is a standard real, then $N = 0$ and the $\operatorname{graph}_\delta(x)$ is a vertical line passing through ${}^\circ x = x$.

The following theorem enables us to represent the Fermat reals geometrically.

Theorem 35. *If $\delta \in \mathbb{R}_{>0}$, then the function $x \in {}^\bullet\mathbb{R} \mapsto \operatorname{graph}_\delta(x) \subset \mathbb{R}^2$ is injective. Moreover if x, $y \in {}^\bullet\mathbb{R}$, then we can find a $\delta \in \mathbb{R}_{>0}$ (depending on x and y) such that $x < y$ if and only if*

$$\forall p, q, t : \ (p, t) \in \operatorname{graph}_\delta(x) \ , \ (q, t) \in \operatorname{graph}_\delta(y) \quad \Longrightarrow \quad p < q. \tag{33}$$

Proof. The application $\rho(x) := \operatorname{graph}_\delta(x)$ for $x \in {}^\bullet\mathbb{R}$ is well defined because it depends on the terms ${}^\circ x$, ${}^\circ x_i$, and $\omega_i(x)$ of the decomposition of x (see Theorem 5 and Definition 8). Suppose now that $\operatorname{graph}_\delta(x) = \operatorname{graph}_\delta(y)$. Then

$$\forall t \in [0, \delta) : \ {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} = {}^\circ y + \sum_{j=1}^M {}^\circ y_j \cdot t^{1/\omega_j(y)}. \tag{34}$$

Consider the Fermat reals generated by these functions, i.e.,

$$x' := \left[t \geqslant 0 \mapsto {}^\circ x + \sum_{i=1}^N {}^\circ x_i \cdot t^{1/\omega_i(x)} \right]_{=\text{ in } {}^\bullet\mathbb{R}}, \qquad y' := \left[t \geqslant 0 \mapsto {}^\circ y + \sum_{j=1}^M {}^\circ y_j \cdot t^{1/\omega_j(y)} \right]_{=\text{ in } {}^\bullet\mathbb{R}}.$$

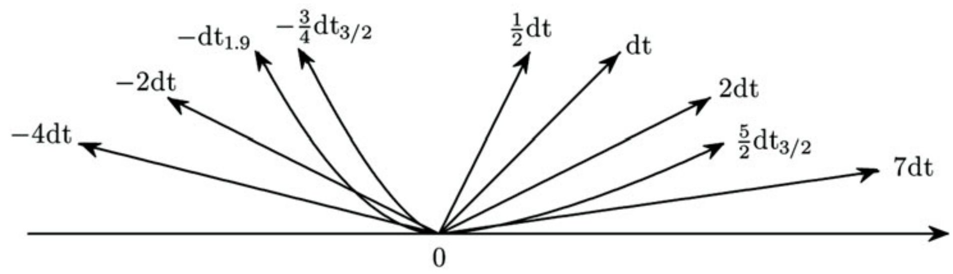

Fig. 2. Some first-order infinitesimals

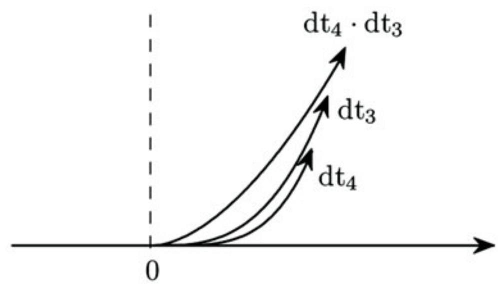

Fig. 3. The product of two infinitesimals

Then the decompositions of x' and y' are exactly the decompositions of x and y,

$$x' = {}^\circ x + \sum_{i=1}^{N} {}^\circ x_i \, \mathrm{d}t_{\omega_i(x)} = x, \tag{35}$$

$$y' = {}^\circ y + \sum_{j=1}^{M} {}^\circ y_j \, \mathrm{d}t_{\omega_j(y)} = y. \tag{36}$$

It follows from (34) that $x' = y'$ in ${}^\bullet\mathbb{R}$, and hence also $x = y$ by (35) and (36).

Suppose now that $x < y$. Then, using the notation of the previous part of the proof, we have $x' = x$ and $y' = y$, and hence,

$$x' = {}^\circ x + \sum_{i=1}^{N} {}^\circ x_i \cdot t^{1/\omega_i(x)} < {}^\circ y + \sum_{j=1}^{M} {}^\circ y_j \cdot t^{1/\omega_j(y)} = y'.$$

Applying Theorem 28 shows that $x'_t < y'_t$ locally, i.e.,

$$\exists \delta > 0 : \ \forall^0 t \geqslant 0 : \ {}^\circ x + \sum_{i=1}^{N} {}^\circ x_i \cdot t^{1/\omega_i(x)} < {}^\circ y + \sum_{j=1}^{M} {}^\circ y_j \cdot t^{1/\omega_j(y)}.$$

This is equivalent to (33) and, by Theorem 28, this is equivalent to $x' = x < y' = y$. $\square$

Example. In Fig. 2, we have a representation of some first-order infinitesimals.

The arrows are justified by the fact that the representing function (32) is defined on $\mathbb{R}_{\geqslant 0}$, and hence, has a clear first point and a direction. The smaller is $\alpha \in (0, 1)$, the nearer is the representation of the product $\alpha \, \mathrm{d}t$ to the vertical line passing through zero, which is the representation of the standard real $x = 0$. Finally, recall that $\mathrm{d}t_k \in D$ if and only if $1 \leqslant k < 2$.

Multiplying two infinitesimals, we obtain a smaller number, and hence, one whose representation is nearer to the vertical line passing through zero, as represented in Fig. 3.

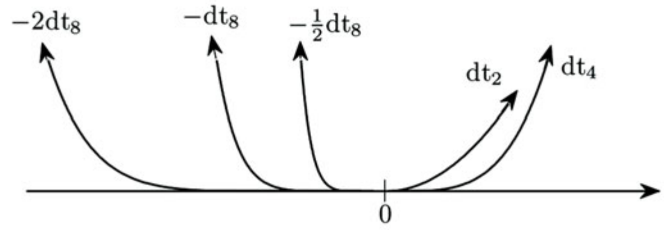

Fig. 4. Some higher-order infinitesimals

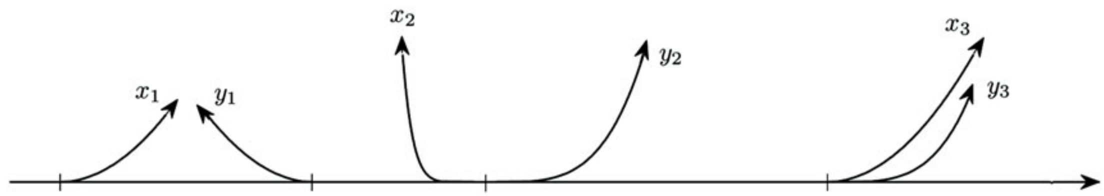

Fig. 5. Different cases in which $x_i < y_i$

In Fig. 4, we have a representation of some infinitesimals of order greater than 1. We can see that the greater is the infinitesimal $h \in D_a$ (with respect to the order relation $\leqslant$ defined in ${}^\bullet\mathbb{R}$), the higher is the order of intersection of the corresponding line $\mathrm{graph}_\delta(h)$.

Finally, in Fig. 5, we represent the order relation on the basis of Theorem 35. Intuitively, the method to see whether or not $x < y$ is to look at a suitably small neighborhood (i.e., at a suitably small $\delta > 0$) at $t = 0$ of their representing lines $\mathrm{graph}_\delta(x)$ and $\mathrm{graph}_\delta(y)$: if the curve $\mathrm{graph}_\delta(x)$ comes before the curve $\mathrm{graph}_\delta(y)$ with respect to the horizontal directed line, then x is less than y.

14. SOME ELEMENTARY EXAMPLES

The elementary examples presented in this section intend to show, in a few lines, the simplicity of the algebraic calculus of nilpotent infinitesimals. Here "simplicity" means that the dialectic with the corresponding informal calculations used, e.g., in engineering or in physics, is really faithful. The importance of this dialectic can be glimpsed both as a proof of the flexibility of the new language, but also for researches in artificial intelligence like automatic differentiation theories (see, e.g., Griewank [18] and the references therein). Last but not least, it may also be important for didactic or historical researches. Several examples are directly taken from those of Bell [3], and the reader is strongly invited to compare the two theories in these cases. In particular, in our point of view, it is not reasonable, like in some parts of Bell [3], to return back to a nonrigorous use of infinitesimals. Mathematical theories of infinitesimals, like our ring of Fermat reals, NSA, or SIA, are great opportunities to avoid several fallacies of the informal approach (our discussion in Section 10 is a clear example), and to advance further, with the new knowledge originating from the rigorous theory, opening the possibility of using infinitesimal methods in more general and less intuitive frameworks (like, e.g., infinite-dimensional spaces of mappings, see Giordano [16]). Once again, the key point is the dialectic between formal and informal thought rather than a single part only.

14.1. *Heat Equation*

In this and the following section, we simply use the language of ${}^\bullet\mathbb{R}$ to reformulate the corresponding considerations of Vladimirov [28]. Consider a body $B \subseteq \mathbb{R}^3$ (identified with its localization) and denote by $I_B := \mathrm{int}(B)$ its interior. Smooth functions $\rho\colon I_B \longrightarrow \mathbb{R}$, $c\colon I_B \longrightarrow \mathbb{R}$, and $k\colon I_B \longrightarrow \mathbb{R}$ are given and can be interpreted as the mass density, the specific heat capacity, and the thermal conductivity coefficient, respectively. Note that assuming these functions as defined on I_B without any favored direction corresponds physically to the isotropy condition for B. Moreover, let us consider $u\colon I_B \times [0, +\infty) \longrightarrow \mathbb{R}$, a smooth function representing the temperature of the body B at each point $x \in I_B$ and time $t \in [0, +\infty)$. To derive the heat diffusion equation, choose an internal point $x \in I_B$ and an infinitesimal volume V. More precisely, a subset of ${}^\bullet\mathbb{R}^3$ of the form

$$V = V(x, \delta\underline{x}) = \left\{ y \in {}^\bullet\mathbb{R}^3 \mid -\delta x_i \leqslant 2(y - x) \cdot \vec{e}_i \leqslant \delta x_i \quad \forall i = 1, 2, 3 \right\} \tag{37}$$

is said to be an infinitesimal parallelepiped if $\delta v := \delta x_1 \cdot \delta x_2 \cdot \delta x_3 \in D_\infty$, i.e., if the corresponding volume is an infinitesimal of some order. Here $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ is the natural basis of $\mathbb{R}^3$, and symbols of the form $\delta y \in {}^\bullet\mathbb{R}$ stress that the infinitesimal increment is associated to the variable y; here δ is not an operator, and we use it instead of the common dy to avoid confusion with our $\mathrm{d}y$ introduced in Definition 5. Because $x \in I_B$, the inclusion $V \subseteq {}^\bullet B$ follows, and thus V can be regarded as the subbody of B corresponding to the infinitesimal parallelepiped centered at x whose sides are parallel to the coordinate axes. This subbody V interacts thermally with its complement $\mathcal{C}V := {}^\bullet B \setminus V$ and with external sources of heat. During the infinitesimal time interval $\delta t \in D_\infty$, the heat flowing perpendicularly to the surface of V (Fourier's law) defines the exchange between the subbody V and its complement $\mathcal{C}V$,

$$Q_{\mathcal{C}V,V} = \delta t \cdot \sum_{i=1}^{3} \delta s_i \cdot \left[k(x + \delta\vec{h}_i) \cdot \frac{\partial u}{\partial \vec{e}_i}(x + \delta\vec{h}_i, t) - k(x - \delta\vec{h}_i) \cdot \frac{\partial u}{\partial \vec{e}_i}(x - \delta\vec{h}_i, t) \right], \tag{38}$$

where $\delta\vec{h}_i := \frac{1}{2}\delta x_i \cdot \vec{e}_i \in {}^\bullet\mathbb{R}^3$ and $\delta s_i := \prod_{j \neq i} \delta x_j \in {}^\bullet\mathbb{R}$. Choosing the infinitesimals in such a way that $\delta v \cdot \delta t \in D$, we obtain $\delta t \cdot \delta s_i \cdot (\delta x_i)^2 = \delta t \cdot \delta v \cdot \delta x_i = 0$ by Theorem 12 (e.g., we can choose $\delta x_i = \mathrm{d}t_6$ and $\delta t = \mathrm{d}t_2$). Simple manipulations using the infinitesimal Taylor's formula in (38) give

$$Q_{\mathcal{C}V,V} = \operatorname{div}\left[k \cdot \operatorname{grad}(u)\right](x, t) \cdot \delta v \cdot \delta t. \tag{39}$$

Of course, these calculations correspond to the infinitesimal version of the Gauss–Ostrogradskii theorem. Interacting thermally with external sources, the subbody V exchanges the heat

$$Q_{\mathrm{ext},V} = F(x, t) \cdot \delta v \cdot \delta t, \tag{40}$$

where $F\colon I_B \times [0, +\infty) \longrightarrow \mathbb{R}$ is a smooth function representing the intensity of the thermal sources. The total heat $Q_{\mathcal{C}V,V} + Q_{\mathrm{ext},V}$ corresponds to the increment $u(x, t+\delta t) - u(x, t)$ of the temperature of V, and hence, to an exchange of heat with the environment, $Q_{\mathrm{env},V}$,

$$Q_{\mathrm{env},V} = [u(x, t + \delta t) - u(x, t)] \cdot c(x) \cdot \rho(x) \cdot \delta v = Q_{\mathcal{C}V,V} + Q_{\mathrm{ext},V}. \tag{41}$$

This, together with (39), (40), the infinitesimal Taylor's formula, and the cancellation law, gives the desired formula $c(x) \cdot \rho(x) \cdot \frac{\partial u}{\partial t}(x, t) = \operatorname{div}\left[k \cdot \operatorname{grad}(u)\right](x, t) + F(x, t)$. To stress that the above proof is completely rigorous, we state the following theorem, without any reference to the physical interpretation.

Theorem 36. *Let $B \subseteq \mathbb{R}^d$, and let $I_B := \operatorname{int}(B)$ be the interior of B. Consider smooth functions $\rho\colon I_B \longrightarrow \mathbb{R}$, $c\colon I_B \longrightarrow \mathbb{R}$, $k\colon I_B \longrightarrow \mathbb{R}$, $u\colon I_B \times [0, +\infty) \longrightarrow \mathbb{R}$, and $F\colon I_B \times [0, +\infty) \longrightarrow \mathbb{R}$. Take a point $(x, t) \in I_B \times [0, +\infty)$ and define V, $Q_{\mathcal{C}V,V}$, $Q_{\mathrm{ext},V}$, and $Q_{\mathrm{env},V}$ as in* (37), (38), (40), *and* (41), *where $\delta v \cdot \delta t \in D$. In this case, $Q_{\mathrm{env},V} = Q_{\mathcal{C}V,V} + Q_{\mathrm{ext},V}$ if and only if*

$$c(x) \cdot \rho(x) \cdot \frac{\partial u}{\partial t}(x, t) = \operatorname{div}\left[k \cdot \operatorname{grad}(u)\right](x, t) + F(x, t).$$

Unfortunately, this statement insufficiently stresses the difference between the physical content of the definition of $Q_{\mathcal{C}V,V}$ (Fourier's law) and that of the definition of $Q_{\mathrm{ext},V}$. In an axiomatic framework for thermodynamics (see, e.g., Truesdell [27]), the notion of heat flux Q_{AB} going from a body A to a body B can be taken as primitive; in that case, (38) becomes an important assumption, whereas (40) is simply the definition of the intensity $F(x, t) = Q_{\mathrm{ext},V}/(\delta v \cdot \delta t)$.

14.2. *Electric Dipole*

In elementary physics, an electric dipole is usually defined as "*a pair of charges with opposite sign placed at a distance d very less than the distance r from the observer.*" Conditions like $r \gg d$ are frequently used in physics, and we often obtain a correct formalization assuming that $d \in {}^\bullet\mathbb{R}$ is infinitesimal and $r \in \mathbb{R} \setminus \{0\}$, i.e., r is finite. Thus, we can define an electric dipole as a pair (p_1, p_2) of electric particles with charges of equal intensity and of opposite sign such that their mutual distance at every time t is a first-order infinitesimal,

$$\forall t: \ |p_1(t) - p_2(t)| =: |\vec{d_t}| =: d_t \in D. \tag{42}$$

In this way, we can evaluate the potential at a point x using the properties of D and the assumption that r is finite and nonzero. In fact, we have $\varphi(x) = (q/(4\pi\epsilon_0))\cdot(1/r_1 - 1/r_2)$, $\vec{r_i} := x - p_i$, and, if $\vec{r} := \vec{r}_2 - \vec{d}/2$, then $1/r_2 = \left(r^2 + d^2/4 + \vec{r}\cdot\vec{d}\right)^{-1/2} = r^{-1}\cdot\left(1 + \vec{r}\cdot\vec{d}/r^2\right)^{-1/2}$ because $d^2 = 0$ for (42). Under our assumptions on d and r, we have $\vec{r}\cdot\vec{d}/r^2 \in D$, and hence, by the derivation formula, $\left(1 + \vec{r}\cdot\vec{d}/r^2\right)^{-1/2} = 1 - \vec{r}\cdot\vec{d}/(2r^2)$. We can proceed for $1/r_1$ in the same way; hence $\varphi(x) = (q/(4\pi\epsilon_0))\cdot(1/r)\cdot\left(1 + \vec{r}\cdot\vec{d}/(2r^2) - 1 + \vec{r}\cdot\vec{d}/(2r^2)\right) = (q/(4\pi\epsilon_0))\cdot\vec{r}\cdot\vec{d}/r^3$. The property $d^2 = 0$ is also used in the calculation of the electric field and for the moment of momentum.

14.3. *Newtonian Limit in Relativity.*

Another example in which we can formalize a condition of the form $r \gg d$ by using the above ideas is the Newtonian limit in relativity. Suppose that

(1) $\forall t: \ v_t \in D_2 \quad \text{and} \quad c \in \mathbb{R}$,

(2) $\forall x \in M_4: \ g_{ij}(x) = \eta_{ij} + h_{ij}(x) \quad \text{with} \quad h_{ij}(x) \in D$,

where $(\eta_{ij})_{ij}$ stands for the matrix of Minkowski's metric. These conditions can be interpreted as $v_t \ll c$ and $h_{ij}(x) \ll 1$ (low speed with respect to the speed of light and weak gravitational field). In this way, we have, e.g., the relations $1/\sqrt{1 - v^2/c^2} = 1 + v^2/(2c^2)$ and $\sqrt{1 - h_{44}(x)} = 1 - (1/2)\, h_{44}(x)$.

14.4. *Linear Differential Equations*

Let $L(y) := A_0\, \mathrm{d}^N y/\mathrm{d}t^N + \dots + A_{N-1}\, \mathrm{d}y/\mathrm{d}t + A_N \cdot y = 0$ be a linear differential equation with constant coefficients. Once again, we want to *discover* independent solutions in the case in which the characteristic polynomial has multiple roots, e.g., $(r-r_1)^2\cdot(r-r_3)\cdots(r-r_N) = 0$. The idea is that, in ${}^\bullet\mathbb{R}$, we have $(r-r_1)^2 = 0$ if $r = r_1+h$ with $h \in D$. Thus, $y(t) = \mathrm{e}^{(r_1+h)t}$ is a solution as well. However, $\mathrm{e}^{(r_1+h)t} = \mathrm{e}^{r_1 t} + ht\cdot\mathrm{e}^{r_1 t}$, and hence $L\left[\mathrm{e}^{(r_1+h)t}\right] = 0 = L\left[\mathrm{e}^{r_1 t} + ht\cdot\mathrm{e}^{r_1 t}\right] = L\left[\mathrm{e}^{r_1 t}\right] + h\cdot L\left[t\cdot\mathrm{e}^{r_1 t}\right]$. We obtain $L\left[t\cdot\mathrm{e}^{r_1 t}\right] = 0$, i.e., $y_1(t) = t\cdot\mathrm{e}^{r_1 t}$ must be a solution. Using infinitesimals of order k, we can deal with other multiple roots in a similar way.

14.5. *Circle of curvature*

A simple application of the infinitesimal Taylor's formula is the parametric equation for the circle of curvature, i.e., the circle of osculating order two for a curve $\gamma\colon [0,1] \longrightarrow \mathbb{R}^3$. In fact, if $r \in (0,1)$ and if $\dot\gamma_r$ is a unit vector, then, by the second-order infinitesimal Taylor's formula,

$$\forall h \in D_2: \ \gamma(r+h) = \gamma_r + h\,\dot\gamma_r + \frac{h^2}{2}\,\ddot\gamma_r = \gamma_r + h\,\vec{t}_r + \frac{h^2}{2} c_r\,\vec{n}_r, \tag{43}$$

where $\vec{n}$ stands for the unit normal vector, $\vec{t}$ for the tangent one, and c_r for the curvature. Once again, $\sin(ch) = ch$ and $\cos(ch) = 1 - \frac{c^2h^2}{2}$ by Taylor's formula. It now suffices to substitute h and $h^2/2$ from these formulas into (43) to obtain the conclusion

$$\forall h \in D_2: \ \gamma(r+h) = \left(\gamma_r + \vec{n}_r/c_r\right) + (1/c_r)\cdot[\sin(c_r h)\vec{t}_r - \cos(c_r h)\vec{n}_r].$$

We can prove in a similar way that any $f \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ can be written for any $h \in D_k$ in the form $f(h) = \sum_{n=0}^k a_n\cdot\cos(nh) + \sum_{n=0}^k b_n\cdot\sin(nh)$, and the idea of Fourier series comes out in a natural way.

14.6. *Commutation of Differentiation and Integration*

This example originates from Kock [20] and Lavendhomme [22]. Suppose we want to *discover* the derivative of the function $g(x) := \int_{\alpha(x)}^{\beta(x)} f(x,t)\,\mathrm{d}t$ for any $x \in \mathbb{R}$, where α, β, and f are smooth functions. We can regard g as a composition of smooth functions, and hence, we can apply the derivation formula, i.e., Theorem 21, which gives

$$g(x+h) = \int_{\alpha(x+h)}^{\beta(x+h)} f(x+h,t)\,\mathrm{d}t = \int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} f(x,t)\,\mathrm{d}t + h\cdot\int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} \partial f/\partial x(x,t)\,\mathrm{d}t$$
$$+\int_{\alpha(x)}^{\beta(x)} f(x,t)\,\mathrm{d}t + h\cdot\int_{\alpha(x)}^{\beta(x)} \partial f/\partial x(x,t)\,\mathrm{d}t + \int_{\beta(x)}^{\beta(x)+h\beta'(x)} f(x,t)\,\mathrm{d}t + h\cdot\int_{\beta(x)}^{\beta(x)+h\beta'(x)} \partial f/\partial x(x,t)\,\mathrm{d}t.$$

Now we use $h^2 = 0$ to obtain, e.g. (see Corollary 23),

$$h\cdot\int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} \partial f/\partial x(x,t)\,\mathrm{d}t = -h^2\cdot\alpha'(x)\cdot\partial f/\partial x(\alpha(x),t) = 0,$$
$$\int_{\alpha(x)+h\alpha'(x)}^{\alpha(x)} f(x,t)\,\mathrm{d}t = -h\cdot\alpha'(x)\cdot f(\alpha(x),t).$$

Calculating similar terms in an analogous way, we finally obtain the well-known conclusion. Note that the final formula comes out by itself, and thus we have "*discovered*" it rather than simply proved it. From the point of view of artificial intelligence or from the didactic point of view, this discovering is surely a nontrivial result.

14.7. *Schwarz' Theorem*

Using nilpotent infinitesimals, we can obtain a simple and meaningful proof of Schwarz' theorem. This simple example aims to show how to manage some differences between our setting and SDG. Let $f\colon V \longrightarrow E$ be a $\mathcal{C}^2$ function between spaces of type $V = \mathbb{R}^m$ and $E = \mathbb{R}^n$, and let $a \in V$. We want to prove that $\mathrm{d}^2 f(a)\colon V \times V \longrightarrow E$ is symmetric. Take $k \in D_2$ and $h, j \in \mathcal{D}_\infty$ such that $jkh \in D_{\neq 0}$ (e.g., we can take $k_t = \mathrm{d}t_2$ and $h_t = j_t = \mathrm{d}t_4$ in such a way that $jkh = \mathrm{d}t$; see also Theorem 12). Using $k \in D_2$, we obtain

$$\begin{aligned} j\cdot f(x+hu+kv) &= j\cdot\left[f(x+hu) + k\,\partial_v f(x+hu) + \frac{k^2}{2}\partial_v^2 f(x+hu)\right] \\ &= j\cdot f(x+hu) + jk\cdot\partial_v f(x+hu), \end{aligned} \tag{44}$$

where we have used the fact that $k^2 \in D$ and j infinitesimal imply $jk^2 = 0$. Since $jkh \in D$, any product of the type $jkhi$ is zero for every $i \in D_\infty$, and thus

$$jk\cdot\partial_v f(x+hu) = jk\cdot\partial_v f(x) + jkh\cdot\partial_u(\partial_v f)(x). \tag{45}$$

However, $k \in D_2$ and $jk^2 = 0$. Hence $j\cdot f(x+kv) - j\cdot f(x) = jk\cdot\partial_v f(x)$. Substituting this formula into (45), and hence into (44), we obtain

$$j\cdot[f(x+hu+kv) - f(x+hu) - f(x+kv) + f(x)] = jkh\cdot\partial_u(\partial_v f)(x). \tag{46}$$

The left-hand side of this equality is symmetric with respect to u and v. Hence, transposing them, we obtain $jkh\cdot\partial_u(\partial_v f)(x) = jkh\cdot\partial_v(\partial_u f)(x)$, as was to be proved, because $jkh \neq 0$ and $\partial_u(\partial_v f)(x)$, $\partial_v(\partial_u f)(x) \in E$. The classical limit relation

$$\lim_{t\to 0^+}\,(f(x+h_t u+k_t v) - f(x+h_t u) - f(x+k_t v) + f(x))/(h_t k_t) = \partial_u\partial_v f(x)$$

immediately follows from (46).

14.8. *Area of the Circle and Volumes of Revolution*

A more or less meaningful proof of the familiar formula for the area of a circle depends on the axioms assumed and on the generality of definitions. In this example, we show the possibility to define suitable smooth functions using an infinitesimal property. Assume the axioms for the real field $\mathbb{R}$, use them to prove the existence of the smooth functions sin and cos, define π as a suitable zero of these functions (see, e.g., Prodi [24] and Šilov (Shilov) [26]), and define the length of an arc of circle of radius r, parametrized by $x(\theta) = r \cdot \cos(\theta)$ and $y(\theta) = r \cdot \sin(\theta)$, as a unique function s for which

$$[s(\theta+k) - s(\theta)]^2 = [x(\theta+k) - x(\theta)]^2 + [y(\theta+k) - y(\theta)]^2 \qquad \forall\theta \in \mathbb{R} \quad \forall k \in D_2, \tag{47}$$

$$s(0) = 0. \tag{48}$$

This definition can be justified in the usual way by using a (second-order!) infinitesimal right-angled triangle. The uniqueness of s follows from (47) and (48) by the smoothness of x and y, the second-order infinitesimal Taylor's formula, and the cancellation law $k^2 \cdot \dot{s}(\theta) = \dot{x}(\theta) \cdot k^2 + \dot{y}(\theta) \cdot k^2$ for any $k \in D_2$ (Theorem 17). This, together with (48), yields the ordinary formula for s, which gives $s(\theta) = r \cdot \theta$ in our particular case. We can now regard the area $A(\theta + h) - A(\theta)$ of a first-order infinitesimal sector of the circle as the area of the isosceles triangle with sides of length r and with base $s(\theta + h) - s(\theta)$. In fact, if $P(\theta) = (r\sin\theta, r\cos\theta)$, then $P(\theta + h) = P(\theta) + h \cdot \vec{t}(\theta)$, where $\vec{t}$ is the tangent vector, and thus in $[\theta, \theta + h]$, $h \in D$, the circle is made of linear segments. Therefore, the area $A(\theta)$ can be defined as a unique function such that $A(\theta+h) - A(\theta) = (1/2)\,[s(\theta + h) - s(\theta)] \cdot r\cos(h/2)$ for any $\theta \in \mathbb{R}$ and any $h \in D$ for which $A(0) = 0$. This, together with the derivation formula, gives $h \cdot A'(\theta) = (1/2)hr \cdot s'(\theta)$ and $A(\theta) = (1/2)\int_0^\theta r \cdot s(u)\,\mathrm{d}u$. In our case, $A(\theta) = (1/2)r^2 \cdot \theta$, which proves the desired formula for $\theta = 2\pi$.

Similarly, we can prove the familiar formula for the volumes of revolution of parametrized curves of the form $\gamma(u) = (x(u), y(u))$, $u \in [a, b]$, around the x axis. Define the volume as a unique smooth function V such that

$$V(u+h) - V(u) = h \cdot \pi \cdot y(u)^2 + \frac{1}{2}\left[h \cdot \pi \cdot y(u+h)^2 - h \cdot \pi \cdot y(u)^2\right], \tag{49}$$

$$V(0) = 0, \tag{50}$$

for every $u \in [a, b]$ and $h \in D$. This definition can be intuitively justified by saying that the volume of the sector of revolution between u and $u + h$ can be evaluated as the sum of the cylinder of radius $y(u)$ and height h plus the halved difference between the cylinder of radius $y(u+h)$ and of height h and that of radius $y(u)$ and of the same height. Implicitly, we are using the straightness of the curve γ in $[u, u+h]$. By (49) and by the property $h^2 = 0$, we readily obtain $V'(u) = \pi \cdot y(u)^2$, and hence, the ordinary formula, using (50).

14.9. *Curvature*

Let us consider an ordinary smooth parametrized curve $\gamma(u) = (x(u), y(u))$ for $u \in [a, b]$. Let $\varphi(u) \in [0, \pi]$ be the nonoriented angle (i.e., the one defined by the scalar product) between the tangent vector $\vec{t} = (\dot{x}, \dot{y})$ and the unit vector $\vec{i}$ of the x axis. Thus, $\sqrt{\dot{x}^2 + \dot{y}^2} \cdot \cos\varphi = \dot{x}$. Multiplying this equality by $\sin\varphi$ gives

$$\dot{y} \cdot \cos\varphi = \dot{x} \cdot \sin\varphi. \tag{51}$$

As is well known, the curvature of γ at the point $u \in [a, b]$ can be evaluated as the rate of change of the nonoriented angle $\varphi(u)$ with respect to an infinitesimal variation in arc length $s(u)$ defined by analogues of (47) and (48). These "rates of changes" can be defined in ${}^\bullet\mathbb{R}$ as a unique (if exists) standard $c(u) \in \mathbb{R}$ defined by $c(u) \cdot [s(u+h) - s(u)] = \varphi(u+h) - \varphi(u)$ for any $h \in D$. Indeed, by the cancellation law, i.e., by Theorem 17, there exists at most one such $c(u) \in \mathbb{R}$ verifying this property. By this uniqueness, we can also use the notation

$$c(u) = \frac{\varphi(u+h) - \varphi(u)}{s(u+h) - s(u)}. \tag{52}$$

These ratios generalize the standard ratios for reals (see Giordano [16] for details). It follows from (52) and from the derivation formula that $c(u) = (h \cdot \varphi'(u))/(h \cdot s'(u)) = \varphi'(u)/s'(u)$ whatever $h \in D_{\neq 0}$ we choose. This and relation (51) (using standard differential calculus rather than infinitesimals) implies the ordinary formula $c = (\dot{x}\ddot{y} - \dot{y}\ddot{x})/\big((\dot{x}^2 + \dot{y}^2)^{3/2}\big)$ at each point $u \in [a, b]$ where $\varphi(u) \neq \pi/2$ and $\dot{\gamma}(u) \neq \underline{0}$.

14.10. *Stretching of a Spring* (*and of the Center of Pressure*)

If $f\colon [a, b] \longrightarrow \mathbb{R}$ is a smooth function and if $J(x) := \int_0^x f(s)\,\mathrm{d}s$, then Corollary 23 and a trivial calculation with the derivation formula give

$$J(x + h) - J(x) = (1/2)\,[f(x + h) + f(x)] \quad \forall h \in D. \tag{53}$$

The right-hand side of (14.17) can be interpreted as the average value of f on the infinitesimal interval $[x, x+h]$. Analogous equalities can be obtained in the d-dimensional case by using suitable generalizations of the above corollary; e.g., if $d = 2$, then we must use

$$\int_0^h \int_0^k f(x, y)\,\mathrm{d}x\,\mathrm{d}y = hk \cdot f(0, 0) \quad \text{for any} \quad h, k \in D_\infty \quad \text{such that} \quad h \cdot k \in D.$$

These relations are used by Bell [3] to calculate the center of pressure of a plane area and the work done when stretching a spring. The meaningfulness of such examples is, however, doubtful, because they can be summarized as follows: assume that there is a smooth J satisfying (53); derive from this fact and from the assumption $J(0) = 0$ that $J'(x) = f(x)$. There is no real use of infinitesimals in this type of reasoning in any case for which the definition $J(x) := \int_0^x f(s)\,\mathrm{d}s$ is customary, like in the above examples.

14.11. *Wave Equation*

The derivation of the wave equation in the framework of Fermat reals is very interesting for two main reasons. Firstly, in the classical deduction (see, e.g., Vladimirov [28]), there are some approximations tied with Hooke's law. Is it possible to make them rigorous by using ${}^\bullet\mathbb{R}$? Do we gain something using this increased rigor? For example, how can we formalize the approximated equations used in the classical derivation? In what sense the wave equation is an approximate relation which holds for small oscillations only?

Secondly, at the end of our derivation, we shall stress the physical principles as important mathematical assumptions of a suitable theorem. We are thus naturally taken to ask whether or not these natural assumptions (some of which are formulated by using the infinitesimals of ${}^\bullet\mathbb{R}$) really have a model. In this way, we shall see that no standard smooth function can satisfy these assumptions; however, we are forced to consider a nonstandard function. For example, $f(x) = h \cdot \sin(x)$ for $x \in {}^\bullet\mathbb{R}$ and $h \in D_\infty$ is an example of a nonstandard smooth function; note that it is obtained from the standard smooth function $g(y, x) := y \cdot \sin(x)$, x, $y \in \mathbb{R}$, by an extension to ${}^\bullet\mathbb{R}^2$ and by fixing one of its variables to be a nonstandard parameter $h \in D_\infty$, namely, $f(x) = {}^\bullet g(h, x)$ for any $x \in {}^\bullet\mathbb{R}$. This motivates the further development of the theory of Fermat reals strongly, in the direction of a more general theory including also these new smooth nonstandard functions.

Let us begin with considering a string making small transversal oscillations around its equilibrium position located on the interval $[a, b]$ of the x axis for a, $b \in \mathbb{R}$, $a < b$. By assumption, the position $s_t \subseteq {}^\bullet\mathbb{R}^2$ of the string is always represented by the graph of a curve $\gamma\colon [a, b] \times [0, +\infty) \longrightarrow {}^\bullet\mathbb{R}^2$ (where $[a, b] = \{x \in {}^\bullet\mathbb{R} \mid a \leqslant x \leqslant b\}$ and $[0, +\infty) = \{x \in {}^\bullet\mathbb{R} \mid 0 \leqslant x\}$; in the following, we always use this notation for the above intervals to identify the corresponding subsets of ${}^\bullet\mathbb{R}$ rather than those of $\mathbb{R}$, and we also use the notation $\gamma_{xt} := \gamma(x, t))$, $s_t = \big\{\gamma_{xt} \in {}^\bullet\mathbb{R}^2 \mid a \leqslant x \leqslant b\big\} \quad \forall t \in [0, +\infty)$. Moreover, the curve γ is assumed to be injective with respect to the parameter $x \in (a, b)$, $\gamma_{x_1 t} \neq \gamma_{x_2 t} \quad \forall t \in [0, +\infty)\ \forall x_1, x_2 \in (a, b) :\ x_1 \neq x_2$, and thus, the order relation on (a, b) implies an order relation on the support s_t. For every pair of points $p = \gamma_{x_p t}$, $q = \gamma_{x_q t} \in s_t$ on the string at time t, we can define the subbodies $\overrightarrow{p} := \{\gamma_{xt} \mid x_p \leqslant x \leqslant b\}$, $\overleftarrow{p} := \{\gamma_{xt} \mid a \leqslant x \leqslant x_p\}$, and

$\overrightarrow{pq} := \{\gamma_{xt} \mid x_p \leqslant x \leqslant x_q\}$ corresponding to the parts of the string after the point $p \in s_t$, before the same point, and between the points $p \in s_t$ and $q \in s_t$, respectively. It is usually tacitly clear that, e.g., every subbody of the form $\overrightarrow{p}$ exerts a force on each subbody with which it is in contact, i.e., of the form $\overrightarrow{pq}$ or $\overleftarrow{p}$. Moreover, the force $\mathbf{F}(A,B) \in {}^\bullet\mathbb{R}^2$ exerted by the subbody A on the subbody B verifies the equalities (see, e.g., Truesdell [27])

$$\mathbf{F}(\overrightarrow{pq}, \overleftarrow{p}) = \mathbf{F}(\overrightarrow{p}, \overleftarrow{p}), \tag{54}$$

$$\mathbf{F}(\overrightarrow{q}, \overrightarrow{pq}) = \mathbf{F}(\overrightarrow{q}, \overleftarrow{q}), \tag{55}$$

$$\mathbf{F}(\overleftarrow{p}, \overrightarrow{pq}) = -\mathbf{F}(\overrightarrow{pq}, \overleftarrow{p}) \quad \text{(action-reaction principle)}, \tag{56}$$

for every pair of points p, $q \in s_t$ and every time $t \in [0,+\infty)$. Using this formalism, the tension at the point $\gamma_{xt} \in s_t$ at time $t \in [0,+\infty)$ can now be defined as follows:

$$\mathbf{T}(x,t) := \mathbf{F}(\overrightarrow{\gamma_{xt}}, \overleftarrow{\gamma_{xt}}). \tag{57}$$

Now consider the infinitesimal subbody $\overrightarrow{x, x+\delta x} := \overrightarrow{\gamma_{xt}\gamma_{x+\delta x,t}} \subseteq s_t$ located at time t between the points $\gamma_{xt} \in s_t$ and $\gamma_{x+\delta x,t} \in s_t$, where $\delta x \in D$ is a generic first-order infinitesimal. Mass forces of linear density $\mathbf{G}\colon [a,b] \times [0,+\infty) \longrightarrow {}^\bullet\mathbb{R}^2$ act on this infinitesimal subbody, and thus Newton's law can be represented as

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} = \mathbf{F}(\overleftarrow{\gamma_{xt}}, \overrightarrow{x, x+\delta x}) + \mathbf{F}(\overrightarrow{\gamma_{x+\delta x,t}}, \overrightarrow{x, x+\delta x}) + \mathbf{G} \cdot \rho \cdot \delta x, \tag{58}$$

where $\rho\colon [a,b] \times [0,+\infty) \longrightarrow {}^\bullet\mathbb{R}$ stands for the linear mass density and where, unless otherwise stated, all functions are evaluated at $(x,t) \in (a,b) \times [0,+\infty)$. Of course, the contact forces appearing in Newton's law are due to the interaction of the infinitesimal subbody with other subbodies contacting along the border $\partial \left[\overrightarrow{x, x+\delta x}\right] = \{\gamma_{xt}, \gamma_{x+\delta x,t}\} \subseteq {}^\bullet\mathbb{R}^2$. Using the action-reaction principle (56) and relation (55) with $q = \gamma_{x+\delta x,t}$ and $p = \gamma_{xt}$ such that $\overrightarrow{pq} = \overrightarrow{x, x+\delta x}$, we see by (58) that $\rho \cdot \delta x \cdot \partial^2\gamma/\partial t^2 = -\mathbf{F}(\overrightarrow{x, x+\delta x}, \overleftarrow{\gamma_{xt}}) + \mathbf{F}(\overrightarrow{\gamma_{x+\delta x,t}}, \overleftarrow{\gamma_{x+\delta x,t}}) + \mathbf{G} \cdot \rho \cdot \delta x$. Using (54) and the definition of tension in (67), we obtain

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} = -\mathbf{F}(\overrightarrow{\gamma_{xt}}, \overleftarrow{\gamma_{xt}}) + \mathbf{F}(\overrightarrow{\gamma_{x+\delta x,t}}, \overleftarrow{\gamma_{x+\delta x,t}}) + \mathbf{G} \cdot \rho \cdot \delta x = -\mathbf{T}(x,t) + \mathbf{T}(x+\delta x, t) + \mathbf{G} \cdot \rho \cdot \delta x. \tag{59}$$

Up to this point of the proof, we have used neither the small oscillations hypothesis nor the transversal oscillations hypothesis. The second one can readily be introduced with the hypotheses

$$\mathbf{G}(x,t) \cdot \vec{e}_1 = 0 \quad \forall x, t, \tag{60}$$

where $(\vec{e}_1, \vec{e}_2)$ are the axial unit vectors. Using the notation $\varphi(x,t)$ for the nonoriented angle between the tangent unit vector $\mathbf{t}(x,t)$ at the point γ_{xt} and the x axis (see (51)), the small oscillations hypothesis can be formalized with the assumption

$$\varphi(x,t) \in D \quad \forall x, t. \tag{61}$$

This enables us to reproduce the classical derivation in the most faithful way (even if a weaker assumption can be considered, see below). Moreover, in the classical derivation of the wave equation, one considers only curves of the form $\gamma_{xt} = (x, u(x,t))$. In this way, by (51) and by the derivation formula, we have $\partial\gamma_2/\partial x \cdot \cos\varphi = \sin\varphi$ and $\partial\gamma_2/\partial x = \varphi \in D$. Hence, $(\partial\gamma_2/\partial x)^2 = 0$, and the total length of the string becomes

$$L = \int_a^b \sqrt{1 + [\partial\gamma_2/\partial x(x,t)]^2}\, \mathrm{d}x = b - a \quad \forall t \in [0,+\infty). \tag{62}$$

By Hooke's law, this proves that the tension can be assumed to have constant modulus, T, depending on neither the position x nor the time t,

$$\mathbf{T}(x,t) = T \cdot \mathbf{t}(x,t) \quad \forall x \in (a,b) \quad \forall t \in [0,+\infty). \tag{63}$$

A tension $\mathbf{T}$ parallel to the tangent vector is the second part of the hypothesis about nontransversal oscillations of the string. Let us note explicitly that the only standard continuous function verifying the equality $L = b - a$ is constant, and thus, the function $u\colon [a,b] \times [0,+\infty) \longrightarrow {}^\bullet\mathbb{R}$ must be treated as a nonstandard one. Below we make further remarks concerning this important point. Projecting equation (59) to the y axis, we obtain

$$\begin{aligned}\rho\cdot \delta x\cdot \partial^2 u/\partial t^2 &= -T \cdot \mathbf{t}(x,t)\cdot \vec{e}_2 + T\cdot \mathbf{t}(x+\delta x,t)\cdot \vec{e}_2 + \mathbf{G}\cdot \vec{e}_2 \cdot \rho \cdot \delta x\\ &= -T\sin\varphi(x,t) + T\cdot \sin\varphi(x+\delta x,t) + G\cdot\rho\cdot\delta x.\end{aligned}$$

However, $\sin\varphi = \varphi = \partial u/\partial x$ because $\varphi \in D$ is a first-order infinitesimal, and hence

$$\rho\cdot\delta x\cdot \partial^2 u/\partial t^2 = T\cdot[\partial u/\partial x(x+\delta x,t) - \partial u/\partial x(x,t)] + G\cdot\rho\cdot\delta x = \left[T\cdot \partial^2 u/\partial x^2(x,t) + G\cdot\rho\right]\cdot\delta x. \tag{64}$$

We cannot use the cancellation law with $\delta x \in D$ to obtain the final result, because, as was mentioned above, the function $u(x,t) \in {}^\bullet\mathbb{R}$ can take nonstandard values. We are to clarify some points here. As mentioned above, there is no standard smooth function verifying the assumptions or the physical principles we have used. Of course, everything depends on the formalization of the classical informal derivation used in elementary physics; e.g., we have chosen to use the equality symbol in (62) instead of an approximate equality. Anyway, we should note that, if we use the symbol $\simeq$ in (62), then the problem becomes how to make this approximation more precise (physically, numerically, or mathematically). Moreover, if we use an approximation symbol in (62), then we must use the same symbol in (63), and therefore, in the final wave equation as well. Nevertheless, smooth nonstandard functions can verify all assumptions and physical principles under consideration; e.g., the function $u(x,t) := u_0\sin(x+\omega\cdot t)$ is one of these nonstandard functions if the maximum amplitude u_0 is in D and if ρ is constant, $G = 0$, and $T = \omega^2\rho$.

Definition 37. If $X \subseteq {}^\bullet\mathbb{R}^{\mathbf{x}}$ and $Y \subseteq {}^\bullet\mathbb{R}^{\mathbf{y}}$, then we say that $f\colon X \to Y$ is (*nonstandard*) *smooth* if and only if f takes X to Y and, for every $x_0 \in X$,

$$f(x) = {}^\bullet g\langle p, x\rangle \quad \forall x \in {}^\bullet V \cap X \tag{65}$$

for some V open in $\mathbb{R}^{\mathbf{x}}$ such that $x_0 \in {}^\bullet V$, $p \in {}^\bullet U$, where U is open in $\mathbb{R}^{\mathbf{p}}$, and $g \in \mathcal{C}^\infty(U\times V, \mathbb{R}^{\mathbf{y}})$, where $\langle -,-\rangle : ([x]_\sim, [y]_\sim) \in {}^\bullet U \times {}^\bullet V \longmapsto [(x,y)]_\sim \in {}^\bullet(U\times V)$ (for the relation $\sim$, see Definition 4).

In other words, locally, a smooth function $f\colon X \longrightarrow Y$ from $X \subseteq {}^\bullet\mathbb{R}^{\mathbf{x}}$ to $Y \subseteq {}^\bullet\mathbb{R}^{\mathbf{y}}$ is constructed as follows.

(1) Begin with an ordinary standard function $g \in \mathcal{C}^\infty(U\times V, \mathbb{R}^{\mathbf{y}})$ with U open in $\mathbb{R}^{\mathbf{p}}$ and V open in $\mathbb{R}^{\mathbf{x}}$. The space $\mathbb{R}^{\mathbf{p}}$ must be regarded as a space of parameters for the function g.
(2) Consider the Fermat extension of g giving ${}^\bullet g\colon {}^\bullet(U\times V) \longrightarrow {}^\bullet\mathbb{R}^{\mathbf{y}}$.
(3) Consider the composition ${}^\bullet g \circ \langle -,-\rangle\colon {}^\bullet U \times {}^\bullet V \longrightarrow {}^\bullet\mathbb{R}^{\mathbf{y}}$, where $\langle -,-\rangle$ is the isomorphism ${}^\bullet U \times {}^\bullet V \simeq {}^\bullet(U\times V)$ defined by $\langle [x]_\sim, [y]_\sim\rangle = [(x,y)]_\sim$; we always use the identification ${}^\bullet U \times {}^\bullet V = {}^\bullet(U\times V)$, and thus, we simply write ${}^\bullet g(p,x)$ instead of ${}^\bullet g\langle p,x\rangle$.
(4) Choose a parameter $p \in {}^\bullet U$ as a first variable of the previous composition, i.e., consider ${}^\bullet g\langle p,-\rangle : {}^\bullet V \longrightarrow {}^\bullet\mathbb{R}^{\mathbf{y}}$. Locally, the mapping f is of the form $f = {}^\bullet g\langle p,-\rangle = {}^\bullet g(p,-)$.

Because $p = {}^\circ p + h$, with $h \in D_\infty$, applying the infinitesimal Taylor's formula to the variable p for the function ${}^\bullet g(p,x)$, we can readily prove the following theorem clarifying further the form of these nonstandard smooth functions, because it claims that these functions can locally be regarded as "infinitesimal polynomials with smooth coefficients."

Theorem 38. *Let $X \subseteq {}^\bullet\mathbb{R}^{\mathrm{x}}$, and let $f\colon X \longrightarrow {}^\bullet\mathbb{R}^n$ be a mapping. In this case, the function $f\colon X \longrightarrow {}^\bullet\mathbb{R}^n$ is nonstandard smooth if and only if, for every $x_0 \in X$,*

$$f(x) = \sum_{\substack{|q|\leqslant k\\ q\in\mathbb{N}^d}} a_q(x)\cdot p^q \quad \forall x \in {}^\bullet V \cap X, \tag{66}$$

for the following suitable objects: (1) d, $k \in \mathbb{N}$, (2) V *is an open subset of* $\mathbb{R}^{\mathrm{x}}$ *such that* $x_0 \in {}^\bullet V$, (3) $(a_q)_{\substack{|q|\leqslant k\\ q\in\mathbb{N}^d}}$ *is a family in* $\mathcal{C}^\infty(V, \mathbb{R}^n)$.

In other words, every smooth function $f\colon X \longrightarrow {}^\bullet\mathbb{R}^n$ can be constructed locally, starting from some "infinitesimal parameters" $p_1, \dots, p_d \in D_k$ and from ordinary smooth functions $a_q \in \mathcal{C}^\infty(V, \mathbb{R}^n)$ and using polynomial operations only with $p_1, \dots, p_d$ and with the coefficients $a_q(-)$. Roughly speaking, we can say that they are "infinitesimal polynomials with smooth coefficients, the variables of the polynomials act as parameters only."

As is natural to expect, several notions of differential and integral calculus (including their infinitesimal versions) can be extended to this type of new smooth functions (for more details, see the preprint by Giordano [16]), and these results will be presented in subsequent works. In this sense, this derivation of the wave equation strongly motivates the future development of the theory of Fermat reals.

On the other hand, we must understand what type of cancellation law can be applied to (64). To this end, we must define the notion of equality up to kth-order infinitesimals.

Definition 39. Let $m = {}^\circ m + \sum_{i=1}^N {}^\circ m_i \cdot \mathrm{d}t_{\omega_i(m)}$ be the decomposition of $m \in {}^\bullet\mathbb{R}$ and $k \in \mathbb{R}_{\geqslant 0} \cup \{\infty\}$, then

$$\iota_k m := \iota_k(m) := {}^\circ m + \sum_{\substack{i=1\\ \omega_i(m)>k}}^N {}^\circ m_i \cdot \mathrm{d}t_{\omega_i(m)}.$$

Finally, if x, $y \in {}^\bullet\mathbb{R}$, then, by definition, $x =_k y$ if and only if $\iota_k x = \iota_k y$ in ${}^\bullet\mathbb{R}$, and we read it as *x is equal to y up to kth order infinitesimals.*

In other words, as is easy to prove, $x =_k y \iff {}^\circ x = {}^\circ y$ and $\omega(x-y) \leqslant k$. Therefore, if we write $I_k := \{x \in D_\infty \mid \omega(x) \leqslant k\}$, for the set of infinitesimals of order less than or equal to k (note that $I_k \subset D_k$), then the condition $x =_k y$ holds if and only if $x - y \in I_k$. Equality up to kth order infinitesimals is of course an equivalence relation, and it preserves the ring operations of ${}^\bullet\mathbb{R}$. Moreover, in general, these equalities are preserved by smooth functions $f\colon {}^\bullet\mathbb{R} \longrightarrow {}^\bullet\mathbb{R}$, i.e., $x =_k y$ implies $f(x) =_k f(y)$. Using this notion, one can readily prove the following cancellation law up to kth-order infinitesimals.

Theorem 40. *Let $m \in {}^\bullet\mathbb{R}$, $n \in \mathbb{N}_{>0}$, $j \in \mathbb{N}^n \setminus \{\underline{0}\}$, and $\alpha \in \mathbb{R}^n_{>0}$. Moreover, consider $k \in \mathbb{R}$ defined by*

$$\frac{1}{k} + \sum_{i=1}^n \frac{j_i}{\alpha_i + 1} = 1. \tag{67}$$

In this case, the following assertions hold.

(1) $\forall h \in D_{\alpha_1} \times \cdots \times D_{\alpha_n} :\ h^j \cdot m = h^j \cdot \iota_k m.$
(2) *If $h^j \cdot m = 0$ for every $h \in D_{\alpha_1} \times \cdots \times D_{\alpha_n}$, then $m =_k 0$.*

For example, if $n = 1$ and $\alpha_1 = j_1 = 1$, then $k = 2$, and hence, $\forall h \in D :\ h \cdot m = h \cdot \iota_2 m$

$$(\forall h \in D :\ h \cdot m = 0) \iff m =_2 0. \tag{68}$$

Using (68) in (64), we obtain the final conclusion

$$\rho \cdot \partial^2 u/\partial t^2 =_2 T \cdot \partial^2 u/\partial x^2 + G \cdot \rho \qquad \forall x \in (a, b) \quad \forall t \in (0, +\infty). \tag{69}$$

It is also interesting to note that not only small oscillations of the string imply (69) but also the converse is true, namely, equation (69) implies that we must necessarily have small oscillations of the string, i.e., that $\varphi(x,t) \in D_\infty$. Moreover, using the equality $=_2$ up to second-order infinitesimals, the classical approximation tied with Hooke's law now become clearer. Indeed, we have the following assertion.

Theorem 41. *Let $a, b \in \mathbb{R}$, with $a < b$; let $\gamma\colon [a,b]\times[0,+\infty) \longrightarrow {}^\bullet\mathbb{R}^2$, $\rho\colon [a,b]\times[0,+\infty) \longrightarrow {}^\bullet\mathbb{R}$ and $\mathbf{G}, \mathbf{T}\colon [a,b]\times[0,+\infty) \longrightarrow {}^\bullet\mathbb{R}^2$ be nonstandard smooth functions, and let $T \in {}^\bullet\mathbb{R}$ be an invertible Fermat real. Suppose that the first component γ_1 of the curve is of the form*

$$\gamma_1(x,t) = [1+\alpha(t)]\cdot x + \beta(t) \quad \forall x, t, \tag{70}$$

with $\alpha(t) \in I_2$. Then the unit tangent vector $\mathbf{t}(x,t)$ to the curve γ exists, and we can further suppose that the relations

$$\mathbf{T}(x,t) =_2 T\cdot\mathbf{t}(x,t), \tag{71}$$

$$\rho\cdot\delta x\cdot\partial^2\gamma_{xt}/\partial t^2 = \mathbf{T}(x+\delta x,t) - \mathbf{T}(x,t) + \mathbf{G}\cdot\rho\cdot\delta x, \tag{72}$$

hold for every point $(x,t) \in (a,b)\times[0,+\infty)$ and for every $\delta x \in D$. Finally, suppose that $\partial\varphi/\partial x(x,t)$ is invertible. Then the following assertions are equivalent:

(1) $\rho(x,t)\cdot\partial^2\gamma_2/\partial t^2(x,t) =_2 T\cdot\partial^2\gamma_2/\partial x^2(x,t) + G_2(x,t)\cdot\rho(x,t)$,
(2) $\varphi(x,t) \in I_4$.

Finally, if (2) holds for every $(x,t) \in (a,b)\times[0,+\infty)$, then length$(\gamma_{-,t}) =_2 b-a$.

To simplify the proof of this result, we need two lemmas.

Lemma 42. *Let $a, b \in \mathbb{R}$ with $a < b$, and let $f, g\colon (a,b) \longrightarrow {}^\bullet\mathbb{R}$ be nonstandard smooth functions such that $f(x) =_2 g(x)$ for any $x \in (a,b)$. Then $f(x+h) - f(x) = g(x+h) - g(x)$ for any $h \in D$ and any $x \in (a,b)$.*

Lemma 43. *Let $m, h \in {}^\bullet\mathbb{R}$. Suppose that m is invertible and $0 \leqslant h \leqslant \pi$. Then the following properties are equivalent*:

(1) $m\cdot\cos^3 h =_2 m$,
(2) $h \in I_4$.

Proof of Theorem 41. We first note that, if (70) holds, then the tangent vector $\mathbf{t}(x,t)$ exists in ${}^\bullet\mathbb{R}$. In fact, since $\partial\gamma_1/\partial x(x,t) = 1+\alpha(t)$, it follows that both the elements $\partial\gamma_1/\partial x(x,t)$ and $[\partial\gamma_1/\partial x(x,t)]^2 + [\partial\gamma_2/\partial x(x,t)]^2$ are invertible; hence, we can take their square roots and then the inverse to define the unit tangent vector. Let us now prove that (1) implies (2). Take a generic $\delta x \in D$. Projecting (72) to $\vec{e}_2$, we obtain

$$\rho\cdot\delta x\cdot\partial^2\gamma_2/\partial t^2 = \mathbf{T}(x+\delta x,t)\cdot\vec{e}_2 - \mathbf{T}(x,t)\cdot\vec{e}_2 + G_2\cdot\rho\cdot\delta x.$$

However, it follows from (71), because smooth operations preserve $=_2$, that $\mathbf{T}\cdot\vec{e}_2 =_2 T\cdot\mathbf{t}\cdot\vec{e}_2$. Therefore, by Lemma 42, we obtain

$$\mathbf{T}(x+\delta x,t)\cdot\vec{e}_2 - \mathbf{T}(x,t)\cdot\vec{e}_2 = T\cdot\mathbf{t}(x+\delta x,t)\cdot\vec{e}_2 - T\cdot\mathbf{t}(x,t)\cdot\vec{e}_2 = T\cdot\sin\varphi(x+\delta x,t) - T\cdot\sin\varphi(x,t),$$

$$\rho\cdot\delta x\cdot\partial^2\gamma_2/\partial t^2 = T\cdot\sin\varphi(x+\delta x,t) - T\cdot\sin\varphi(x,t) + G_2\cdot\rho\cdot\delta x. \tag{73}$$

On the other hand, we can multiply (1) by δx (so that $=_2$ becomes $=$, see Theorem 40) and obtain

$$\begin{aligned}\rho\cdot\delta x\cdot\partial^2\gamma_2/\partial t^2 &= T\cdot[\partial\gamma_2/\partial x(x+\delta x,t) - \partial\gamma_2/\partial x(x,t)] + G_2\cdot\rho\cdot\delta x \qquad (74)\\ &= T\cdot\tan\varphi(x+\delta x,t)\cdot\partial\gamma_1/\partial x(x+\delta x,t) - T\tan\varphi(x,t)\cdot\partial\gamma_1/\partial x(x,t) + G_2\cdot\rho\cdot\delta x.\end{aligned}$$

Equating (73) and (74) and cancelling T, we see that

$$\begin{aligned}\sin\varphi(x+\delta x,t)-\sin\varphi(x,t)&=\tan\varphi(x+\delta x,t)\cdot\partial\gamma_1/\partial x(x+\delta x,t)-\tan\varphi(x,t)\cdot\partial\gamma_1/\partial x(x,t),\\ \delta x\cdot\cos\varphi\cdot\partial\varphi/\partial x&=\delta x\cdot(1/\cos^2\varphi)\cdot\partial\varphi/\partial x\cdot\partial\gamma_1/\partial x(x,t)+\tan\varphi\cdot\partial^2\gamma_1/\partial x^2(x,t)\\ &=\delta x\cdot(1/\cos^2\varphi)\cdot\partial\varphi/\partial x\cdot[1+\alpha(t)]=\delta x\cdot(1/\cos^2\varphi)\cdot\partial\varphi/\partial x,\end{aligned} \tag{75}$$

where every function is evaluated at (x,t), unless otherwise stated. Note that, in (75), we have used the property $\delta x\cdot\alpha(t)=0$ which follows from $\delta x\in D$ and $\alpha(t)\in I_2$; moreover, it follows from (51) (for $\varphi=\pi/2$) that we would have $\partial\gamma_2/\partial x\cdot\cos\varphi=0=\partial\gamma_1/\partial x\cdot\sin\varphi=1+\alpha(t)$, which is impossible because $\alpha(t)\in D_\infty$. By setting $m:=\partial\varphi/\partial x(x,t)\in{}^\bullet\mathbb{R}$ for simplicity, using (75), and cancelling δx, we obtain

$$m\cdot\cos^3\varphi=_2 m. \tag{76}$$

By Lemma 43, this implies the desired conclusion.

Vice versa, if φ is an infinitesimal of order less than or equal to 4, then, by Lemma 43, we obtain (76), and we can again go over the previous paragraphs in the opposite direction to prove part (1).

Suppose that $\varphi(x,t)\in I_4$ for every $(x,t)\in(a,b)\times[0,+\infty)$. Then

$$\mathrm{length}(\gamma_{-,t})=\int_a^b\sqrt{[1+\alpha(t)]^2+[\partial\gamma_2/\partial x(x,t)]^2}\,\mathrm{d}x=\int_a^b\sqrt{1+2\alpha(t)+[\partial\gamma_2/\partial x(x,t)]^2}\,\mathrm{d}x, \tag{77}$$

because $\alpha(t)\in I_2$, and hence, $\alpha(t)^2=0$. However, $[1+\alpha(t)]\cdot\sin\varphi=\partial\gamma_2/\partial x(x,t)\cdot\cos\varphi$, and thus, $\partial\gamma_2/\partial x(x,t)=[1+\alpha(t)]\tan\varphi=[1+\alpha(t)]\left(\varphi+\varphi^3/3\right)=\varphi+\varphi^3/3+\alpha(t)\cdot\varphi$, because $\alpha(t)\in I_2$ and $\varphi\in I_4$, and hence, $\alpha(t)\cdot\varphi^3=0$. Substituting this into (77) and using the derivation formula for the function $x\mapsto\sqrt{1+x}$, we obtain

$$\begin{aligned}\sqrt{1+2\alpha(t)+[\partial\gamma_2/\partial x(x,t)]^2}&=1+(1/2)\cdot\left\{2\alpha(t)+[\partial\gamma_2/\partial x(x,t)]^2\right\}\\ &=1+\alpha(t)+(1/2)\left[\varphi+\varphi^3/3+\alpha(t)\cdot\varphi\right]^2=1+\alpha(t)+\varphi^2/2+\varphi^4/3+\alpha(t)\cdot\varphi^2.\end{aligned}$$

Therefore

$$\begin{aligned}\mathrm{length}(\gamma_{-,t})&=\int_a^b\left[1+\alpha(t)+\frac{\varphi(x,t)^2}{2}+\frac{\varphi(x,t)^4}{3}+\alpha(t)\cdot\varphi(x,t)^2\right]\mathrm{d}x\\ &=b-a+\alpha(t)\cdot(b-a)+\int_a^b\left[\frac{\varphi(x,t)^2}{2}+\frac{\varphi(x,t)^4}{3}+\alpha(t)\cdot\varphi(x,t)^2\right]\mathrm{d}x.\end{aligned} \tag{78}$$

Using Theorem 38, we can readily prove that the last integral in (78) is an infinitesimal of order less than or equal to 2, and thus the conclusion follows from the assumption $\alpha(t)\in I_2$. $\square$

Proof of Lemma 42. First of all, it follows from the assumption $f(x)=_2 g(x)$ for every $x\in(a,b)$ that

$$^\circ f(x)={}^\circ g(x)\quad\forall x\in(a,b). \tag{79}$$

Choose a point $x\in(a,b)$. By Theorem 38, we can write $f(x_1)=a_0(x_1)+\sum_i p_i\cdot a_i(x_1)$ and $g(x_1)=b_0(x_1)+\sum_j q_j\cdot b_j(x_1)$ for every $x_1\in(x-\delta,x+\delta)\subseteq(a,b)$, where p_i,$q_j\in D_\infty$ and a_i, b_j are ordinary smooth functions defined on an open neighborhood V of $^\circ x\in(a,b)\cap\mathbb{R}$. By (79), $a_0(^\circ x_1)=b_0(^\circ x_1)$ for every $x_1\in{}^\bullet V$, and thus $a_0=b_0$ on V, and hence, also ${}^\bullet a_0={}^\bullet b_0$ on ${}^\bullet V$. Therefore,

$$f(r)-g(r)=\sum_i p_i\cdot a_i(r)-\sum_j q_j\cdot b_j(r)\quad\forall r\in(a,b)\cap\mathbb{R}. \tag{80}$$

The order of this difference must be less than or equal to 2 because $f(r) =_2 g(r)$, and thus, we have $\omega\Big[\sum_i p_i \cdot a_i(r) - \sum_j q_j \cdot b_j(r)\Big] = \max_i \omega\left[p_i \cdot a_i(r)\right] \vee \max_j \omega\left[q_j \cdot b_j(r)\right] \leqslant 2$. Suppose, for simplicity, that $\omega(p_1 \cdot a_1(r))$ is the term of maximum order. Since $a_1(r) \in \mathbb{R}$, we must have $\omega(p_1) \leqslant 2$, and hence, also $\omega(p_i) \leqslant \omega(p_1) \leqslant 2$ and $\omega(q_j) \leqslant \omega(p_1) \leqslant 2$. Finally,

$$f(x+h) - f(x) = h \cdot f'(x) = h \cdot a_0'(x) + \sum_i h \cdot p_i \cdot a_i'(x),$$

where $a_0'(x) = b_0'(x)$ because $a_0 = b_0$ and $h \cdot p_i = 0$ because $\omega(h) < 2$ and $\omega(p_i) \leqslant 2$; thus, $f(x+h) - f(x) = h \cdot b_0'(x) = h \cdot b_0'(x) + \sum h \cdot q_j \cdot b_j'(x) = h \cdot g'(x) = g(x+h) - g(x)$. $\square$

Proof of Lemma 43. If $m \cdot \cos^3 h =_2 m$, then the standard parts of both sides must be equal, $^\circ\left(m \cdot \cos^3 h\right) = {}^\circ m$ and $^\circ m \cdot \cos^3\left({}^\circ\varphi\right) = {}^\circ m$. By assumption, m is invertible, and hence, $^\circ m \neq 0$. We obtain $^\circ h = 0$ because $0 \leqslant h \leqslant \pi$, i.e., $h \in D_\infty$. Moreover, since $m \cdot \cos^3 h =_2 m$, it follows from the infinitesimal Taylor's formula applied to $\cos h$ that

$$m \cdot \Big(1 - \sum_{1 \leqslant i < (\omega(h)+1)/2} (-1)^i c \cdot h^{2i}/(2i)!\Big)^3 =_2 m, \qquad m \cdot \left(1 + a \cdot h^2\right)^3 =_2 m,$$

$$m \cdot \left(1 + a^3 h^6 + 3ah^2 + 3a^2 h^2\right) =_2 m, \qquad m \cdot \left(1 + \alpha \cdot h^2\right) =_2 m,$$

where $a := -\sum_{1 \leqslant i < \frac{\omega(h)+1}{2}} (-1)^i \frac{h^{2i-2}}{(2i)!} \in {}^\bullet\mathbb{R}$ and $\alpha := 3a^2 + 3a + a^3 h^4$ are invertible Fermat reals. This gives $m \cdot \alpha \cdot h^2 =_2 0$, and hence, $h^2 =_2 0$, i.e., $\omega(h^2) \leqslant 2$, and therefore, $\omega(h) \leqslant 4$.

Vice versa, if h is an infinitesimal of order less than or equal to 4 (and thus, $\varphi^n = 0$ for $n \geqslant 5$), then $\cos^3 h = \left(1 - h^2/2 + h^4/4!\right)^3 = 1 - 3h^2/2 + 3h^4/4!$. Hence $m \cdot \cos^3 h = m - 3mh^2 \cdot \left(\frac{1}{2} - 3\frac{h^2}{4!}\right)$, and thus, $m \cdot \cos^3 h - m = -3mh^2 \cdot \left(\frac{1}{2} - 3\frac{h^2}{4!}\right)$ is an infinitesimal of order $\omega(h^2) \leqslant 2$, i.e., $m \cos^3 h =_2 m$. $\square$

The reader with a certain knowledge of SDG had surely noted that this derivation of the wave equation cannot be reproduced in SDG because of the use of nonstandard smooth functions, of equalities up to kth order infinitesimals, and of the frequent use of the useful statement in Theorem 12 to study products of powers of nilpotent infinitesimals.

15. CONCLUSIONS

The problem of transforming informal infinitesimal methods into a rigorous theory has been addressed by several authors. The most used theories, i.e., NSA and SDG, require a good knowledge of mathematical logic and a strong formal control. Some others, like Weil functors (see, e.g., Kriegl and Michor [21]) or the Levi-Civita field (see, e.g., Shamseddine [25]) are mainly based on formal/algebraic methods and sometimes lack intuitive meaning. In this initial work, we have shown that it is possible to bypass the inconsistency of SIA with classical logic by modifying the Kock–Lawvere axiom (see, e.g., Lavendhomme [22]) while always keeping a very good intuitive meaning. We have seen how to define the algebraic operations between this type of nilpotent infinitesimals, infinitesimal Taylor formula, and order properties. In the final part, we have seen several elementary examples of the use of these infinitesimals, some of them taken from classical derivations in elementary physics. In our opinion, these examples are able to show that some results that frequently may appear as unnatural in the standard context, can be discovered by using Fermat reals, even by suitably designed algorithms. Moreover, our generalization of the classical proof of the wave equation shows that a rigorous theory of infinitesimals enables one to obtain results that are inaccessible when using an intuitive approach only.

REFERENCES

1. S. Albeverio, J.E. Fenstad, R. Høegh-Krohn, and T. Lindstrøm, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics* (Pure and Applied Mathematics. Academic Press, 1988, 2nd ed., Dover, 2009).
2. E.T. Bell, *Men of Mathematics* (Simon and Schuster, New York, 1937).
3. J.L. Bell, *A Primer of Infinitesimal Analysis* (Cambridge University Press, 1998).
4. V. Benci and M. Di Nasso, "A Ring Homomorphism is Enough to Get Nonstandard Analysis." Bull. Belg. Math. Soc. Simon Stevin **10**, 481–490, 2003.
5. V. Benci and M. Di Nasso, "A Purely Algebraic Characterization of the Hyperreal Numbers," Proc. Amer. Math. Soc. **133** (9), 2501–05 (2005).
6. W. Bertram, *Differential Geometry, Lie Groups and Symmetric Spaces over General Base Fields and Rings* (American Mathematical Society, Providence, 2008).
7. M. Berz, "Analysis on a Nonarchimedean Extension of the Real Numbers," Mathematics Summer Graduate School of the German National Merit Foundation, MSUCL-933, Department of Physics, Michigan State University, 1992 and 1995 edition, 1994.
8. T. Bröcker, *Differentiable Germs and Catastrophes*, Vol. 17 of LMS Lect. Note Series (Cambridge University Press, Cambridge, 1975).
9. J.H. Conway, *On Numbers and Games,* Number 6 in LMS Monographs (Academic Press, London & New York, 1976).
10. P.A.M. Dirac, *General Theory of Relativity* (John Wiley and Sons, 1975).
11. C.H. Edwards, *The Historical Development of the Calculus* (Springer-Verlag, New York, 1979).
12. C. Ehresmann, "Les prolongements d'une variété différentiable: Calculus des jets, prolongement principal," C. R. Acad. Sci. Paris **233**, 598–600 (1951).
13. A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover, 1926).
14. H. Eves, *An Introduction to the History of Mathematics* (Saunders College Publishing, Fort Worth, TX, 1990).
15. P. Giordano, "Infinitesimal Differential Geometry," Acta Math. Univ. Comenian. **LXIII** (2), 235–278 (2004).
16. P. Giordano, "Fermat Reals: Nilpotent Infinitesimals and Infinite Dimensional Spaces," arXiv:0907.1872 (July 2009).
17. M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities,* Vol. 14 of Grad. Texts in Math. (Springer, Berlin, 1973).
18. A. Griewank, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation,* Vol. 19 of Frontiers Appl. Math. (SIAM, 2000).
19. C.W. Henson, Foundations of Nonstandard Analysis. A Gentle Introduction to Nonstandard Extension. In L.O. Arkeryd, N.J. Cutland, and C.W. Henson, editors, *Nonstandard Analysis: Theory and Applications* (Edinburgh, 1996), pp. 1–49, Dordrecht, 1997. NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., Vol. 493, Kluwer Acad. Publ.
20. A. Kock, *Synthetic Differential Geometry,* Vol 51 of *LMS Lect. Note Series* (Cambridge Univ. Press, 1981).
21. A. Kriegl and P.W. Michor, "Product Preserving Functors of Infinite Dimensional Manifolds," Arch. Math. (Brno) **32** (4), 289–306 (1996).
22. R. Lavendhomme, *Basic Concepts of Synthetic Differential Geometry* (Kluwer Academic Publishers, Dordrecht, 1996).
23. I. Moerdijk and G.E. Reyes, *Models for Smooth Infinitesimal Analysis* (Springer, Berlin, 1991).
24. G. Prodi, *Analisi matematica*, Ed. Bollati Boringhieri (Torino, 1970).
25. K. Shamseddine, *New Elements of Analysis on the Levi-Civita Field,* PhD thesis, Michigan State University, East Lansing, Michigan (USA, 1999).
26. G.E. Silov, *Analisi matematica. Funzioni di una variabile* (it. transl.) (Mir, Mosca, 1978).
27. C. Truesdell, *A First Course in Rational Continuum Mechanics:* V.1 *General Concepts,* 2nd ed., Vol. 71 of Pure Appl. Math. (Academic Press Inc., 1991).
28. V.S. Vladimirov, *Equazioni della fisica matematica* (MIR, 1987).