INFINITESIMAL AND INFINITE NUMBERS IN APPLIED MATHEMATICS

ALEKSANDR BRYZGALOV, KEVIN ISLAMI, AND PAOLO GIORDANO

Abstract. The need to describe abrupt changes or response of nonlinear systems to impulsive stimuli is ubiquitous in applications. Also the informal use of infinitesimal and infinite quantities is still a method used to construct idealized but tractable models within the famous J. von Neumann *reasonably wide area* of applicability. We review the theory of *generalized smooth functions* as a candidate to address both these needs: a rigorous but simple language of infinitesimal and infinite quantities, and the possibility to deal with continuous and generalized function as if they were smooth maps: with pointwise values, free composition and hence nonlinear operations, all the classical theorems of calculus, a good integration theory, and new existence results for differential equations. We exemplify the applications of this theory through several models of singular dynamical systems: deduction of the heat and wave equations extended to generalized functions, a singular variable length pendulum wrapping on a parallelepiped, the oscillation of a pendulum damped by different media, a nonlinear stress-strain model of steel, singular Lagrangians as used in optics, and some examples from quantum mechanics.

Contents

2020 *Mathematics Subject Classification.* 46Fxx, 46F30, 12J25, 37Nxx.

Key words and phrases. Schwartz distributions, generalized functions for nonlinear analysis, non-Archimedean analysis, singular dynamical systems.

This research was funded in whole or in part by the Austrian Science Fund (FWF) P30407, P34113, P33945, 10.55776/P33538, 10.55776/PAT9221023. For open access purposes, the author has applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

1. Introduction: informal uses of infinitesimals and infinities in applied mathematics

Even if infinitesimal numbers have been banished by modern mathematics, several physicists, engineers and mathematicians still profitably continue to use them. Usually, this is in order to simplify calculations, to construct idealized but notwithstanding interesting models of physical systems, or to relate different parts of physics, such as in passing from quantum to classical mechanics if $\hbar$ is infinitesimal. An authoritative example in this direction is given by A. Einstein when he writes

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = 1 + \frac{v^2}{2c^2} \qquad \sqrt{1-h_{44}(x)} = 1 - \frac{1}{2}h_{44}(x) \tag{1.1}$$

with explicit use of infinitesimals $v/c \ll 1$ or $h_{44}(x) \ll 1$ such that e.g. $h_{44}(x)^2 = 0$. More generally, in [22] Einstein writes the formula (using the equality sign and not the approximate equality sign $\simeq$)

$$f(x, t+\tau) = f(x,t) + \tau \cdot \frac{\partial f}{\partial t}(x,t) \tag{1.2}$$

justifying it with the words "*since τ is very small*"; note that (1.1) are a particular case of the general (1.2). Also P.A.M. Dirac in [19] writes an analogous equality studying the Newtonian approximation in general relativity.

A certain degree of inconsistency appears also at the level of elementary topics, e.g. in the deduction of the wave and heat equations, see e.g. [77]. For example, if $u(x,t)$ is the string displacement, then formula (1.2) is once again used e.g. "to ignore magnitudes of order greater than $\frac{\partial u}{\partial x}$". This means that we need to have $\left(\frac{\partial u}{\partial x}\right)^2 = 0$ to arrive at the final equation with an equality sign and not with some kind of approximation $\simeq$. But then the length of the string becomes

$$L = \int_a^b \sqrt{1 + \left[\frac{\partial u}{\partial x}(x,t)\right]^2}\,\mathrm{d}x = b - a,$$

and it is clear that this necessarily yields that the function u is constant. It clearly does not really help to use $\simeq$ when we have a contradiction, but then to change it into $=$ when we need the final equation. It is for this type of motivations that V.I. Arnol'd in [2] wrote: *Nowadays, when teaching analysis, it is not very popular to talk about infinitesimal quantities. Consequently present-day students are* *not fully in command of this language. Nevertheless, it is still necessary to have command of it.*

A similar, but sometimes more troublesome, practice concerns the use of infinite numbers. A typical example is given by Heisenberg's uncertainty principle

$$\Delta p_x \Delta x \geq \frac{\hbar}{2}. \tag{1.3}$$

It is frequently informally argued that if the position x is measured by a Dirac delta, then $\Delta x \approx 0$ is infinitesimal; thereby, (1.3) necessarily implies that Δp_x must be an infinite number.

Another classical example of informal use of infinite numbers concerns Schwartz distributions and their point values. Many relevant physical systems are in fact described by singular Hamiltonians. Among them, we can e.g. list:

(i) Non smooth classical mechanics. For example, a classical particle (or a high-frequency wave) moving through discontinuous media containing barriers or interfaces where the Hamiltonian is discontinuous: see e.g. [7, 8, 17, 75, 40, 69, 72] and references therein.
(ii) Discontinuous Lagrangian in geometrical optics: see e.g. [47].
(iii) Nonlinear deformation in continuum mechanics, which includes non differentiable stress-strain relations: see e.g. [76, 10].
(iv) The use of infinite quantities in quantum mechanics. An elementary example is given by the solution of the stationary Schrödinger equation for an infinite rectangular potential well (a case that cannot be formalized using Schwartz distributions, see e.g. [25]).

This type of problems is hence widely studied from the mathematical point of view (see e.g. [15, 53, 51, 73, 45, 44, 52]), even if the presented solutions are not general and hold only in special conditions. In this sense, the fact that J.D. Marsden's works [55, 56] did not start a consolidated research thread to study singular Hamiltonian mechanics using Schwartz distributions, can be considered as a clue that the classical distributional framework is not well suited to face this problem in general terms.

A related problem concerns nonlinear operations on Schwartz distributions, which can be simply presented as follows. Let A be an associative and commutative algebra endowed with a derivation (satisfying the Leibniz rule). Then any element H of A such that $H \cdot H = H$ is necessarily a constant, that is, $H' = 0$. Indeed, $(H^2)' = 2HH'$ and $(H^3)' = 3H^2H'$. Now $H = H^2 = H^3$, so this implies $2HH' = H' = 3HH'$. Therefore, $HH' = 0$ and hence also $H' = 2HH' = 0$. Even worse, we also recall that Dirac in [20] uses terms of the form $\sqrt{\delta}$ in his proposal for the foundation of quantum mechanics.

There are obviously many possibilities to formalize this kind of intuitive reasonings, obtaining a more or less good dialectic between informal and formal thinking, and nowadays there are indeed several rigorous theories of infinitesimals, infinities and generalized functions. Concerning the notion of infinitesimal, we can distinguish two definitions: in the first one we have at least a ring R containing the real field $\mathbb{R}$ and infinitesimals are elements $h \in R$ such that $-r < h < r$ for every positive standard real $r \in \mathbb{R}_{>0}$. The second type of infinitesimal is defined using some algebraic property of nilpotency, i.e. $h^n = 0$ for some natural number $n \in \mathbb{N}$ (therefore, in this case h cannot be trivially invertible and we cannot form infinities as reciprocal of infinitesimals). For some ring R these definitions can coincide, but anyway they lead, of course, only to the trivial infinitesimal $h = 0$ if $R = \mathbb{R}$ is the real field.

Mathematical theories of infinitesimals can also be classified as belonging to two different classes. In the first one, we have theories needing a certain amount of non trivial results of mathematical logic, whereas in the second one we have attempts to define sufficiently strong theories of infinitesimals without the use of non trivial results of mathematical logic. In the first class, we can list *nonstandard analysis* and *synthetic differential geometry* (also called *smooth infinitesimal analysis*, see e.g. [3, 43, 49, 58]), in the second one we have, e.g., *Weil functors* (see [46]), *Levi-Civita field* (see [71]), *surreal numbers* (see [14]), *Fermat reals* (see [27]), *Colombeau's generalized numbers* (see [13] and [21, 41] for a general survey). More precisely, we can say that to work both in nonstandard analysis and in synthetic differential geometry, one needs a formal control stronger than the one used in "standard mathematics". Indeed, to use nonstandard analysis one has to be able to formally write sentences in order to apply the transfer theorem. Whereas synthetic differential geometry does not admit models in classical logic, but in intuitionistic logic only, and hence we have to be sure that in our proofs there is no use of the law of the excluded middle, or e.g. of the classical part of De Morgan's law or of some form of the axiom of choice or of the implication of double negation toward affirmation and any other logical principle which do not hold in intuitionistic logic. Physicists, engineers, but also the greatest part of mathematicians are not used to have this strong formal control in their work, and it is for this reason that there are attempts to present both nonstandard analysis and synthetic differential geometry reducing as much as possible the necessary formal control, even if at some level this is technically impossible (see e.g. [42, 36], and [4, 5] for nonstandard analysis; [3] and [49] for synthetic differential geometry).

On the other hand, nonstandard analysis is surely the best known theory of invertible infinitesimals with results in several areas of mathematics and its applications, see e.g. [1]. Synthetic differential geometry is a theory of nilpotent infinitesimals with non trivial results of differential geometry in infinite dimensional spaces.

Concerning mathematical theories of generalized functions, the difficulties stemmed from dealing with the lacking of well-posedness in PDE initial value problems led to a zoo of spaces of generalized functions. In an incomplete list we can mention: Schwartz distributions, Colombeau generalized functions, ultradistributions, hyperfunctions, nonstandard theory of Colombeau generalized functions, ultrafunctions, etc. See e.g. [39] for a survey, and the International Conference on Generalized Functions series, e.g. https://ps-mathematik.univie.ac.at/e/index.php?event=GF2022.

Unfortunately, there is a certain lacking of dialog between the most used theory of generalized functions, i.e. Schwartz distributions, and the actual use of generalized functions in physics and engineering, where e.g. point values and nonlinear operations are frequently needed, see e.g. [13].

In the present paper, we describe the main results of *generalized smooth functions* (GSF) theory and some of its applications in applied mathematics. GSF are an extension of classical distribution theory and of Colombeau theory, which makes it possible to model nonlinear singular problems, while at the same time sharing a number of fundamental properties with ordinary smooth functions, such as the closure with respect to composition and several non trivial classical theorems of the calculus, see [31, 32, 29, 28, 54, 33]. One could describe GSF as a methodological restoration of Cauchy-Dirac's original conception of generalized function (GF), see [48]. In essence, the idea of Cauchy and Dirac (but also of Poisson, Kirchhoff, Helmholtz, Kelvin and Heaviside) was to view generalized functions as suitable types of smooth set-theoretical maps obtained from ordinary smooth maps depending on suitable infinitesimal or infinite parameters.

The calculus of GSF is closely related to classical analysis, in fact:

(i) GSF are set-theoretical maps defined on, and attaining values in the non-Archimedean ring ${}^{\rho}\widetilde{\mathbb{R}}$ of Robinson-Colombeau. Therefore, in ${}^{\rho}\widetilde{\mathbb{R}}$ we have infinitesimals, infinities and also a suitable language of nilpotent infinitesimals, see Sec. 2, Sec. 3.

(ii) GSF include all Colombeau generalized functions and hence also all Schwartz distributions, see Sec. 3.

(iii) They allow nonlinear operations on GF and to compose them unrestrictedly, so that terms such as $\delta^2(x)$ or even $\delta(\delta(x))$ are possible, see Sec. 3.

(iv) GSF allow us to prove a number of analogues of theorems of classical analysis: e.g., mean value theorem, intermediate value theorem, extreme value theorem, Taylor's theorem, local and global inverse function theorem, integrals via primitives, multidimensional integrals, theory of compactly supported GF. Therefore, this approach to GF results in a flexible and rich framework which allows both the formalization of calculations appearing in physics and engineering and the development of new applications in mathematics and mathematical physics. Some of these results are presented in Sec. 3.

(v) Several results of the classical theory of calculus of variations and optimal control can be developed for GSF: the fundamental lemma, second variation and minimizers, necessary Legendre condition, Jacobi fields, conjugate points and Jacobi's theorem, Noether's theorem, see [50, 24].

(vi) The closure with respect to composition leads to a solution concept of differential equations close to the classical one. In GSF theory, we have a non-Archimedean version of the Banach fixed point theorem, a Picard-Lindelöf theorem for both ODE and PDE, results about the maximal set of existence, Gronwall theorem, flux properties, continuous dependence on initial conditions, full compatibility with classical smooth solutions, etc., see Sec. 4.

As we will see in Sec. 5 and Sec. 6, using GSF theory, we have a rigorous theory of infinitesimal and infinite numbers that can be used to develop mathematical models of physical problems. On the other hand, it is also a flexible theory of GF that can be used to model situations with singular (non smooth) physical quantities. One of the main aim of this paper is to show that within GSF theory several informal calculations of physics or engineering now become perfectly rigorous without detaching too much from the original deduction.

The structure of the paper is as follows: in Sec. 2, we introduce our new ring of scalars ${}^{\rho}\widetilde{\mathbb{R}}$, the ring of Robinson-Colombeau. In Sec. 3, we define GSF as suitable set-theoretical functions defined and valued in the new ring of scalars; we will also see the relationships with Colombeau GF and hence with Schwartz distributions. In Sec. 4, we see the Picard-Lindelöf theorem for singular nonlinear ODE involving GSF. In Sec. 5, we see how to transform the classical deductions of the wave and heat equations into formal mathematical theorems whose scope includes GSF. Finally, in

Sec. 6 we see applications to non-smooth mechanics, an empirical non-linear stress-strain model for steel, some applications in optics with discontinuous Lagrangians, how to see the classical finite and infinite potential wells of QM within GSF theory.

The paper is a review of GSF theory, so it is self-contained in the sense that it contains all the statements required for the proofs of Sec. 5 and Sec. 6. We also introduce clear intuitions about the new mathematical objects of this theory and references for the complete proofs. Therefore, to understand this paper, only a basic knowledge of distribution theory is needed.

2. Numbers: The ring of Robinson-Colombeau

A natural method sometimes used in applied mathematics, mathematical physics and engineering in order to deal with non differentiable functions at singularities is to introduce a new parameter $\varepsilon \in (0,1]$ and to approximate the singular function with a smooth one at distance $d(\varepsilon) \to 0^+$ as $\varepsilon \to 0^+$. Instead of dealing with the non differentiable function, we then consider a different model substituting the singular map with this ε-regularized version. For example, if $N_{0,\sigma}(t)$ is a Gaussian with zero mean and standard deviation σ, then $N_{0,\varepsilon}$ and $H(x) = \int_{-\infty}^{x} N_{0,\varepsilon}(t)\,\mathrm{d}t$ can be used to construct an approximate model of Dirac delta and Heaviside step function. Of course, this method depends on the new parameter ε, which usually has no physical meaning, so that the final claims are frequently assessed as $\varepsilon \to 0^+$. However, if the final solution is still a GF, e.g. the dynamics during a collision, the limit as $\varepsilon \to 0^+$ may not exist because certain derivatives become infinite.

Similarly, the values of this ε-regularized functions can be infinitesimal or infinite quantities as $\varepsilon \to 0^+$, e.g. $N_{0,\varepsilon}(0) \to +\infty$, whereas all the calculations in this type of model are meaningful only for a fixed $\varepsilon \in (0,1]$, where these values are still classical real numbers, and hence they cannot be infinitesimal or infinite quantities.

The idea of the ring of Robinson-Colombeau is to create a simple and intuitively clear mathematical setting where this informal idea is fully rigorous and where a full language of infinitesimal or infinite numbers is available.

We can motivate the new ring of scalars as follows: Exactly as real numbers can be seen as equivalence classes of sequences $(q_n)_{n\in\mathbb{N}}$ of rationals[1], it is very natural to consider a non-Archimedean extension of $\mathbb{R}$ defined by a quotient ring $\widetilde{\mathbb{R}} := \mathcal{R}/\sim$, where $\mathcal{R} \subseteq \mathbb{R}^I$ is a subset of all the functions $I \longrightarrow \mathbb{R}$ defined on $I := (0,1]$ which is closed with respect to pointwise algebraic operations (i.e. $\mathcal{R}$ is an algebra). We always call *net* any function in the independent variable ε, and we use for them notations of the type $(x_\varepsilon)_{\varepsilon\in I} \in \mathbb{R}^I$ or simply $(x_\varepsilon) := (x_\varepsilon)_{\varepsilon\in I}$. For simplicity and for historical reasons, we consider $I = (0,1]$, corresponding to $\varepsilon \to 0^+$, $\varepsilon \in I$, as in the previous informal method. In this work, we will denote equivalence classes simply by $[x_\varepsilon] := [(x_\varepsilon)]_\sim \in \widetilde{\mathbb{R}}$. The basic problem is hence to understand when we have to say that two nets (x_ε), $(x'_\varepsilon) \in \mathcal{R}$ are equivalent $(x_\varepsilon) \sim (x'_\varepsilon)$ (i.e. they are two different representatives of the new generalized number $[x_\varepsilon] = [x'_\varepsilon] \in \widetilde{\mathbb{R}}$) in order to obtain a ring containing also infinitesimal and infinite numbers corresponding to ordinary infinitesimal and infinite functions (nets) as $\varepsilon \to 0^+$.

The following observation points to a natural way of achieving this goal. Let us assume that $[z_\varepsilon] = 0 \in \widetilde{\mathbb{R}}$ and $[J_\varepsilon] \in \widetilde{\mathbb{R}}$ is generated by an infinite net (J_ε), i.e. such

[1] In the naturals $\mathbb{N} = \{0, 1, 2, \dots\}$ we include zero.

that $\lim_{\varepsilon\to0^+}|J_\varepsilon| = +\infty$. Then we would have

$$\begin{aligned}[z_\varepsilon]\cdot[J_\varepsilon] &= 0\cdot[J_\varepsilon] = 0\\ &= [z_\varepsilon\cdot J_\varepsilon].\end{aligned} \tag{2.1}$$

Finally, let us assume that

$$\forall[w_\varepsilon]\in\widetilde{\mathbb{R}}:\ [w_\varepsilon]=0\ \Rightarrow\ \lim_{\varepsilon\to0^+} w_\varepsilon = 0. \tag{2.2}$$

Under these assumptions, (2.1) yields $\lim_{\varepsilon\to0^+} z_\varepsilon\cdot J_\varepsilon = 0$, and hence

$$\exists\varepsilon_0\in I\,\forall\varepsilon\in(0,\varepsilon_0]:\ |z_\varepsilon|\le\left|J_\varepsilon^{-1}\right|. \tag{2.3}$$

Consequently, *the nets* (z_ε) *representing* 0*, i.e. such that* $(z_\varepsilon)\sim0$*, must be dominated by the reciprocals of every infinite number* $[J_\varepsilon]\in\widetilde{\mathbb{R}}$. It is not hard to prove that if *every* infinite net (J_ε) is in $\mathcal{R}$, then (2.3) implies that the equivalence relation $\sim$ must be trivial:

$$\exists\varepsilon_0\in I\,\forall\varepsilon\in(0,\varepsilon_0]:\ z_\varepsilon=0. \tag{2.4}$$

This situation corresponds to the Schmieden-Laugwitz model, [70].

If we do not want to have the trivial model (2.4), we can hence either negate the natural property (2.2) (this is the case of nonstandard analysis; see [16] for more details) or to restrict the class of all the nets (J_ε) generating infinite numbers in $\widetilde{\mathbb{R}}$. Since we want to start from an algebra $\mathcal{R}\subseteq\mathbb{R}^I$, a *first* natural idea is to fix an infinitesimal net (ρ_ε), $\rho_\varepsilon\to0^+$, and to consider the following class of infinite nets

$$\mathcal{I} := \left\{(\rho_\varepsilon^{-a})\mid a\in\mathbb{R}_{>0}\right\}. \tag{2.5}$$

and hence to consider the algebra $\mathcal{R}\subseteq\mathbb{R}^I$ containing nets $(b_\varepsilon)\in\mathbb{R}^I$ bounded by some $(J_\varepsilon)\in\mathcal{I}$. Therefore, the fixed net (ρ_ε) yields a measure of the strongest infinite nets representing numbers in $\widetilde{\mathbb{R}}=\mathcal{R}/\sim$.

This idea is generalized in the following definition, where we take exactly (2.3) as the widest possible definition of $(z_\varepsilon)\sim0$. More in depth details about these basic notions and the omitted proofs as well can be found in [32, 31].

Definition 1. Let $I:=(0,1]$ and $\rho=(\rho_\varepsilon)\in(0,1]^I$ be a net such that $(\rho_\varepsilon)\to0$ as $\varepsilon\to0^+$ (in the following, such a net will be called a *gauge*), then

(i) $\mathcal{I}(\rho):=\{(\rho_\varepsilon^{-a})\mid a\in\mathbb{R}_{>0}\}$ is called the *asymptotic gauge* generated by ρ.

(ii) If $\mathcal{P}(\varepsilon)$ is a property of $\varepsilon\in I$, we use the notation $\forall^0\varepsilon:\ \mathcal{P}(\varepsilon)$ to denote $\exists\varepsilon_0\in I\,\forall\varepsilon\in(0,\varepsilon_0]:\ \mathcal{P}(\varepsilon)$. We can read $\forall^0\varepsilon$ as: *"for* ε *small"*.

(iii) We say that a net $(x_\varepsilon)\in\mathbb{R}^I$ *is* ρ*-moderate*, and we write $(x_\varepsilon)\in\mathbb{R}_\rho$, if

$$\exists(J_\varepsilon)\in\mathcal{I}(\rho):\ x_\varepsilon=O(J_\varepsilon)\text{ as }\varepsilon\to0^+, \tag{2.6}$$

i.e., if

$$\exists N\in\mathbb{N}\,\forall^0\varepsilon:\ |x_\varepsilon|\le\rho_\varepsilon^{-N}.$$

(iv) Let (x_ε), $(y_\varepsilon)\in\mathbb{R}^I$, then we say that $(x_\varepsilon)\sim_\rho(y_\varepsilon)$ if

$$\forall(J_\varepsilon)\in\mathcal{I}(\rho):\ x_\varepsilon=y_\varepsilon+O(J_\varepsilon^{-1})\text{ as }\varepsilon\to0^+,$$

that is if

$$\forall n\in\mathbb{N}\,\forall^0\varepsilon:\ |x_\varepsilon-y_\varepsilon|\le\rho_\varepsilon^n. \tag{2.7}$$

This is a congruence relation on the ring $\mathbb{R}_\rho$ of ρ-moderate nets with respect to pointwise operations (i.e. the pointwise algebraic operations $[x_\varepsilon]+[y_\varepsilon] := [x_\varepsilon + y_\varepsilon]$ and $[x_\varepsilon]\cdot[y_\varepsilon] := [x_\varepsilon \cdot y_\varepsilon]$ are well-defined), and we can hence define

$$ {}^{\rho}\widetilde{\mathbb{R}} := \mathbb{R}_\rho / \sim_\rho, \tag{2.8}$$

which we call *Robinson-Colombeau ring of generalized numbers*.

This name is justified by [67, 13]: Indeed, in [67] A. Robinson introduced the notion of moderate and negligible nets depending on an arbitrary fixed infinitesimal ρ (in the framework of nonstandard analysis); independently, J.F. Colombeau, cf. e.g. [13] and references therein, studied the same concepts without using nonstandard analysis, but considering only the particular infinitesimal $\rho_\varepsilon = \varepsilon$. Equivalence classes of the quotient ring (2.8) are simply denoted with $[x_\varepsilon] := [(x_\varepsilon)_\varepsilon]_{\sim_\rho} \in {}^{\rho}\widetilde{\mathbb{R}}$.

Considering constant net $x_\varepsilon = r \in \mathbb{R}$ we have the embedding $\mathbb{R} \subset {}^{\rho}\widetilde{\mathbb{R}}$. We define $[x_\varepsilon] \le [y_\varepsilon]$ if there exists $(z_\varepsilon) \in \mathbb{R}^I$ such that $(z_\varepsilon) \sim_\rho 0$ (we then say that (z_ε) is *ρ-negligible*) and $x_\varepsilon \le y_\varepsilon + z_\varepsilon$ for ε small. Equivalently, we have that $x \le y$ if and only if there exist representatives $[x_\varepsilon] = x$ and $[y_\varepsilon] = y$ such that $x_\varepsilon \le y_\varepsilon$ for all ε.

A proficient intuitive point of view on these generalized numbers is to think at $[x_\varepsilon] \in {}^{\rho}\widetilde{\mathbb{R}}$ as a *dynamic point in the time $\varepsilon \to 0^+$*; classical real numbers are hence *static points*. This corresponds to the informal method presented just before Def. 1 above. Morever, we say that $x = [x_\varepsilon] \in {}^{\rho}\widetilde{\mathbb{R}}$ is *near-standard* if $\exists \lim_{\varepsilon\to 0^+} x_\varepsilon =: x^\circ \in \mathbb{R}$.

Even though the order $\le$ is not total, we still have the possibility to define the infimum $[x_\varepsilon] \wedge [y_\varepsilon] := [\min(x_\varepsilon, y_\varepsilon)]$, the supremum $[x_\varepsilon] \vee [y_\varepsilon] := [\max(x_\varepsilon, y_\varepsilon)]$ of a finite amount of generalized numbers. Henceforth, we will also use the customary notation ${}^{\rho}\widetilde{\mathbb{R}}^*$ for the set of invertible (we recall that $x \in {}^{\rho}\widetilde{\mathbb{R}}$ is invertible if $\exists y \in {}^{\rho}\widetilde{\mathbb{R}} : x\cdot y = 1$) generalized numbers, and we write $x < y$ to say that $x \le y$ and $x - y \in {}^{\rho}\widetilde{\mathbb{R}}^*$, i.e. if x is less of equal to y and $x - y$ is invertible. The intervals are denoted by: $[a, b] := \{x \in {}^{\rho}\widetilde{\mathbb{R}} \mid a \le x \le b\}$, $[a, b]_{\mathbb{R}} := [a, b] \cap \mathbb{R}$. Finally, we set $\mathrm{d}\rho := [\rho_\varepsilon] \in {}^{\rho}\widetilde{\mathbb{R}}$, which is a positive invertible infinitesimal, whose reciprocal is $\mathrm{d}\rho^{-1} = [\rho_\varepsilon^{-1}]$, which is necessarily a strictly positive infinite number. It is remarkable to note that $x = [x_\varepsilon] \in {}^{\rho}\widetilde{\mathbb{R}}$ is an infinitesimal number, i.e. $|x| \le r$ for all $r \in \mathbb{R}_{>0}$, denoted by $x \approx 0$, if and only if $\lim_{\varepsilon\to 0^+} x_\varepsilon = 0$; similarly, x is an infinite number, i.e. $|x| \ge r$ for all $r \in \mathbb{R}_{>0}$, if and only if $\lim_{\varepsilon\to 0^+} |x_\varepsilon| = +\infty$. This intuitively clear result is not possible neither in nonstandard analysis nor in synthetic differential geometry, see [27, 36, 43].

The following result proves to be useful in dealing with positive and invertible generalized numbers. For its proof, see e.g. [35].

Lemma 2. *Let $x \in {}^{\rho}\widetilde{\mathbb{R}}$. Then the following are equivalent:*

(i) *x is invertible and $x \ge 0$, i.e. $x > 0$.*
(ii) *For each representative $(x_\varepsilon) \in \mathbb{R}_\rho$ of x we have $\forall^0\varepsilon : \ x_\varepsilon > 0$.*
(iii) *For each representative $(x_\varepsilon) \in \mathbb{R}_\rho$ of x we have $\exists m \in \mathbb{N}\, \forall^0\varepsilon : \ x_\varepsilon > \rho_\varepsilon^m$.*
(iv) *There exists a representative $(x_\varepsilon) \in \mathbb{R}_\rho$ of x such that $\exists m \in \mathbb{N}\, \forall^0\varepsilon : \ x_\varepsilon > \rho_\varepsilon^m$.*

One can clearly feel insecure in working with a ring of scalar which is not a totally ordered field (i.e. it does not hold that $x \le y$ or $x \ge y$ and we can have $x \ne 0$ but $\nexists y \in {}^{\rho}\widetilde{\mathbb{R}} : \ x\cdot y = 1$). On the one hand, we can reread the list of results presented in Sec. 1 to get a reassurance that these properties are actually not indispensable to obtain all these well-known classical results. On the other hand, using the notion of subpoint (e.g. a meaningful case is given by a subpoint $[x_{\varepsilon_n}]$ of $[x_\varepsilon]$ which is considered only on a sequence $(\varepsilon_n)_{n\in\mathbb{N}} \to 0^+$), see [59], we developed very practical substitutes of both the field and the total order property.

2.1. **Topologies on ${}^\rho\widetilde{\mathbb{R}}^n$.** A first non-trivial conceptual step is to consider ${}^\rho\widetilde{\mathbb{R}}$ as our new ring of scalar. The natural extension of the Euclidean norm on the ${}^\rho\widetilde{\mathbb{R}}$-module ${}^\rho\widetilde{\mathbb{R}}^n$, i.e. $|[x_\varepsilon]| := [|x_\varepsilon|] \in {}^\rho\widetilde{\mathbb{R}}$, where $[x_\varepsilon] \in {}^\rho\widetilde{\mathbb{R}}^n$, goes exactly in this direction. In fact, even if this generalized norm takes values in ${}^\rho\widetilde{\mathbb{R}}$, and not in the old $\mathbb{R}$, it shares some essential properties with classical norms:

$$
\begin{aligned}
&|x| = x \vee (-x) \\
&|x| \geq 0 \\
&|x| = 0 \Rightarrow x = 0 \\
&|y \cdot x| = |y| \cdot |x| \\
&|x + y| \leq |x| + |y| \\
&||x| - |y|| \leq |x - y|.
\end{aligned}
$$

It is therefore natural to consider on ${}^\rho\widetilde{\mathbb{R}}^n$ topologies generated by balls defined by this generalized norm and a set of radii. A second non-trivial step is to understand that the meaningful set of radii we need to have continuity of our class of generalized function is the set ${}^\rho\widetilde{\mathbb{R}}^*_{\geq 0} = {}^\rho\widetilde{\mathbb{R}}_{>0}$ of positive and invertible generalized numbers:

Definition 3. We define

(i) $B_r(x) := \left\{ y \in {}^\rho\widetilde{\mathbb{R}}^n \mid |y - x| < r \right\}$ for any $r \in {}^\rho\widetilde{\mathbb{R}}_{>0}$.

(ii) $B^{\mathrm{E}}_r(x) := \{ y \in \mathbb{R}^n \mid |y - x| < r \}$, for any $r \in \mathbb{R}_{>0}$, denotes an ordinary Euclidean ball in $\mathbb{R}^n$.

The relation $<$ has more beneficial topological properties as compared to the usual strict order relation $x \leq y$ and $x \neq y$ (a relation that we will therefore *never* use) due to the property that the set of balls $\left\{ B_r(x) \mid r \in {}^\rho\widetilde{\mathbb{R}}_{>0},\ x \in {}^\rho\widetilde{\mathbb{R}}^n \right\}$ is a base for a topology on ${}^\rho\widetilde{\mathbb{R}}^n$ called *sharp topology*, and we call *sharply open set* any open set in this topology. Therefore, $A \subseteq {}^\rho\widetilde{\mathbb{R}}^n$ is a sharply open set if for each $a \in A$ there exists a radius $r \in {}^\rho\widetilde{\mathbb{R}}_{>0}$ such that $B_r(a) \subseteq A$.
We also recall that the sharp topology on ${}^\rho\widetilde{\mathbb{R}}^n$ is Hausdorff and Cauchy complete, see e.g. [32, 31]. A peculiar property of the sharp topology is that it is also generated by all the infinitesimal balls of the form $B_{\mathrm{d}\rho^q}(x)$, where $q \in \mathbb{N}_{>0}$. The necessity to consider infinitesimal neighborhoods occurs in any theory containing continuous GF which have infinite derivatives. Indeed, from the mean value theorem Thm. 27.(i) below, we have $f(x) - f(x_0) = f'(c) \cdot (x - x_0)$ for some $c \in [x, x_0]$. Therefore, we have $f(x) \in B_r(f(x_0))$, for a given $r \in {}^\rho\widetilde{\mathbb{R}}_{>0}$, if and only if $|x - x_0| \cdot |f'(c)| < r$, which yields an infinitesimal neighborhood of x_0 in case $f'(c)$ is infinite; see [30, 31] for precise statements and proofs corresponding to this intuition. On the other hand, the existence of infinitesimal neighborhoods implies that the sharp topology induces the discrete topology on $\mathbb{R}$; once again, this is a general result that occurs in all the theories of infinitesimals, see [30].

A natural way to obtain sharply open, closed and bounded sets in ${}^\rho\widetilde{\mathbb{R}}^n$ is by using a net (A_ε) of subsets $A_\varepsilon \subseteq \mathbb{R}^n$. Once again, thinking at $[x_\varepsilon]$ and (A_ε) as a dynamic point and set as the time $\varepsilon \to 0^+$, we have two ways of extending the membership relation $x_\varepsilon \in A_\varepsilon$ to generalized points $[x_\varepsilon] \in {}^\rho\widetilde{\mathbb{R}}^n$:

Definition 4. Let (A_ε) be a net of subsets of $\mathbb{R}^n$, then

(i) $[A_\varepsilon] := \left\{ [x_\varepsilon] \in {}^\rho\widetilde{\mathbb{R}}^n \mid \forall^0\varepsilon : \ x_\varepsilon \in A_\varepsilon \right\}$ is called the *internal set* generated by the net (A_ε).

(ii) Let (x_ε) be a net of points of $\mathbb{R}^n$, then we say that $x_\varepsilon \in_\varepsilon A_\varepsilon$, and we read it as (x_ε) *strongly belongs to* (A_ε), if
   (i) $\forall^0\varepsilon : \ x_\varepsilon \in A_\varepsilon$.
   (ii) If $(x'_\varepsilon) \sim_\rho (x_\varepsilon)$, then also $x'_\varepsilon \in A_\varepsilon$ for ε small.

   Moreover, we set $\langle A_\varepsilon \rangle := \left\{ [x_\varepsilon] \in {}^\rho\widetilde{\mathbb{R}}^n \mid x_\varepsilon \in_\varepsilon A_\varepsilon \right\}$, and we call it the *strongly internal set* generated by the net (A_ε).

Therefore, $x \in [A_\varepsilon]$ if *there exists a representative* $[x_\varepsilon] = x$ such that $x_\varepsilon \in A_\varepsilon$ for ε small, whereas this membership is independent from the chosen representative in case of strongly internal sets: if $[x'_\varepsilon] = [x_\varepsilon] \in \langle A_\varepsilon \rangle$, then for ε sufficiently small *both representatives satisfy* x'_ε, $x_\varepsilon \in A_\varepsilon$. An internal set generated by a constant net $A_\varepsilon = A \subseteq \mathbb{R}^n$ will simply be denoted by $[A]$.

The following theorem shows that internal and strongly internal sets have dual topological properties:

Theorem 5. *For $\varepsilon \in I$, let $A_\varepsilon \subseteq \mathbb{R}^n$ and let $x_\varepsilon \in \mathbb{R}^n$. Then we have*

(i) $[x_\varepsilon] \in [A_\varepsilon]$ *if and only if* $\forall q \in \mathbb{R}_{>0}\, \forall^0\varepsilon : \ d(x_\varepsilon, A_\varepsilon) \le \rho_\varepsilon^q$. *Therefore* $[x_\varepsilon] \in [A_\varepsilon]$ *if and only if* $[d(x_\varepsilon, A_\varepsilon)] = 0 \in {}^\rho\widetilde{\mathbb{R}}$.

(ii) $[x_\varepsilon] \in \langle A_\varepsilon \rangle$ *if and only if* $\exists q \in \mathbb{R}_{>0}\, \forall^0\varepsilon : \ d(x_\varepsilon, A_\varepsilon^c) > \rho_\varepsilon^q$, *where* $A_\varepsilon^c := \mathbb{R}^n \setminus A_\varepsilon$. *Therefore, if* $(d(x_\varepsilon, A_\varepsilon^c)) \in \mathbb{R}_\rho$, *then* $[x_\varepsilon] \in \langle A_\varepsilon \rangle$ *if and only if* $[d(x_\varepsilon, A_\varepsilon^c)] > 0$.

(iii) $[A_\varepsilon]$ *is sharply closed.*

(iv) $\langle A_\varepsilon \rangle$ *is sharply open.*

(v) $[A_\varepsilon] = [\mathrm{cl}\,(A_\varepsilon)]$, *where* $\mathrm{cl}\,(S)$ *is the closure of* $S \subseteq \mathbb{R}^n$.

(vi) $\langle A_\varepsilon \rangle = \langle \mathrm{int}(A_\varepsilon) \rangle$, *where* $\mathrm{int}\,(S)$ *is the interior of* $S \subseteq \mathbb{R}^n$.

For example, it is not hard to show that the closure in the sharp topology of a ball of center $c = [c_\varepsilon]$ and radius $r = [r_\varepsilon] > 0$ is

$$\overline{B_r(c)} = \left\{ x \in {}^\rho\widetilde{\mathbb{R}}^d \mid |x - c| \le r \right\} = \left[\overline{B^{\mathrm{E}}_{r_\varepsilon}(c_\varepsilon)} \right], \tag{2.9}$$

whereas

$$B_r(c) = \left\{ x \in {}^\rho\widetilde{\mathbb{R}}^d \mid |x - c| < r \right\} = \langle B^{\mathrm{E}}_{r_\varepsilon}(c_\varepsilon) \rangle.$$

The reader can be concerned with the fact that the ring of scalar ${}^\rho\widetilde{\mathbb{R}}$ is not a totally ordered field. Besides the language of subpoints (see [59]) that allows one to proceed alternatively when total order or invertibility properties are in play, the following result is also useful:

Lemma 6. *Invertible elements of ${}^\rho\widetilde{\mathbb{R}}$ are dense in the sharp topology, i.e.*

$$\forall h \in {}^\rho\widetilde{\mathbb{R}}\, \forall \delta \in {}^\rho\widetilde{\mathbb{R}}_{>0}\, \exists k \in (h - \delta, h + \delta) : \ k \text{ is invertible.}$$

This is even more important since our GSF are continuous in the sharp topology, as we will see in the next section.

3. Functions and distributions: generalized smooth functions

After the introduction of numbers, their sets and topologies, we introduce the notion of function.

3.1. **Definition of GSF and sharp continuity.** Using the ring ${}^{\rho}\widetilde{\mathbb{R}}$, it is easy to consider a Gaussian with an infinitesimal standard deviation. If we denote this probability density by $f(x,\sigma)$, and if we set $\sigma=[\sigma_\varepsilon]\in{}^{\rho}\widetilde{\mathbb{R}}_{>0}$, where $\sigma\approx0$, we obtain the net of smooth functions $(f(-,\sigma_\varepsilon))_{\varepsilon\in I}$. This is the basic idea we are going to develop in the following definitions and it corresponds to informal method we explained just before Def. 1. We will first introduce the notion of a net of functions (f_ε) defining a generalized smooth function of the type $X\longrightarrow Y$, where $X\subseteq{}^{\rho}\widetilde{\mathbb{R}}^n$ and $Y\subseteq{}^{\rho}\widetilde{\mathbb{R}}^d$. This is a net of smooth functions $f_\varepsilon\in\mathcal{C}^\infty(\Omega_\varepsilon,\mathbb{R}^d)$ that induces well-defined maps of the form $[\partial^\alpha f_\varepsilon(-)]:\langle\Omega_\varepsilon\rangle\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}^d$, for every multi-index $\alpha\in\mathbb{N}^n$.

Definition 7. Let $X\subseteq{}^{\rho}\widetilde{\mathbb{R}}^n$ and $Y\subseteq{}^{\rho}\widetilde{\mathbb{R}}^d$ be arbitrary subsets of generalized points. Let (Ω_ε) be a net of open subsets of $\mathbb{R}^n$, and (f_ε) be a net of smooth functions, with $f_\varepsilon\in\mathcal{C}^\infty(\Omega_\varepsilon,\mathbb{R}^d)$. Then, we say that

$$(f_\varepsilon)\ \textit{defines a generalized smooth function}:X\longrightarrow Y$$

if:

(i) $X\subseteq\langle\Omega_\varepsilon\rangle$ and $[f_\varepsilon(x_\varepsilon)]\in Y$ for all $[x_\varepsilon]\in X$.
(ii) $\forall[x_\varepsilon]\in X\,\forall\alpha\in\mathbb{N}^n:\ (\partial^\alpha f_\varepsilon(x_\varepsilon))\in\mathbb{R}^d_\rho$.

Where the notation

$$\forall[x_\varepsilon]\in X:\ \mathcal{P}\{(x_\varepsilon)\}$$

means

$$\forall(x_\varepsilon)\in\mathbb{R}^n_\rho:\ [x_\varepsilon]\in X\ \Rightarrow\ \mathcal{P}\{(x_\varepsilon)\},$$

i.e. for all representatives (x_ε) generating a point $[x_\varepsilon]\in X$, the property $\mathcal{P}\{(x_\varepsilon)\}$ holds.

A generalized smooth function (or map, in this paper these terms are used as synonymous) is simply a function of the form $f=[f_\varepsilon(-)]|_X$:

Definition 8. Let $X\subseteq{}^{\rho}\widetilde{\mathbb{R}}^n$ and $Y\subseteq{}^{\rho}\widetilde{\mathbb{R}}^d$ be arbitrary subsets of generalized points, then we say that

$$f:X\longrightarrow Y\text{ is a }\textit{generalized smooth function}$$

if $f\in\mathbf{Set}(X,Y)$ and there exists a net $f_\varepsilon\in\mathcal{C}^\infty(\Omega_\varepsilon,\mathbb{R}^d)$ defining a generalized smooth map of type $X\longrightarrow Y$, in the sense of Def. 7, such that

$$\forall[x_\varepsilon]\in X:\ f([x_\varepsilon])=[f_\varepsilon(x_\varepsilon)]\,. \tag{3.1}$$

We will also say that f *is defined by the net of smooth functions* (f_ε) or that the net (f_ε) *represents* f. The set of all these GSF will be denoted by ${}^{\rho}\mathcal{GC}^\infty(X,Y)$.

Let us note explicitly that definitions 7 and 8 state minimal logical conditions to obtain a set-theoretical map from X into Y which is defined by a net of smooth functions such that all the derivatives still lie in our ring of scalars for condition Def. 7.(ii). In particular, the following Thm. 9 states that in equality (3.1) we have independence from the representatives for all derivatives $[x_\varepsilon]\in X\mapsto[\partial^\alpha f_\varepsilon(x_\varepsilon)]\in{}^{\rho}\widetilde{\mathbb{R}}^d$, $\alpha\in\mathbb{N}^n$.

Theorem 9. *Let $X \subseteq {}^\rho\widetilde{\mathbb{R}}^n$ and $Y \subseteq {}^\rho\widetilde{\mathbb{R}}^d$ be arbitrary subsets of generalized points. Let (Ω_ε) be a net of open subsets of $\mathbb{R}^n$, and (f_ε) be a net of smooth functions, with $f_\varepsilon \in \mathcal{C}^\infty(\Omega_\varepsilon, \mathbb{R}^d)$. Assume that (f_ε) defines a generalized smooth map of the type $X \longrightarrow Y$, then*

$$\forall \alpha \in \mathbb{N}^n \, \forall (x_\varepsilon), (x'_\varepsilon) \in \mathbb{R}^n_\rho : \; [x_\varepsilon] = [x'_\varepsilon] \in X \;\Rightarrow\; (\partial^\alpha f_\varepsilon(x_\varepsilon)) \sim_\rho (\partial^\alpha f_\varepsilon(x'_\varepsilon)).$$

Note that taking arbitrary subsets $X \subseteq {}^\rho\widetilde{\mathbb{R}}^n$ in Def. 7, we can also consider GSF defined on closed sets, like the set of all infinitesimals (which is also open, like in all non trivial theories of infinitesimals), or like a closed interval $[a, b] \subseteq {}^\rho\widetilde{\mathbb{R}}$. We can also consider GSF defined at infinite generalized points. A simple case is the exponential map

$$e^{(-)} : [x_\varepsilon] \in \left\{ x \in {}^\rho\widetilde{\mathbb{R}} \mid \exists z \in {}^\rho\widetilde{\mathbb{R}}_{>0} : \; x \le \log z \right\} \mapsto [e^{x_\varepsilon}] \in {}^\rho\widetilde{\mathbb{R}}. \tag{3.2}$$

The domain of this map depends on the infinitesimal net ρ. For instance, if $\rho = (\varepsilon)$ then all its points are bounded by generalized numbers of the form $[-N \log \varepsilon]$, $N \in \mathbb{N}$; whereas if $\rho = \left(e^{-\frac{1}{\varepsilon}}\right)$, all points are bounded by $[N\varepsilon^{-1}]$, $N \in \mathbb{N}$. This underscores the importance to choose different gauges ρ depending on our needs.

A first regularity property of GSF is the above cited continuity with respect to the sharp topology, as proved in the following

Theorem 10. *Let $X \subseteq {}^\rho\widetilde{\mathbb{R}}^n$, $Y \subseteq {}^\rho\widetilde{\mathbb{R}}^d$ and $f_\varepsilon \in \mathcal{C}^\infty(\Omega_\varepsilon, \mathbb{R}^d)$ be a net of smooth functions that defines a GSF of the type $X \longrightarrow Y$. Then*

(i) *For all $\alpha \in \mathbb{N}^n$, the GSF $g : [x_\varepsilon] \in X \mapsto [\partial^\alpha f_\varepsilon(x_\varepsilon)] \in \widetilde{\mathbb{R}}^d$ is locally Lipschitz in the sharp topology, i.e. each $x \in X$ possesses a sharp neighborhood U such that $|g(x) - g(y)| \le L|x - y|$ for all x, $y \in U$ and some $L \in {}^\rho\widetilde{\mathbb{R}}$.*

(ii) *Each $f \in {}^\rho\mathcal{GC}^\infty(X, Y)$ is continuous with respect to the sharp topologies induced on X, Y.*

(iii) *$f : X \longrightarrow Y$ is a GSF if and only if there exists a net $v_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^d)$ defining a generalized smooth map of type $X \longrightarrow Y$ such that $f = [v_\varepsilon(-)]|_X$.*

3.2. **Embedding of Schwartz distributions.** Among the re-occurring themes of this work are the choices which the solution of a given problem within our framework may depend upon. For instance, (3.2) shows that the domain of a GSF depends on the infinitesimal net ρ. It is also easy to show that the trivial Cauchy problem

$$\begin{cases} x'(t) = [\varepsilon^{-1}] \cdot x(t) \\ x(0) = 1 \end{cases}$$

has no solution in ${}^\rho\mathcal{GC}^\infty(\mathbb{R}, \mathbb{R})$ if $\rho_\varepsilon = \varepsilon$ because the solution is not moderate e.g. at $t = 1$. Nevertheless, it has the unique solution $x(t) = \left[e^{\frac{1}{\varepsilon}t}\right] \in {}^\rho\mathcal{GC}^\infty(\mathbb{R}, \mathbb{R})$ if $\rho_\varepsilon = e^{-\frac{1}{\varepsilon}}$. Therefore, the choice of the infinitesimal net ρ is closely tied to the possibility of solving a given class of differential equations in *non infinitesimal* intervals (a solution in a suitable *infinitesimal* interval always exists, see Sec. 4). This illustrates the dependence of the theory on the infinitesimal net ρ.

Further choices concern the embedding of Schwartz distributions: Since we need to associate a net of smooth functions (f_ε) to a given distribution $T \in \mathcal{D}'(\Omega)$ (e.g. T can be any continuous non-differentiable function defined on Ω), this embedding is naturally built upon a regularization process: this corresponds to the informal method explained in Sec. 2. In our approach, this regularization will depend on an infinite number $b \in {}^{\rho}\widetilde{\mathbb{R}}$, and the choice of b depends on what properties we need from the embedding. For example, if δ is the (embedding of the) one-dimensional Dirac delta, then we have the property

$$\delta(0) = b, \tag{3.3}$$

We can also choose the embedding so as to get the property

$$H(0) = \frac{1}{2}, \tag{3.4}$$

where H is the (embedding of the) Heaviside step function. Equalities like these are used in diverse applications (see, e.g., [13] and references therein). In fact, we are going to construct a family of structures depending on a linear embedding of Schwartz distributions $\iota_\Omega : \mathcal{D}'(\Omega) \longrightarrow {}^{\rho}\mathcal{GC}^\infty(\mathrm{c}(\Omega), {}^{\rho}\widetilde{\mathbb{R}})$ (see below, Thm. 12). The particular structure we need to consider depends on the problem we have to solve. Of course, one may be more interested in having an intrinsic embedding of distributions. This can be done by following the ideas of the full Colombeau algebra (see e.g. [35]). Nevertheless, this choice decreases the simplicity of the present approach and is incompatible with properties like (3.3) and (3.4).

If $\varphi \in \mathcal{D}(\mathbb{R}^n)$, $r \in \mathbb{R}_{>0}$ and $x \in \mathbb{R}^n$, we use the notation $r \odot \varphi$ for the function $x \in \mathbb{R}^n \mapsto \frac{1}{r^n} \cdot \varphi\left(\frac{x}{r}\right) \in \mathbb{R}$. Our embedding procedure will ultimately rely on convolution with suitable mollifiers. In fact, it is well-known that if $\varphi \in \mathcal{D}(\mathbb{R}^n)$ is a compactly supported smooth function and $f \in \mathcal{C}^0(\mathbb{R}, \mathbb{R})$ is only a continuous function, then the net of smooth functions $\left(\varepsilon^{-1} \odot \varphi\right)$ (which is called a *mollifier*) behaves like the Dirac delta and the convolution $f * \left(\varepsilon^{-1} \odot \varphi\right)$ is an ε-net of *smooth* functions which regularizes singular points of f and converges to f as $\varepsilon \to 0^+$: it is hence a good candidate to replace f as GSF. This is one of the basic ideas to formalize the intuitive method presented just before Def. 1.

To construct the mollifiers which fully preserve smooth functions, we need some technical preparations.

Lemma 11. *For any $n \in \mathbb{N}_{>0}$ there exists some $\mu_n \in \mathcal{S}(\mathbb{R})$ with the following properties:*

(i) $\int \mu_n(x)\,\mathrm{d}x = 1$.
(ii) $\int_0^\infty x^{\frac{j}{n}} \mu_n(x)\,\mathrm{d}x = 0$ *for all* $j \in \mathbb{N}_{>0}$.
(iii) $\mu_n(0) = 1$.
(iv) μ_n *is even.*
(v) $\mu_n(k) = 0$ *for all* $k \in \mathbb{Z} \setminus \{0\}$.

We call *Colombeau mollifier* (for a fixed dimension n) any function μ that satisfies the properties of the previous lemma. Concerning embeddings of Schwartz distributions, the idea is classically to regularize distributions using a mollifier. The use of a Colombeau mollifier allows us, on the one hand, to identify the distribution $\varphi \in \mathcal{D}(\Omega) \mapsto \int f\varphi$ with the GSF $f \in \mathcal{C}^\infty(\Omega) \subseteq {}^{\rho}\mathcal{GC}^\infty(\Omega, \mathbb{R})$ (thanks to property (ii)); on the other hand, it allows us to explicitly calculate compositions such as $\delta \circ \delta$, $H \circ \delta$, $\delta \circ H$ (see below).

As a final preparation for the embedding of $\mathcal{D}'(\Omega)$ into ${}^{\rho}\mathcal{GC}^\infty(\mathrm{c}(\Omega), {}^{\rho}\widetilde{\mathbb{R}})$ we need to construct suitable n-dimensional mollifiers from a Colombeau mollifier μ as given

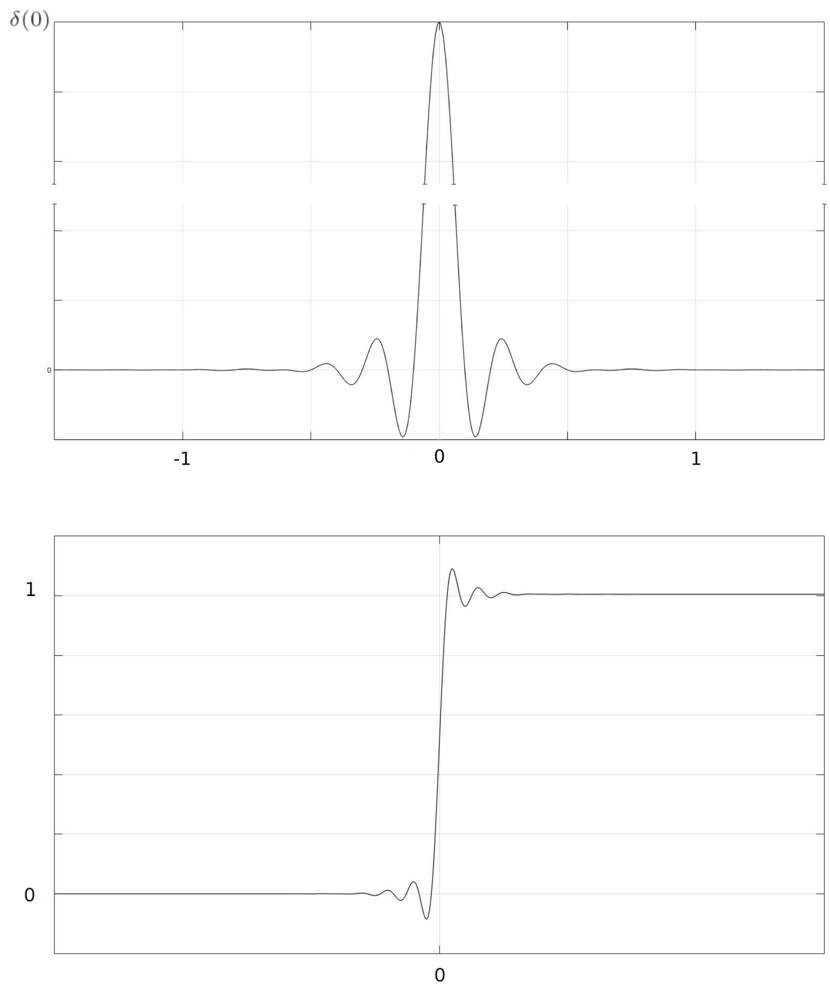

FIGURE 3.1. A representation of Dirac delta and Heaviside function. A Colombeau mollifier has a representation similar to Dirac delta (but with finite values).

by Lemma 11. To this end, let ω_n denote the surface area of S^{n-1} and set

$$c_n := \begin{cases} \frac{2n}{\omega_n} & \text{for } n > 1 \\ 1 & \text{for } n = 1. \end{cases}$$

Then let $\tilde{\mu} : \mathbb{R}^n \to \mathbb{R}$, $\tilde{\mu}(x) := c_n \mu(|x|^n)$. Since μ is even, $\tilde{\mu}$ is smooth. Moreover, by Lemma 11, it has unit integral and all its higher moments $\int x^\alpha \tilde{\mu}(x)\, dx$ vanish ($|\alpha| \geq 1$).

Schwartz distributions are naturally defined only on *finite* points of $\langle\Omega\rangle$ (also called *compactly supported points*), i.e. on the set

$$\mathrm{c}(\Omega) := \{x \in \langle\Omega\rangle \mid \exists R \in \mathbb{R}_{>0} : \ |x| \leq R, \ d(x, \partial\Omega) \in \mathbb{R}_{>0}\}$$

of finite points that remain sufficiently far from the boundary. This underscores an important difference between this type of GF and GSF, since the latter can also be defined on purely infinitesimal domains (note that $\Omega \subseteq \mathrm{c}(\Omega)$) or on infinite points.

Theorem 12. *Let $(\emptyset \neq)\Omega \subseteq \mathbb{R}^n$ be an open set. Set*

$$\Omega_\varepsilon := \left\{ x \in \Omega \mid d(x, \Omega^c) \geq \varepsilon,\ |x| \leq \frac{1}{\varepsilon} \right\}$$

and fix some $\chi \in \mathcal{D}(\mathbb{R}^n)$, $\chi = 1$ on $\overline{B_1^{\mathrm{E}}(0)}$, $0 \leq \chi \leq 1$ and $\chi = 0$ on $\mathbb{R}^n \setminus B_2^{\mathrm{E}}(0)$. Take $\kappa_\varepsilon \in \mathcal{D}(\Omega)$ such that $\kappa_\varepsilon = 1$ on a neighborhood L_ε of Ω_ε. Also, let $b = [b_\varepsilon] \in {}^\rho\widetilde{\mathbb{R}}$ be an infinite positive number, i.e. $\lim_{\varepsilon \to 0^+} b_\varepsilon = +\infty$. Set

$$\mu_\varepsilon^b(x) := (b_\varepsilon^{-1} \odot \tilde{\mu})(x)\chi(x|\log(b_\varepsilon)|) = b_\varepsilon^n \tilde{\mu}(b_\varepsilon x)\chi(x|\log(b_\varepsilon)|). \tag{3.5}$$

Then the map

$$\iota_\Omega^b : T \in \mathcal{D}'(\Omega) \mapsto \left[\left((\kappa_\varepsilon \cdot T) * \mu_\varepsilon^b \right)(-) \right] \in {}^\rho\mathcal{GC}^\infty(\mathrm{c}(\Omega), {}^\rho\widetilde{\mathbb{R}}). \tag{3.6}$$

satisfies:

- *(i)* *$\iota^b : \mathcal{D}' \longrightarrow {}^\rho\mathcal{GC}^\infty(\mathrm{c}(-), {}^\rho\widetilde{\mathbb{R}})$ is a sheaf-morphism of real vector spaces, i.e. if $\Omega' \subseteq \Omega$ is another open set and $T \in \mathcal{D}'(\Omega)$, then $\iota_\Omega^b(T)|_{\mathrm{c}(\Omega')} = \iota_{\Omega'}^b(T|_{\Omega'})$ and ι_Ω^b is a linear injective map.*
- *(ii)* *Any $f \in \mathcal{C}^\infty(\Omega)$ can naturally be considered an element of ${}^\rho\mathcal{GC}^\infty(\mathrm{c}(\Omega), {}^\rho\widetilde{\mathbb{R}})$ via $[x_\varepsilon] \mapsto [f(x_\varepsilon)]$. Moreover, $\forall q \in \mathbb{N}_{>0}\ \forall x \in \mathrm{c}(\Omega) :\ \left|\iota_\Omega^b(f)(x) - f(x)\right| \leq b^{-q}$.*
- *(iii)* *If $f \in \mathcal{C}^\infty(\Omega)$ and if $b \geq \mathrm{d}\rho^{-a}$ for some $a \in \mathbb{R}_{>0}$, then $\iota_\Omega^b(f) = f$. In particular, the embedding ι^b preserves multiplication of smooth functions.*
- *(iv)* *For any $T \in \mathcal{D}'(\Omega)$ and any $\alpha \in \mathbb{N}^n$, $\iota_\Omega^b(\partial^\alpha T) = \partial^\alpha \iota_\Omega^b(T)$, i.e. ι_Ω^b preserves partial derivatives of distributions.*
- *(v)* *Let $b \geq \mathrm{d}\rho^{-a}$ for some $a \in \mathbb{R}_{>0}$. Then for any $\varphi \in \mathcal{D}(\Omega)$ and any $T \in \mathcal{D}'(\Omega)$,*

$$\left[\int_\Omega \iota_\Omega^b(T)_\varepsilon(x) \cdot \varphi(x)\, \mathrm{d}x \right] = \langle T, \varphi \rangle \quad \text{in } {}^\rho\widetilde{\mathbb{R}}.$$

- *(vi)* *$\iota_{\mathbb{R}^n}^b(\delta)(0) = c_n b^n$ and if $b \geq \mathrm{d}\rho^{-a}$ for some $a \in \mathbb{R}_{>0}$, then $\iota_{\mathbb{R}}^b(H)(0) = \frac{1}{2}$.*
- *(vii)* *The embedding ι^b does not depend on the particular choice of (κ_ε) and (if $b \geq \mathrm{d}\rho^{-a}$ for some $a \in \mathbb{R}_{>0}$) χ as above.*
- *(viii)* *ι^b does not depend on the representative (b_ε) of b.*

Whenever we use the notation ι^b for an embedding, we assume that $b \in {}^\rho\widetilde{\mathbb{R}}$ satisfies the overall assumptions of Thm. 12 and of (iii) in that Theorem, and that ι^b has been defined as in (3.6) using a Colombeau mollifier μ for the given dimension.

Remark 13.

- (i) Let δ, $H \in {}^\rho\mathcal{GC}^\infty({}^\rho\widetilde{\mathbb{R}}, {}^\rho\widetilde{\mathbb{R}})$ be the corresponding ι^b-embeddings of the Dirac delta and of the Heaviside function. Then $\delta(x) = b \cdot \mu(b \cdot x)$ and $\delta(x) = 0$ if x is near-standard and $x^\circ \neq 0$ or if x is infinite because $\mu \in \mathcal{S}(\mathbb{R})$. Also, by construction of μ_ε^b, δ can be represented like in the first diagram of Fig. 3.1. E.g., $\delta(k/b) = 0$ for each $k \in \mathbb{Z} \setminus \{0\}$, and each $\frac{k}{b}$ is a nonzero infinitesimal. Similar properties can be stated e.g. for $\delta^2(x) = b^2 \cdot \mu(b \cdot x)^2$.
- (ii) Analogously, we have $H(x) = 1$ if x is near-standard and $x^\circ > 0$ or if $x > 0$ is infinite; $H(x) = 0$ if x is near-standard and $x^\circ < 0$ or if $x < 0$ is infinite.

(iii) Let $\mathrm{vp}(\frac{1}{x}) \in \mathcal{D}'(\mathbb{R})$ be the Cauchy principal value. If $x = [x_\varepsilon]$ is far from the origin, in the sense that $|x| \ge r$ for some $r \in \mathbb{R}_{>0}$. Then $\iota^b_{\mathbb{R}}(\mathrm{vp}(\frac{1}{x}))(x) = \frac{1}{x}$. The behavior of the GSF $\iota^b_{\mathbb{R}}(\mathrm{vp}(\frac{1}{x}))(-)$ in an infinitesimal neighborhood of the origin depends on the Colombeau mollifier μ. For example, if in Lem. 11 we add the linear condition $\int \frac{\mu_n(x)}{x}\,\mathrm{d}x = 0$, then also $\iota^b_{\mathbb{R}}(\mathrm{vp}(\frac{1}{x}))(0) = 0$.

3.3. **Closure with respect to composition.** In contrast to the case of distributions, there is no problem in considering the composition of two GSF. This property opens new interesting possibilities, e.g. in considering differential equations $y' = f(y,t)$, where y and f are GSF. For instance, there is no problem in studying $y' = \delta(y)$ (see Sec. 4).

Theorem 14. *Let $X \subseteq {}^\rho\widetilde{\mathbb{R}}^n$, $Y \subseteq {}^\rho\widetilde{\mathbb{R}}^d$ and $Z \subseteq {}^\rho\widetilde{\mathbb{R}}^k$. If $f \in {}^\rho\mathcal{GC}^\infty(X,Y)$ and $g \in {}^\rho\mathcal{GC}^\infty(Y,Z)$, then $g \circ f \in {}^\rho\mathcal{GC}^\infty(X,Z)$, i.e. GSF are closed with respect to composition.*

For instance, we can think of the Dirac delta as a map of the form $\delta : {}^\rho\widetilde{\mathbb{R}} \longrightarrow {}^\rho\widetilde{\mathbb{R}}$, and therefore the composition e^δ is defined in $\{x \in {}^\rho\widetilde{\mathbb{R}} \mid \exists z \in {}^\rho\widetilde{\mathbb{R}}_{>0} :\ \delta(x) \le \log z\}$, which of course does not contain $x = 0$ but only suitable non zero infinitesimals. On the other hand, $\delta \circ \delta : {}^\rho\widetilde{\mathbb{R}} \longrightarrow {}^\rho\widetilde{\mathbb{R}}$. Moreover, from the inclusion of ordinary smooth functions (Thm. 12) and the closure with respect to composition, it directly follows that every ${}^\rho\mathcal{GC}^\infty(U, {}^\rho\widetilde{\mathbb{R}})$ is an algebra with pointwise operations for every subset $U \subseteq {}^\rho\widetilde{\mathbb{R}}^n$.

Example 15. The composition $\delta \circ \delta \in {}^\rho\mathcal{GC}^\infty({}^\rho\widetilde{\mathbb{R}}, {}^\rho\widetilde{\mathbb{R}})$ is given by $(\delta \circ \delta)(x) = b\mu\left(b^2\mu(bx)\right)$ and is an even function. If x is near-standard and $x^\circ \ne 0$, or x is infinite, then $(\delta \circ \delta)(x) = b$. Since $(\delta \circ \delta)(0) = 0$, by the intermediate value theorem (see Cor. 26 below), we have that $\delta \circ \delta$ attains any value in the interval $[0,b] \subseteq {}^\rho\widetilde{\mathbb{R}}$. If $0 \le x \le \frac{1}{2b}$, then (for a μ as in Fig. 3.2) x is infinitesimal and $(\delta \circ \delta)(x) = 0$ because $\delta(x) \ge b\mu\left(\frac{1}{k}\right)$ is an infinite number. If $x = \frac{k}{b}$ for some $k \in \mathbb{N}_{>0}$, then x is still infinitesimal but $(\delta \circ \delta)(x) = b$ because $\mu(bx) = 0$. A representation of $\delta \circ \delta$ is given in Fig. 3.2. Analogously, one can deal with $H \circ \delta$ and $\delta \circ H$.

Similarly, we can define generalized functions of class ${}^\rho\mathcal{GC}^k$, with $k \le +\infty$:

Definition 16. Let $X \subseteq {}^\rho\widetilde{\mathbb{R}}^n$ and $Y \subseteq {}^\rho\widetilde{\mathbb{R}}^d$ be arbitrary subsets of generalized points and $k \in \mathbb{N} \cup \{+\infty\}$. Then we say that

$$f : X \longrightarrow Y \text{ is a } \textit{generalized } \mathcal{C}^k \textit{ function}$$

if there exists a net $f_\varepsilon \in \mathcal{C}^k(\Omega_\varepsilon, \mathbb{R}^d)$ defining f in the sense that

(i) $X \subseteq \langle \Omega_\varepsilon \rangle$,

(ii) $f([x_\varepsilon]) = [f_\varepsilon(x_\varepsilon)] \in Y$ for all $x = [x_\varepsilon] \in X$,

(iii) $(\partial^\alpha f_\varepsilon(x_\varepsilon)) \in \mathbb{R}^d_\rho$ for all $x = [x_\varepsilon] \in X$ and all $\alpha \in \mathbb{N}^n$ such that $|\alpha| \le k$.

(iv) $\forall \alpha \in \mathbb{N}^n\, \forall [x_\varepsilon], [x'_\varepsilon] \in X :\ |\alpha| = k,\ [x_\varepsilon] = [x'_\varepsilon] \ \Rightarrow\ [\partial^\alpha f_\varepsilon(x_\varepsilon)] = [\partial^\alpha f_\varepsilon(x'_\varepsilon)]$.

(v) For all $\alpha \in \mathbb{N}^n$, with $|\alpha| = k$, the map $[x_\varepsilon] \in X \mapsto [\partial^\alpha f_\varepsilon(x_\varepsilon)] \in {}^\rho\widetilde{\mathbb{R}}^d$ is continuous in the sharp topology.

The space of generalized $\mathcal{C}^k$ functions from X to Y is denoted by ${}^\rho\mathcal{GC}^k(X,Y)$.

Note that properties (iv), (v) are required only for $|\alpha| = k$ because for lower length they can be proved using property (iii) and the classical mean value theorem for

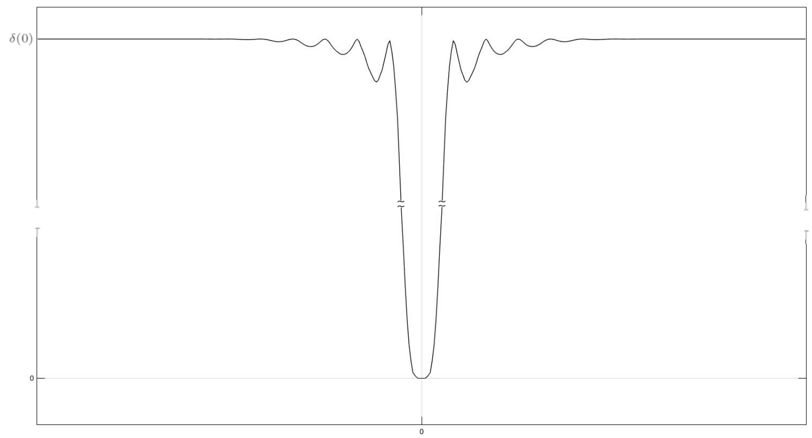

FIGURE 3.2. A representation of $\delta\circ\delta$

f_ε (see e.g. [32]). From Thm. 9 and Thm. 10.(ii) it follows that this definition of ${}^\rho\mathcal{GC}^k$ is equivalent to Def. 7 if $k=+\infty$. Moreover, properties similar to (iii) and Thm. 14 can also be proved for ${}^\rho\mathcal{GC}^k$.

Note that the absolute value function $|-|:{}^\rho\widetilde{\mathbb{R}}\longrightarrow{}^\rho\widetilde{\mathbb{R}}$ is not a GSF because its derivative is not sharply continuous at the origin; clearly, it is a ${}^\rho\mathcal{GC}^0$ function.

3.4. **Differential calculus of GSF.** In this section we show how the derivatives of a GSF can be calculated using a form of incremental ratio. The idea is to prove the Fermat-Reyes theorem for GSF (see [32, 30, 43]). Essentially, this theorem shows the existence and uniqueness of another GSF serving as incremental ratio. This is the first of a long list of results demonstrating the close similarities between ordinary smooth functions and GSF.
In the present setting, the Fermat-Reyes theorem (also called Carathéodory definition of derivative) is the following.

Theorem 17. *Let $U\subseteq{}^\rho\widetilde{\mathbb{R}}^n$ be a sharply open set, let $v=[v_\varepsilon]\in{}^\rho\widetilde{\mathbb{R}}^n$, $k\in\mathbb{N}\cup\{+\infty\}$, and let $f\in{}^\rho\mathcal{GC}^{k+1}(U,{}^\rho\widetilde{\mathbb{R}})$ be a ${}^\rho\mathcal{GC}^{k+1}$map generated by the net of functions $f_\varepsilon\in\mathcal{C}^{k+1}(\Omega_\varepsilon,\mathbb{R})$. Then*

(i) *There exists a sharp neighborhood T of $U\times\{0\}$ and a map $r\in{}^\rho\mathcal{GC}^k(T,{}^\rho\widetilde{\mathbb{R}})$, called the* generalized incremental ratio *of f along v, such that*
$$\forall(x,h)\in T:\ f(x+hv)=f(x)+h\cdot r(x,h).$$
(ii) *Any two generalized incremental ratios coincide on a sharp neighborhood of $U\times\{0\}$, so that we can use the notation $\frac{\partial f}{\partial v}[x;h]:=r(x,h)$ if (x,h) are sufficiently small.*
(iii) *We have $\frac{\partial f}{\partial v}[x;0]=\left[\frac{\partial f_\varepsilon}{\partial v_\varepsilon}(x_\varepsilon)\right]$ for every $x\in U$ and we can thus define $df(x)\cdot v:=\frac{\partial f}{\partial v}(x):=\frac{\partial f}{\partial v}[x;0]$, so that $\frac{\partial f}{\partial v}\in{}^\rho\mathcal{GC}^k(U,{}^\rho\widetilde{\mathbb{R}})$.*

Note that this result allows us to consider the partial derivative of f with respect to an arbitrary generalized vector $v\in{}^\rho\widetilde{\mathbb{R}}^n$ which can be, e.g., near-standard or infinite. Since any partial derivative of a GSF is still a GSF, higher order derivatives $\frac{\partial^\alpha f}{\partial v^\alpha} \in {}^\rho\mathcal{GC}^\infty(U, {}^\rho\widetilde{\mathbb{R}})$ are simply defined recursively.

As follows from Thm. 17.(i) and Thm. 12.(iv), the concept of derivative defined using the Fermat-Reyes theorem is compatible with the classical derivative of Schwartz distributions via the embeddings ι^b from Thm. 12. The following result follows from the analogous properties for the nets of smooth functions defining f and g or directly from the Fermat-Reyes Thm. 17.

Theorem 18. *Let $U \subseteq {}^\rho\widetilde{\mathbb{R}}^n$ be an open subset in the sharp topology, let $v \in {}^\rho\widetilde{\mathbb{R}}^n$ and $f,\ g : U \longrightarrow {}^\rho\widetilde{\mathbb{R}}$ be generalized smooth maps. Then*

(i) $\frac{\partial(f+g)}{\partial v} = \frac{\partial f}{\partial v} + \frac{\partial g}{\partial v}$

(ii) $\frac{\partial(r\cdot f)}{\partial v} = r \cdot \frac{\partial f}{\partial v} \quad \forall r \in {}^\rho\widetilde{\mathbb{R}}$

(iii) $\frac{\partial(f\cdot g)}{\partial v} = \frac{\partial f}{\partial v} \cdot g + f \cdot \frac{\partial g}{\partial v}$

(iv) *For each $x \in U$, the map $\mathrm{d}f(x).v := \frac{\partial f}{\partial v}(x) \in {}^\rho\widetilde{\mathbb{R}}$ is ${}^\rho\widetilde{\mathbb{R}}$-linear in $v \in {}^\rho\widetilde{\mathbb{R}}^n$.*

(v) *Let $V \subseteq {}^\rho\widetilde{\mathbb{R}}^d$ be open subsets in the sharp topology and $h \in {}^\rho\mathcal{GC}^\infty(V, U)$ be a generalized smooth maps. Then for all $x \in V$ and all $v \in {}^\rho\widetilde{\mathbb{R}}^d$*

$$\frac{\partial\,(f \circ h)}{\partial v}(x) = \mathrm{d}f\,(h(x))\,.\frac{\partial h}{\partial v}(x)$$
$$\mathrm{d}(f \circ h)\,(x) = \mathrm{d}f\,(h(x)) \circ \mathrm{d}h(x).$$

3.5. **Integral calculus using primitives.** In this section, we inquire existence and uniqueness of primitives F of a GSF $f \in {}^\rho\mathcal{GC}^\infty([a, b], {}^\rho\widetilde{\mathbb{R}})$. To this end, we shall have to introduce the derivative $F'(x)$ at boundary points $x \in [a, b]$, i.e. such that $x - a$ or $b - x$ is not invertible. Let us note explicitly, in fact, that the Fermat-Reyes Theorem 17 is stated only for sharply open domains.

The following result shows that every GSF can have at most one primitive GSF up to an additive constant.

Theorem 19. *Let $X \subseteq {}^\rho\widetilde{\mathbb{R}}$ and let $f \in {}^\rho\mathcal{GC}^\infty(X, {}^\rho\widetilde{\mathbb{R}})$ be a generalized smooth function. Let $a,\ b \in {}^\rho\widetilde{\mathbb{R}}$, with $a < b$, such that $(a, b) \subseteq X$. If $f'(x) = 0$ for all $x \in \mathrm{int}(a, b)$, then f is constant on (a, b). An analogous statement holds if we take any other type of interval (closed or half closed) instead of (a, b).*

Remark 20. From the Fermat-Reyes Thm. 17 and from Thm. 19, it follows that the function $i(x) := 1$ if $x \approx 0$ and $i(x) := 0$ otherwise cannot be a GSF on any large neighborhood of $x = 0$. This example stems from the property that different standard real numbers can always be separated by infinitesimal balls.

At interior points $x \in [a, b]$ in the sharp topology, the definition of derivative $f^{(k)}(x)$ follows from the Fermat-Reyes Theorem 17. At boundary points, we have the following

Theorem 21. *Let $a,\ b \in {}^\rho\widetilde{\mathbb{R}}$ with $a < b$, and $f \in {}^\rho\mathcal{GC}^\infty([a, b], {}^\rho\widetilde{\mathbb{R}})$ be a generalized smooth function. Then for all $x \in [a, b]$, the following limit exists in the sharp topology*

$$\lim_{\substack{y \to x \\ y \in \mathrm{int}([a,b])}} f^{(k)}(y) =: f^{(k)}(x).$$

Moreover, if the net $f_\varepsilon \in \mathcal{C}^\infty(\Omega_\varepsilon, \mathbb{R})$ defines f and $x = [x_\varepsilon]$, then $f^{(k)}(x) = [f_\varepsilon^{(k)}(x_\varepsilon)]$ and hence $f^{(k)} \in {}^\rho\mathcal{GC}^\infty([a, b], {}^\rho\widetilde{\mathbb{R}})$.

We can now state existence and uniqueness of primitives of GSF:

Theorem 22. *Let $k \in \mathbb{N} \cup \{+\infty\}$ and $f \in {}^\rho\mathcal{GC}^k([a,b], {}^\rho\widetilde{\mathbb{R}})$ be defined in the interval $[a,b] \subseteq {}^\rho\widetilde{\mathbb{R}}$, where $a < b$. Let $c \in [a,b]$. Then, there exists one and only one generalized $\mathcal{C}^{k+1}$ map $F \in {}^\rho\mathcal{GC}^{k+1}([a,b], {}^\rho\widetilde{\mathbb{R}})$ such that $F(c) = 0$ and $F'(x) = f(x)$ for all $x \in [a,b]$. Moreover, if f is defined by the net $f_\varepsilon \in \mathcal{C}^k(\mathbb{R}, \mathbb{R})$ and $c = [c_\varepsilon]$, then $F(x) = \left[\int_{c_\varepsilon}^{x_\varepsilon} f_\varepsilon(s)\mathrm{d}s\right]$ for all $x = [x_\varepsilon] \in [a,b]$.*

Definition 23. Under the assumptions of Theorem 22, we denote by $\int_c^{(-)} f := \int_c^{(-)} f(s)\,\mathrm{d}s \in {}^\rho\mathcal{GC}^\infty([a,b], {}^\rho\widetilde{\mathbb{R}})$ the unique generalized smooth function such that:

(i) $\int_c^c f = 0$

(ii) $\left(\int_c^{(-)} f\right)'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_c^x f(s)\,\mathrm{d}s = f(x)$ for all $x \in [a,b]$.

To consider a generalization of this concept of integration to GSF in several variables and to more general domains of integration $M \subseteq {}^\rho\widetilde{\mathbb{R}}^d$, see [32].

Example 24.

(i) Since ${}^\rho\widetilde{\mathbb{R}}$ contains both infinitesimal and infinite numbers, our notion of definite integral also includes "improper integrals". Let e.g. $f(x) = \frac{1}{x}$ for $x \in {}^\rho\widetilde{\mathbb{R}}_{>0}$ and $a = 1$, $b = \mathrm{d}\rho^{-q}$, $q > 0$. Then

$$\int_a^b f(s)\,\mathrm{d}s = \left[\int_1^{\rho_\varepsilon^{-q}} \frac{1}{s}\,\mathrm{d}s\right] = [\log \rho_\varepsilon^{-q}] - \log 1 = -q \log \mathrm{d}\rho, \tag{3.7}$$

which is, of course, a positive infinite generalized number. This apparently trivial result is closely tied to the possibility to define GSF on arbitrary domains, like $F \in {}^\rho\mathcal{GC}^\infty([a,b], {}^\rho\widetilde{\mathbb{R}})$ in Thm. 22 where b is an infinite number as in (3.7), which is one of the key properties allowing one to get the closure with respect to composition.

(ii) If p, $q \in {}^\rho\widetilde{\mathbb{R}}$, $p < 0 < q$ and both p and q are not infinitesimal, then $\int_p^q \delta(t)\,\mathrm{d}t \approx 1$. If $p \le -r$ and $q \ge s$ where r, $s \in \mathbb{R}_{>0}$, then $\int_p^q \delta(t)\,\mathrm{d}t = 1$.

Theorem 25. *Let $f \in {}^\rho\mathcal{GC}^\infty(X, {}^\rho\widetilde{\mathbb{R}})$ and $g \in {}^\rho\mathcal{GC}^\infty(Y, {}^\rho\widetilde{\mathbb{R}})$ be generalized smooth functions defined on arbitrary domains in ${}^\rho\widetilde{\mathbb{R}}$. Let a, $b \in {}^\rho\widetilde{\mathbb{R}}$ with $a < b$ and $[a,b] \subseteq X \cap Y$, then*

(i) $\int_a^b (f+g) = \int_a^b f + \int_a^b g$

(ii) $\int_a^b \lambda f = \lambda \int_a^b f \quad \forall \lambda \in {}^\rho\widetilde{\mathbb{R}}$

(iii) $\int_a^b f = \int_a^c f + \int_c^b f$ *for all* $c \in [a,b]$

(iv) $\int_a^b f = -\int_b^a f$

(v) $\int_a^b f' = f(b) - f(a)$

(vi) $\int_a^b f' \cdot g = [f \cdot g]_a^b - \int_a^b f \cdot g'$

(vii) If $f(x) \le g(x)$ for all $x \in [a,b]$, then $\int_a^b f \le \int_a^b g$.

Let $f \in {}^\rho\mathcal{GC}^\infty(T, {}^\rho\widetilde{\mathbb{R}})$ and $\varphi \in {}^\rho\mathcal{GC}^\infty(S, T)$ be generalized smooth functions defined on arbitrary domains in ${}^\rho\widetilde{\mathbb{R}}$. Let a, $b \in {}^\rho\widetilde{\mathbb{R}}$, with $a < b$, such that $[a,b] \subseteq S$, $\varphi(a) < \varphi(b)$ and $[\varphi(a), \varphi(b)] \subseteq T$. Finally, assume that $\varphi([a,b]) \subseteq [\varphi(a), \varphi(b)]$.

Then

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,\mathrm{d}t = \int_a^b f\left[\varphi(s)\right]\cdot\varphi'(s)\,\mathrm{d}s.$$

For integration of several variables GSF, see [32].

3.6. **Classical theorems for GSF.** It is natural to expect that several classical theorems of differential and integral calculus can be extended from the ordinary smooth case to the generalized smooth framework. Once again, we underscore that these faithful generalizations are possible because we do not have a priori limitations in the evaluation $f(x)$ for GSF.

We start from the intermediate value theorem.

Corollary 26. *Let $f\in{}^{\rho}\mathcal{GC}^{\infty}(X,{}^{\rho}\widetilde{\mathbb{R}})$ be a generalized smooth function defined on the subset $X\subseteq{}^{\rho}\widetilde{\mathbb{R}}$. Let a, $b\in{}^{\rho}\widetilde{\mathbb{R}}$, with $a<b$, such that $[a,b]\subseteq X$. Assume that $f(a)<f(b)$. Then*

$$\forall y\in{}^{\rho}\widetilde{\mathbb{R}}:\ f(a)\le y\le f(b)\ \Rightarrow\ \exists c\in[a,b]:\ y=f(c).$$

Using this theorem we can conclude that no GSF can assume only a finite number of values which are comparable with respect to the relation $<$ on any nontrivial interval $[a,b]\subseteq X$, unless it is constant. For example, this provides an alternative way of seeing that the function i of Rem. 20 cannot be a generalized smooth map. We note that the solution $c\in[a,b]$ of the previous generalized smooth equation $y=f(x)$ need not even be continuous in ε, see e.g. [32] for an explicit counter example. This allows us to draw the following general conclusion: if we consider generalized numbers as solutions of smooth equations, then we are forced to work on a non-totally ordered ring of scalars derived from discontinuous (in ε) representatives. To put it differently: if we choose a ring of scalars with a total order or continuous representatives, we will not be able to solve every smooth equation, and the given ring can be considered, in some sense, incomplete.

The next theorem deals with different version of the mean value theorem

Theorem 27. *Let $f\in{}^{\rho}\mathcal{GC}^{\infty}(X,{}^{\rho}\widetilde{\mathbb{R}}^{d})$ be a generalized smooth function defined in the sharply open set $X\subseteq{}^{\rho}\widetilde{\mathbb{R}}^{n}$. Let a, $b\in{}^{\rho}\widetilde{\mathbb{R}}^{n}$ such that $[a,b]\subseteq X$. Then*

- *(i)* *If $n=d=1$, then $\exists c\in[a,b]:\ f(b)-f(a)=(b-a)\cdot f'(c)$.*
- *(ii)* *If $n=d=1$, then $\exists c\in[a,b]:\ \int_a^b f(t)\,\mathrm{d}t=(b-a)\cdot f(c)$.*
- *(iii)* *If $d=1$, then $\exists c\in[a,b]:\ f(b)-f(a)=\nabla f(c)\cdot(b-a)$.*
- *(iv)* *Let $h:=b-a$, then $f(a+h)-f(a)=\int_0^1\mathrm{d}f(a+t\cdot h).h\,\mathrm{d}t$.*

Internal and bounded sets generated by a net of compact sets serve as a substitute for compact subsets for GSF, as can be seen from the following extreme value theorem:

Lemma 28. *Let $\emptyset\ne K=[K_\varepsilon]\subseteq{}^{\rho}\widetilde{\mathbb{R}}^{n}$ be an internal set generated by compact sets $K_\varepsilon\Subset\mathbb{R}^n$ such that K is* bounded, *i.e. $K\subseteq B_R(0)$ for some $R\in{}^{\rho}\widetilde{\mathbb{R}}_{>0}$. Assume that $\alpha:K\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}$ is a well-defined map given by $\alpha(x)=[\alpha_\varepsilon(x_\varepsilon)]$ for all $x\in K$, where $\alpha_\varepsilon:K_\varepsilon\longrightarrow\mathbb{R}$ are continuous maps (e.g. $\alpha(x)=|x|$). Then*

$$\exists m,M\in K\,\forall x\in K:\ \alpha(m)\le\alpha(x)\le\alpha(M).$$

Corollary 29. *Let $f \in {}^{\rho}\mathcal{GC}^{\infty}(X, {}^{\rho}\widetilde{\mathbb{R}})$ be a generalized smooth function defined in the subset $X \subseteq {}^{\rho}\widetilde{\mathbb{R}}^n$. Let $\emptyset \neq K = [K_\varepsilon] \subseteq X$ be as above, then*

$$\exists m, M \in K\, \forall x \in K: \ f(m) \le f(x) \le f(M). \tag{3.8}$$

These results motivate the following

Definition 30. A subset K of ${}^{\rho}\widetilde{\mathbb{R}}^n$ is called *functionally compact*, denoted by $K \Subset_{\mathrm{f}} {}^{\rho}\widetilde{\mathbb{R}}^n$, if there exists a net (K_ε) such that

(i) $K = [K_\varepsilon] \subseteq {}^{\rho}\widetilde{\mathbb{R}}^n$

(ii) K is bounded, i.e. $\exists R \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}: \ K \subseteq B_R(0)$

(iii) $\forall \varepsilon \in I: \ K_\varepsilon \Subset \mathbb{R}^n$

If, in addition, $K \subseteq U \subseteq {}^{\rho}\widetilde{\mathbb{R}}^n$ then we write $K \Subset_{\mathrm{f}} U$. Any net (K_ε) such that $[K_\varepsilon] = K$ is called a *representative* of K.

We motivate the name *functionally compact subset* by noting that on this type of subsets, GSF have properties very similar to those that ordinary smooth functions have on standard compact sets.

Remark 31.

(i) By Thm. 5.(iii), any internal set $K = [K_\varepsilon]$ is closed in the sharp topology. Therefore, functionally compact sets are sharply closed and bounded subsets of ${}^{\rho}\widetilde{\mathbb{R}}^n$. In particular, the open interval $(0,1) \subseteq \widetilde{\mathbb{R}}$ is not functionally compact since it is not closed.

(ii) If $H \Subset \mathbb{R}^n$ is a non-empty ordinary compact set, then the internal set $[H]$ is functionally compact. In particular, $[0,1] = [[0,1]_{\mathbb{R}}]$ is functionally compact.

(iii) The empty set $\emptyset = \widetilde{\emptyset} \Subset_{\mathrm{f}} \widetilde{\mathbb{R}}$.

(iv) $\widetilde{\mathbb{R}}^n$ is not functionally compact since it is not bounded.

(v) The set of finite points $\mathrm{c}(\mathbb{R})$ is not functionally compact because the GSF $f(x) = x$ does not satisfy the conclusion (3.8) of Cor. 29.

We also underscore the following properties of functionally compact sets.

Theorem 32. *Let $K \subseteq X \subseteq \widetilde{\mathbb{R}}^n$, $f \in {}^{\rho}\mathcal{GC}^k(X, \widetilde{\mathbb{R}}^d)$. Then $K \Subset_f \widetilde{\mathbb{R}}^n$ implies $f(K) \Subset_f \widetilde{\mathbb{R}}^d$.*

As a corollary of this theorem and Rem. (31).(ii) we get

Corollary 33. *If a, $b \in \widetilde{\mathbb{R}}$ and $a \le b$, then $[a,b] \Subset_f \widetilde{\mathbb{R}}$.*

Let us note that a, $b \in \widetilde{\mathbb{R}}$ can also be infinite numbers, e.g. $a = \mathrm{d}\rho^{-N}$, $b = \mathrm{d}\rho^{-M}$ or $a = -\mathrm{d}\rho^{-N}$, $b = \mathrm{d}\rho^{-M}$ with M, $N \in \mathbb{N}_{>0}$, so that e.g. $[-\mathrm{d}\rho^{-N}, \mathrm{d}\rho^{M}] \supseteq \mathbb{R}$. Therefore, despite very similar properties shared by functionally compact sets and classical compact sets, the former can also be unbounded from the classical point of view.
Finally, in the following result we consider the product of functionally compact sets:

Theorem 34. *Let $K \Subset_f \widetilde{\mathbb{R}}^n$ and $H \Subset_f \widetilde{\mathbb{R}}^d$, then $K \times H \Subset_f \widetilde{\mathbb{R}}^{n+d}$. In particular, if $a_i \le b_i$ for $i = 1, \ldots, n$, then $\prod_{i=1}^n [a_i, b_i] \Subset_f \widetilde{\mathbb{R}}^n$.*

A theory of compactly supported GSF has been developed in [29], and it closely resembles the classical theory of LF-spaces of compactly supported smooth functions. It establishes that for suitable functionally compact subsets, the corresponding space of compactly supported GSF contains all Schwartz distributions.

Note also that any interval $[a,b]\subseteq {}^{\rho}\widetilde{\mathbb{R}}$ with $b-a\in\mathbb{R}_{>0}$, is functionally compact but *not* connected: in fact if $c\in(a,b)$, then both $c+D_{\infty}$ and $[a,b]\setminus(c+D_{\infty})$ are sharply open in $[a,b]$. Once again, this is a general property in several non-Archimedean frameworks (see e.g. [67, 43]). On the other hand, as in the case of functionally compact sets, GSF behave on intervals as if they were connected, in the sense that both the intermediate value theorem Cor. 26 and the extreme value theorem Cor. 29 hold for them (therefore, $f([a,b])=[f(m),f(M)]$, where we used the notations from the results just mentioned).

We close this section with generalizations of Taylor's theorem in various forms. In the following statement, $\mathrm{d}^k f(x):{}^{\rho}\widetilde{\mathbb{R}}^{dk}\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}$ is the k-th differential of the GSF f, viewed as an ${}^{\rho}\widetilde{\mathbb{R}}$-multilinear map ${}^{\rho}\widetilde{\mathbb{R}}^d\times\overset{k}{\ldots\ldots}\times{}^{\rho}\widetilde{\mathbb{R}}^d\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}$, and we use the common notation $\mathrm{d}^k f(x)\cdot h^k:=\mathrm{d}^k f(x)(h,\ldots,h)$. Clearly, $\mathrm{d}^k f(x)\in{}^{\rho}\mathcal{GC}^{\infty}({}^{\rho}\widetilde{\mathbb{R}}^{dk},{}^{\rho}\widetilde{\mathbb{R}})$. For multilinear maps $A:{}^{\rho}\widetilde{\mathbb{R}}^p\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}^q$, we set $|A|:=[|A_{\varepsilon}|]\in{}^{\rho}\widetilde{\mathbb{R}}$, the generalized number defined by the norms of the operators $A_{\varepsilon}:\mathbb{R}^p\longrightarrow\mathbb{R}^q$.

Theorem 35. *Let $f\in{}^{\rho}\mathcal{GC}^{\infty}(U,{}^{\rho}\widetilde{\mathbb{R}})$ be a generalized smooth function defined in the sharply open set $U\subseteq{}^{\rho}\widetilde{\mathbb{R}}^d$. Let a, $b\in{}^{\rho}\widetilde{\mathbb{R}}^d$ such that the line segment $[a,b]\subseteq U$, and set $h:=b-a$. Then, for all $n\in\mathbb{N}$ we have*

(i) $\exists\xi\in[a,b]:\ f(a+h)=\sum_{j=0}^{n}\frac{\mathrm{d}^j f(a)}{j!}\cdot h^j+\frac{\mathrm{d}^{n+1}f(\xi)}{(n+1)!}\cdot h^{n+1}$.

(ii) $f(a+h)=\sum_{j=0}^{n}\frac{\mathrm{d}^j f(a)}{j!}\cdot h^j+\frac{1}{n!}\cdot\int_0^1(1-t)^n\,\mathrm{d}^{n+1}f(a+th)\cdot h^{n+1}\,\mathrm{d}t$.

Moreover, there exists some $R\in{}^{\rho}\widetilde{\mathbb{R}}_{>0}$ such that

$$\forall k\in B_R(0)\,\exists\xi\in[a,a+k]:\ f(a+k)=\sum_{j=0}^{n}\frac{\mathrm{d}^j f(a)}{j!}\cdot k^j+\frac{\mathrm{d}^{n+1}f(\xi)}{(n+1)!}\cdot k^{n+1}\tag{3.9}$$

$$\frac{\mathrm{d}^{n+1}f(\xi)}{(n+1)!}\cdot k^{n+1}=\frac{1}{n!}\cdot\int_0^1(1-t)^n\,\mathrm{d}^{n+1}f(a+tk)\cdot k^{n+1}\,\mathrm{d}t\approx 0.\tag{3.10}$$

Formulas (i) and (ii) correspond to a plain generalization of Taylor's theorem for ordinary smooth functions with Lagrange and integral remainder, respectively. Dealing with GF, it is important to note that this direct statement also includes the possibility that the differential $\mathrm{d}^{n+1}f(\xi)$ may be infinite at some point. For this reason, in (3.9) and (3.10), considering a sufficiently small increment k, we get more classical infinitesimal remainders $\mathrm{d}^{n+1}f(\xi)\cdot k^{n+1}\approx 0$.

The following definitions allow us to state Taylor formulas in Peano and in infinitesimal form. The latter has no remainder term thanks to the use of an equivalence relation that permits the introduction of a language of nilpotent infinitesimals, see e.g. [26, 27] for a similar formulation. For simplicity, we only present the 1-dimentional case.

Definition 36. (i) Let $U\subseteq{}^{\rho}\widetilde{\mathbb{R}}$ be a sharp neighborhood of 0 and P, $Q:U\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}$ be maps defined on U. Then we say that

$$P(u)=o(Q(u))\quad\text{as }u\to 0$$

if there exists a function $R:U\longrightarrow{}^{\rho}\widetilde{\mathbb{R}}$ such that

$$\forall u\in U:\ P(u)=R(u)\cdot Q(u)\quad\text{and}\quad\lim_{u\to 0}R(u)=0,$$

where the limit is taken in the sharp topology.

(ii) Let $x, y \in {}^{\rho}\widetilde{\mathbb{R}}$ and $k, j \in \mathbb{R}_{>0}$, then we write $x =_j y$ if there exist representatives (x_ε), (y_ε) of x, y, respectively, such that

$$|x_\varepsilon - y_\varepsilon| = O(\rho_\varepsilon^{\frac{1}{j}}). \tag{3.11}$$

We will read $x =_j y$ as *x is equal to y up to j-th order infinitesimals*. Finally, if $k \in \mathbb{N}_{>0}$, we set $D_{kj} := \left\{x \in {}^{\rho}\widetilde{\mathbb{R}} \mid x^{k+1} =_j 0\right\}$, which is called the *set of k-th order infinitesimals for the equality $=_j$*, and

$$D_{\infty j} := \left\{x \in {}^{\rho}\widetilde{\mathbb{R}} \mid \exists k \in \mathbb{N}_{>0} : \ x^{k+1} =_j 0\right\}$$

which is called the *set of infinitesimals for the equality $=_j$*.

Of course, the reformulation of Def. 36 (i) for the classical Landau's little-oh is particularly suited to the case of a ring like ${}^{\rho}\widetilde{\mathbb{R}}$, instead of a field. The intuitive interpretation of $x =_j y$ is that for particular (e.g. physics-related) problems one is not interested in distinguishing quantities whose difference $|x - y|$ is less than an infinitesimal of order j. In fact, if $x =_j y$ we can write $x_\varepsilon = y_\varepsilon + r_\varepsilon$ with $r_\varepsilon \to 0$ of order at most $\rho_\varepsilon^{\frac{1}{j}}$. The idea behind taking $\frac{1}{j}$ in (3.11) is to obtain the property that the greater the order j of the infinitesimal error, the greater the difference $|x - y|$ is allowed to be. This is a typical property in rings with nilpotent infinitesimals (see e.g. [26, 43]). The set D_{ki} represents the neighborhood of infinitesimals of k-th order for the equality $=_j$. Once again, the greater the order k, the bigger is the neighborhood (see Theorem 37.(ix) below). Note that if $x =_j y$, then $x_\varepsilon = y_\varepsilon + o\left(\rho_\varepsilon^{\frac{1}{j}-a}\right)$ for all $a \in (0, 1/j]_{\mathbb{R}}$. In particular, $x_\varepsilon = y_\varepsilon + o\left(\rho_\varepsilon\right)$ implies $x =_1 y$, whereas $x =_1 y$ yields only $x_\varepsilon = y_\varepsilon + o\left(\rho_\varepsilon^{1-a}\right)$ for all $a \in (0, 1]_{\mathbb{R}}$. Finally, note that $x =_j y$ is equivalent to $|x - y| \le C \mathrm{d}\rho^{\frac{1}{j}}$ for some $C \in \mathbb{R}_{\ge 0}$.

Theorem 37. *Let $f \in {}^{\rho}\mathcal{GC}^\infty(U, {}^{\rho}\widetilde{\mathbb{R}})$ be a generalized smooth function defined in the sharply open set $U \subseteq {}^{\rho}\widetilde{\mathbb{R}}$. Let $x, \delta \in {}^{\rho}\widetilde{\mathbb{R}}$, with $\delta > 0$ and $[x - \delta, x + \delta] \subseteq U$. Let k, l, $j \in \mathbb{R}_{>0}$. Then*

(i) $\forall n \in \mathbb{N} : \ f(x + u) = \sum_{r=0}^{n} \frac{f^{(r)}(x)}{r!} u^r + o(u^n)$ *as* $u \to 0$.
(ii) *The definition of $x =_j y$ does not depend on the representatives of x, y.*
(iii) *$=_j$ is an equivalence relation on ${}^{\rho}\widetilde{\mathbb{R}}$.*
(iv) *If $x =_j y$ and $l \ge j$, then $x =_l y$. Therefore, $D_{nj} \subseteq D_{nl}$.*
(v) *If $x =_j y$ for all $j \in \mathbb{R}_{>0}$ sufficiently small, then $x = y$.*
(vi) *If $x =_j y$ and $z =_j w$ then $x + z =_j y + w$. If x and z are finite, then $x \cdot z =_j y \cdot w$.*
(vii) *If $x =_j y$, $f \in {}^{\rho}\mathcal{GC}^\infty([a, b], {}^{\rho}\widetilde{\mathbb{R}})$, $x, y \in [a, b]$, and $f'(c)$ is finite for all $c \in [a, b]$, then $f(x) =_j f(y)$.*
(viii) $\forall h \in D_{kj} : \ h \approx 0$.
(ix) *$D_{mj} \subseteq D_{kj} \subseteq D_{\infty j}$ if $m \le k$.*
(x) *D_{kj} is a subring of ${}^{\rho}\widetilde{\mathbb{R}}$. For all $h \in D_{kj}$ and all finite $x \in {}^{\rho}\widetilde{\mathbb{R}}$, we have $x \cdot h \in D_{kj}$.*
(xi) *Let $n \in \mathbb{N}_{>0}$ and assume that j, k and f satisfy*

$$\forall z \in {}^{\rho}\widetilde{\mathbb{R}} \, \forall \xi \in [x - \delta, x + \delta] : \ z =_j 0 \ \Rightarrow \ z \cdot f^{(n+1)}(\xi) =_k 0. \tag{3.12}$$

Then, we have

$$\forall u \in D_{nj} : \ f(x+u) =_k \sum_{r=0}^{n} \frac{f^{(r)}(x)}{r!} u^r.$$

(xii) *For all* $n \in \mathbb{N}_{>0}$ *there exist* $e \in \mathbb{R}_{>0}$ *such that* $e \le j$, *and* $\forall u \in D_{ne}$: $f(x+u) =_j \sum_{r=0}^{n} \frac{f^{(r)}(x)}{r!} u^r$.

We shall use the nilpotent Taylor formula (xii) in Sec. 5 for the deduction of the heat and wave equation for GSF; we therefore note here that the index e depends on the GSF f: in that case, we say that *the nilpotent Taylor formula of order* n *holds for* f *on* D_{ne}. From (iv) it hence follows that it also holds on $D_{ne'}$ for all $e' \le e$.

4. Differential equations: the Picard-Lindelöf theorem for ODE

As in the classical case, thanks to the extreme value Lem. 28 and the properties of functionally compact sets K, we can naturally define a topology on the space ${}^{\rho}\mathcal{GC}^k(K, {}^{\rho}\widetilde{\mathbb{R}}^d)$:

Definition 38. Let $K \Subset_{\mathrm{f}} {}^{\rho}\widetilde{\mathbb{R}}^n$ be a functionally compact set such that $K = \overline{\mathring{K}}$ (so that partial derivatives at boundary points can be defined as limits of partial derivatives at interior points; such K are called *solid* sets). Let $l \in \mathbb{N}_{\le k}$ and $v \in {}^{\rho}\mathcal{GC}^k(K, {}^{\rho}\widetilde{\mathbb{R}}^d)$. Then

$$\|v\|_l := \max_{\substack{|\alpha| \le l \\ 1 \le i \le d}} \max\left(\left|\partial^\alpha v^i(M_{ni})\right|, \left|\partial^\alpha v^i(m_{ni})\right| \right) \in {}^{\rho}\widetilde{\mathbb{R}},$$

where M_{ni}, $m_{ni} \in K$ satisfy

$$\forall x \in K : \ \partial^\alpha v^i(m_{ni}) \le \partial^\alpha v^i(x) \le \partial^\alpha v^i(M_{ni}).$$

The following result permits us to calculate the (generalized) norm $\|v\|_l$ using any net (v_ε) that defines v.

Lemma 39. *Under the assumptions of Def. 38, let* $[K_\varepsilon] = K \Subset_f {}^{\rho}\widetilde{\mathbb{R}}^n$ *be any representative of* K. *Then we have:*

(i) *If the net* (v_ε) *defines* v, *then* $\|v\|_l = \left[\max_{\substack{|\alpha| \le l \\ 1 \le i \le d}} \max_{x \in K_\varepsilon} \left| \partial^\alpha v^i_\varepsilon(x) \right| \right] \in {}^{\rho}\widetilde{\mathbb{R}}$;

(ii) $\|v\|_l \ge 0$;

(iii) $\|v\|_l = 0$ *if and only if* $v = 0$;

(iv) $\forall c \in {}^{\rho}\widetilde{\mathbb{R}} : \ \|c \cdot v\|_l = |c| \cdot \|v\|_l$;

(v) *For all* $u \in {}^{\rho}\mathcal{GC}^k(K, {}^{\rho}\widetilde{\mathbb{R}}^d)$, *we have* $\|u+v\|_l \le \|u\|_l + \|v\|_l$ *and* $\|u \cdot v\|_l \le c_l \cdot \|u\|_l \cdot \|v\|_l$ *for some* $c_l \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}$.

Using these ${}^{\rho}\widetilde{\mathbb{R}}$-valued norms, we can naturally define a topology on the space ${}^{\rho}\mathcal{GC}^k(K, {}^{\rho}\widetilde{\mathbb{R}}^d)$.

Definition 40. Let $K \Subset_{\mathrm{f}} {}^{\rho}\widetilde{\mathbb{R}}^n$ be a solid set. Let $l \in \mathbb{N}_{\le k}$, $u \in {}^{\rho}\mathcal{GC}^k(K, {}^{\rho}\widetilde{\mathbb{R}}^d)$, $r \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}$, then

(i) $B^l_r(u) := \left\{ v \in {}^{\rho}\mathcal{GC}^k(K, {}^{\rho}\widetilde{\mathbb{R}}^d) \mid \|v-u\|_l < r \right\}$

(ii) If $U \subseteq {}^\rho\mathcal{GC}^k(K, {}^\rho\widetilde{\mathbb{R}}^d)$, then we say that U is a *sharply open set* if

$$\forall u \in U\, \exists l \in \mathbb{N}_{\le k}\, \exists r \in {}^\rho\widetilde{\mathbb{R}}_{>0}:\ B^l_r(u) \subseteq U.$$

One can easily prove that sharply open sets form a sequentially Cauchy complete topology on ${}^\rho\mathcal{GC}^k(K, {}^\rho\widetilde{\mathbb{R}}^d)$, see e.g. [29, 54].

The Banach fixed point theorem can be easily generalized to spaces of generalized continuous functions with the sup-norm $\|-\|_0$ (see Def. 38). As a consequence, we have the following Picard-Lindelöf theorem for ODE in the ${}^\rho\mathcal{GC}^k$ setting, see also [23, 54].

Theorem 41. *Let $t_0 \in {}^\rho\widetilde{\mathbb{R}}$, $y_0 \in {}^\rho\widetilde{\mathbb{R}}^d$, α, $r \in {}^\rho\widetilde{\mathbb{R}}_{>0}$. Let $F \in {}^\rho\mathcal{GC}^k([t_0-\alpha, t_0+\alpha] \times \overline{B_r(y_0)}, {}^\rho\widetilde{\mathbb{R}}^d)$. Set $M := \max_{\substack{t_0-\alpha\le t\le t_0+\alpha\\ |y-y_0|\le r}} |F(t,y)|$, $L := \max_{\substack{t_0-\alpha\le t\le t_0+\alpha\\ |y-y_0|\le r}} |\partial_y F(t,y)| \in {}^\rho\widetilde{\mathbb{R}}$ and assume that*

$$\alpha \cdot M \le r,$$

$$\lim_{n\to+\infty} \alpha^n L^n = 0, \tag{4.1}$$

where the limit in (4.1) is clearly taken in the sharp topology. Then there exists a unique solution $y \in {}^\rho\mathcal{GC}^{k+1}\left([t_0-\alpha, t_0+\alpha], {}^\rho\widetilde{\mathbb{R}}^d\right)$ of the Cauchy problem

$$\begin{cases} y'(t) = F(t, y(t)) \\ y(t_0) = y_0. \end{cases} \tag{4.2}$$

This solution is given by

$$y = \lim_{n\to+\infty} P^n(y_0)$$

$$P(y)(t) := y_0 + \int_{t_0}^t F(s, y(s))\,\mathrm{d}s \quad \forall t \in [t_0-\alpha, t_0+\alpha],$$

and for all $n \in \mathbb{N}$ satisfies $\|y - P^n(y_0)\|_0 \le \alpha M \sum_{k=n}^{+\infty} \frac{\alpha^n L^n}{n!}$ and $\|y - y_0\|_0 \le r$.

Finally, we have the following Grönwall-Bellman inequality in integral form:

Theorem 42. *Let $\alpha \in {}^\rho\widetilde{\mathbb{R}}_{>0}$. Let u, a, $b \in {}^\rho\mathcal{GC}^k\left([0,\alpha], {}^\rho\widetilde{\mathbb{R}}\right)$ and assume that $\|a\|_0 \cdot \alpha < N \cdot \log\left(\mathrm{d}\rho^{-1}\right)$ for some $N \in \mathbb{N}$. Assume that $a(t) \ge 0$ for all $t \in [0,\alpha]$, and that $u(t) \le b(t) + \int_0^t a(s)u(s)\,\mathrm{d}s$. Then*

(i) For every $t \in [0,\alpha]$ we have

$$u(t) \le b(t) + \int_0^t a(s)b(s)e^{\int_s^t a(r)\,\mathrm{d}r}\,\mathrm{d}s.$$

(ii) If $b(t) \le b(s)$ for all $t \le s$, i.e. if b is non-decreasing, then for every $t \in [0,\alpha]$ we have

$$u(t) \le b(t)e^{\int_0^t a(s)\,\mathrm{d}s}.$$

Finally, the following theorem considers global solutions of homogeneous linear ODE:

Theorem 43 (Solution of homogeneous linear ODE)**.** *Let $A \in {}^{\rho}\mathcal{GC}^{\infty}([a,b], {}^{\rho}\widetilde{\mathbb{R}}^{d\times d})$, where a, $b \in {}^{\rho}\widetilde{\mathbb{R}}$, $a < b$, and $t_0 \in [a,b]$, $y_0 \in {}^{\rho}\widetilde{\mathbb{R}}^{d}$. Assume that*

$$\left|\int_{t_0}^{t} A(s)\mathrm{d}s\right| \le -C \cdot \log \mathrm{d}\rho \quad \forall t \in [a,b], \tag{4.3}$$

where $C \in \mathbb{R}_{>0}$. Then there exists one and only one $y \in {}^{\rho}\mathcal{GC}^{\infty}([a,b], {}^{\rho}\widetilde{\mathbb{R}}^{d})$ such that

$$\begin{cases} y'(t) = A(t)\cdot y(t) & \text{if } t \in [a,b] \\ y(t_0) = y_0 \end{cases} \tag{4.4}$$

Moreover, this y is given by $y(t) = \exp\left(\int_{t_0}^{t} A(s)\mathrm{d}s\right)\cdot y_0$ for all $t \in [a,b]$.

In general, the solution of a differential equation in a non-Archimedean setting is defined on an infinitesimal neighborhood of the initial condition. This is a general fact of every non-Archimedean theory having at least one positive and invertible infinitesimal h. If fact, the Cauchy problem

$$\begin{cases} y' = -\frac{t}{1+y}\cdot\frac{1}{h} \\ y(0) = 0 \end{cases} \tag{4.5}$$

has solution $y(t) = -1 + \sqrt{1 - \frac{t^2}{h}}$ which is defined and smooth only in the infinitesimal interval $(-\sqrt{h}, \sqrt{h})$. Moreover, we have that $\lim_{t\to\pm\sqrt{h}} y'(t) = +\infty$ (in the sharp topology) and this clearly shows that the solution cannot be extended. However, very general sufficient conditions to have non-infinitesimal domains can be proved, considering e.g. the case where the right hand side F in (4.2) is an ordinary smooth function, or when we extend the theory of Picard iterations P^n to an infinite natural number $n = [n_\varepsilon] \in {}^{\rho}\widetilde{\mathbb{R}}$, $n_\varepsilon \in \mathbb{N}$, see [54]. We also finally state that a very general Picard-Lindelöf theorem can also be proved for PDE, see [33, 34, 18].

5. Formal deductions corresponding to informal reasonings

In the previous sections, we reviewed GSF theory and we hope we persuaded the reader that a meaningful and sufficiently complete theory containing infinitesimal and infinite numbers is possible. This non-Archimedean theory does not require any background in mathematical logic, has clear connections with the usual standard calculus, is intuitively clear, but also solves non trivial problems such as the possibility to consider generalized functions with infinite derivatives, making non-linear operations on Schwartz distributions and sharing several results of ordinary smooth functions.

Now, the framework of GSF theory allows one to formalize several informal reasonings with the intuitive use of infinitesimal and infinite numbers we can find in physics, engineering and even in mathematics. The main goal is absolutely not the empty searching for the mathematical rigour, but the learning of the true rules of infinitesimal calculus instead of unclear foggy explanations and, mainly, the flexibility to create new and simpler mathematical models of real-world problems. As a trivial example, using the Taylor formula with nilpotent infinitesimals Thm. 37, if $\frac{v^2}{c^2} \in D_{1j}$, we can write (1.1) as $1/\sqrt{1 - v^2/c^2} =_j 1 + \frac{v^2}{2c^2}$ for all $j \in \mathbb{R}_{>0}$ and Einstein calculations remain essentially unchanged. In the next sections, we will see that this method not only allows one to obtain a rigorous version of the usual informal deductions of the heat and wave equations, but that these same proofs show the validity of these equations for GSF, opening new applications for example to optics of different materials and geophysics.

A frequently underestimated consequence of seeing generalized functions, e.g. any Schwartz distribution T, as set-theoretical functions is that pointwise values $T(x_0)$ are now always well-defined. Therefore, non-linear boundary value problems are now conceivable (see e.g. (4.2)), and this is a solution of a non trivial drawback of Schwartz theory having important consequences for mathematical modeling.

5.1. **Derivation of the heat equation for GSF.** In this section, we derive the heat equation in a similar way to [77, 27], with the difference that here we extend the applicability to GSF and not only to smooth functions. Let $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ denotes the standard basis of $\mathbb{R}^3$, so that any vector $a \in {}^\rho\widetilde{\mathbb{R}}^3$ is of the form $a = \lambda_1 \cdot \vec{e}_1 + \lambda_2 \cdot \vec{e}_2 + \lambda_3 \cdot \vec{e}_3$ for λ_1, λ_2, $\lambda_3 \in {}^\rho\widetilde{\mathbb{R}}$. In the following, a symbol of the form $\delta y \in {}^\rho\widetilde{\mathbb{R}}$ intuitively means that the infinitesimal increment δy is associated to the variable y.

Let us consider a body $B \subseteq {}^\rho\widetilde{\mathbb{R}}^3$ represented by a solid set, i.e. $B = \overline{\mathring{B}}$, so that values of GSF on the boundary of B can be computed as limits of values at interior points. We consider the following GSF:

- $\varrho : B \to {}^\rho\widetilde{\mathbb{R}}$ (mass density) ,
- $c : B \to {}^\rho\widetilde{\mathbb{R}}$ (heat capacity),
- $k : B \to {}^\rho\widetilde{\mathbb{R}}$ (thermal conductivity coefficient).

Note that we do not make any assumptions on the favoured directions of these functions on their domain B. This assumption corresponds to the isotropy condition for B. The next GSF we need represents the temperature of the body B at each point $x \in B$ and time $t \in [0, \infty)$ and is denoted by $u : B \times [0, \infty) \to {}^\rho\widetilde{\mathbb{R}}$.
We choose an interior point $x \in \mathring{B}$ and an infinitesimal volume $V \subset {}^\rho\widetilde{\mathbb{R}}^3$ of the form

$$V = V(x, \delta\bar{x}) = \{y \in {}^\rho\widetilde{\mathbb{R}}^3 \mid -\delta x_i \le 2(y - x) \cdot \vec{e}_i \le \delta x_i \ \forall i = 1, 2, 3\}, \tag{5.1}$$

where $\delta x_i \in {}^\rho\widetilde{\mathbb{R}}_{>0}$ and $\delta\bar{x} := (\delta x_1, \delta x_2, \delta x_3)$. Such a set is said to be an infinitesimal parallelepiped if $\delta v := \delta x_1 \cdot \delta x_2 \cdot \delta x_3 \approx 0$, that is, if the corresponding volume is infinitesimal. Note that since $x \in \mathring{B}$, we have $\exists \delta\bar{x} \in {}^\rho\widetilde{\mathbb{R}}^3_{>0} : \ V = V(x, \delta\bar{x}) \subseteq B$, and hence we can view V as the subbody of B corresponding to the infinitesimal parallelepiped centered at x with sides parallel to the coordinate axes. This subbody interacts thermally with its complement $CV := B \setminus V$ and with external heat sources. In this type of deductions, the physical part frequently consists, from the mathematical point of view, in physically meaningful definitions or assumptions corresponding to physical principles or constitutive relations. For example, we now recall Fourier's law, which states that during the infinitesimal time interval δt the heat $Q_{CV,V}$ flowing perpendicularly to the surface of V defines the exchange between V and CV, and this yields the following

$$Q_{CV,V} := Q_{CV,V}(x, t, \delta t, \delta\bar{x}) \tag{5.2}$$

$$= \delta t \cdot \sum_{i=1}^{3} \delta s_i \cdot [k(x + \delta\vec{h}_i) \cdot \frac{\partial u}{\partial x_i}(x + \delta\vec{h}_i, t) - k(x - \delta\vec{h}_i) \cdot \frac{\partial u}{\partial x_i}(x - \delta\vec{h}_i, t)], \tag{5.3}$$

where $\delta\vec{h}_i := \dfrac{1}{2}\delta x_i \cdot \vec{e_i} \in {}^\rho\widetilde{\mathbb{R}}^3$ and $\delta s_i := \prod_{j\neq i} \delta x_j \in {}^\rho\widetilde{\mathbb{R}}$. Note explicitly that $Q_{CV,V}$ depends on x, t, δt, δx_i. The heat exchange of V due to thermal interactions with external sources is given by the expression

$$Q_{\text{ext},V} := Q_{\text{ext},V}(x,t,\delta t,\delta\bar{x}) = F(x,t)\cdot\delta v\cdot\delta t, \tag{5.4}$$

where $F(x,t) : B \to {}^\rho\widetilde{\mathbb{R}}$ is a GSF representing the intensity of the thermal sources. The total heat is $Q_{CV,V} + Q_{\text{ext},V}$ and it 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 $Q_{\text{env},V}$ that reads

$$\begin{aligned} Q_{\text{env},V} &:= Q_{\text{env},V}(x,t,\delta t,\delta\bar{x}) = [u(x,t+\delta t) - u(x,t)]\cdot c(x)\cdot\varrho(x)\cdot\delta v, && (5.5)\\ &= Q_{CV,V} + Q_{\text{ext},V}. && (5.6)\end{aligned}$$

We now want to apply the first order nilpotent Taylor formula Thm. 37.(xii), at (5.2) and (5.5), i.e. at the GSF k, $\frac{\partial u}{\partial x_i}(-,t)$ and $u(x,-)$. From (xii) and (iv) of Thm. 37, if these formulas hold respectively on $D_{1e'}$, $D_{1e''}$ and $D_{1\bar{e}}$, then they also hold on D_{1e}, where $e = \min(e', e'', \bar{e}, j)$. We choose our infinitesimals in such a way that $\delta v \cdot \delta t \in D_{1e}$, $\delta t \cdot \delta s_i \cdot (\delta x_i)^2 =_j 0$ and $(\delta t)^2 \delta v =_j 0$. Using these infinitesimals, second order terms using nilpotent Taylor formula Thm. 37.(xi) in (5.2) and (5.5) will not give a contribution if we use the equality $=_j$. We will see later that infinitesimals δt and δx_i satisfying all the needed conditions actually exist.
This allows us to rewrite (5.2) and (5.5) as follows

$$Q_{CV,V} =_j \operatorname{div}[k\cdot\operatorname{grad}(u)](x,t)\cdot\delta v\cdot\delta t, \tag{5.7}$$

$$Q_{\text{env},V} =_j c(x)\cdot\varrho(x)\cdot\frac{\partial u}{\partial t}(x,t)\cdot\delta v\cdot\delta t. \tag{5.8}$$

Note that the calculations with the nilpotent Taylor formula to get (5.7) correspond to the divergence theorem. From (5.7), (5.4) and (5.8) we therefore get that the equality $Q_{\text{env},V} =_j Q_{CV,V} + Q_{\text{ext},V}$ is equivalent to

$$c(x)\cdot\varrho(x)\cdot\frac{\partial u}{\partial t}(x,t)\delta t\delta v =_j [\operatorname{div}[k\cdot\operatorname{grad}(u)](x,t) + F(x,t)]\,\delta t\delta v. \tag{5.9}$$

More precisely: (5.6) implies (5.9), and the latter implies the former but with $=_j$ replacing $=$. The following theorem allows us to cancel the nilpotent factor $\delta t\delta v$ in (5.9):

Theorem 44. *Let $x, r, s \in {}^\rho\widetilde{\mathbb{R}}$, $|x| \geq \mathrm{d}\rho^q$, $j \in \mathbb{R}_{>0}$. Assume that $x\cdot r =_j x\cdot s$ and $\frac{1}{j} - q =: \frac{1}{k} > 0$. Then $r =_k s$. Vice versa, if $r =_k s$, and x is finite, then $x\cdot r =_k x\cdot s$.*

Proof. Assume that $x\cdot r =_j x\cdot s$. Then $|xr - xs| \leq C\mathrm{d}\rho^{\frac{1}{j}}$, with $C \in \mathbb{R}_{\geq 0}$. Then , $|r-s| = |x|\cdot|r-s|\cdot\dfrac{1}{|x|} \leq \dfrac{C\cdot\mathrm{d}\rho^{\frac{1}{j}}}{\mathrm{d}\rho^q} = C\cdot\mathrm{d}\rho^{\frac{1}{j}-q} = C\mathrm{d}\rho^{\frac{1}{k}}$ since $\frac{1}{k} = \frac{1}{j} - q$. For the second part of the conclusion, x finite means $|x| \leq K \in \mathbb{R}_{>0}$, so that $|r-s| \leq C\mathrm{d}\rho^{\frac{1}{k}}$ implies $|xr - xs| \leq KC\mathrm{d}\rho^{\frac{1}{k}}$. □

This derivation is summed up in the following Lemma which we just have proven.

Lemma 45. *Let* $B \subseteq {}^\rho\widetilde{\mathbb{R}}^3$, $B = \overline{\mathring{B}}$, *and consider the GSF* ϱ, c, $k : B \to {}^\rho\widetilde{\mathbb{R}}$, u, $F : B \times [0, \infty) \to {}^\rho\widetilde{\mathbb{R}}$. *Take a point* $(x, t) \in \mathring{B} \times [0, \infty)$ *and define* V, $Q_{CV,V}$, $Q_{\text{ext},V}$ *and* $Q_{\text{env},V}$ *as in* (5.2), (5.4), (5.5), *where the infinitesimals* δt, $\delta x_i \in {}^\rho\widetilde{\mathbb{R}}_{>0}$ *satisfy*

$$\begin{aligned} &\delta v \cdot \delta t \in D_{1e},\ \delta t \cdot \delta s_i \cdot (\delta x_i)^2 =_j 0,\ (\delta t)^2 \delta v =_j 0 \\ &\delta v \cdot \delta t \ge \mathrm{d}\rho^q,\ \frac{1}{k} = \frac{1}{j} - q, \end{aligned} \tag{5.10}$$

and where the first order nilpotent Taylor formula for k, $\frac{\partial u}{\partial x_i}(-, t)$ *and* $u(x, -)$ *holds in* D_{1e}. *Then the following are equivalent:*

(i) $Q_{\text{env},V} =_j Q_{CV,V} + Q_{\text{ext},V}$,

(ii) $c(x) \cdot \varrho(x) \cdot \dfrac{\partial u}{\partial t}(x, t) =_k \operatorname{div}[k \cdot \operatorname{grad}(u)](x, t) + F(x, t)$.

Note that this result corresponds to the usual informal derivation, but it is now stated as a formal theorem where the use of nilpotent infinitesimals and the corresponding Taylor formula is now precise.

The next natural steps thus concern the existence of infinitesimals satisfying (5.10) and how to obtain a true equality $=$ in the final heat equation for GSF. Conditions (5.10) hold if e.g. we choose $\delta t = \mathrm{d}\rho^{\frac{1}{3e}}$ and $\delta x_i = \mathrm{d}\rho^{\frac{1}{5e}}$ (recall that $e \le j$ and note that these infinitesimals depend on j); thereby, it easily follows that we can take $q = q(j) = \frac{14}{15j}$ and hence $k = k(j) = 15j$.

Finally, assume that $Q_{\text{env},V}(x, t, \delta t, \delta\bar{x}) = Q_{CV,V}(x, t, \delta t, \delta\bar{x}) + Q_{\text{ext},V}(x, t, \delta t, \delta\bar{x})$ holds at (x, t) and for all infinitesimals δt, $\delta\bar{x}$. Thereby (using simplified notations)

$$Q_{\text{env},V} =_j Q_{CV,V} + Q_{\text{ext},V} \quad \forall j \in \mathbb{R}_{>0}. \tag{5.11}$$

Lemma 45 yields the heat equation with equality up to order $k(j) = 15j$. If we now let $j \to 0^+$, then also $k(j) \to 0^+$ and hence Thm. 37.(v) proves the heat equation with $=$.

Even if it is true that the full equality $=$ implies $=_{k(j)}$ in the heat equation, the opposite implication (i.e. that (ii) above but with $=$ instead of $=_k$, implies (i) above with $=$ instead of $=_j$) cannot be proved simply by reversing the previous steps because we would arrive at (5.11) with infinitesimals $\delta t = \delta t(j)$, $\delta\bar{x} = \delta\bar{x}(j)$ satisfying (5.10) that would depend on j: taking $j \to 0^+$ in (5.11) would not get anything meaningful because $\delta t(j)$, $\delta\bar{x}(j) \to 0$.

The final result is then stated as follows:

Theorem 46. *Let* $B \subseteq {}^\rho\widetilde{\mathbb{R}}^3$, $B = \overline{\mathring{B}}$, *and consider the GSF* ϱ, c, $k : B \to {}^\rho\widetilde{\mathbb{R}}$, u, $F : B \times [0, \infty) \to {}^\rho\widetilde{\mathbb{R}}$. *Take a point* $(x, t) \in \mathring{B} \times [0, \infty)$ *and define* V, $Q_{CV,V}$, $Q_{\text{ext},V}$ *and* $Q_{\text{env},V}$ *as in* (5.2), (5.4), (5.5). *Finally assume that* $Q_{env,V}(x, t, \delta t, \delta\bar{x}) = Q_{CV,V}(x, t, \delta t, \delta\bar{x}) + Q_{ext,V}(x, t, \delta t, \delta\bar{x})$ *holds at* (x, t) *and for all infinitesimals* δt, $\delta\bar{x}$. *Then*

$$c(x) \cdot \varrho(x) \cdot \frac{\partial u}{\partial t}(x, t) = \operatorname{div}[k \cdot \operatorname{grad}(u)](x, t) + F(x, t). \tag{5.12}$$

Moreover, if these conditions hold at all points $x \in \mathring{B}$, *then equation* (5.12) *holds on the entire* B *because* $B = \overline{\mathring{B}}$.

5.2. **Derivation of the wave equation for GSF.** In this section, we derive the wave equation in a similar way to [77, 27], with the difference that we extend its applicability to GSF and not only to smooth functions. Consider a string with given equilibrium position located on an interval $[a,b] \subseteq {}^\rho\widetilde{\mathbb{R}}$ for $a, b \in {}^\rho\widetilde{\mathbb{R}}$, $a<b$. Let this string now make small transversal oscillations around its equilibrium position. The position $s_t \subseteq {}^\rho\widetilde{\mathbb{R}}^2$ of the string is always represented by the graph of a curve $\gamma : [a,b] \times [0,\infty) \to {}^\rho\widetilde{\mathbb{R}}^2$. Furthermore, we set $\gamma_{xt} := \gamma(x,t)$, $s_t := \{\gamma_{xt} \in {}^\rho\widetilde{\mathbb{R}}^2 \mid a \le x \le b\}$ for all $t \in [0,\infty)$. The curve γ is supposed to be injective with respect to $x \in (a,b)$, i.e. $\gamma_{x_1 t} \ne \gamma_{x_2 t}$ for all $t \in [0,+\infty)$ and all x_1, $x_2 \in (a,b)$ such that $x_1 \ne x_2$; therefore, the order relation on (a,b) implies an order relation on the support s_t. For all pairs 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 $[p := \{\gamma_{xt} \mid x_p \le x \le b\}$, $p] := \{\gamma_{xt} \mid a \le x \le x_p\}$ and $[p|q] := \{\gamma_{xt} \mid x_p \le x \le x_q\}$ corresponding to the parts of the string after the point $p \in s_t$, before the same point and between the points $p, q \in s_t$. Clearly, every subbody of the form $p]$ exerts a force on every other subbody it is in contact with, i.e. $[p|q]$ and $p]$. Moreover, the force $\vec{F}(A,B) \in {}^\rho\widetilde{\mathbb{R}}$ exerted by the subbody A on the subbody B satisfies the following equalities:

$$\vec{F}([p|q],p]) = \vec{F}([p,p]) \tag{5.13}$$
$$\vec{F}([q,[p|q]) = \vec{F}([q,q]) \tag{5.14}$$
$$\vec{F}(p],[p|q]) = -\vec{F}([p|q],p]), \tag{5.15}$$

for all pairs of points p, $q \in s_t$ and time $t \in [0,\infty)$. The third equation (5.15) corresponds to the action-reaction principle.

We can now define the tension $\vec{T}$ at the point $\gamma_{xt} \in s_t$ and time $t \in [0,\infty)$ as

$$\vec{T}(x,t) := \vec{F}([\gamma_{xt},\gamma_{xt}]). \tag{5.16}$$

Consider now the infinitesimal subbody $[x|x+\delta x] := [\gamma_{xt}|\gamma_{x+\delta x,t}] \subseteq s_t$ located at time t between the points $\gamma_{xt} \in s_t$ and $\gamma_{x+\delta x,t} \in s_t$, and defined by the first order infinitesimal $\delta x \in D_{1j}$, $\delta x > 0$. We have an action on this infinitesimal subbody due to mass forces of linear density $\vec{G} : [a,b] \times [0,\infty) \to {}^\rho\widetilde{\mathbb{R}}^2$ that allows us to represent Newton's law as follows:

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} = \vec{F}(\gamma_{xt}],[x|x+\delta x]) + \vec{F}([\gamma_{x+\delta x,t},[x|x+\delta x]) + \vec{G} \cdot \varrho \cdot \delta x, \tag{5.17}$$

where $\varrho : [a,b] \times [0,\infty) \to {}^\rho\widetilde{\mathbb{R}}^2$ is the linear mass density, and all functions, unless stated otherwise, are evaluated at $(x,t) \in (a,b) \times [0,\infty)$.

The contact forces appearing in Newton's law are caused by the interaction of the infinitesimal subbody with other contacting subbodies along the border $\partial[x|x+\delta x] = \{\gamma_{xt},\gamma_{x+\delta x,t}\} \subseteq {}^\rho\widetilde{\mathbb{R}}^2$. Using now relations (5.14) and (5.15) with $q = \gamma_{x+\delta x,t}$ and $p = \gamma_{xt}$, so that $[p|q] = [x|x+\delta x]$, we see by (5.17) that

$$\rho \cdot \delta x \cdot \frac{\partial^{2\gamma}}{\partial t^2} = -\vec{F}([x|x+\delta x],\gamma_{xt}]) + \vec{F}([\gamma_{x+\delta x,t},\gamma_{x+\delta x,t}]) + \vec{G} \cdot \rho \cdot \delta x. \tag{5.18}$$

By (5.13), the definition of tension (5.16) and inserting it in (5.18), we obtain

$$\begin{aligned} \rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2} &= -\vec{F}([\gamma_{xt},\gamma_{xt}]) + \vec{F}([\gamma_{x+\delta x,t},\gamma_{x+\delta x,t}]) + \overrightarrow{G} \cdot \rho \cdot \delta x \\ &= -\vec{T}(x,t) + \vec{T}(x+\delta x,t) + \vec{G} \cdot \rho \cdot \delta x. \end{aligned} \tag{5.19}$$

Note that, up to now, we have not used neither the small oscillation nor the transversal oscillation hypothesis of the force $\vec{G}$. As for the latter, it can be introduced with the assumption

$$\vec{G}(x,t)\cdot\vec{e}_1 = 0 \quad \forall x,t, \tag{5.20}$$

where $(\vec{e}_1,\vec{e}_2)$ are the axial unit vectors. Let now $\varphi(x,t)$ denote the non-oriented angle between the tangent unit vector $\vec{t}(x,t) := \frac{\partial\gamma}{\partial x}(x,t)/\left|\frac{\partial\gamma}{\partial x}(x,t)\right|$ (a subsequent assumption will guarantee that $\vec{t}$ always exists) at the point $\gamma_{x,t}$ and the x-axis, i.e. the unique $\varphi(x,t)\in[0,\pi]\subseteq {}^{\rho}\widetilde{\mathbb{R}}$ defined by

$$\vec{t}(x,t) = \cos(\varphi(x,t))\vec{e}_1 + \sin(\varphi(x,t))\vec{e}_2. \tag{5.21}$$

Setting $(\gamma_1,\gamma_2)=\gamma$ for the two components of the curve γ, from this equality directly follows

$$\frac{\partial\gamma_1}{\partial x}\sin\varphi = \frac{\partial\gamma_2}{\partial x}\cos\varphi \tag{5.22}$$

The small oscillation hypothesis can then be formalized with the assumption that this angle $\varphi(x,t)$ is a first order infinitesimal (in the following Thm. 47, we will assume a weaker assumption), i.e.

$$\varphi(x,t)\in D_{1j} \quad \forall x,t. \tag{5.23}$$

This allows us to recreate the classical derivation in the most faithful possible way. Furthermore, in the standard proof of the wave equation, only curves of the specific form $\gamma_{xt}=(x,u(x,t))$ are considered (this implies that the tangent unit vector $\vec{t}(x,t)$ always exists). The Tayor-formula for nilpotent infinitesimals Thm. 37.(xi) yields $\sin(\varphi)=_j\varphi\in D_{1j}$ and $\cos(\varphi)=_j 1$ (note that assumption (3.12) holds for any j and k for both $\sin(x)$ and $\cos(x)$), and hence $\dfrac{\partial\gamma_2}{\partial x}=_j\varphi$ from (5.22). Therefore, $\left(\dfrac{\partial\gamma_2}{\partial x}\right)^2=_j 0$ and the total length of the string becomes

$$L=\int_a^b\sqrt{1+\left[\frac{\partial\gamma_2}{\partial x}(x,t)\right]^2}\,\mathrm{d}x =_j b-a \quad \forall t\in[0,\infty). \tag{5.24}$$

Following Hooke's law, this allows us to assume that the tension is of constant modulus $T=|\vec{T}(x,t)|$ that is neither depending on the position x nor on the time t, i.e.

$$\vec{T}(x,t)=T\cdot\vec{t}(x,t) \quad \forall x\in(a,b)\,\forall t\in[0,\infty). \tag{5.25}$$

Note that, as a second part of the hypothesis about nontransversal oscillations of the string, we have that the tension $\vec{T}$ is parallel to the tangent vector. We then project equation 5.19 to the y-axis and obtain

$$\begin{aligned}\rho\cdot\delta x\cdot\frac{\partial^2u}{\partial t^2} &= -T\cdot\vec{t}(x,t)\cdot\vec{e}_2 + T\cdot\vec{t}(x+\delta x,t)\cdot\vec{e}_2+\vec{G}\cdot\vec{e}_2\cdot\rho\cdot\delta x\\ &= -T\cdot\sin(\varphi(x,t))+T\cdot\sin(\varphi(x+\delta x),t)+G_2\cdot\rho\cdot\delta x\\ &= T\cdot\left[\frac{\partial u}{\partial x}(x+\delta x,t)\cos(\varphi(x+\delta x,t))-\frac{\partial u}{\partial x}(x,t)\cos(\varphi(x,t))\right] \qquad (5.26)\\ &\quad +G_2\cdot\rho\cdot\delta x, \qquad (5.27)\end{aligned}$$

where $G_2=\vec{G}\cdot\vec{e}_2$ is the second component of $\vec{G}$. Now, assume that the first order Taylor formula for $\frac{\partial u}{\partial x}(-,t)$ holds on D_{1e}, with $e\le j$, and take $\delta x\in D_{1e}$,

$\delta x \geq \mathrm{d}\rho^q$ (e.g. $\delta x = \mathrm{d}\rho^{\frac{1}{2e}+\frac{1}{2}}$). Then, $\cos(\varphi(x+\delta x, t)) =_j 1 =_j \cos(\varphi(x,t))$ and $\frac{\partial u}{\partial x}(x+\delta x, t) - \frac{\partial u}{\partial x}(x,t) =_j \frac{\partial^2 u}{\partial x^2}(x,t)\delta x$, and from (5.26) we hence get

$$\rho \cdot \delta x \cdot \frac{\partial^2 u}{\partial t^2} =_j T \frac{\partial^2 u}{\partial x^2}(x,t)\delta x + G_2 \cdot \rho \cdot \delta x.$$

We can now use the cancellation law Thm. 44 to cancel out the δx obtaining

$$\rho \frac{\partial^2 u}{\partial t^2} =_k T \frac{\partial^2 u}{\partial x^2}(x,t) + G_2 \rho, \tag{5.28}$$

for $\frac{1}{k} = \frac{1}{j} - q$.

Can we take $j \to 0^+$ (and hence $k \to 0^+$) in (5.28)? Actually no, because all this deduction depends on the small oscillations assumption (5.23), and the only $\varphi \in D_{1j}$ for all j is $\varphi = 0$, i.e. the string is not oscillating at all. In order to underscore that this classical deduction of the wave equation leads to an approximate equality only, we generalize the previous proof in the following

Theorem 47. *Let $a, b \in {}^\rho\widetilde{\mathbb{R}}$, with $a < b$, $\gamma : [a,b] \times [0,\infty) \to {}^\rho\widetilde{\mathbb{R}}^2$, $\rho : [a,b] \times [0,\infty) \to {}^\rho\widetilde{\mathbb{R}}$, $\vec{G}$, $\vec{T} : [a,b] \times [0,\infty) \to {}^\rho\widetilde{\mathbb{R}}^2$ be GSF, and let $T \in {}^\rho\widetilde{\mathbb{R}}$ be an invertible generalized number such that both T and $\frac{1}{T}$ are finite. Suppose that $\gamma(x,t) = (x, u(x,t))$ for all x, t, and let $\vec{t}(x,t)$ be the unit tangent vector to γ. Assume that at least an approximate version of Hooke's law and the second Newton's law*

$$\vec{T}(x,t) =_j T \cdot \vec{t}(x,t), \tag{Hooke}$$

$$\rho \cdot \delta x \cdot \frac{\partial^2 \gamma}{\partial t^2}(x,t) = \vec{T}(x+\delta x, t) - \vec{T}(x,t) + \vec{G} \cdot \rho \cdot \delta x, \tag{II Newton}$$

hold for every point $(x,t) \in (a,b) \times [0,\infty)$ and for an infinitesimal $\delta x = \mathrm{d}\rho^q$ such that $\delta x \in D_{1e}$, where the first order Taylor formula for $\frac{\partial u}{\partial x}(-,t)$ holds on D_{1e} and $e \leq j$. Finally, let $\varphi(x,t)$ be the non-ordered angle between $\vec{t}(x,t)$ and the x-axis, and suppose that $\dfrac{\partial \varphi}{\partial x}(x,t) \geq \mathrm{d}\rho^p$, $\varphi(x,t) < \frac{\pi}{2}$. Then we have:

(i) *If $\rho(x,t) \cdot \dfrac{\partial^2 u}{\partial t^2}(x,t) =_j T \cdot \dfrac{\partial^2 u}{\partial x^2}(x,t) + G_2(x,t) \cdot \rho(x,t)$, then $\cos^3(\varphi(x,t)) =_h 1$, where $\frac{1}{h} = \frac{1}{j} - p - 2q$.*

(ii) *If $\cos^3(\varphi(x,t)) =_j 1$, then $\rho(x,t) \cdot \dfrac{\partial^2 u}{\partial t^2}(x,t) =_k T \cdot \dfrac{\partial^2 u}{\partial x^2}(x,t) + G_2(x,t) \cdot \rho(x,t)$, where $\frac{1}{k} = \frac{1}{j} - q$.*

For example, the assumption of (ii) holds if $\varphi(x,t) \in D_{k\hat{\jmath}}$ and $\frac{(k+1)}{2}\hat{\jmath} = j$. Finally, if $\varphi(x,t) \in D_{3j}$ for all x, t, and $b - a$ is finite, then $\mathrm{length}(\gamma(-,t)) =_{2j} b - a$.

Proof. As usual, if the arguments of a function are missing, we mean they are evaluated at (x,t).

(i): Projecting (II Newton) on $\vec{e}_2$ and using (Hooke) and (5.21) we get

$$\rho \delta x \frac{\partial^2 u}{\partial t^2} = T \sin(\varphi(x+\delta x, t)) - T \sin(\varphi(x,t)) + G_2 \rho \delta x.$$

Therefore, the assumption of (i) implies

$$T \sin(\varphi(x+\delta x, t)) - T \sin(\varphi(x,t)) + G_2 \rho \delta x =_j T \frac{\partial^2 u}{\partial x^2} \delta x + G_2 \rho \delta x.$$

Since $\delta x \in D_{1e}$ and $e \le j$, we can use the first order Taylor formula with $\frac{\partial u}{\partial x}(-,t)$ to get

$$T\sin(\varphi(x+\delta x,t)) - T\sin(\varphi(x,t)) + G_2\rho\delta x =_j$$
$$T\delta x\left[\frac{\partial u}{\partial x}(x+\delta x,t) - \frac{\partial u}{\partial x}(x,t)\right] + G_2\rho\delta x.$$

Multiply by $\frac{1}{T}$ (which is finite, see Thm. 44) and use (5.22) considering that $\varphi(x,t) < \frac{\pi}{2}$ to obtain

$$\left[\sin(\varphi(x+\delta x,t)) - \sin(\varphi(x,t))\right]\delta x =_j \left[\tan(\varphi(x+\delta x,t)) - \tan(\varphi(x,t))\right]\delta x.$$

Using the cancellation law Thm. 44 with $\frac{1}{k} := \frac{1}{j} - q$, this yields

$$\sin(\varphi(x+\delta x,t)) - \sin(\varphi(x,t)) =_k \tan(\varphi(x+\delta x,t)) - \tan(\varphi(x,t)).$$

We can use the first order Taylor formula Thm. 37.(xi) both with $\sin(\varphi(-,t))$ and $\tan(\varphi(-,t))$ because $e \le j$ and hence $\delta x \in D_{1e} \subseteq D_{1j} \subseteq D_{1k}$ (note that the derivatives of these functions are always finite because $\varphi(x,t) < \frac{\pi}{2}$)

$$\delta x \cdot \cos(\varphi) \cdot \frac{\partial \varphi}{\partial x} =_k \delta x \frac{1}{\cos^2(\varphi)} \cdot \frac{\partial \varphi}{\partial x}.$$

Simplifying by $\delta x \cdot \frac{\partial \varphi}{\partial x} \ge \mathrm{d}\rho^{q+p}$, we obtain $\cos(\varphi) =_h \frac{1}{\cos^2(\varphi)}$, where $\frac{1}{h} := \frac{1}{k} - p - q = \frac{1}{j} - p - 2q$. Since $\cos^2(\varphi)$ is finite, using Thm. 44 we obtain the conclusion.
(ii): It suffices to invert all the previous steps starting from $\cos^3(\varphi) =_j 1$ and considering that we always have to multiply by finite numbers. Only in the last step we need to simplify by δx and hence we switch from $=_j$ to $=_k$.

From Taylor formula with Peano remainder Thm. 37.(i) we have $\cos^3(\varphi) = \left(1 - \frac{\varphi^2}{2} + o(\varphi^3)\right)^3 = 1 - \frac{3}{2}\varphi^2 + o(\varphi^3)$. If $\varphi \in D_{k\hat{j}}$, then $|\varphi^{k+1}| \le C\mathrm{d}\rho^{\frac{1}{\hat{j}}}$ and hence $\varphi^2 \le C\mathrm{d}\rho^{\frac{2}{(k+1)\hat{j}}} = C\mathrm{d}\rho^{\frac{1}{j}}$ and $\left|\cos^3(\varphi) - 1\right| = \left|\frac{3}{2}\varphi^2 + o(\varphi^3)\right| \le \bar{C}\mathrm{d}\rho^{\frac{1}{j}}$. Note that this property includes the classical case $\varphi \in D_{1j}$, but also e.g. $\varphi \in D_{2j-1,1}$.

Finally, assume that $\varphi(x,t) \in D_{3j}$ for all x, t. From Taylor formula $\sin(\varphi) =_j \varphi - \frac{\varphi^3}{6}$ and $\cos(\varphi) =_j 1 - \frac{\varphi^2}{2}$. Therefore, (5.22) yields $\varphi - \frac{\varphi^3}{6} =_j \frac{\partial u}{\partial x}\left(1 - \frac{\varphi^2}{2}\right)$. Taking the square and considering that $\varphi^4 =_j 0$, this implies $\varphi^2 =_j \left(\frac{\partial u}{\partial x}\right)^2\left(1 - \varphi^2\right)$. Multiplying both sides by $1 + \varphi^2$ and using again that $\varphi^4 =_j 0$ we obtain $\left(\frac{\partial u}{\partial x}\right)^2(x,t) =_j \varphi^2(x,t)$ for all x, t. The mean value theorem Thm. 27.(ii) and Thm. 37.(vii) yield $\mathrm{length}(\gamma(-,t)) = \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2(c,t)} \cdot (b-a) =_j \sqrt{1 + \varphi^2(c,t)} \cdot (b-a)$ for some $c \in [a,b]$. The Taylor formula with Peano remainder applied to the function $\sqrt{1+x}$ gives $\mathrm{length}(\gamma(-,t)) =_j b - a + \frac{b-a}{2}\varphi^2(c,t) + o(\varphi)$, which implies the conclusion because $\left|\frac{b-a}{2}\right|\varphi^2(c,t) \le C\mathrm{d}\rho^{\frac{1}{2j}}$. □

This theorem suggests the following comments and potential applications:

(i) It highlights that the wave equation is intrinsically approximated because it implies $\cos^3(\varphi) =_h 1$, which is necessarily only an approximated relation.

(ii) It is formulated as a general mathematical theorem depending on two assumptions corresponding to physical laws.

(iii) In our deduction, we do not conclude by “magically” transforming approximate equalities $\simeq$ into true equalities $=$ or neglecting little-oh terms despite keeping true equalities.

(iv) The validity of the wave equation for GSF can find possible applications in geophysics. In seismology, we have for example elastodynamical oscillations after earthquakes or simply the elastodynamical properties of materials that have a rapid change in density like the seabed or earth's crust. This leads to the seismological base equations of elastodynamics with a special case being the isotropic wave equation where the setting of GSF could be used to treat the special case with non-smooth coefficients. A motivation for this topic can be found in [6, 9].

(v) Other potential applications can also be considered in global seismology, where one is dealing with seismic wave propagation. In fact, hyberbolic PDE in global seismology do have generalized functions as coefficients, together with a singular structure created by geological and physical processes. These processes are supposed to behave in a fractal way. In the so-called *seismic transmission problem*, we want to diagonalize a first order system of PDE and then transform it to the second order wave equation. This requires us to differentiate the coefficients, which means that even though the original model medium varies continuously, coefficients that are (highly) discontinuous will naturally appear in this procedure. A possible way to deal with this is to embed the fractal coefficients into GSF or in a Colombeau algebra. See e.g. [37].

(vi) We can finally think at using GSF in mathematical general relativity, where one considers wave equations on Lorentzian manifolds with non-smooth metric, i.e. non-smooth coefficients in the corresponding wave equation, see for example [38]. Colombeau generalized functions is already a tool used to prove local well-posedness of the wave equation in space times that are of conical type. Cosmic strings are e.g. objects that can be treated within this theory. There has even been a generalization of this result to a class of locally bounded space-times with discussion of a static case and an extension to non-scalar equations. Similar applications can hence be considered using GSF, because of their better properties with respect to Colombeau theory.

6. Examples of applications

Nature is made up of different bodies, having boundaries and frequently interacting in a non-smooth way. Even the simple motion of an elastic bouncing ball seems to be more easily modeled using non-differentiable functions than classical $\mathcal{C}^2$ ones, at least if we are not interested to model the non-trivial behavior at the collision times. Therefore, the motivation to introduce a suitable kind of generalized functions formalism in a mathematical model is clear, and this would undoubtedly be of an applicable advantage, since many relevant systems are described by singular mathematical objects: non-smooth constraints, collisions between two or more bodies, motion in different or in granular media, discontinuous propagation of rays of light, even turning on the switch of an electrical circuit, to name but a few, and only in the framework of classical physics. In this section we show several applications of the theory of GSF we reviewed above.

We will *not* consider mathematical models of singular dynamical systems *at* the times when singularities occur. Indeed, this would clearly require new physical ideas, e.g. in order to consider the nonlinear behavior of objects or materials for the entire duration of the singularity. Like in every mathematical model, the correct point of view concerns J. von Neumann's *reasonably wide area* of applicability of a mathematical model, i.e. the range of phenomena where our model is expected to work (see [78, pag. 492]). Therefore, it is *not* epistemologically correct to use the theory described in the present article to deduce a physical property of our modeled systems when a singularity occurs. Stating it with a language typically used in physics, *we consider physical systems where the duration of the singularity is negligible with respect to the durations of the other phenomena that take place in the system*. Mathematically, this means to consider as infinitesimal the duration of the singularities. As a consequence, several quantities changing during this infinitesimal interval of time have infinite derivatives. We can hence paraphrase the latter sentence saying that the amplitude (of the derivatives) of these physical quantities is much larger than all the other (finite) quantities we can estimate in the system. However, this is a logical consequence of our lacking of interest to include in our mathematical model what happens during the singularity, constructing at the same time a beautiful and sufficiently powerful mathematical model, and not because these quantities really become infinite. Thereby, it is *not* epistemologically correct to state that, e.g., if a speed is infinite at some singularity, this means that we must use relativity theory: on the contrary, relativity theory is exactly a modeling setting where infinite speeds are impossible!

On the other hand, the aforementioned "wide area" is now able to include in a single equation the dynamical properties of our modeled systems, without being forced to subdivide into cases of the type "before/after the occurrence of each singularity". Which can be considered as not reasonable in several cases, e.g. in the motion of a particle in a granular medium or of a ray of light in an optical fiber.

Finally, note that remaining far from the singularity (from the point of view of the physical interpretation), is what allow us to state that in several cases this kind of models are already experimentally validated.

Moreover, the applications we are going to present always end up with an ODE. Existence and uniqueness of the solution is therefore guaranteed by Thm. 41. Clearly, if an explicit analytic solution is possible, this is preferable, but this is a rare event, and frequently we have to opt for a numerical solution, usually simply solving the corresponding ε-wise ODE, for several values of sufficiently small ε. This mean that we are considering numerical solutions of our differential equations as empirical laboratories helping us to guess suitable properties and hence *conjecture* on the solutions. In principle, these properties *must* be justified by corresponding theorems. From this point of view, the fact that GSF share with ordinary smooth functions a lot of classical theorems (such as the intermediate value, the extreme value, the mean value, Taylor theorems, etc.) is usually of great help. For example, pictures of Heaviside's function and Dirac's delta in Fig. 3.1 are clearly obtained in the same way by numerical methods, but their properties can be fully justified by suitable theorems, see e.g. Rem. 13.(i) and (ii) or Example 15.

Finally, we already saw in Sec. 3.2 that if μ is a 1-dimensional Colombeau mollifier, and δ is the ι^b-embedding of the Dirac delta, then $\delta(x) = b\mu(bx)$ for all $x \in {}^\rho\widetilde{\mathbb{R}}$. Thereby, the Heaviside function is $H(x) = \int_c^x \delta(t)\,\mathrm{d}t = \int_{bc}^{bx} \mu(t)\,\mathrm{d}t$, for all $x \in {}^\rho\widetilde{\mathbb{R}}$ and all $c \in {}^\rho\widetilde{\mathbb{R}}$ sufficiently far from 0, i.e. such that $c < r < 0$ for some $r \in \mathbb{R}_{<0}$. If the oscillations in an infinitesimal neighborhood of 0 shown in Fig. 3.1 have no modelling meaning, one can easily implement e.g. a non-decreasing Heaviside-like function by smoothly interpolating the constant functions $y = 0$ and $y = 1$ in the intervals $(-\infty, a_\varepsilon]_{\mathbb{R}}$ and $[b_\varepsilon, +\infty)_{\mathbb{R}}$, where $a = [a_\varepsilon] < 0 < b = [b_\varepsilon]$ are chosen depending on the model requirements.

By looking at the examples of this section, some readers may argue that it is already know how to solve these problems the classical way. However, we can reply as follows:

(i) No other mathematical theory of GF allows one to write non linear operations such as those presented in the following, e.g. see (6.1), (6.9), the *general* Euler-Lagrange equation (6.4) *without limitations* thanks to the closure with respect to composition, i.e. Thm. 14.

(ii) If one thinks that treating informally GF suffices for applications, the formal calculations starting from $H \cdot H = H$ and arriving at $\delta = 0$ we presented in Sec. 1 are an embarrassing drawback that needs a clear solution.

(iii) We already tried to convince the reader that the natural idea to insert a new parameter $\varepsilon \in (0, 1]$, regularizing the singularities and then taking $\varepsilon \to 0^+$ sometimes does not work because the limit may not exist in the class of smooth functions.

(iv) The lacking of an intuitively clear theory of infinitesimal and infinite quantities represents a missing useful language to construct simple models, as clearly stated by V.I. Arnol'd in [2], see also Sec. 1. This has also clear disadvantages in teaching.

(v) GSF theory allows us to treat GF as if they were smooth with a lot of shared properties and results. A consequence of this is the possibility to recognize rigorous and clear deduction starting from informal ones and discovering their real range of applications, frequently including new GF, see Sec. 5 for examples in this direction.

(vi) The literature recognizes that dealing with non-differentiable Lagrangians in optics as if they were smooth, see e.g. [79], leads to a theory with a lot of incorrect steps, see Sec. 6.4.

(vii) Also the authors of [11] recognize that their justifications of several steps in the treatment of finite and infinite step potential in QM is not completely clear and better mathematical deductions are needed, see Sec. 6.5.

(viii) The lacking of infinitesimal and infinite quantities in the understanding of Heisenberg uncertainty principle, see Sec. 6.6, allows one to understand this principle only at an intuitive level, and this can be surely judged as a drawback.

Summarizing these motivations, we can say that GSF theory justifies the use of several informal calculations with GF and non-smooth functions, and infinitesimal or infinite quantities, or to understand the behaviour of solution near singularities. This is already a positive feature because it justifies the freedom of applied mathematicians, physicists and engineers in this kind of calculations. It also opens the possibility to learn a rigorous mathematical theory of these notions, and this proved to be useful in several fields when the foundational problems are particularly insidious, like in QM, continuum mechanics, thermodynamics, medicine, biology, information science, economics, social sciences, and urban studies, to name but a few. It is clear that when problems get deceitful, models based on a strong theory reduce uncertain steps and allow to acquire a more secure knowledge.

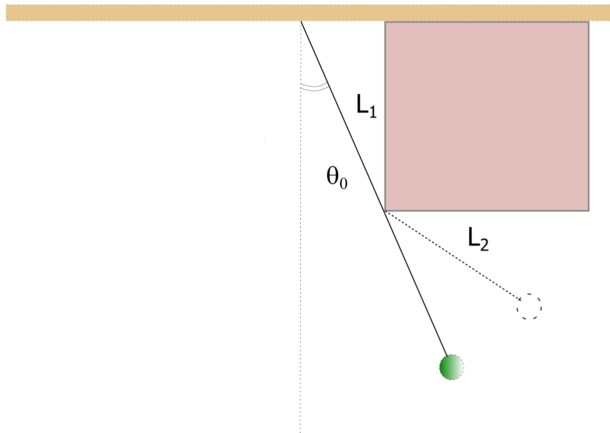

FIGURE 6.1. Oscillations of a pendulum wrapping on a parallelepiped

6.1. **Singular variable length pendulum.** As a first example, we want to study the dynamics of a pendulum with singularly variable length, e.g. because it is wrapping on a parallelepiped (see Fig. 6.1; see [57] for a similar but non-singular case, and [63] for a similar problem of jumps in the Lagrangian, but without the explicit use of infinitesimals and generalized functions).

The pendulum length function is therefore $\Lambda(\theta) = H(\theta_0 - \theta)L_1 + L_2$, where H is the (embedding of the) Heaviside function. We always assume that $L_1, L_2 \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}$ are finite and non-infinitesimal numbers. From this it follows that for all θ, $H(\theta_0 - \theta) > \frac{\mathrm{d}\rho - L_2}{L_1} \approx -\frac{L_2}{L_1}$ and hence that also $\Lambda(\theta) > \mathrm{d}\rho$ is invertible. The kinetic energy is given by:

$$T(\theta, \dot{\theta}) = \frac{1}{2} m \dot{\theta}^2 \Lambda(\theta)^2. \tag{6.1}$$

The potential energy (the zero level being the suspension point of the pendulum) is:

$$U(\theta) = -mg\Lambda(\theta)\cos\theta - mg(1 - H(\theta_0 - \theta))L_1 \cos\theta_0. \tag{6.2}$$

Let us define the Lagrangian L for this problem as

$$L(\theta, \dot{\theta}) := T(\theta, \dot{\theta}) - U(\theta). \tag{6.3}$$

The equation of motion is assumed to satisfy the Euler–Lagrange equation, see also [24], and can be written as:

$$\frac{\partial L}{\partial \theta} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\theta}}. \tag{6.4}$$

Thereby

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\theta}} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{\theta}} \left(\frac{1}{2} m \dot{\theta}^2 \Lambda(\theta)^2 \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(m \dot{\theta} \Lambda(\theta)^2 \right) = m\Lambda(\theta)^2 \ddot{\theta} + 2m\dot{\theta}\Lambda(\theta)\dot{\Lambda}(\theta), \tag{6.5}$$

where $\dot{\Lambda}(\theta) := \frac{\mathrm{d}}{\mathrm{d}t}\Lambda(\theta(t))$. From (6.2), the left side of the Euler–Lagrange equation (6.4) reduces to

$$\frac{\partial L}{\partial \theta} = \frac{\partial T}{\partial \theta} + \frac{\partial(-U)}{\partial \theta} = m\dot{\theta}^2\Lambda(\theta)\Lambda'(\theta) + mg\Lambda'(\theta)\left(\cos\theta - \cos\theta_0\right) - mg\Lambda(\theta)\sin\theta, \tag{6.6}$$

where

$$\Lambda'(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}\left(H(\theta_0-\theta)L_1 + L_2\right) = -\delta(\theta_0-\theta)L_1, \tag{6.7}$$

and δ is the Dirac delta function. We then obtain the following equation of motion:

$$m\dot{\theta}^2\Lambda(\theta)\Lambda'(\theta) + mg\Lambda'(\theta)\left(\cos\theta - \cos\theta_0\right) - mg\Lambda(\theta)\sin\theta = m\Lambda(\theta)^2\ddot{\theta} + 2m\dot{\theta}\Lambda(\theta)\dot{\Lambda}(\theta). \tag{6.8}$$

Taking into account that $\dot{\Lambda}(\theta) = \Lambda'(\theta)\dot{\theta}$, we finally obtain the equation of motion for the variable length pendulum:

$$\ddot{\theta} + \dot{\theta}\frac{\dot{\Lambda}(\theta)}{\Lambda(\theta)} - g\frac{\dot{\Lambda}(\theta)}{\dot{\theta}\Lambda(\theta)^2}\left(\cos\theta - \cos\theta_0\right) + \frac{g}{\Lambda(\theta)}\sin\theta = 0. \tag{6.9}$$

Note in (6.9) the nonlinear operations on the Schwartz distribution Λ, on the GSF θ and the composition $t \mapsto \Lambda(\theta(t))$. Before showing the numerical solution of (6.9), let us consider the simplest case of the dynamics far from the singularity and that of small oscillations. The former, as we mentioned above, is the only physically meaningful one.

6.1.1. *Description far from singularity and small oscillations.* For simplicity, let us consider the simplest case $\theta_0 = 0$. Furthermore, we consider that the pendulum is initially at rest and starts its movement at $t_1 \in {}^\rho\widetilde{\mathbb{R}}$. The initial conditions we use are hence:

$$\begin{cases} \theta(t_1) = \theta_1; \\ \dot{\theta}(t_1) = 0, \end{cases} \tag{6.10}$$

with $\theta_1 < 0$. Assuming that at some time $t_3 \in {}^\rho\widetilde{\mathbb{R}}$ we have $\theta(t_3) > 0$, by the intermediate value theorem for GSF, there exists $t_2 \in {}^\rho\widetilde{\mathbb{R}}$ where we have the singularity, i.e. $\theta(t_2) = 0$ and the length of the pendulum *smoothly* (in the sharp topology) changes from $L_1 + L_2$ to L_2 after the rope touches the parallelepiped. This change happens in an infinitesimal interval, because by contradiction it is possible to prove that if $\Lambda(\theta) \in (L_2, L_1 + L_2)$, then $|\theta| \le \frac{-1}{\log \mathrm{d}\rho} \approx 0$.

Definition 48. Let x, $y \in {}^\rho\widetilde{\mathbb{R}}$. We say that x *is far from* y if $|x-y| \ge \mathrm{d}\rho^a$ for all $a \in \mathbb{R}_{>0}$. More generally, we say that x *is far from* y *with respect to the class of infinitesimals* $\mathcal{I} \subset {}^\rho\widetilde{\mathbb{R}}$, if $|x-y| \ge i$ for all $i \in \mathcal{I}$.

For example, if $|x| \ge r$ for some $r \in \mathbb{R}_{>0}$, then x is far from 0, but also the infinitesimal number $x = \frac{-1}{k \log \mathrm{d}\rho}$ ($k \in \mathbb{R}_{>0}$) is far from 0; similarly, the infinitesimal $x = \frac{-1}{k \log\log \mathrm{d}\rho}$ if far from 0 with respect to all the infinitesimals of the type $\frac{-1}{h \log \mathrm{d}\rho}$ for $h \in \mathbb{R}_{>0}$.

If θ is far from 0 and $b \ge \mathrm{d}\rho^{-a}$, $a \in \mathbb{R}_{>0}$, then $|b\theta| \ge \mathrm{d}\rho^{-a}|\theta| \ge \mathrm{d}\rho^{-a/2} \ge 1$. Therefore, $H(-\theta) \in \{0,1\}$ and hence $\dot{\Lambda}(\theta(t)) = 0$. Equation (6.9) becomes

$$\theta(t) \text{ is far from } 0 \ \Rightarrow \begin{cases} \ddot{\theta} + \frac{g}{L_1+L_2}\sin\theta(t) = 0 & \text{if } \theta(t) < 0, \\ \ddot{\theta} + \frac{g}{L_2}\sin\theta(t) = 0 & \text{if } \theta(t) > 0. \end{cases} \tag{6.11}$$

If we assume that $\theta(t_1) = \theta_1 < 0$ and $\theta(t_3) > 0$ are far from 0, the sharp continuity of θ yields the existence of δ_1, $\delta_3 \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}$ such that

$$\begin{aligned}
&\forall t \in [t_1, t_1 + \delta_1) \cup (t_3 - \delta_3, t_3] : \ \theta(t) \text{ is far from } 0 \\
&\forall t \in [t_1, t_1 + \delta_1) : \ \theta(t) < 0 \\
&\forall t \in (t_3 - \delta_3, t_3] : \ \theta(t) > 0
\end{aligned} \tag{6.12}$$

(and hence $t_2 \notin [t_1, t_1 + \delta_1) \cup (t_3 - \delta_3, t_3]$ because $\theta(t_2) = 0$). Assuming that t_1, t_3 are far from t_2, without loss of generality we can also assume to have taken δ_i so small that also $t_1 + \delta_1$ and $t_3 - \delta_3$ are far from t_2.

We now employ the non Archimedean framework of ${}^{\rho}\widetilde{\mathbb{R}}$ in order to formally consider small oscillations, i.e. $\theta_1 \approx 0$. We first note that we cannot only assume θ_1 infinitesimal, because if θ_1 is not far from 0 then our solution will not be physically meaningful. However, we already have seen that we can take θ_1 far from 0 and infinitesimal at the same time, e.g. $\theta_1 = \frac{-1}{\log \mathrm{d}\rho}$. In other words, θ_1 is a "large" infinitesimal with respect to all the infinitesimals of the form $\mathrm{d}\rho^a$. Let ϑ_1, ϑ_3 be the solution of the linearized problems

$$\begin{cases} \ddot{\vartheta}_1 + \frac{g}{L_1+L_2}\vartheta_1 = 0, & t_1 \le t < t_1 + \delta_1 \\ \dot{\vartheta}_1(t_1) = 0, \ \vartheta_1(t_1) = \theta_1 \end{cases} \tag{6.13}$$

$$\begin{cases} \ddot{\vartheta}_3 + \frac{g}{L_2}\vartheta_3 = 0, & t_3 - \delta_3 < t \le t_3 \\ \dot{\vartheta}_1(t_3) = \dot{\theta}(t_3), \ \vartheta_1(t_3) = \theta(t_3), \end{cases}$$

i.e. $\vartheta_1(t) = \theta_1 \cos\left(\omega(t - t_1)\right)$, $\omega := \sqrt{\frac{g}{L_1+L_2}}$, and $\vartheta_3(t) = \theta(t_3)\cos\left(\omega'(t_3 - t)\right) - \frac{\dot{\theta}(t_3)}{\omega'}\sin\left(\omega'(t_3 - t)\right)$, $\omega' = \sqrt{\frac{g}{L_2}}$. We want to show that $\theta(t) \approx \vartheta_i(t)$ at least in an infinitesimal neighborhood of t_1 and t_3 exactly because $\theta_1 \approx 0$. For simplicity, we proceed only for ϑ_1, the other case being similar. For any $t \in [t_1, t_1 + \delta_1)$, we have that $\theta(t) < 0$ is far from 0 from (6.12), and hence $\ddot{\theta} + \frac{g}{L_1+L_2}\sin\theta(t) = 0$ from (6.11). Recalling the initial conditions, we obtain

$$\theta(t_1 + h) - \theta_1 = -\omega^2 \int_{t_1}^{t_1+h} \sin\theta(s)\,\mathrm{d}s \quad \forall h \in (0, \delta_1).$$

Similarly, integrating (6.13), we get

$$\vartheta_1(t_1 + h) - \theta_1 = -\omega^2 \int_{t_1}^{t_1+h} \vartheta(s)\,\mathrm{d}s \quad \forall h \in (0, \delta_1).$$

Using Taylor Thm. 35 at t_1 with increment h of these integral GSF, we obtain

$$\begin{aligned}
\theta(t_1 + h) - \vartheta(t_1 + h) &= -\omega^2 \left\{\sin\theta_1 - \theta_1 + h\cos\theta_1 \cdot \dot{\theta}(t_1) - h\dot{\vartheta}_1(t_1) + h^2 R(h)\right\} = \\
&= -\omega^2 \left\{\sin\theta_1 - \theta_1 + h^2 R(h)\right\},
\end{aligned}$$

where $R(-)$ is a suitable GSF. Thereby, $\theta(t_1 + h) - \vartheta(t_1 + h) \approx -\omega^2 h^2 R(h) \approx 0$ for all $h \approx 0$ sufficiently small because $\sin\theta_1 \approx \theta_1$ since $\theta_1 \approx 0$.

Since each $t \in [t_1, t_1 + \delta_1) \cup (t_3 - \delta_3, t_3]$ is far from t_2, we can also formally join the two solutions ϑ_i using the Heaviside's function:

$$\theta(t) \approx \vartheta_1(t) + H(t_2 - t)\left(\vartheta_3(t) - \vartheta_1(t)\right) \quad \forall t \in [t_1, t_1 + h) \cup (t_3 - h, t_3]. \tag{6.14}$$

For the epistemological motivations previously stated, this infinitesimal approximation cannot be extended to a neighborhood of t_2.

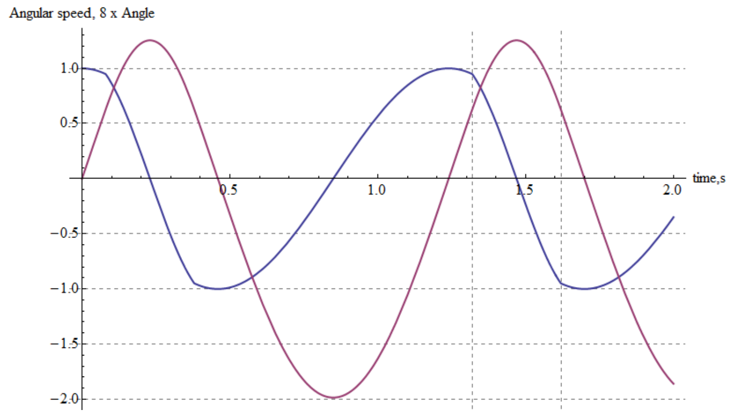

FIGURE 6.2. 8 times re-scaled solution (violet line) in radians and its derivative in rad/s (at $\theta = \theta_0 = \pi/40$ rad we can see a corner point). Parameters used: $L_1 = 0.4$ m, $L_2 = 0.2$ m, $g = 9.8$ m/s^2.

We close this section noting that all these deductions can be repeated using any GSF $H \in {}^{\rho}\mathcal{GC}^{\infty}({}^{\rho}\widetilde{\mathbb{R}}, {}^{\rho}\widetilde{\mathbb{R}})$ satisfying for all x far from zero $H(x) = 1$ if $x > 0$ and $H(x) = 0$ if $x < 0$.

6.1.2. *Numerical Solution.* The numerical solution of equation (6.9) has been computed using Mathematica Solver NDSolve (see [80]). Initial conditions we used are:

$$\begin{cases} \theta(0) = 0 \text{ rad}, \\ \dot{\theta}(0) = 1 \text{ rad/s}. \end{cases} \tag{6.15}$$

The graph of $\theta(t)$, and its derivative $\dot{\theta}(t)$, based on the Mathematica definitions of $H(x)$ and $\delta(x)$ (see [81]) are shown in Figure 6.2.

In Figure 6.3, we show the second derivative graph. Directly from (6.9) and (6.7) we can prove that when $\theta(t) = \theta_0$, $\ddot{\theta}(t)$ is an infinite number and hence $\dot{\theta}(t)$ has a corner point. Because of the classical Mathematica implementation of H and δ we can say that these graphs represent the solution far from the singularities.

6.2. **Oscillations damped by two media.** The second example concerns oscillations of a pendulum in the interface of two media. Since we are not interested at the dynamics occurring at singular times (i.e. at the changing of the medium), this can be considered only a toy model approximating the case of a very small but sufficiently heavy moving particle.

We hence want to model the system employing a “jump” in the damping coefficient β, i.e. a finite change occurring in an infinitesimal interval of time, see Fig. 6.4. Since the frictional forces acting in this case are not conservative, it is well-known that the Euler-Lagrange equations cannot be assumed to describe the dynamics of the system and we have to use the D’Alembert principle, see [24] for details.

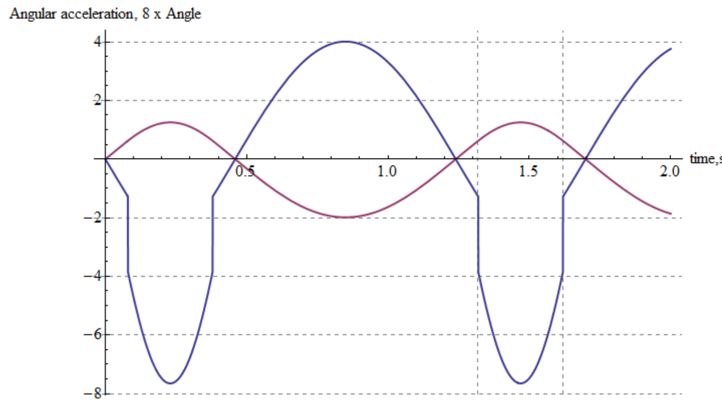

FIGURE 6.3. 8 times re-scaled solution in radians (violet line) and its second derivative in rad/s^2

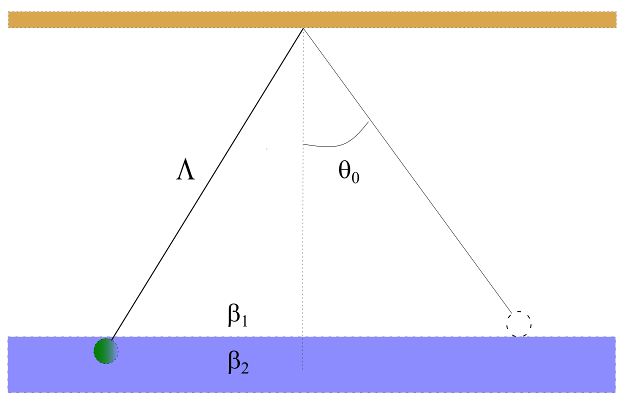

FIGURE 6.4. Simple pendulum moving in two media

The kinetic energy is given by:

$$T(\dot{\theta}) = \frac{1}{2} m \dot{\theta}^2 \Lambda^2, \tag{6.16}$$

and the potential energy (the zero level is the suspension point of the pendulum) is:

$$U(\theta) = -mg\Lambda \cos\theta. \tag{6.17}$$

In case of fluid resistance proportional to the velocity, we can introduce the generalized forces Q as:

$$Q(\dot{\theta}) = -r\Lambda^2 \dot{\theta}, \tag{6.18}$$

where r is a proportional coefficient depending on the media. Let's define the Lagrangian L as

$$L(\theta, \dot{\theta}) := T(\dot{\theta}) - U(\theta). \tag{6.19}$$

We hence assume that the equation of motion for this non-conservative system is given by the D'Alembert's principle, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = Q. \tag{6.20}$$

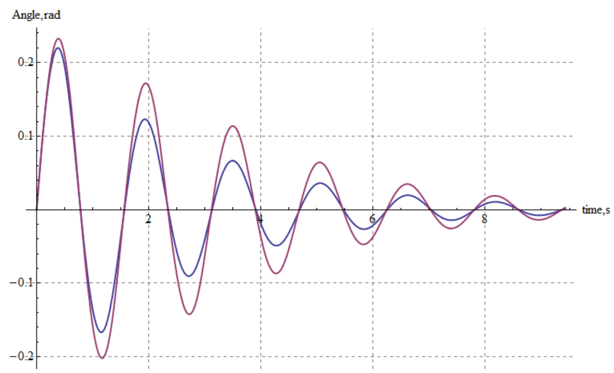

FIGURE 6.5. Solution θ of (6.22) (blue line). For comparison, the violet line is the case $\beta = \text{const.} = \beta_1$. Used parameters: $\beta_1 = 0.0064$ (air), $\beta_2 = 0.3859$ (water), $\theta_0 = \pi/40$ rad, $\Lambda = 0.6$ m, $g = 9.8$ m/s^2.

Inserting (6.16), (6.17) and (6.18) into (6.20) we obtain the following equation of motion:

$$m\Lambda^2\ddot{\theta} + mg\Lambda\sin\theta = -r\Lambda^2\dot{\theta}. \tag{6.21}$$

By introducing the damping coefficient $\beta(\theta) := r(\theta)/(2m)$ (we clearly assume that the mass $m \in {}^\rho\widetilde{\mathbb{R}} > 0$ is invertible) we obtain the classical form of the equation of motion for damped oscillations:

$$\ddot{\theta} + 2\beta(\theta)\dot{\theta} + \frac{g\sin\theta}{\Lambda} = 0. \tag{6.22}$$

If the pendulum crosses the boundary between two media with damping coefficients β_1 and β_2, we can model the system using the Heaviside function H:

$$\beta(\theta) = \beta_1 + \left(H(\theta+\theta_0) - H(\theta-\theta_0)\right)(\beta_2 - \beta_1), \tag{6.23}$$

where $\theta = \pm\theta_0$ are the angles at which we have the changing of the medium (singularities).

The numerical solution of (6.22) with β defined by (6.23) and initial conditions (6.15) is presented in Fig. 6.5. The numerical solution has been computed using Mathematica Solver NDSolve, but with an implementation of the Heaviside's function H corresponding to Thm. 12, i.e. as represented in Fig. 3.1.

We also include the graphs of the angular frequency $\dot{\theta}$ (which shows corner points) and of the angular acceleration $\ddot{\theta}$ (which shows "jumps", i.e. infinite derivatives at singular times, as we can directly see from (6.22) and (6.23)).

6.3. **Non linear strain-stress model.** In this section, we want to show how to construct a mathematical model starting from an empirical function (the strain-stress relation for a steel sample) and representing it as a GSF. Starting from Newton's second law, we hence arrive at a single nonlinear equation describing the behaviour of the steel sample. Since the empirical function is not differentiable at the end of the linear part, the use of GSF is therefore essential.

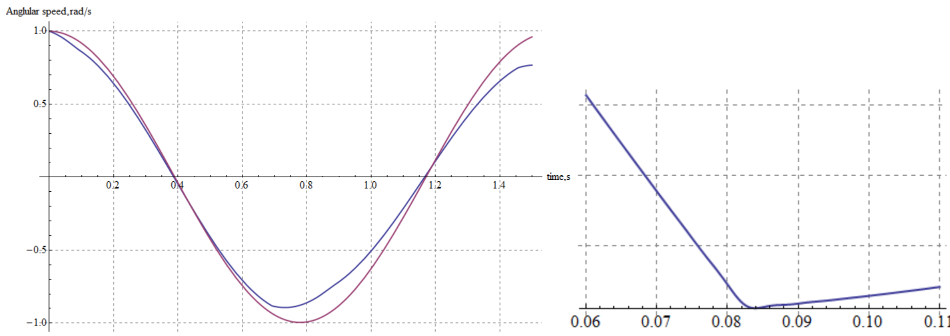

FIGURE 6.6. First derivative $\dot{\theta}$ of the solution of (6.22) (blue line). The case with $\beta = \text{const} = \beta_1$ is also shown for comparison (violet line). Note the corner points at the singular moments, for example at $t = 0.083$ s (scaled in the right figure).

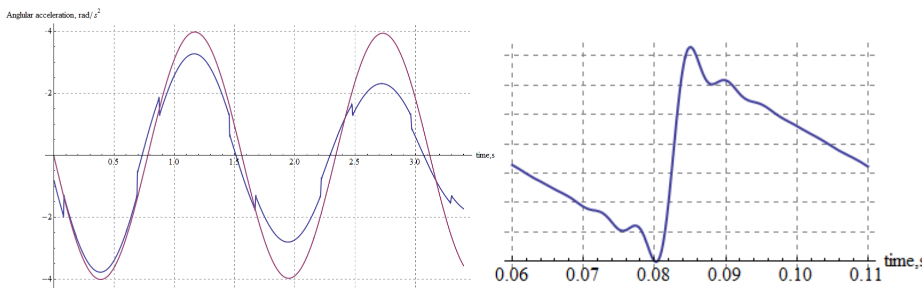

FIGURE 6.7. Second derivative $\ddot{\theta}$ of the solution of (6.22) (blue line). The case with $\beta = \text{const} = \beta_1$ is also shown for comparison (violet line). Note the "jumps" at the singular moments, for example at $t = 0.083$ s (scaled in the right figure). The infinitesimal oscillations are caused by the embedding as GSF of the Heaviside function.

The strain-stress curve we consider is shown in Fig. 6.8.
We recall that stress can be defined as $\sigma = \frac{F}{S_0}$, where F is the force applied to the sample, and S_0 is the initial cross-section of the cylindrical sample. The strain ε is usually introduced as $\varepsilon = \frac{L-L_0}{L_0}$, where L_0 is the unstressed length and L is the length after force application. In order to reproduce the experimental dependence of Fig. 6.8 we considered the parameters $d = 0.37\,\text{mm}$ for the diameter of the steel cilinder, and $L_0 = 2.2\,\text{m}$ for the unstressed length of the sample. Thus, during the elastic behaviour (linear part) we have a Young's modulus $E = \frac{\sigma}{\varepsilon} = 2.13 \cdot 10^{11}\,\text{Pa}$, a stiffness $k = \frac{ES_0}{L_0} = 10423\,\frac{\text{N}}{\text{m}}$, and hence the magnitude of the linear part of the force is given by $F_1(x) = kx$. For the nonlinear part of the empirical law, we use the Mathematica built-in function NonlinearModelFit, see [82]. The result is shown in Fig. 6.9.
The resulting expression is

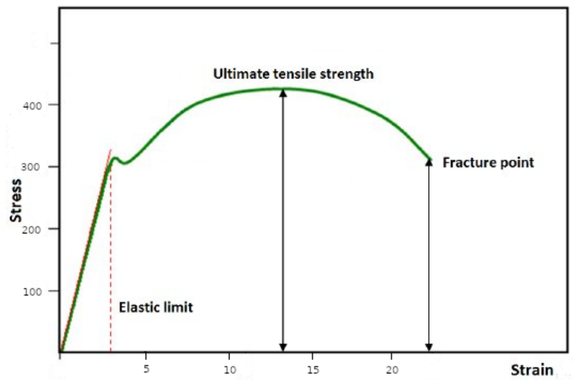

FIGURE 6.8. Strain-stress empirical model of the steel (see [62]).

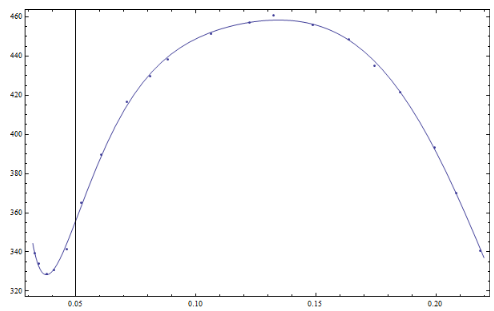

FIGURE 6.9. Non-linear part modelling of the strain-stress curve using the NonlinearModelFit. The dots are values extracted from the original data.

$a_1 = 1.5 \cdot 10^3$	$a_5 = -9.9 \cdot 10^4$	$a_9 = -1.8 \cdot 10^9$
$a_2 = 3.9$	$a_6 = 2.8 \cdot 10^6$	$a_{10} = 4.8 \cdot 10^9$
$a_3 = 3.0$	$a_7 = -4.4 \cdot 10^7$	$a_{11} = -5.1 \cdot 10^9$
$a_4 = 1.0 \cdot 10^2$	$a_8 = 3.8 \cdot 10^8$	

TABLE 1. Coefficients used in non-linear part F_n of the force.

$$F_\text{n}(x) = a_1 \exp(a_2 x) + a_3 \cos(a_4 x) + a_5 x + a_6 x^2 + a_7 x^3 + a_8 x^4 + a_9 x^5 + a_{10} x^6 + a_{11} x^7, \tag{6.24}$$

where the coefficients a_k are given in Tab. 1.

Using the Heaviside function, we can hence write the force F acting on the steel sample as

$$F(x) = -F_\text{l}(x) - (F_\text{n}(x) - F_\text{l}(x))H(x - x_0), \tag{6.25}$$

and it is represented in Fig. 6.10. Note that the negative sign is due to the fact that the force is directed opposite to the elongation of the sample; moreover, $x_0 =$

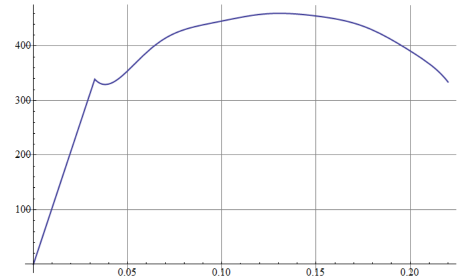

FIGURE 6.10. Recomputed strain-stress model.

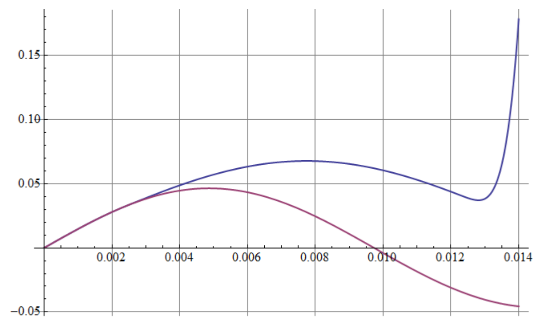

FIGURE 6.11. The solution $x(t)$ of equation 6.26 with initial conditions $x(0) = 0.0\,\mathrm{m}$,$\dot{x}(0) = 15\,\frac{\mathrm{m}}{\mathrm{s}}$ (blue line) in comparison with the solution in the linear setting for $x(0) = 0.0\,\mathrm{m}$, $\dot{x}(0) = 5\,\frac{\mathrm{m}}{\mathrm{s}}$ (violet line).

0.033 according to Fig. 6.10. Thus, the position x of the steel sample satisfies the differential equation

$$\ddot{x} - \frac{F(x)}{m} = 0. \tag{6.26}$$

Once again, note that the GSF F is nonlinear and the term $F(x(t))$ is a composition of GSF.
Far from the singularity $x = x_0$, the validity of (6.26) can also be seen using the conservation of the mechanical energy. In fact, if x is far from x_0, in the sense that $x \le x_1$ for some $x_1 \in \mathbb{R}_{<0}$, then $F(x) = -F_1(x)$ and we are in the zone of Hooke's law; we thus have the potential energy: $U(x) = \frac{kx^2}{2}$. Similarly, if $x \ge x_2$ for some $x_2 \in \mathbb{R}_{>0}$, then $F(x) = -F_{\mathrm{n}}(x)$ and $U(x) = \frac{a_1}{a_2}\exp(a_2 x) + \frac{a_3}{a_4}\sin(a_4 x) + \frac{a_5}{2}x^2 + \frac{a_6}{3}x^3 + \frac{a_7}{4}x^4 + \frac{a_8}{5}x^5 + \frac{a_9}{6}x^6 + \frac{a_{10}}{7}x^7 + \frac{a_{11}}{8}x^8$. Therefore, far from the singularity, the conservation of the mechanical energy is equivalent to (6.26).

Clearly, the nonlinear behaviour depends on the initial conditions: if $x(0) = 0.0\,\mathrm{m}$ and $\dot{x}(0) = 5\,\frac{\mathrm{m}}{\mathrm{s}}$, we remain in the linear setting, whereas for $x(0) = 0.0\,\mathrm{m}$, $\dot{x}(0) = 15\,\frac{\mathrm{m}}{\mathrm{s}}$ we enter into the nonlinear one, see Fig. 6.11.
See also [13, 64, 65, 66] for more complete models of this type in the setting of Colombeau theory.

6.4. **Discontinuous Lagrangians in optics (Snell's law derivation).** A typical example where one would like to use the usual results of calculus despite dealing with non differentiable functions, is geometrical optics at the interface of two media, where usually the Lagrangian function is not smooth. It is well-know, e.g., that the rigorous derivation of Snell's law is a paradigmatic example, see e.g. [79]. The main aim of this section is to show the features of the nonlinear calculus of GSF by deriving Snell's law for plane stratified media from the classical Fermat's principle. For example, in the following deduction, the refraction index $n(x)$ can be any GSF, e.g. the embedding of a locally integrable function. For more general versions of Snell's law in different media, see e.g. [12, 68]. For the classical deduction where the refraction index $n(x)$ and the light path $x = x(s)$ are $\mathcal{C}^2$ functions, see e.g. [47, 61].

We assume that we are considering a body $B := \langle B_\varepsilon \rangle \subseteq \mathbb{R}^3$ represented by the strongly internal set generated by the net $B_\varepsilon \subseteq \mathbb{R}^3$, and that our light path satisfies the classical Fermat principle, i.e. the path of light between two given points P, $Q \in B$ is the one which minimizes the travel time. In order to mathematically state this principle for GSF, we introduce the space of paths (see [50, 24] for details):

$$^{\rho}\mathcal{GC}^\infty_{\text{bd}}(P,Q) := \left\{ r \in {}^{\rho}\mathcal{GC}^\infty([0,1],B) \mid r(0) = P,\ r(1) = Q \right\}, \tag{6.27}$$

and the travel time functional

$$T[r] := \frac{1}{c}\int_0^1 n(r)\,\mathrm{d}r := \frac{1}{c}\int_0^1 n(r(s))\,|\dot{r}(s)|\,\mathrm{d}s \quad \forall r \in {}^{\rho}\mathcal{GC}^\infty_{\text{bd}}(P,Q). \tag{6.28}$$

As usual, c is the speed of light in vacuum, and $n \in {}^{\rho}\mathcal{GC}^\infty(B, {}^{\rho}\widetilde{\mathbb{R}}_{\ge 0})$ is the refraction index of the media B we are considering. The Fermat principle hence implies that the light path $r \in {}^{\rho}\mathcal{GC}^\infty_{\text{bd}}(P,Q)$ is a weak extremal of the travel time functional $T[-]$, i.e. it satisfies

$$\delta T(r;h) := \left.\frac{\mathrm{d}}{\mathrm{d}x} T[r + xh]\right|_{x=0} = 0 \quad \forall h \in {}^{\rho}\mathcal{GC}^\infty_{\text{bd}}(0,0). \tag{6.29}$$

Note that, since in (6.27) we consider only paths $r \in {}^{\rho}\mathcal{GC}^\infty([0,1],B)$ valued in the strongly internal set $B = \langle B_\varepsilon \rangle$, Thm. 5 implies

$$\forall h \in {}^{\rho}\mathcal{GC}^\infty_{\text{bd}}(0,0)\, \exists \delta \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}\, \forall x \in (-\delta,\delta) :\ r + xh \in {}^{\rho}\mathcal{GC}^\infty([0,1],B),$$

and therefore, it is correct to consider the derivative in (6.29). Physically this means that we are considering only paths which lay completely inside the body $B = \langle B_\varepsilon \rangle$. The weak extremal condition (6.29) is equivalent to the Euler-Lagrange equations (see [50]) for the Lagrangian $(r,v) \mapsto L(r,v) := n(r)\sqrt{v \cdot v} \in {}^{\rho}\mathcal{GC}^\infty(B \times {}^{\rho}\widetilde{\mathbb{R}}^3, {}^{\rho}\widetilde{\mathbb{R}})$. We use the notation $L(r,v) = L(r_1,r_2,r_3,v_1,v_2,v_3)$ for the variables of L. We also explicitly note the nonlinear operations in this Lagrangian, and the composition $n(r(s))$ of GSF. We use the customary notations $\vec{v}(s) := \frac{\mathrm{d}r}{\mathrm{d}s}(s) \in {}^{\rho}\widetilde{\mathbb{R}}^3$, $v(s) := |\vec{v}(s)| \in {}^{\rho}\widetilde{\mathbb{R}}$, and $L[r](s) := L(r(s),\vec{v}(s))$. We hence get

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial L}{\partial v_j}[r](s)\right) = \frac{\partial L}{\partial r_j}[r](s), \quad \forall j = 1,2,3,\ \forall s \in [0,1]. \tag{6.30}$$

We always assume that the frame of reference is chosen so that the light path satisfies $v(s) \in {}^\rho\widetilde{\mathbb{R}}_{>0}$. Calculating the derivatives $\partial L/\partial v_j$ and $\partial L/\partial r_j$ in Euler-Lagrange equations, we get

$$\frac{\partial L}{\partial v_j}(r,v) = n(r)\frac{v_j}{\sqrt{v\cdot v}}$$
$$\frac{\partial L}{\partial r_j}(r,v) = \frac{\partial n}{\partial r_j}(r)\sqrt{v\cdot v},$$

for all $(r,v) \in B \times {}^\rho\widetilde{\mathbb{R}}^3$ and all $j = 1,2,3$. Substituting in (6.30), we obtain the eikonal equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n(r(s))\frac{\vec{v}(s)}{v(s)}\right) = \nabla n(r(s))v(s), \quad \forall s \in [0,1]. \tag{6.31}$$

We now consider the case of a plane stratified media, i.e. where n changes only along one direction $\vec{k} \in {}^\rho\widetilde{\mathbb{R}}^3$, $|\vec{k}| = 1$, so that

$$\nabla n(r(s)) \parallel \vec{k} \quad \forall s \in [0,1]. \tag{6.32}$$

For simplicity, where it is clear, we omit the evaluation at s. Thus, the cross product of the two vectors $\nabla n(r)$ and $\vec{k}$ is $\vec{k} \times \nabla n(r) = 0 = \vec{k} \times \nabla n(r)v$. Using (6.31) we get $\vec{k} \times \frac{\mathrm{d}}{\mathrm{d}s}(n(r)\frac{\vec{v}}{v}) = 0$, and hence $\frac{\mathrm{d}}{\mathrm{d}s}(\vec{k} \times n(r(s))\frac{\vec{v}(s)}{v(s)}) = 0$ for all $s \in [0,1]$, i.e. the function $s \in [0,1] \mapsto \vec{k} \times n(r(s))\frac{\vec{v}(s)}{v(s)} \in {}^\rho\widetilde{\mathbb{R}}^3$ is a constant $\vec{C} \in {}^\rho\widetilde{\mathbb{R}}^3$. Taking its magnitude $C = |\vec{C}|$

$$\forall s \in [0,1]: \ C = |\vec{k}|\,|n(r(s))|\left|\frac{\vec{v}(s)}{v(s)}\right|\sin\varphi(s),$$

where $\varphi(s)$ is the angle between $\vec{k}$ and $\vec{v}(s) = \frac{\mathrm{d}r}{\mathrm{d}s}(s)$. We proved Snell's law for GSF:

Theorem 49. *Let $B_\varepsilon \subseteq \mathbb{R}^3$, $B := \langle B_\varepsilon\rangle$, P, $Q \in \langle B_\varepsilon\rangle$, $n \in {}^\rho\mathcal{GC}^\infty(B, {}^\rho\widetilde{\mathbb{R}}_{\ge 0})$. Assume that $r \in {}^\rho\mathcal{GC}^\infty_{\mathrm{bd}}(P,Q)$ is a weak extremal of the travel time functional (6.28), i.e. it satisfies (6.29). Set $\vec{v}(s) := \frac{\mathrm{d}r}{\mathrm{d}s}(s) \in {}^\rho\widetilde{\mathbb{R}}^3$, $v(s) := |\vec{v}(s)| \in {}^\rho\widetilde{\mathbb{R}}$ and assume that $v(s) > 0$ for all $s \in [0,1]$. Then the eikonal equation (6.31) holds. Moreover, if $\nabla n(r(s)) \parallel \vec{k}$, where $\vec{k} \in {}^\rho\widetilde{\mathbb{R}}^3$, $|\vec{k}| = 1$, and $\varphi(s)$ is the angle between $\vec{k}$ and $\vec{v}(s)$,then the quantity $n(r(s)) \cdot \sin\varphi(s)$ is constant for all $s \in [0,1]$.*

6.5. **Finite and infinite step potential.** Models of quantum mechanics such as the potential well or the step potential with finite or infinite walls are clear and simple examples showing features of various quantum mechanical effects, see e.g. [11]. However, the mathematics of such models is not very clear sometimes, see again e.g. [11, pag. 34-40, pag. 68] and authors' comments about mathematical rigour. Once again, in this section we see how the formalism of GSF theory allows one to completely recover a mathematically and physically clear proof by formalizing the intuitive steps of [11, pag. 68]. We consider the step-potential problem, where the high of the potential can be any finite or infinite generalized number; a similar approach can be used for the rectangular potential wells.

In the following, we write $x \ll 0$ if $\exists r \in \mathbb{R}_{<0}: \ x < r$, and similarly for $x \gg 0$, and we simply say that x *is far from* 0. The step function potential for the one-dimensional stationary Schrodinger equation is a GSF $U \in {}^\rho\mathcal{GC}^\infty({}^\rho\widetilde{\mathbb{R}}, {}^\rho\widetilde{\mathbb{R}})$ such that

$$U(x)=\begin{cases}0 & x\ll 0,\\ U_0 & x\gg 0,\end{cases} \tag{6.33}$$

where $U_0\in{}^\rho\widetilde{\mathbb{R}}_{>0}$ is an arbitrary generalized number (finite or infinite). For example, $U(x)=H(x)\cdot U_0$ satisfies these conditions. However, as stated in [11, pag. 34], this is actually an idealized model of the potential, and we cannot say it is a physically meaningful model for infinitesimal $x\approx 0$ (similarly to what we have already seen e.g. in Sec. 6.1 for the singular variable length pendulum).

The system satisfies the stationary Schrodinger's equation

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+U(x)\right]\psi(x)=E\psi(x), \tag{6.34}$$

where $\hbar$ is the Planck's constant, $m\in\mathbb{R}_{>0}$ the mass of the particle, E the energy, and $\psi(x)$ the wave function. Using Thm. 43, we can state that there exists a $\psi\in{}^\rho\mathcal{GC}^\infty({}^\rho\widetilde{\mathbb{R}},{}^\rho\widetilde{\mathbb{R}})$ satisfying (6.34). Repeating exactly the usual calculations, for $x\ll 0$ we have:

$$\psi(x)=\frac{1}{\sqrt{k_1}}\left(A_1e^{ik_1x}+A_2e^{-ik_1x}\right), \tag{6.35}$$

where $k_1:=\sqrt{2mE/\hbar^2}$ and A_1, $A_2\in{}^\rho\widetilde{\mathbb{R}}$ are undefined constants. For $x\gg 0$, we have:

$$\psi(x)=\frac{1}{\sqrt{k_2}}\left(B_1e^{ik_2x}+B_2e^{-ik_2x}\right), \tag{6.36}$$

where $k_2=\sqrt{2m(E-V_0)/\hbar^2}$, B_1, $B_2\in{}^\rho\widetilde{\mathbb{R}}$ are undefined constants. As stated in [11, pag. 68], in order to find these constants, we need some mathematically careful steps to justify the corresponding initial conditions. Take any *standard real number* $\eta\in\mathbb{R}_{>0}$ and integrate (6.34) on $[-\eta,\eta]\subseteq{}^\rho\widetilde{\mathbb{R}}$ to get

$$\frac{\mathrm{d}\psi}{\mathrm{d}x}(\eta)-\frac{\mathrm{d}\psi}{\mathrm{d}x}(-\eta)=\frac{2m}{\hbar^2}\int_{-\eta}^{\eta}\left[U(x)-E\right]\psi(x)\,\mathrm{d}x. \tag{6.37}$$

As in [11, pag. 68], we assume that

$$U(x)-E \text{ is finite for all finite } x\in{}^\rho\widetilde{\mathbb{R}} \tag{6.38}$$

$$\frac{\mathrm{d}\psi}{\mathrm{d}x}(\eta) \text{ is finite for all } \eta\in\mathbb{R}_{>0}. \tag{6.39}$$

From (6.37) and the first of these assumptions, we obtain

$$\left|\frac{\mathrm{d}\psi}{\mathrm{d}x}(\eta)-\frac{\mathrm{d}\psi}{\mathrm{d}x}(-\eta)\right|\le\frac{4m}{\hbar^2}\cdot C\cdot\eta$$

for some $C\in\mathbb{R}_{>0}$ (coming from (6.38)), i.e.

$$\lim_{\substack{\eta\to 0^+\\ \eta\in\mathbb{R}_{>0}}}\left|\frac{\mathrm{d}\psi}{\mathrm{d}x}(\eta)-\frac{\mathrm{d}\psi}{\mathrm{d}x}(-\eta)\right|=0. \tag{6.40}$$

Similarly, from (6.39) and the fundamental theorem of calculus for GSF Thm. 25.(v), we have

$$|\psi(\eta)-\psi(0)|=\left|\int_0^\eta\frac{\mathrm{d}\psi}{\mathrm{d}x}(x)\,\mathrm{d}x\right|\le\bar{C}\cdot\eta,$$

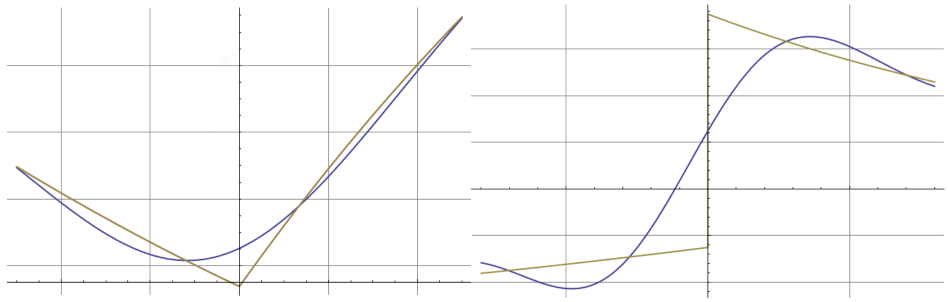

FIGURE 6.12. The fist derivative (left) and the second derivative (right) of the wave function ψ in an infinitesimal neighborhood of 0 (blue lines). Green lines are the same derivatives for $x \ll 0$ and $x \gg 0$.

for some $\bar{C} \in \mathbb{R}_{>0}$ (coming from (6.39)), and hence

$$\lim_{\substack{\eta \to 0^+ \\ \eta \in \mathbb{R}_{>0}}} |\psi(\eta) - \psi(0)| = 0. \tag{6.41}$$

Recall that from Thm. 10.(ii) and Thm. 17 it directly follows that both $\psi(x)$ and its derivative $\frac{\mathrm{d}\psi}{\mathrm{d}x}(x)$ are GSF, and hence they are continuous in the sharp topology at each point $x \in {}^{\rho}\widetilde{\mathbb{R}}$. Stated explicitly at $x = 0$, this means that

$$\exists \lim_{\substack{x \to 0 \\ x \in {}^{\rho}\widetilde{\mathbb{R}}}} \frac{\mathrm{d}\psi}{\mathrm{d}x}(x) = \frac{\mathrm{d}\psi}{\mathrm{d}x}(0),$$
$$\exists \lim_{\substack{x \to 0 \\ x \in {}^{\rho}\widetilde{\mathbb{R}}}} \psi(x) = \psi(0),$$

and these are different than (6.40) and (6.41), where $\eta \in \mathbb{R}_{>0}$. Indeed, balls $B_\eta(c) \subseteq {}^{\rho}\widetilde{\mathbb{R}}$, for radii $\eta \in \mathbb{R}_{>0}$, generate a different topology on ${}^{\rho}\widetilde{\mathbb{R}}$ (called *Fermat topology*, see e.g. [31, 32]). See Fig. 6.12 for an intuitive diagram of the solution ψ in an infinitesimal neighborhood of $x = 0$: whereas $\frac{\mathrm{d}\psi}{\mathrm{d}x}(\eta)$ is continuous for $\eta \to 0$, $\eta \in \mathbb{R}_{>0}$, it is well-known (see [11]) that the same property does not hold for the second derivative. In Fig. 6.12, the green lines represents the solution for $x \ll 0$ or $x \gg 0$, and the blue one the GSF function $\psi = [\psi_\varepsilon(-)]$ (we actually represented ψ_ε for ε sufficiently small); we therefore have to think as infinitesimal the differences between blue and green lines, and hence as infinite the second derivative at $x = 0$.

From (6.35), (6.36) and (6.40), (6.41) we obtain that the constants are uniquely determined by the system

$$\begin{cases} (A_1 + A_2) = (B_1 + B_2) \\ k_1 (A_1 - A_2) = k_2 (B_1 - B_2) \end{cases} \tag{6.42}$$

6.6. **Heisenberg uncertainty principle.** We close this section of applications by mentioning how we can use infinitesimal and infinite numbers and GSF theory to fully justify the most frequent example of the Heisenberg uncertainty principle.

We do not have sufficient space here to present a complete list of result about the so called *hyperfinite Fourier transform* (see [60]), but we can surely present the main ideas:

(i) Since in the ring ${}^{\rho}\widetilde{\mathbb{R}}$ we have infinite numbers $k \in {}^{\rho}\widetilde{\mathbb{R}}_{>0}$, e.g. $k = \mathrm{d}\rho^{-1}$, and since every GSF is always integrable on a functionally compact set of the form $K := [-k, k]^n \subseteq {}^{\rho}\widetilde{\mathbb{R}}^n$, we can simply define the *hyperfinite Fourier transform* $\mathcal{F}_k(f)$ of any $f \in {}^{\rho}\mathcal{GC}^{\infty}(K, {}^{\rho}\widetilde{\mathbb{C}})$ as

$$\mathcal{F}_k(f)(\omega) := \int_K f(x) e^{-ix\cdot\omega} \, \mathrm{d}x = \int_{-k}^{k} \mathrm{d}x_1 \ldots \int_{-k}^{k} f(x_1, \ldots, x_n) e^{-ix\cdot\omega} \, \mathrm{d}x_n. \quad (6.43)$$

(ii) The main feature of this transform is that, despite the fact that essentially all the usual classical properties of the Fourier transform can be proved for $\mathcal{F}_k(f)$, it is well-defined for *all* GSF $f \in {}^{\rho}\mathcal{GC}^{\infty}(K, {}^{\rho}\widetilde{\mathbb{C}})$, even if they are not of tempered type. Clearly, this allows one to use the Fourier method to find non-tempered solutions of differential equations. For example, let $f(x) = e^x$ for all $|x| \le k$, where $k := -\log(\mathrm{d}\rho)$. The hyperfinite Fourier transform $\mathcal{F}_k$ of f is

$$\mathcal{F}_k(f)(\omega) = \frac{1}{1 - i\omega}\left(\frac{\mathrm{d}\rho^{i\omega}}{\mathrm{d}\rho} - \frac{\mathrm{d}\rho}{\mathrm{d}\rho^{i\omega}}\right) \quad \forall \omega \in {}^{\rho}\widetilde{\mathbb{R}}.$$

Therefore, $\mathcal{F}_k(f)(\omega)$ is always an infinite complex number for all finite numbers ω and hence the non-Archimedean language is essential here.

(iii) The set $\mathrm{supp}(f) := \overline{\{x \in X \mid |f(x)| > 0\}}$, where $\overline{(\cdot)}$ denotes the relative closure in X with respect to the sharp topology, is called the *support* of f. Let $H \Subset_{\mathrm{f}} {}^{\rho}\widetilde{\mathbb{R}}^n$ be a functionally compact set (see Def. 30), we say that $f \in {}^{\rho}\mathcal{GD}(H)$ if $f \in {}^{\rho}\mathcal{GC}^{\infty}({}^{\rho}\widetilde{\mathbb{R}}^n, {}^{\rho}\widetilde{\mathbb{C}})$ and $\mathrm{supp}(f) \subseteq H$. Such an f is called *compactly supported*.

We can now state the uncertainty principle (see [59] for the proof):

Theorem 50. *If* $\psi \in {}^{\rho}\mathcal{GD}({}^{\rho}\widetilde{\mathbb{R}})$, *then* $\left(\int x^2 |\psi(x)|^2 \, \mathrm{d}x\right)\left(\int \omega^2 |\mathcal{F}(\psi)(\omega)|^2 \, \mathrm{d}\omega\right) \ge \frac{1}{4}\|\psi\|_2 \|\mathcal{F}(\psi)\|_2$.

On the contrary with respect the classical formulation in $L^2(\mathbb{R})$ of the uncertainty principle, in Thm. 50 we can e.g. consider $\psi = \delta \in {}^{\rho}\mathcal{GD}({}^{\rho}\widetilde{\mathbb{R}})$, and we have

$$\int x^2 \delta(x)^2 \, \mathrm{d}x = \left[\int_{-1}^{1} x^2 b_\varepsilon^2 \mu_\varepsilon(b_\varepsilon x)^2 \, \mathrm{d}x\right]$$

where $\mu(x) = [\mu_\varepsilon(x_\varepsilon)]$ is a Colombeau mollifier and $b = [b_\varepsilon] \in {}^{\rho}\widetilde{\mathbb{R}}$ satisfies $b \ge \mathrm{d}\rho^{-a}$ for some $a \in \mathbb{R}_{>0}$ (see embedding Thm. 12). Since normalizing the function $\varepsilon \mapsto b_\varepsilon^2 \mu_\varepsilon(b_\varepsilon x)^2$ we get an approximate identity, we have $\lim_{\varepsilon \to 0^+} \int_{-1}^{1} x^2 b_\varepsilon^2 \mu_\varepsilon(b_\varepsilon x)^2 \, \mathrm{d}x = 0$, and hence $\int x^2 \delta(x)^2 \, \mathrm{d}x \approx 0$ is an infinitesimal. The uncertainty principle Thm. 50 implies that it is an invertible infinitesimal. Considering the HFT $\mathbb{1} := \mathcal{F}(\delta)$, we have

$$\int \omega^2 \mathbb{1}(\omega)^2 \, \mathrm{d}\omega \ge \int_{-r}^{r} \omega^2 \, \mathrm{d}\omega = 2\frac{r^3}{3} \quad \forall r \in \mathbb{R}_{>0}.$$

Consequently, $\int \omega^2 \mathbb{1}(\omega)^2 \, \mathrm{d}\omega$ is an infinite number.

7. Conclusions

In all the presented examples, the model describes some kind of singular dynamical system including abrupt changes, impulsive stimuli, nonlinear discontinuities, infinite barriers, etc. This kind of problems are ubiquitous in applied mathematics, essentially because the real world is made of different bodies, having boundaries and frequently interacting in a non-smooth way. In constructing a model for these systems is hence important to achieve mathematical simplicity but, at the same time, a physical reasonably high faithfulness of description.

On the one hand, the use of infinitesimal and infinite numbers has always been a method to simplify a given problem. Unfortunately, frequently this technique remains only informal, using "sufficiently small quantities" or "taking the limit for $\varepsilon \to 0$", and then transforming approximated equalities into true ones. As motivational thoughts, these remain wonderful methods. In our examples, we tried to show that a corresponding simple and intuitively clear mathematical theory of these infinitesimal and infinite quantities is possible. Surprisingly, this theory allows one to arrive at very similar, but clear and rigorous, thoughts. Therefore, the risk of doing mistakes is quite lower, and its teaching is also way more clear.

On the other hand, physical systems with singularities are naturally represented by non-smooth functions. We presented a theory that allows one to deal with such functions as if they were smooth, thanks to a lot of properties that GSF share with ordinary smooth functions. This is as generalized functions are still informally used in physics and engineering, despite the fact that Schwartz theory of distributions is quite old nowadays. Using GSF theory, we can therefore state that the searched mathematical simplicity in models of singular systems, possibly with a clear use of infinitesimal or infinite quantities, is really achievable.

Declarations.

Author contributions. A. Bryzgalov developed Sec. 6.
K. Islami developed Sec. 5 and some examples in Sec. 6.
P. Giordano coordinated the development and formally checked and fixed all the results.

Conflict of interest. The authors declare that they have no relevant financial or non-financial interests to disclose.

Ethical approval. We certify that this manuscript is original and has not been published and will not be submitted elsewhere for publication while being considered by Nonlinear Dynamics. And the study is not split up into several parts to increase the quantity of submissions and submitted to various journals or to one journal over time. No data have been fabricated or manipulated (including images) to support our conclusions. No data, text, or theories by others are presented as if they were our own. And authors whose names appear on the submission have contributed sufficiently to the scientific work and therefore share collective responsibility and accountability for the results. Finally, this article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Albeverio, S., Fenstad, J.E., Høegh-Krohn, R., Lindstrøm, T., *Nonstandard Methods in Stochastic Analysis and Mathematical Physics* (Pure and Applied Mathematics. Academic Press, 1988, 2nd ed., Dover, 2009).

[2] Arnol´d, V.I., Huygens and Barrow, Newton and Hooke. Pioneers in mathematical analysis and catastrophe theory from evolvents to quasicrystals. Translated from the Russian by Eric J. F. Primrose. Birkhäuser Verlag, Basel, 1990.

[3] Bell, J.L., *A Primer of Infinitesimal Analysis*, Cambridge University Press, 1998.

[4] Benci, V., Di Nasso, M., A Ring Homomorphism is Enough to Get Nonstandard Analysis, Bull. Belg. Math. Soc. Simon Stevin 10, 481–490, 2003. 5.

[5] Benci, V., Di Nasso, M., A Purely Algebraic Characterization of the Hyperreal Numbers, Proc. Amer. Math. Soc. 133 (9), 2501–05 (2005).

[6] Brazda, K., V. de Hoop, M., Hoermann G., Variational formulation of the earth's elastic-gravitational deformations under low regularity conditions, 2017. See https://arxiv.org/abs/1702.04741

[7] Brogliato, B., *Nonsmooth impacts mechanics. Models, dynamics and control.* Springer-Verlag, 1996.

[8] Brogliato, B., *Impacts in mechanical systems. Analysis and modelling.* Springer-Verlag, 2000.

[9] Burridge, R., Keller, J.B., Peeling, Slipping and Cracking-Some One-Dimensional Free-Boundary Problems in Mechanics, SIAM Review, Vol. 20, No. 1, 31-61, 1978, DOI:10.1137/1020003.

[10] Champion, R., Champion, W.L., Departure from linear mechanical behaviour of a helical spring, Mathematical and Computer Modelling 53, 915-926, 2011.

[11] Cohen-Tannoudji, C., Diu, B., Laloe, F., *Quantum Mechanics*, Vol. 1, 1st Edition, Wiley, 1991.

[12] Coleman, C.J., On the generalization of Snell's law, Radio Science, VOL. 39, RS2005, doi:10.1029/2003RS003002, 2004.

[13] Colombeau, J.F., *New generalized functions and multiplication of distributions.* North-Holland, Amsterdam, 1984.

[14] Conway, J.H., *On Numbers and Games*, Number 6 in LMS Monographs (Academic Press, London & New York, 1976).

[15] Crandall, M.G., Lions, P.L. Viscosity solutions of Hamilton-Jacobi equations. Trans. Am. Math. Soc., 277, 1-42, 1983.

[16] Cutland, N., Kessler, C., Kopp, E., Ross, D., On Cauchy's notion of infinitesimal. *British J. Philos. Sci.*, 39(3):375–378, 1988.

[17] da Silva, G.T., Martins, R.M., Dynamics and stability of non-smooth dynamical systems with two switches, Nonlinear Dyn (2022) 108:3157–3184, https://doi.org/10.1007/s11071-022-07384-9.

[18] Deguchi, H., Oberguggenberger, M., Propagation of Singularities for Generalized Solutions to Wave Equations with Discontinuous Coefficients, SIAM Journal on Mathematical Analysis, Vol. 48, No. 1, 397-442, 2016, DOI:10.1137/15M1032661.

[19] Dirac, P.A.M., *General Theory of Relativity*, John Wiley and Sons, 1975.

[20] Dirac, P.A.M., *The Principles of Quantum Mechanics*, Clarendon Press, Oxford University Press, Oxford, 2010.

[21] Ehrlich, P., Contemporary Infinitesimalist Theories of Continua and their late 19th- and early 20th-century forerunners, arXiv:1808.03345 [math.HO].

[22] Einstein, A., *Investigations on the Theory of the Brownian Movement*, Dover, 1926.

[23] Erlacher, E., Grosser, M., Ordinary Differential Equations in Algebras of Generalized Functions, in Pseudo-Differential Operators, Generalized Functions and Asymptotics, S. Molahajloo, S. Pilipović, J. Toft, M. W. Wong eds, Operator Theory: Advances and Applications Volume 231, 2013, 253–270.

[24] Frederico, G.S.F., Giordano, P., Bryzgalov, A.A., Lazo, M.J., Calculus of variations and optimal control for generalized functions, Nonlinear Analysis, Volume 216, 2022. See also arxiv.org/abs/2011.09660

[25] Galindo, A., Pascual, P., *Quantum Mechanics I*, Springer-Verlag Berlin Heidelberg, 1990.

[26] Giordano, P., The ring of Fermat reals. *Advances in Mathematics*, **225** (4): 2050–2075, 2010.

[27] Giordano, P., Infinitesimals without logic, Russian Journal of Mathematical Physics, 17(2), pp.159-191, 2010.

[28] Giordano, P., Kunzinger, M., Inverse Function Theorems for Generalized Smooth Functions. Invited paper for the Special issue ISAAC - Dedicated to Prof. Stevan Pilipovic for his 65 birthday. Eds. M. Oberguggenberger, J. Toft, J. Vindas and P. Wahlberg, Springer series Operator Theory: Advances and Applications, Birkhaeuser Basel, 2016.

[29] Giordano, P., Kunzinger, M., A convenient notion of compact sets for generalized functions. Proceedings of the Edinburgh Mathematical Society, Volume 61, Issue 1, February 2018, pp. 57-92.

[30] Giordano, P., Kunzinger, M., New topologies on Colombeau generalized numbers and the Fermat-Reyes theorem, *Journal of Mathematical Analysis and Applications* 399 (2013) 229–238. http://dx.doi.org/10.1016/j.jmaa.2012.10.005

[31] Giordano, P., Kunzinger, M., Vernaeve, H., Strongly internal sets and generalized smooth functions. Journal of Mathematical Analysis and Applications, volume 422, issue 1, 2015, pp. 56-71.

[32] Giordano, P., Kunzinger, M., Vernaeve, H., A Grothendieck topos of generalized functions I: basic theory. Dissertationes Mathematicae, 592 (2024), 74 pp. SDOI: 10.4064/dm230920-7-3

[33] Giordano, P., Luperi Baglini, L., The Grothendieck topos of generalized smooth functions III: normal PDE. Preprint, see http://www.mat.univie.ac.at/~giordap7/ToposIII.pdf

[34] Giordano, P., Luperi Baglini, L., Beyond Cauchy-Kowalevski: A Picard-Lindelöf theorem for smooth PDE. Under referee evaluation for Acta Mathematica, 2023. See http://dx.doi.org/10.13140/RG.2.2.13861.99045

[35] Grosser, M., Kunzinger, M., Oberguggenberger, M., Steinbauer, R., *Geometric theory of generalized functions*, Kluwer, Dordrecht, 2001.

[36] Henson, C.W., 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.

[37] Hörmann, G., V. de Hoop, M., Geophysical modelling with Colombeau functions: Microlocal properties and Zygmund regularity, In *Nonlinear Algebraic Analysis*, eds. A. Delcroix, M. Hasler, J.-A. Marti, V. Valmorin, Cambridge Scientific Publishers, Cambridge, 2004. See also https://arxiv.org/abs/math/0104007

[38] Hörmann, G., Kunzinger, M., Steinbauer, R., Wave equations on non-smooth space-times, Evolution Equations of Hyperbolic and Schrödinger Type, Progress in Mathematics, volume 301, Birkhaeuser 2012, pp. 163-186. See also https://arxiv.org/abs/1107.0014

[39] Hoskins, R.F., Pinto, J.S., *Theories of Generalised Functions. Distributions, Ultradistributions and Other Generalised Functions*, Woodhead Pub. 2005.

[40] Jin, S., Wu, H., Huang, Z., A Hybrid Phase-Flow Method for Hamiltonian Systems with Discontinuous Hamiltonians, SIAM J. Sci. Comput., 31(2), 1303-1321, 2008.

[41] Kaplan, W., Introduction to the Theory of Infinitesimals (J. D. Stroyan and W. A. J. Luxemburg), SIAM Review, Vol. 20, No. 2, 403-403, 1978, DOI:10.1137/1020056.

[42] Keisler, H.J., *Elementary Calculus: An Infinitesimal Approach* (2nd ed.), New York: Dover Publications, (2011).

[43] Kock, A., *Synthetic Differential Geometry*, Vol 51 of LMS Lect. Note Series (Cambridge Univ. Press, 1981).

[44] Kunzinger, M., Oberguggenberger, M., Steinbauer, R., Vickers, J.A., Generalized Flows and Singular ODEs on Dierentiable Manifolds, Acta Applicandae Mathematicae, 80(2), pp 221-241, 2004.

[45] Kunzle, A.F., Singular Hamiltonian systems and symplectic capacities, Singularities and Differential Equations Banach Center Publications, 33, pp.171-187, 1996.

[46] Kriegl, A., Michor, P.W., Product Preserving Functors of Infinite Dimensional Manifolds, Arch. Math. (Brno) 32 (4), 289–306 (1996).

[47] Lakshminarayanan, V., Ghatak, A.K., Thyagarajan, K., *Lagrangian Optics*, Springer Netherlands, 2011.

[48] Laugwitz, D., Definite values of infinite sums: aspects of the foundations of infinitesimal analysis around 1820, Arch. Hist. Exact Sci. 39 (3): 195-245, 1989.

[49] Lavendhomme, R., *Basic Concepts of Synthetic Differential Geometry* (Kluwer Academic Publishers, Dordrecht, 1996). 23.

[50] Lecke, A., Luperi Baglini, L., Giordano, P., The classical theory of calculus of variations for generalized functions. Advances in Nonlinear Analysis, vol. 8, no. 1, 2019, pp. 779-808. https://doi.org/10.1515/anona-2017-0150.

[51] Lerman, E., Montgomery, R., Sjamaar, R., Examples of singular reduction. Symplectic geometry, London Math. Soc. Lecture Note Ser., 192, Cambridge Univ. Press, Cambridge, 1993.

[52] Li, F., Hua, Q., Zhang, S., New periodic solutions of singular Hamiltonian systems with xed energies, J Inequal Appl 2014: 400.

[53] Lim, C.C., On singular Hamiltonians: the existence of quasi-periodic solutions and nonlinear stability, Bull. Amer. Math. Soc. (N.S.) 20(1) 35-40, 1989.

[54] Luperi Baglini, L., Giordano, P., The Grothendieck topos of generalized smooth functions II: ODE, Preprint see http://www.mat.univie.ac.at/~giordap7/ToposII.pdf.

[55] Marsden, J.E., Generalized Hamiltonian mechanics, Arch. Rat. Mech. Anal., 28(4), pp. 323-361, 1968.

[56] Marsden, J.E., Non-smooth geodesic ows and classical mechanics, Canad. Math. Bull., 12, pp. 209-212, 1969.

[57] Mazaheri, H., Hosseinzadeh, A., Ahmadian, M.T. , Nonlinear oscillation analysis of a pendulum wrapping on a cylinder. Scientia Iranica Transactions B: Mechanical Engineering, 19 (2), pp. 335–340, 2012.

[58] Moerdijk, I., Reyes, G.E., *Models for Smooth Infinitesimal Analysis* (Springer, Berlin, 1991).

[59] Mukhammadiev, A., Tiwari, D., Apaaboah, G. et al. Supremum, infimum and hyperlimits in the non-Archimedean ring of Colombeau generalized numbers. Monatshefte für Mathematik, Vol. 196, pages 163-190, 2021.

[60] Mukhammadiev, A., Tiwari, D., Giordano, P., A Fourier transform for all generalized functions. Accepted in Dissertationes Mathematicae, 2023. See https://arxiv.org/abs/2111.15408

[61] Noorbala, M., Sepehrinia, R., Is $n \sin \theta$ conserved along the light path?, Eur. J. Phys. 37 (2016) 025301, doi:10.1088/0143-0807/37/2/025301

[62] Otani, L.B., Pereira, A.H.A., Melo, J.D.D., Amico, S.C., Elastic Moduli characterization of composites using the Impulse Excitation, White paper: Technical-scientific Informative ATCP Physical Engineering, ITC-06 / ATCP, 2014.

[63] Murphy, B., Glendinning, P., Piecewise Smooth Models of Pumping a Child's Swing, SIAM Review 2023 65:3, 887-902.

[64] Průša, V., Rajagopal, K.R. Jump conditions in stress relaxation and creep experiments of Burgers type fluids: a study in the application of Colombeau algebra of generalized functions. Z. Angew. Math. Phys. 62, 707–740 (2011).

[65] Průša V. and K. R. Rajagopal On the response of physical systems governed by nonlinear ordinary differential equations to step input. Int. J. nonlinearMech. 81 (2016), 207–221

[66] V. Průša, M. Řehoř, K. Tůma, Colombeau algebra as a mathematical tool for investigating step load and step deformation of systems of nonlinear springs and dashpots, Zeitschrift für angewandte Mathematik und Physik, 2017.

[67] Robinson, A., Function theory on some nonarchimedean fields, Amer. Math. Monthly 80 (6) (1973) 87–109; Part II: Papers in the Foundations of Mathematics.

[68] Rousseau, E., Felbacq, D., Detailed derivation of the generalized Snell–Descartes laws from Fermat's principle. Journal of the Optical Society of America. A Optics, Image Science, and Vision, 2023, 40 (4), pp.676. 10.1364/JOSAA.478123. hal-04073052

[69] Saha, T., Pal, P.J., Malay Banerjee, M., Slow–fast analysis of a modified Leslie–Gower model with Holling type I functional response, Nonlinear Dyn (2022) 108:4531–4555, https://doi.org/10.1007/s11071-022-07370-1

[70] Schmieden, C., Laugwitz, D., Eine Erweiterung der Infinitesimalrechnung, *Math. Zeitschr.*, 69, l–39, 1958.

[71] Shamseddine, K., Berz, M., Analysis on the Levi-Civita Field: A Brief Overview", Contemporary Mathematics, 508 pp 215-237 (2010).

[72] Sieniutycz, S., Jezowski, J., *Energy Optimization in Process Systems and Fuel Cells*. Elsevier, 2013.

[73] Tanaka, K., A Prescribed Energy Problem for a Singular Hamiltonian System with a Weak Force, Journal of Functional Analysis 113(2), pp 351-390, 1993.

[74] Tiwari, D., Mukhammadiev, A., Giordano, P., Hyper-power series and generalized real analytic functions. In revision for Monatshefte für Mathematik, 2023. See arXiv 2212.04757.

[75] Teodorescu, P.P., Kecs, W.W., Toma, A., *Distribution Theory. With Applications in Engineering and Physics*. Wiley-VCH Verlag, 2013.

[76] Ugural, A.C., Fenster, S.K., *Advanced mechanics of materials and elasticity*, Prentice Hall International Series in the Physical and Chemical Engineering Sciences, Prentice Hall; 5 edition, 2011.

[77] Vladimirov, V.S., Equations of mathematical physics, Marcel Dekker Inc., New York, 1971.

[78] von Neumann, J., Method in the Physical Sciences. Collected Works Vol. VI. Theory of Games, Astrophysics, Hydro-dynamics and Meteorology, A.H. Taub (ed.), Pergamon Press, Oxford, 1961.

[79] Zatkis, H., Snell's law and the calculus of variations. Journal of the optical society of America, Vol. 55, n. 1, 1965.

[80] See https://reference.wolfram.com/language/ref/NDSolve.html, Wolfram Research, Inc., Mathematica, Champaign, IL.

[81] See https://mathworld.wolfram.com/HeavisideStepFunction.html and https://mathworld.wolfram.com/DeltaFunction.html, Wolfram Research, Inc., Mathematica, Champaign, IL.

[82] See https://reference.wolfram.com/language/ref/NonlinearModelFit.html, Wolfram Research, Inc., Mathematica, Champaign, IL.

The Martin Luther University Halle-Wittenberg (MLU), Faculty of Medicine, Institute of Medical Epidemiology, Biometry and Informatics (IMEBI), Magdeburger Strasse 8, 06112 Halle (Saale), Germany

Email address: aleksandr.bryzgalov@uk-halle.de

Faculty of Mathematics, University of Vienna, Austria, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

Email address: kevin.islami@univie.ac.at

Faculty of Mathematics, University of Vienna, Austria, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

Email address: paolo.giordano@univie.ac.at