# Distributional Methods in General Relativity

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
an der Formal- und Naturwissenschaftlichen Fakultät
der Universität Wien

eingereicht von

ROLAND STEINBAUER

Juli 2000

El mundo era tan reciente, que muchas cosas careción de nombre,
y para mencionarles había que señalarles con el dedo.

Gabriel García Márquez
Cien años de soledad

ABSTRACT

The aim of this work is a detailed study of applicability and applications of distributional concepts and methods—with a special focus on the theory of algebras of generalized functions—in the theory of general relativity.

Idealizations play a crucial role in modelling physical phenomena: In many cases, they are indispensable for making the latter accessible to a theoretical treatment. As typical examples, think of point particles and point charges. On describing these idealizations mathematically one is naturally led to L. Schwartz' theory of distributions. Unfortunately this theory is only linear, a fact that seriously limitates its range of applicability in nonlinear physical theories. In the present work, after reviewing the theory of distribution valued sections in vector bundles (chapter 1), we investigate its usefulness in the inherently nonlinear theory of general relativity. Following Geroch and Traschen, in chapter 2 we draw the conclusion that a mathematically rigorous and physically sensible framework based upon linear distribution theory excludes the description of such interesting spacetimes as cosmic strings and impulsive gravitational waves.

At this stage the theory of algebras of generalized functions as developed by J. F. Colombeau throughout the 1980s enters the field. In this approach one constructs associative and commutative differential algebras canonically containing the vector space of distributions as a subspace and the algebra of smooth functions as a faithful subalgebra. Hence, according to L. Schwartz' so-called "impossibility result," it combines all favorable differential algebraic properties with a maximum of consistency properties with respect to classical operations. Apart from being a valuable tool in the analysis of nonlinear partial differential equations involving singular data or coefficients, the usefulness of algebras of generalized functions for geometric applications in the beginning was seriously restricted due to its lack of diffeomorphism invariance; a flaw that has ultimately been removed only recently. In this work we introduce algebras of generalized functions in chapter 3 and devote the entire chapter 4 to the construction of generalized sections in vector bundles. In particular, we construct a generalized curvature framework well suited to the needs of general relativity.

The final chapter 5 provides a detailed distributional description of the geometry of impulsive gravitational waves. We treat the geodesic as well as the geodesic deviation equation for this class of singular spacetimes in the previously developed generalized setting. Moreover, we carry out a detailed mathematical analysis of the discontinuous change of coordinates frequently applied to the impulsive wave metric in physical literature. We conclude this work with an outlook to promising lines of further research.

*2000 Mathematics Subject Classification: 46F30, 83C35, 83C15, 46F10, 46F05, 83C75*

*1999 Physics and Astronomy Classification Scheme: 04.20.Cv, 02.30.Sa, 04.20.Jb, 04.30.-w*

# ZUSAMMENFASSUNG

Die vorliegende Arbeit enthält eine detaillierte Studie über die Anwendbarkeit und die Anwendungen distributioneller Methoden—vor allem der von J. F. Colombeau entwickelten Theorie der Algebren verallgemeinerter Fuktionen—in der Allgemeinen Relativitätstheorie.

Idealisierungen spielen eine herausragende Rolle bei der Modellierung physikalischer Phänomene. Oft machen sie diese erst einer theoretischen Beschreibung zugänglich. Besonders nützliche Idealisierungen sind z.B. Punktteilchen und Punktladungen. Ihre mathematische Beschreibung führt in natürlicher Weise auf die von L. Schwartz begründete Theorie der Distributionen. Diese ist von Natur aus eine lineare Theorie; das Produkt zweier Distributionen kann im allgemeinen nicht wieder als Distribution aufgefaßt werden. Daher sind Anwendungen distributioneller Methoden in nichtlinearen physikalischen Theorien enge Grenzen gesetzt.
In dieser Arbeit gehen wir nach einer Zusammenfassung der Theorie der distributionswertigen Schnitte in Vektorbündeln (Kapitel 1) den Grenzen ihrer Anwendbarkeit in der imanent nichtlinearen Allgemeinen Relativitätstheorie nach. Einer Arbeit von Geroch und Traschen folgend, kommen wir im 2. Kapitel zum Schluß, daß ein mathematisch konsistenter und physikalisch vernünftiger, auf der linearen Distributionentheorie aufbauender Formalismus die Beschreibung von so interessanten Raumzeiten wie kosmischen Strings und impulsiven Graviationswellen ausschließt.
An diesem Punkt tritt die von J. F. Colombeau in den 1980er Jahren entwickelten Theorie der Algebren verallgemeinerter Funktionen auf den Plan. Dieser Zugang ermöglicht die Konstruktion kommutativer und assoziativer Differentialalgebren, in die der Vektorraum der Distributionen kanonisch als Teilraum und die Algebra der glatten Funktionen als treue Teilalgebra eingebettet sind. Im Lichte des sogenannten ,,Unmöglichkeitsresultats" von L. Schwartz vereinigt sie damit die maximal möglichen Konsistenzeigenschaften in bezug auf klassische Operationen mit allen wünschenswerten differential-algebraischen Eigenschaften. Obwohl ein wertvolles Werkzeug zur Behandlung nichtlinearer partieller Differentialgleichungen mit singulären Daten oder Koeffizienten, waren Algebren verallgemeinerter Funktionen aufgrund fehlender Diffeomorphismeninvarianz geometrischen Anwednungen anfänglich nur schwer zugänglich. Dieser Nachteil konnte erst unlängst vollständig ausgeräumt werden. Die vorliegende Arbeit beschäftigt sich nach einer Einführung in die Theorie der Algebren verallgemeinerter Funktionen (Kaptiel 3) ausführlich mit der Konstruktion verallgemeinerter Schnitte in Vektorbündeln (Kapitel 4). Insbesondere konstruieren wir eine Theorie verallgemeinerter Krümmungsgrößen, die auf Anwendungen in der Allgemeinen Relativitätstheorie zugeschnitten ist.
Im abschließenden 5. Kapitel beschreiben wir detailliert die distributionelle Geometrie impulsiver Gravitationswellen. Wir behandeln die Geodäten- sowie die geodätische Deviationsgleichung im zuvor entwickelten Formalismus. Außerdem wird eine mathematisch korrekte Beschreibung des in der physikalischen Literatur verwendeten, unstetigen Koordinatenwechsels für die Metrik impulsiver Gravitationswellen vorgestellt. Wir schließen mit einem Ausblick auf erfolgversprechende weitere Forschungsprogramme.

**PREFACE**

The line of research finally leading to the present work was actually started a long time ago. It was in the autumn of 1993 when I asked my favorite theoretical physics teacher Prof. Helmuth Urbantke to propose a theme for my diploma thesis. He came up with several suggestions but I kept asking him for "more mathematical" topics. At last he told me that he knew of a construction—called Colombeau algebras—that would allow to multiply distributions and he conjectured it could be an adequate tool to compute the ultrarelativistic limit of the Reissner-Nordstrøm black hole. After a short time of reflection I decided to devote my thesis to it. By that time I neither had a conception of how lucky that choice was nor how far it would carry me.
After some lonely months studying Colombeau's books I attended a lecture course on (classical) distribution theory given by Michael Grosser at the Department of Mathematics. There I got to know that he was currently supervising the Ph.D. thesis of Michael Kunzinger on Lie transformation groups in Colombeau algebras (a booklet that quickly became one of my favorite requisites). Shortly after that we had our first joint seminar on Colombeau algebras and ever since that our cooperation has intensified. Not only that both Michaels helped me a lot finishing the more mathematical parts of my diploma thesis, they also took me with them on their regular visits to Innsbruck, thereby introducing me to Michael Oberguggenberger. In fact, it was during my first stay at Innsbruck when a lot of the ideas now forming the backbone of this work came upon me. Eventually I became a member of the DIANA (=DIfferential Algebras and Nonlinear Analysis) research group (see `http://www.mat.univie.ac.at/~diana`) that had been formed shortly before by the three Michaels and Günther Hörmann.

So many people have contributed substantially to the genesis of this work that it seems hopeless to do justice and to pay tribute to them all; nevertheless it seems worth trying to.
Right from the beginning I enjoyed having two supervisors directing my thesis. Prof. Helmuth Urbantke was in charge of the physics part, but also backed up the whole project by his general overview and helpful advice. Michael Grosser kept me on the tracks of exactness and rigor. He aided a lot to my general mathematical education impressing me time and again by his ability to understand *and* answer my questions on a level much more fundamental than I had ever expected; everyone who knows his clear wording will hopefully recognize its traces in this work. Over the past two years he devoted much of his time and energy to the construction of diffeomorphism invariant Colombeau algebras breaking through even the hardest roadblocks in a style that filled people much more experienced than me with enthusiasm.
The most important contributions I probably owe to Michael Kunzinger. His ongoing willingness to listen to my ideas at a stage too vague to allow any reasonable mathematical formulation finally encouraged their rigorous restatement. Patiently he checked pages and pages of my scribbled handwritten estimates conferring them with the (high-rated; not only by me!) "Kunzinger-seal" of approval. As the "PR-manager" of the DIANA-group he aided my formulations with numerous precise terms and catch phrases. Finally he always provided the criticism of a good friend.
Michael Oberguggenberger—without question the main protagonist of DIANA—backed up the whole work with his universal knowledge of the field; no question could seriously embarrass him. His hardiness and staying power is exemplary. Often he astonished us in the morning presenting the solutions to the most serious problems of the previous day he had just worked out overnight. All this together with the warm hospitality of him and Elisabeth Oberguggenberger turned my stays in Innsbruck into extraordinary experiences.
Though Günther Hörmann's influence on this work might seem not to be an imediate one, nevertheless a lot of things I learned in our joint seminars (which I *really* enjoyed) finally entered here in one way or another. Moreover he always impressed me by keeping perspective in the most critical situations.
During a project on the "Nonlinear Theory of Generalized Functions" organized by Michael Oberguggenberger at the Erwin Schrödinger Institute I got into contact with a lot of researchers working in the field. In particular, I am grateful to James Vickers for numerous stimulating discussions and his tireless efforts during our collaborations. I also wish to thank Prof. Stevan Pilipovic and his group at the University of Novi Sad for their kind hospitality and their support.
Finally I have to thank Prof. Peter C. Aichelburg and H. Balasin for their comments and for many discussions.

I often had the privilege and—at the same time—the handicap to be regarded as a physicist by mathematicians and as a mathematician by physicists, while I enjoyed voyaging between both worlds. Even though it finally seems that I have found my home at the Department of Mathematics approximately the first half of this work was carried out at the Institute for Theoretical Physics. There I enjoyed having a desk in the big

working room of the gravity group with a lot of people around, always ready to start all kinds of (scientific) discussions. I owe a lot of thanks to them all. Sascha Husa who (almost) never was at a loss for an answer taught me (among a lot of other things) to "think like a physicist." It was also him who got me on the tracks of computer system administration and shared some extraordinary travel-experiences with me, in particular to India. Together with him and Barbara Pocza I enjoyed that short science fiction space-walk.
Walter Simon was always willing to lend an open ear to my questions and to go out for a beer afterwards (to be precise, *he* took wine). Often we were accompanied by Marc Mars who really enriched the group during his much-too-short stay in Vienna. I owe special thanks to Franz Embacher for guiding (not only) me on a extraordinary pleasure trip through the realms of quantum gravity. Fritz Schein worked next desk to me and answered many of my (most stupid—as it seemed to me) questions. Cristina Stanciulescu guided Sascha and me safely through Romania. Also I remember a lot of nice occasions like joint dinners, skiing-excursions and the annual "Gschnas."

Since the summer of 1998 I hold a teaching position at the Department of Mathematics. There I benefited from the tireless personal efforts of our head of department, Harald Rindler, who kept (and still keeps) us going. Michael Grosser, Günther Hörmann, Eva Farkas and Michael Kunzinger provided me with the best possible working environment I can imagine. Moreover, they encouraged me every time the burden of administrative work seemed to be overwhelming. Susanne Kandl, Florian Wisser and Andreas Ulovec aided much to the friendly atmosphere at our floor and essentially supported (not only) me in all circumstances of computer system administration. Finally I'd like to thank the secretaries of our Department who always processed my requests in the most helpful way.

The writing of this thesis has been made possible by the Austrian Academy of Science which provided me with a grant (DOC-338) during the first one-and-a-half year of my Ph.D. studies. Later I was founded by project P-12023 MAT (led by H. Urbantke) of the Austrian Science Foundation (FWF). Mark J. Heinzle heroically entered into the FWF-project during the times I was not at disposal and contributed a considerable lot of material to it.

Special thanks go to Roman Ortner whom I got to know at the very beginning of my studies and who has been a very special friend ever since. Thanks to Anita for all her support during the years. Finally I'd like to express my gratitude towards my parents to whom I owe so much more than just guiding me through my Ph.D. years.

Roland Steinbauer

Vienna, July 2000

CONTENTS

## 0. INTRODUCTION

The aim of this work is a detailed study of applicability and applications of distributional concepts and methods—with a special focus on the theory of algebras of generalized functions—in the theory of general relativity.

In linear field theories distributions arise naturally in a number of different ways. They are used to mathematically describe idealized situations where matter or the source of a field is confined to some lower dimensional submanifold of space; typical examples being point particles, point charges, strings, layers and shells of matter. Furthermore distributions are used to model such different phenomena and situations as shock waves, impulsive waves, junction conditions, idealized potentials in elementary quantum mechanics, etc.
*General relativity*—the theory of space, time and gravitation as formulated by A. Einstein 85 years ago—dramatically differs from other field theories. Spacetime is no longer given beforehand but rather it is described by a 4-dimensional manifold together with a Lorentzian metric, which itself is subject to field equations. More precisely, the curvature of the metric is related to the energy momentum content of the spacetime via the so-called Einstein equations, which—written in coordinates—form a complicated hyperbolic system of 10 nonlinear partial differential equations of second order for the coefficients of the metric tensor; hence one has to assume the metric to satisfy sufficient differentiability conditions. Usually one supposes the metric to be smooth; $\mathcal{C}^2$ suffices to do classical differential geometry, whereas $\mathcal{C}^3$ is needed to ensure energy conservation. $\mathcal{C}^{2-}$ (i.e., the first derivatives Lipschitz continuous) guarantees at least unique solvability of the geodesic equations and local boundedness of the Riemann curvature tensor.
A singularity in a general relativistic spacetime intuitively is a "place" where curvature blows up or some other "pathological behavior" of the metric occurs. There are two main obstacles, however, to make this notion precise. The first originates from the fact that we only may speak of an event at all (i.e., of a point in spacetime) if—according to the above—the metric is, say, $C^2$ "there." Hence a singularity has to be viewed as a "singular boundary point" rather than a point in spacetime. Since a Lorentzian metric—contrary to a Riemannian one—does not give rise to a distance function, the construction of a topological boundary is a non-trivial matter. In fact, no fully satisfactory general notion of a singular boundary of spacetime exists. The second difficulty in defining a singularity in general relativity is deeply connected with another of the theory's main principles: diffeomorphism invariance. The fact that the components of, say, the Riemann tensor blow up along a curve may simply be due to a bad choice of coordinates in the following sense. It might be possible to find a different coordinate system which allows to extend the spacetime beyond the "critical point" with the Riemann tensor perfectly well-behaved. Moreover, there is a big variety of examples showing that the blowup of some curvature quantity is not an adequate tool to detect singularities.
From the above it should become clear why it is tempting to characterize singularities in general relativity by their geometrical properties rather than by their analytical ones. In fact, by the standard definition a spacetime is called singular if there exist incomplete geodesics, i.e., geodesics of finite affine parameter length which may not be extended. Obviously designed to capture the intuitive notion of a "hole in spacetime," there are, however, also some problems associated with this "geometric" approach. First, it does not provide an ultimate answer to the question what a singularity actually is; note that we have only defined the notion of a singular spacetime. Instead there is a quite lengthy catalogue of possible ways in which a spacetime could "break down" (non-smoothness, unboundedness or local non-integrability of the Riemann tensor, spacetimes "created" with a primordial singularity and the like). Moreover, by the singularity theorems of Penrose and Hawking (see e.g. [80], chap. 8) many physically reasonable spacetimes (in particular, all realistic models of an expanding universe and of gravitational collapse) are singular with respect to this definition.
Consequently, the recent development of the study of spacetime singularities has focussed more upon a study of the field equations. General relativity as a physical theory is governed by particular physical equations; what is of primarily interest is the breakdown of physics which may, or may not, result in a breakdown of geometry. Unfortunately, there is somehow a conflict between the mathematical contexts appropriate to, on the one hand, partial differential equations and, on the other hand, geometry. In the differential geometric study of singularities one deals with geodesic equations which are uniquely solvable provided the metric is $\mathcal{C}^{2-}$ (as already remarked above); beyond this, the differenatibility of the metric is of little geometrical significance. By contrast, in the study of hyperbolic PDEs the question of differentiability is crucial; the differentiability chosen determines the character of the solutions allowed. By choosing low differentiability one admits solutions like shock waves or impulsive waves, which, on the other hand, are ruled out as "singular" when insisting on high differentiability.

Accepting the field equations to play the primary role, they should determine the choice of differentiability. A singularity should be an obstruction to the existence of a solution to the field equations. Clearly, this cannot be determined in general. On the other hand, there are existence theorems that tell when it *is* possible to construct a solution, i.e., we know what is not a singularity. However, in general it is difficult to link back a certain differentiability condition necessary to prove (local) existence of solutions to Einstein's equations (Sobolev conditions, usually imposed on the Riemann tensor on specific hypersurfaces in a particular class of coordinate systems, well suited for the formulation of the Einstein equations as an initial value problem) to geometry, that is the differentiability of the metric on the entire spacetime manifold.
Summing up the entire discussion so far, it would be very desirable to have a description of singularities as internal points of the spacetime manifold where the field equations are satisfied in a weak (probably distributional) sense (see also [35]). Hence one would wish to significantly lower the "geometric differentiability bound" (i.e., $C^{2-}$) on the metric. Indeed, a recent monograph on the subject of spacetime singularities [34] reaches the conclusion that the answer to many of the questions raised above "involve detailed considerations of distributional solutions to Einstein's equations, leading into an area that is only starting to be explored [...]" In particular, one wants to be able to describe spacetimes containing matter whose density function is e.g. unbounded but integrable or confined to a submanifold in spacetime (both scenarios amounting to a finite mass per unit volume). Important examples in this class include thin cosmic strings and impulsive gravitational waves.
Another strong motivation for a study of distributional spacetimes arises from the fact that by an argument of Isham [96] the latter will substantially contribute to a path integral description of (a yet to be formulated theory of) quantum gravity.
The big challenge in setting up a framework that might be called a "distributional geometry" adapted to the needs of general relativity of course resides in the immanent *nonlinearity* of the latter theory; calculating the Riemann tensor from the metric is an essentially nonlinear operation. On the other hand, classical distribution theory as founded by L. Schwartz in his famous book [167] is a *linear* theory. This deep conceptual problem obviously is the main reason why applications of distribution theory to general relativity have been rare and either limited to special situations or lacking the necessary mathematical rigor. Moreover, Geroch and Traschen ( [71]) have shown that a physically sensible and mathematically sound framework based on linear distribution theory cannot handle such interesting sources of the gravitational field as strings and point particles. We shall give a detailed account on the usefulness and limitations of linear distributional methods applied to general relativity in chapter 2 below.

At this point we turn to the second theory mentioned in the very first sentence of this introduction: the *theory of algebras of generalized functions* as developed by J. F. Colombeau in the 1980s ( [38], [39], [41]). In this approach one constructs differential algebras canonically containing the vector space of distributions as a subspace (embedded basically by convolution with a general class of mollifiers) and the space of smooth functions as a faithful subalgebra. As an associative and commutative algebra it combines a maximum of favorable differential algebraic properties with a maximum of consistency properties with respect to classical operations according to the impossibility result of L. Schwartz (cf. [168]). The need for algebras of that type arises, for example, from the necessity of considering nonlinear PDEs where either the respective coefficients, the data or the prospective solutions are singular. Classical linear distribution theory does not permit the treatment of such problems. On the other hand, Colombeau algebras have proven to be a useful tool for analyzing such questions (for applications in nonlinear PDEs, cf. e.g., [23], [24], [40], [42], [43], [106], [144], for applications to numerics, see e.g., [22], [20], [21], for applications in mathematical physics, e.g., [186], [74] as well as the literature cited in these works).
Since Colombeau's first monograph [38], there have been introduced a considerable number of variants of Colombeau algebras, many of them adapted to special purposes. Designed primarily as a tool in the field of nonlinear PDEs, the development of the theory of algebras of generalized functions on manifolds progressed at a much slower pace. Indeed, the question of diffeomorphism invariance of Colombeau algebras—a vital question not only in the context of applications to general relativity—is somehow subtle and has been finally settled only recently.
In fact, the first variants of algebras of generalized functions did not have the property of diffeomorphism invariance. Some of the key ingredients used in the definitions turned out not to be invariant under the natural action of a diffeomorphism. Consequently, Colombeau algebras, although in principle constituting a framework suitable for a "nonlinear distributional geometry" from the beginning, did not provide a ready-to-use tool for relativists. First steps to remove this shortcoming have been undertaken by Colombeau and Meril [44], whereas later also relativists contributed to this line of research (cf. [184], [14]). Based on the

construction of Jelínek [99] the goal of diffeomorphism invariance was reached only recently in a series papers [61], [76], [77]. A local diffeomorphism invariant version of the theory has been first presented in the article [61], while a global ("geometrical") construction on differentiable manifolds was carried out by the authors of [77]. The latter yields a differential (w.r.t the Lie derivative) algebra of generalized functions and a canonical embeddeding of the vector space of distributions on the manifold which commutes with Lie derivatives. In the present work the entire fourth chapter is devoted to the study of Colombeau algebras on manifolds.
Summing up, Colombeau algebras provide a rigorous mathematical framework for simultaneously treating singular (i.e., distributional) objects, nonlinear operations and differentiation, hence providing a setting suitable to the description of general relativistic spacetimes of low differentiability.

First applications of algebras of generalized functions in general relativity appeared in the mid 1990s, including the description of conical singularities [36] and the ultrarelativistic Reissner-Nordstrøm metric [171]. Since then, a growing number of publications have taken up these ideas, applying Colombeau algebras to such different problems in general relativity as Kerr-Newman and ultrarelativistic Kerr-Newman spacetimes ( [16], and the work cited there), further studies of conical singularities and cosmic strings ( [185], [189], [188]), impulsive gravitational waves ( [15], [174], [118], [119]) and signature change ( [137], [104]). A recent overview of applications of nonlinear generalized functions in general relativity may be found in a review article by J. Vickers [186].

This work is organized in the following way. In chapter 1 we review linear distributional geometry, i.e., the theory of distribution-valued sections of vector bundles. Using a unifying approach based upon the concept of $q$-densities, we put main emphasis on revealing the interrelations of the different approaches taken in the literature, each assuming distinct levels of additional structure on the manifold.
Chapter 2 is devoted to a study of the limitations of linear distribution theory. After reviewing the problems generally arising in tip-toeing over the borders of the classical linear theory, in sec. 2.A we study in detail its consequences in the context of general relativity. In particular we review the classical result of Geroch and Traschen [71] (see above) in sec. 2.B.
The following two chapters are devoted to an introduction to the theory of algebras of generalized functions. In chapter 3 we present the so-called "special" version of the theory on open sets of $\mathbb{R}^n$ introducing all of its basic concepts. The embedding of distributions is constructed in sec. 3.B, point values and integration of generalized functions are introduced in sec. 3.C and 3.D, respectively. The central notion of association and its relations to nonlinear modelling is covered in sec. 3.E.
In chapter 4 we treat nonlinear generalized functions on manifolds. The scalar theory is presented in sec. 4.A, whereas a theory of generalized sections in vector bundles is developed in sec. 4.B. The following sec. 4.C is concerned with a nonlinear distributional geometry suited to the needs of general relativity. Finally in sec. 4.D we review the recent development towards the construction of diffeomorphism invariant Colombeau algebras.
The final chapter 5 is devoted to a detailed study of the distributional geometry of impulsive gravitational waves. After introducing this class of singular space times in sec. 5.A, we study the geodesic as well as the geodesic deviation equation (secs. 5.B and 5.C) in the previously introduced generalized geometrical framework. Sec. 5.D is devoted to a mathematical analysis of the discontinuous change of coordinates frequently applied to the impulsive wave metric in the physical literature. We close with an outlook to further lines of research in sec. 5.E.

# 1. DISTRIBUTIONS ON MANIFOLDS

The aim of this chapter is to provide an introduction into "distributions on manifolds," where this term stands for the various distributional spaces (i.e., distributions, distributional densities, etc.) on a differentiable manifold itself, as well as for distribution valued sections in vector bundles. In textbooks one finds several, quite different approaches to this topic some of them using certain levels of additional structure, most notably Dieudonné [53], De Rham [49], Hörmander [88], Choquet-Bruhat et al [31] and Friedlander [65].

We shall give a unified presentation of the subject based upon [169] using the concept of $q$-densities, which are introduced in section 1.B. In this section we also set up the main part of our notation for vector bundles. In sec. 1.C we define the appropriate topologies on the spaces of test fields and give the basic definition of distributional spaces as their topological duals, while sec. 1.D deals with their local description. Finally in sec. 1.E we review Marsden's (cf. [138]) distributional geometry on oriented manifolds. We begin with an introduction providing an outline to the main ideas of this chapter.

## 1.A. Introduction

**1.1** On open sets of $\mathbb{R}^n$ a distribution is defined to be a continuous linear functional on the space of (smooth, compactly supported) test *functions* $\varphi$. Any smooth (even locally integrable) *function* $f$ gives rise to a *regular* distribution via the (natural) assignment

$$\varphi \mapsto \int f(x) \cdot \varphi(x)\, dx \ .$$

On an arbitrary manifold $X$, these two statements cannot hold simultaneously in a meaningful way (with emphasis on "functions") since in this context, there is no notion of an integral of (even smooth, compactly supported) functions.
On a manifold—in absence of a preferred measure—the objects to be integrated are (one-)*densities* rather than functions, since they contain information on the "volume element" with respect to which the integration has to be performed (see sec. 1.B for details). Thus, either the nature of test "functions" $\varphi$ or of regular distributions $f$ or of both has to be changed in such a way that their product $f \cdot \varphi$ becomes a density. Now the product of a density with a smooth function turns out to be a density again. This immediately offers two (in a sense, complementary) ways of proceeding. On one hand, we can replace test *functions* by test *densities* and define a distribution to be a continuous linear functional on the space of these densities. Then again each (say, smooth) *function* can be considered as a distribution. This is in accordance with e.g., Hörmander [88], sec. 6.3. On the other hand, we could keep the function character of the test objects; then the regular objects in the dual space of the space of test *functions* have to be taken as (smooth) *densities* on $X$. This is the definition adopted e.g. by Dieudonné, in [53], chap. XVII.
More generally, the burden of rendering $f \cdot \varphi$ a density can be split up in one part contributed by $f$ and in a (complementary) part contributed by $\varphi$. It is possible to define, for each real $q$, the notion of a $q$-density in such a way that for arbitrary real $q, q'$, the product of a $q$-density with a $q'$-density is a $(q + q')$-density; moreover, one-densities are just densities in the above sense and zero-densities correspond to functions. If we now define the test objects to be (compactly supported, smooth) $q$-densities, the appropriate $(1 - q)$-densities can be embedded in their dual space as regular objects such that, again, $\varphi \mapsto \int f \cdot \varphi$ is meaningful. Setting $q = 0$ gives Dieudonné's choice while $q = 1$ reproduces Hörmander's definition. The general version comprising both these special cases can be found e.g. in [169] or [149].

**1.2** The technical definition of densities as well as the generalization to vector-valued (test objects and) distributions uses the language of vector bundles. We will present the corresponding notation and terminology in the next section. For the moment, let us anticipate that a density (resp. $q$-density) can be realized as a section of the (one-dimensional) so-called volume bundle $\mathrm{Vol}\,(\mathrm{X})$ (resp. $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$). A $q$-density with values in some vector bundle $E$ over $X$ is then defined as a section $u$ of the bundle $E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$. An appropriate regular dual object for such (compactly supported, smooth) sections $u$ obviously would be a smooth section $f$ of the bundle $E^* \otimes \mathrm{Vol}^{\,1-\mathrm{q}}(\mathrm{X})$ where $E^*$ denotes the dual bundle of $E$. The canonical bilinear form $(\,.\,,\,.\,)$ on $E^* \times E$ and the product of densities make $(f, u)$ a 1-density for which the integral is defined. Interchanging $E$ and $E^*$ as well as $q$ and $1 - q$, we finally arrive at the definition of $E$-valued distributions of density character $q$ (which in detail will be given in 1.14) as the dual space of the space of (compactly supported, smooth) sections of the bundle $E^* \otimes \mathrm{Vol}^{\,1-\mathrm{q}}(\mathrm{X})$.

We start our thorough treatment by recalling some basic facts on densities and integration (for more details see e.g. [139], chap. 8).

### 1.B. Densities, Integration, Orientation

**1.3** To begin with let $W$ be a (real) vector space of dimension $n$ and denote by $\wedge^n W$ the $n$-fold antisymmetrized tensor product of $W$. For $q \in \mathbb{R}$ a *$q$-density $\mu$ on $W$* is a map $\mu : \wedge^n W \setminus \{0\} \to \mathbb{C}$ such that for all $0 \neq s \in \mathbb{R}$ and for all $0 \neq \omega \in \wedge^n W$ we have

$$\mu(s\omega) = \mid s \mid^q \mu(\omega)\,.$$

A $q$-density $\mu$ on $W$ is uniquely determined by its value on a particular $\omega \neq 0$. Indeed, since $\wedge^n W$ is one-dimensional any of its elements can be uniquely written as $\omega' = c\omega$, where $c \in \mathbb{R}$, hence $\mu(\omega') = |c|^q \mu(\omega)$. With the obvious operations the space of all $q$-densities on $W$ becomes a (one-dimensional complex) vector space, which we denote by $\mathrm{Vol}^q(\mathrm{W})$. Let $(v_i)_{i=1}^n$ and $(w_i)_{i=1}^n$ be two bases of $W$ and $A = (a_{ij})$ the matrix of the basis change, i.e., $v_i = \sum a_{ij} w_j$ then we clearly have

$$\mu(v_1 \wedge \ldots \wedge v_n) = \mid \det A \mid^q \mu(w_1 \wedge \ldots \wedge w_n)\,. \tag{1}$$

**1.4** Consider now a paracompact smooth Hausdorff manifold of dimension $n$, which from now on, we always denote by $X$. Here we will introduce $q$-densities to be the sections of a (complex) line bundle whose transition functions are the analogue of (1). For the notions from differential geometry used in the sequel cf. [1,53]. In particular, using the notation of [53], vector bundles with base space $X$ will be denoted by $(E, X, \pi)$ and for a chart $(V, \psi)$ in $X$, a *vector bundle chart* $(V, \Psi)$ over $\psi$ will be written in the form ($\mathbb{K} = \mathbb{R}$ for *real* resp. $\mathbb{K} = \mathbb{C}$ for *complex* vector bundles)

$$\begin{aligned} \Psi : \pi^{-1}(V) &\to \psi(V) \times \mathbb{K}^N \\ z &\to (\psi(p), \psi^1 z, \ldots, \psi^N z) \equiv (\psi(p), \boldsymbol{\psi} z)\,, \end{aligned} \tag{2}$$

where $p = \pi(z)$ and the typical fiber is $\mathbb{K}^N$. Let $(V_\alpha, \Psi_\alpha)_\alpha$ denote a vector bundle atlas then we write $\Psi_\alpha \circ \Psi_\beta^{-1}(y, w) = (\psi_{\alpha\beta}(y), \boldsymbol{\psi_{\alpha\beta}}(y)w)$, where $\psi_{\alpha\beta} := \psi_\alpha \circ \psi_\beta^{-1}$ is the change of chart on the base and $\boldsymbol{\psi_{\alpha\beta}}$ : $\psi_\beta(V_\alpha \cap V_\beta) \to \mathrm{GL}(N, \mathbb{K})$ denotes the *transition functions*. Recall that the system of transition functions contains all the relevant information on the respective vector bundle in the following sense: Given a family of functions $\boldsymbol{\psi_{\alpha\beta}} : \psi_\beta(V_\alpha \cap V_\beta) \to \mathrm{GL}(N, \mathbb{K})$ satisfying the *cocycle conditions* (cf. [139], 6.3) then there exists a vector bundle over $X$ having precisely the given $\boldsymbol{\psi_{\alpha\beta}}$ as transition functions (see the proof of thm. 6.3 in [139]).
Spaces of *$\mathcal{C}^k$-sections* ($0 \leq k \leq \infty$) in the bundle $(E, X, \pi)$ (with compact support) will be denoted by $\Gamma^k(E)$ (respectively $\Gamma^k_c(E)$), while local sections over the open set $V \subseteq X$ will be denoted by $\Gamma^k(V, E)$ (respectively $\Gamma^k_c(V, E)$). If $k = \infty$ we often drop it to simplify notation. Let $\varphi : X \to Y$ be a diffeomorphism and $\Phi$ : $(E, X, \pi_E) \to (F, Y, \pi_F)$ a *vector bundle homomorphism along* $\varphi$, i.e., a fiberwise linear smooth map from $E$ to $F$ satisfying $\pi_F \circ \Phi = \varphi \circ \pi_E$. For a section $u \in \Gamma(E)$, we define $\Phi_* u \in \Gamma(F)$ by

$$\Phi_* u := \Phi \circ u \circ \varphi^{-1}\,.$$

Note that the definition of $\Phi_*$ is functorial, i.e., $(\Phi_1 \circ \Phi_2)_* = (\Phi_1)_* \circ (\Phi_2)_*$ and $(\mathrm{id})_* = \mathrm{id}$ (provided that also $\varphi$ is the identity map). If, in addition, $\Phi$ is a diffeomorphism we also write $(\Phi^{-1})^*$ in place of $\Phi_*$ (compare $(\varphi^{-1})^*(\chi) \in \mathcal{C}^\infty(Y)$ for $\chi \in \mathcal{C}^\infty(X)$).
The *local expression* (cf. [53], 16.15.1) of a section $u \in \Gamma(E)$ in a chart $(V_\alpha, \Psi_\alpha)$ with $\psi_\alpha(p) = (x^1(p), \ldots, x^n(p))$ (which in the sequel will be called a chart with coordinates $x^i$) will be written in the form

$$u_\alpha := (\Psi_\alpha^{-1})^* \, u|_{V_\alpha} : \quad x \mapsto (x, u^1_\alpha(x), \ldots, u^N_\alpha(x))\;,$$

where $u^i_\alpha = \psi^i_\alpha \circ u \circ \psi_\alpha^{-1} \in \mathcal{C}^\infty(\psi_\alpha(V_\alpha))$ ($i = 1, \ldots, N$) denote the *components* of the section $u$ with respect to the chart. Clearly $u_\alpha \in \Gamma(\psi_\alpha(V_\alpha), \mathbb{K}^N)$. On the other hand, in the case of a $(r, s)$-tensor field $T$ (i.e., a section in $E = (TX)^r \otimes (T^*X)^s) =: T^r_s(X)$, where in the case $r = 1$, $s = 0$ resp. $r = 0$, $s = 1$ we simply write $\mathfrak{X}(X)$, resp. $\mathfrak{X}^*(X)$) we shall use the term *components* also for the functions $T^{\alpha\,i_1 \ldots i_r}_{\;\;\,j_1 \ldots j_s} :=$

$T(dx^{i_1},\dots,dx^{i_r},\frac{\partial}{\partial x^{j_1}},\dots,\frac{\partial}{\partial x^{j_s}})$ in $\mathcal{C}^\infty(V_\alpha)$ which we distinguish from the previous notion by the *upper* position of the chart index. Note that, apart from the different domain of definition as compared to $u^i_\alpha$, the functions $T^{\alpha\,i_1\dots i_r}_{\;\;j_1\dots j_s}$ involve the use of the *natural basis* in the respective tensor bundle. We are now prepared to define

**1.5 Definition.** *Let $(V_\alpha,\psi_\alpha)_\alpha$ be an atlas for $X$. The* $q$-volume bundle *over $X$ denoted by* $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ *is the one-dimensional (complex) vector bundle (line bundle) given by the following cocycle of transition functions*

$$\begin{aligned}\boldsymbol{\psi_{\alpha\beta}}: &\quad \psi_\beta(V_\alpha\cap V_\beta)\to\mathbb{R}\setminus\{0\}=GL(1,\mathbb{R})\\ \boldsymbol{\psi_{\alpha\beta}}(y) &= |\det D(\psi_\alpha\circ\psi_\beta^{-1})(y)|^{-q} = |\det D(\psi_\beta\circ\psi_\alpha^{-1})(x)\,|^{q}\ ,\end{aligned}\tag{3}$$

*where $x=\psi_{\alpha\beta}(y)$. If $q=1$ we omit it for simplicity. $\mathcal{C}^k$-sections of* $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ *are called* $\mathcal{C}^k$-$q$-densities on $X$.

First note that zero-densities are just functions (i.e., $\mathrm{Vol}^{\,0}(\mathrm{X})=\mathrm{X}\times\mathbb{C}$). Next, in order to give a concrete representation of $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$, set $E=\bigcup_{p\in X}\{p\}\times\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{T_pX})$. For any chart $(V_\alpha,\psi_\alpha)$ of $X$ and any $p\in V_\alpha$ there exists a unique element $|dx^1\wedge\dots\wedge dx^n|^q(p)\in\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{T_pX})$ such that $|dx^1\wedge\dots\wedge dx^n|^q(p)((\partial_{x^1}\wedge\dots\wedge\partial_{x^n})|_p)=1$. Note that $|dx^1\wedge\dots\wedge dx^n|^q(p)$ is just a notation in this context (although one that suggests the correct transformation behavior as, we shall see shortly) and *not* the composition of the $n$-form $dx^1\wedge\dots\wedge dx^n$ with $|\ |^q$. It is straightforward from the definition of $|dx^1\wedge\dots\wedge dx^n|^q$ that if $(V_\beta,\psi_\beta)$ is another chart with coordinates $y^i$ and $p\in V_\alpha\cap V_\beta$ we have

$$|dy^1\wedge\dots\wedge dy^n|^q(p)=\left|\det\frac{\partial y^i}{\partial x^j}\right|^q(p)\ |dx^1\wedge\dots\wedge dx^n|^q(p)\,.\tag{4}$$

For $(p,\mu_p)\in\{p\}\times\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{T_pX})$ $(p\in V_\alpha)$, let $\Psi_\alpha(p,\mu_p)=(\psi_\alpha(p),\mu_p((\partial_{x_1}\wedge\dots\wedge\partial_{x_n})(p)))$. Then $\mu^\alpha_p:=\mu_p((\partial_{x_1}\wedge\dots\wedge\partial_{x_n})(p))$ is the unique complex number such that $\mu_p=\mu^\alpha_p\,|dx^1\wedge\dots\wedge dx^n|^q(p)$. For $(y,w)\in\psi_\beta(V_\alpha\cap V_\beta)\times\mathbb{R}$ we obtain from eq. (4) (setting $x=\psi_{\alpha\beta}(y)$))

$$\Psi_\alpha\circ\Psi_\beta^{-1}(y,w)=(\psi_{\alpha\beta}(y),w\left|\det D(\psi_\beta\circ\psi_\alpha^{-1})(x)\right|^q)\,.\tag{5}$$

In other words, the $\Psi_\alpha$ form a vector bundle atlas for $E$ reproducing exactly the cocycle of transition functions given by (3), so $E$ indeed can be taken to be $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$. To obtain the transformation behavior for the local expression $\mu_\alpha=(\Psi_\alpha^{-1})^*\mu|_{V_\alpha}$ of a $q$-density $\mu\in\Gamma^k(X,\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$, note that $\mu|_{V_\alpha}=(\mu_\alpha\circ\psi_\alpha)\,|dx^1\wedge\dots\wedge dx^n|^q$. Now eq. (5) immediately yields

$$\mu_\beta\circ\psi_{\beta\alpha}(x)\ =\ |\det(D(\psi_\beta\circ\psi_\alpha^{-1}))|^{-q}(x)\mu_\alpha(x)\ =\ |\det\frac{\partial y^i}{\partial x^j}(x)|^{-q}\mu_\alpha(x)\,.$$

From these constructions we also see that $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ is trivial: We can glue together positive local sections using a partition of unity to obtain a $q$-density since positivity is preserved under transitions. However, there is no natural trivialization.

**1.6 Example.** An important example of a *one-density* is the *volume of a Pseudo-Riemannian metric $g$ on $X$*. In a chart $(V_\alpha,\psi_\alpha)$ as above we may write $g=g^\alpha_{ij}dx^idx^j$, where the functions $g^\alpha_{ij}=g(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j})$ on $V_\alpha$ form a non-degenerate symmetric matrix. Hence $|\det(g^\alpha_{ij})|>0$ and we define on $V_\alpha$

$$dg^\alpha:=\sqrt{|\det g^\alpha_{ij}\,|}\,.$$

Consequently, on $V_\alpha\cap V_\beta$ ($(V_\beta,\psi_\beta)$ a chart with coordinates $y^i$) we have

$$\sqrt{\left|\det g(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j})\right|}=\sqrt{\left|\det g(\frac{\partial y^k}{\partial x^i}\frac{\partial}{\partial y^k},\frac{\partial y^l}{\partial x^j}\frac{\partial}{\partial y^l})\right|}=\left|\det\frac{\partial y^k}{\partial x^i}\right|\sqrt{\left|\det g(\frac{\partial}{\partial y^i},\frac{\partial}{\partial y^j})\right|}\,.$$

Hence these local representatives define a section in $\mathrm{Vol}\,(\mathrm{X})$ called the *volume (or density) of the metric $g$*. Next we turn to the notion of the *integral of a one-density.*

**1.7 Definition.** *Let $\mu \in \Gamma_c^0(\mathrm{Vol}\,(\mathrm{X}))$, $(V_\alpha, \psi_\alpha)_\alpha$ an atlas ($V_\alpha$ relatively compact) for $X$ and $(f_\alpha)_\alpha$ a partition of unity subordinate to $(V_\alpha)_\alpha$ (cf. e.g. [49], §2 cor. 2). We define the* integral *of $\mu$ by*

$$\int_X \mu = \sum_\alpha \int_{V_\alpha} f_\alpha \mu := \sum \int_{\psi_\alpha(V_\alpha)} f_\alpha(\psi_\alpha^{-1}(x))\,\mu_\alpha(x)\,dx\ ,$$

*where $\mu_\alpha = \Psi_\alpha \circ \mu \circ \psi_\alpha^{-1}$ again denotes the component of $\mu$ with respect to $(V_\alpha, \Psi_\alpha)$.*
This definition is independent of the choice of the atlas and the partition of unity since the transition functions for one-densities are adapted to the transformation formula for multiple integrals. More precisely, let $(V_\alpha, \Psi_\alpha)_\alpha$ and $(U_\beta, \Phi_\beta)_\beta$ be vector bundle atlasses induced by the atlasses $(V_\alpha, \psi_\alpha)_\alpha$ and $(U_\beta, \phi_\beta)_\beta$ on the base with subordinate partitions of unity $(f_\alpha)_\alpha$ and $(g_\beta)_\beta$ respectively, then we have

$$\begin{aligned}
\sum_\alpha \int\limits_{V_\alpha} f_\alpha \mu &= \sum_\alpha \int\limits_{\psi_\alpha(V_\alpha)} f_\alpha(\psi_\alpha^{-1}(x))\,\boldsymbol{\psi_\alpha}(\mu(\psi_\alpha^{-1}(x)))\,dx \\
&= \sum_\alpha \int\limits_{\psi_\alpha(V_\alpha)} f_\alpha(\psi_\alpha^{-1}(x))\,\sum_\beta g_\beta(\psi_\alpha^{-1}(x))\,\boldsymbol{\psi_\alpha}(\mu(\psi_\alpha^{-1}(x)))\,dx \\
&= \sum_{\alpha\beta} \int\limits_{\psi_\alpha(U_\alpha \cap V_\beta)} f_\alpha \circ \psi_\alpha^{-1}(x)\, g_\beta \circ \psi_\alpha^{-1}(x)\,\boldsymbol{\psi_\alpha}(\mu(\psi_\alpha^{-1}(x)))\,dx \\
&= \sum_{\alpha\beta} \int\limits_{\phi_\beta(U_\alpha \cap V_\beta)} f_\alpha \circ \phi_\beta^{-1}(y)\, g_\beta \circ \phi_\beta^{-1}(y)\,\boldsymbol{\psi_\alpha}(\mu(\phi_\beta^{-1}(y)))\,|\det D(\psi_\alpha \circ \phi_\beta^{-1})(y)|dy \\
&= \sum_{\alpha\beta} \int\limits_{\phi_\beta(U_\alpha \cap V_\beta)} f_\alpha \circ \phi_\beta^{-1}(y)\, g_\beta \circ \phi_\beta^{-1}(y)\,\boldsymbol{\phi_\beta}(\mu(\phi_\beta^{-1}(y)))\,dy \\
&= \sum_\beta \int\limits_{\phi_\beta(U_\beta)} g_\beta(\phi_\beta^{-1}(y))\,\boldsymbol{\phi_\beta}(\mu(\phi_\beta^{-1}(y)))\,dy\ = \\
\sum_\beta \int\limits_{V_\beta} g_\beta \mu\ .
\end{aligned}$$

**1.8** To deal with non-compactly supported densities we first introduce the notion of *non-negativity.* Clearly every $q$-density $\mu$ on a vector space $W$ takes values in some half-line $\{re^{i\gamma} : r \geq 0\}$; we call $\mu$ non-negative if $\gamma = 0 \pmod{2\pi}$. Note that the absolute value $|\mu|$ of $\mu$ is a non-negative $q$-density on $W$. Extending this notion to the case $\mu \in \Gamma^0(\mathrm{Vol}\,(\mathrm{X}))$ in the natural way we may define, for $\mu$ non-negative, $\int \mu := \sum_\alpha \int_{\psi_\alpha(V_\alpha)} f_\alpha \circ \psi_\alpha^{-1}(x)\mu_\alpha(x)dx$ with $(V_\alpha, \psi_\alpha)$ and $f_\alpha$ as above. By the preceding argument $\int \mu\ (\in [0,\infty) \cup \infty)$ is independent of the choice of $(V_\alpha, \psi_\alpha)$ and $f_\alpha$. Finally let $\mu \in \Gamma^0(\mathrm{Vol}\,(\mathrm{X}))$ be arbitrary; if $\int |\mu| < \infty$ we set

$$\int \mu = \sum_\alpha \int_{\psi_\alpha(V_\alpha)} f_\alpha(\psi_\alpha^{-1}(x))\mu_\alpha(x)\,dx\,.$$

Due to $|\int f_\alpha \circ \psi_\alpha^{-1}\mu_\alpha| \leq \int |f_\alpha \circ \psi_\alpha^{-1}||\mu_\alpha|$ the series is absolutely convergent and again $\int \mu$ is independent of the choice of the atlas and partition of unity.
Let $\mu \in \Gamma^0(\mathrm{Vol}\,(\mathrm{X}))$ be an arbitrary one-density and $f$ a continuous function on $X$ with compact support, then we define the *integral of $f$ with respect to $\mu$* by $\int_X f\mu$. Hence $\mu$ defines a Radon measure on $X$. For the converse direction see [139], 8.3.

**1.9** Note that $X$ need not be orientable in order for the integral of a one-density to be well-defined. In most textbooks on general relativity however, $n$-forms are integrated on $n$-dimensional *oriented* manifolds. We close this section by recalling some facts on the notion of orientation in rather abstract language, thereby clarifying the interrelations between one-densities and $n$-forms. For the remainder of this section we assume all bundles to be real.
A manifold $X$ is called *orientable* if there exists an *oriented atlas*, i.e., an atlas $(V_\alpha, \psi_\alpha)_\alpha$ such that $\det(D(\psi_\alpha \circ \psi_\beta^{-1})(y)) > 0$ for all $y \in \psi_\beta(V_\alpha \cap V_\beta)$. More abstractly one defines the *orientation bundle $Or(X)$* over $X$ to be the line bundle given by the transition functions

$$\boldsymbol{\tau_{\alpha\beta}}(x) = \operatorname{sign}\det(D(\psi_\beta \circ \psi_\alpha^{-1})(x)) \,. \tag{6}$$

$X$ is orientable iff the orientation bundle is trivial ( [139], 8.7). In the case of $X$ being connected there are two distinguished global frames in $\mathrm{Or}(X)$ (i.e., sections of absolute value one; observe that the absolute value of a section in $\mathrm{Or}(X)$ is well-defined), corresponding to the two possible *orientations* of $X$; specifying one of these makes $X$ an *oriented manifold*.

On the other hand $n$-forms are sections of the line bundle $\wedge^n T^*X$, which may be characterized by the following cocycle of transitions functions

$$\boldsymbol{\phi_{\alpha\beta}}(x) = \det(D(\psi_\beta \circ \psi_\alpha^{-1})(x)) \,. \tag{7}$$

We denote the space of $n$-forms on $X$ by $\Omega^n(X)$. A brief look at equations (3) (setting $q = 1$), (6) and (7) reveals

$$\boldsymbol{\tau}_{\alpha\beta}\, \boldsymbol{\psi}_{\alpha\beta} = \boldsymbol{\phi}_{\alpha\beta} \,, \tag{8}$$

hence we have for the bundles $\wedge^n T^*X \;=\; \mathrm{Or}(X) \otimes \mathrm{Vol}\,(\mathrm{X})$. Here, equality of vector bundles amounts to equality of the respective transition functions with respect to the "natural" vector bundle atlasses (which is stronger than just stating the vector bundles to be isomorphic).

In the case of an oriented manifold, i.e., $\tau_{\alpha\beta} = 1$ with respect to an oriented atlas we even have

$$\wedge^n T^*X \;=\; \mathrm{Vol}\,(\mathrm{X})$$

(which locally amounts to $|d^n x| = d^n x$). Hence an orientation allows a canonical identification of $n$-forms and one-densities and we may define the integral of $\omega \in \Gamma^0_c(\wedge^n T^*X)$ by the integral of the associated density. Explicitly, in an oriented atlas $(V_\alpha, \psi_\alpha)_\alpha$ we have

$$\int\limits_M \omega = \sum_\alpha \int\limits_{\psi_\alpha(V_\alpha)} \boldsymbol{\phi_\alpha}(\omega(\psi_\alpha^{-1}(x)))\, dx \,.$$

Note that this definition of the integral precisely coincides with the usual one, based on chartwise integration of the components of the $n$-form on $\psi_\alpha(V_\alpha)$. Moreover, in the case of an oriented manifold $X$ with boundary (denoted by $\partial X$) we have Stokes theorem (e.g. [139], 8.11) guaranteeing that for any $n-1$ form $\omega$

$$\int\limits_X d\omega \;=\; \int\limits_{\partial X} \omega \,.$$

Also note that on an orientable manifold $X$ we may use eq. (8) together with the triviality of $\mathrm{Vol}\,(\mathrm{X})$ to construct nowhere vanishing $n$-forms. If $\theta$ is such an $n$-form we may orient $X$ by stipulating $\theta > 0$; in that case we say that $\theta$ is inducing *the orientation* of $X$.

### 1.C. Test Fields and Distributions

**1.10** In this section we are going to define spaces of test objects and to topologize them appropriately to enable the definition of distributional spaces as their topological duals. First note that the family of bundles $\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ has the following properties (for a proof see e.g. [169], prop. 1.4.4)

$$\begin{aligned} &\text{(i)} \quad \mathrm{Vol}^{\,\mathrm{p}}(\mathrm{X}) \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}) = \mathrm{Vol}^{\,\mathrm{p+q}}(\mathrm{X}) \\ &\text{(ii)} \quad (\mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))^* = \mathrm{Vol}^{\,-\mathrm{q}}(\mathrm{X})\,. \end{aligned} \tag{9}$$

Now let $\mu$, $\nu$ be $\mathcal{C}^k$-densities of order $q$ and $1-q$ respectively, at least one of them compactly supported. Their tensor product $\mu\nu$ is a one-density with compact support, hence can be integrated. Therefore it is reasonable (cf. sec. 1.A) to consider compactly supported sections in the bundle $(E^* \otimes \mathrm{Vol}^{\,1-\mathrm{q}}(\mathrm{X}), \mathrm{X}, \pi)$, ($E$ an arbitrary vector bundle over $X$) as test objects.
Our next task is to define a suitable topology for the spaces $\Gamma^k_c(F)$, where $(F, X, \pi)$ is an arbitrary vector bundle over $X$. To this end we closely follow [117], chap. 1.9. A detailed discussion of the topological concepts to be used in the sequel can be found in [87], chaps. 2,3 or [19], chaps. 1,2 (in the special case of functions on (open sets of) $\mathbb{R}^n$).

**1.11** Denote by $\Omega$ an open set in $\mathbb{R}^n$ and recall that the spaces $\mathcal{C}^\infty(\Omega)$ (respectively $\mathcal{C}^k(\Omega)$) are *Fréchet spaces* with respect to the topology of uniform convergence in all (respectively all $k$) derivatives on compact subsets of $\Omega$. Now on $\Gamma^k(V, F)$ (the local sections over the open set $V \subseteq X$) we introduce a locally convex topology by the following notion of convergence: A sequence $(u_l)_l$ of local sections converges to zero iff

$$\forall K \subset\subset \psi(V),\, \forall \nu \in \mathbb{N}_0^n,\, |\nu| \le k \text{ (if } k < \infty) \;\; \partial^\nu(\psi^j \circ u_l \circ \psi^{-1}) \to 0 \text{ uniformly on } K \quad (1 \le j \le N)$$

Such a topology exists ( [53], chap. 17.2) and is generated by the following family of semi-norms: Choose a countable sub family $(V_\alpha, \Psi_\alpha)_\alpha$ of a vector bundle atlas such that $(V_\alpha)_\alpha$ is a locally finite cover of $V$ and denote by $(K^\alpha_m)_m$ a fundamental sequence of compact subsets of $\psi_\alpha(V_\alpha)$. Now we define

$$p_{s,m,\alpha}(u) := \sum_{j=1}^N \sup_{x \in K^\alpha_m, |\nu| \le s} |\partial^\nu(\psi^j_\alpha \circ u|_{V_\alpha} \circ \psi^{-1}_\alpha(x))|\,, \tag{10}$$

where $s \in \mathbb{N}_0$ and $m,\, \alpha \in \mathbb{N}$. Note that $(u_l)_l$ converges to zero in $\Gamma^k(V, F)$ iff $(u_l)_l \to 0$ in all $\Gamma^k(V_\alpha, F)$ and by definition $\Gamma^k(V_\alpha, F)$ is isomorphic to $\mathcal{C}^k(\psi_\alpha(V_\alpha))^N$. Hence convergence of sections just means convergence of coefficient functions.

**1.12** In the case $F$ is some tensor bundle $\mathcal{T}^r_s(X)$ the same topology may also be constructed in an explicitly covariant way ( [31], sec. VII.8). To achieve this endow $X$ with a smooth Riemannian metric $h$ and denote its metric covariant derivative by $\nabla$. We recall that the pointwise norm of a section $u \in \Gamma^k(F)$ with respect to $h$ is defined by $||u(p)|| := |u^{i_1 \dots i_N}(p)\, u_{i_1 \dots i_N}(p)|^{1/2}$, where we have used abstract index notation (cf. [187], pp. 23ff and 2.14 below) and indices are raised and lowered by $h$ (hence $u^{i_1 \cdots i_N} u_{i_1 \dots i_N}$ denotes the total metric contraction of $u$). Now we are ready to define the semi-norms

$$p_{K,s}(u) := \sup_{p \in K, l \le s} ||\nabla^{(l)} u(p)||,\,, \tag{11}$$

where $K$ is compact in $X$ and $l, s$ are nonnegative integers. The locally convex topology induced by this family of semi-norms is independent of the choice of the Riemannian metric $h$. Note that (10) is just the coordinate expression for the particular choice $h = \mathrm{diag}(1, \dots, 1)$ in (11).

**1.13** The topology introduced above turns the spaces $\Gamma^k(V, F)$ into separable Fréchet spaces ( [53], 17.2.2). Note that for the trivial bundle $X \times \mathbb{C}$ we just have $\Gamma^k(V, F) = \mathcal{C}^k(V)$. In the relative topology the subspace of compactly supported sections is dense in $\Gamma^k(V, F)$; therefore we have to find a finer topology turning these spaces into complete ones. Let $K$ denote any compact subset of the base $X$ and set $\Gamma^k_{c,K}(F) := \{u \in \Gamma^k(F) : \mathrm{supp(u)} \subseteq \mathrm{K}\}$. $\Gamma^k_{c,K}(F)$ is a closed subspace of the Fréchet space $\Gamma^k(F)$, hence a Fréchet space itself. Since for compact subsets $K_1 \subseteq K_2$ the spaces of sections $\Gamma^k_{c,K_1}(F)$ are contained continuously in $\Gamma^k_{c,K_2}(F)$ and $\Gamma^k_c(F) = \bigcup_{K \subset\subset X} \Gamma^k_{c,K}(F)$ it is reasonable to endow $\Gamma^k_c(F)$ with the *inductive limit topology* of the $\Gamma^k_{c,K}(F)$. Hence setting

$$\Gamma^k_c(F) \;=\; \varinjlim \Gamma^k_{c,K}(F)$$

turns the space of compactly supported $\mathcal{C}^k$-sections of $F$ into a *strict (LF)-space* (cf. [116], chap. 7).

Finally we are ready to give the "geometric"

**1.14 Definition.** *Let $E$ denote an arbitrary vector bundle over $X$.* Distributions (of order $k$) with values in $E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X})$ *or* $E$-valued distributions of density character $q$ *are continuous linear forms on* $\Gamma_c(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$ *(respectively* $\Gamma_c^k(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$*). We denote the space of $E$-valued distributions of density character $q$ by $\mathcal{D}'(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X}))$ (respectively $\mathcal{D}'^{(k)}(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X}))$), i.e.,*

$$\mathcal{D}'^{(k)}(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X})) := [\Gamma_{\mathrm{c}}^{\mathrm{k}}(\mathrm{X}, \mathrm{E}^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X})]'\,. \tag{12}$$

*If $q = 0$ the corresponding spaces will be called* section distributions *of the bundle $E$ and denoted by $\mathcal{D}'^{(k)}(X, E)$ while in the case $q = 1$ the term* section distribution densities *and the notation $\mathcal{D}_d'^{(k)}(X, E)$ will be used.*

Some comments on notation and special cases of this definition are in order. If we set $E = \Omega \times \mathbb{C}$, $X = \Omega$ an open set in $\mathbb{R}^n$, we recover the usual distributional spaces which we simply denote by $\mathcal{D}'^{(k)}(\Omega)$, the choice of $q$ being irrelevant here (due to the existence of a natural chart). If $E = X \times \mathbb{C}$, $k = \infty$ and $q = 0$ we have the space of *distributions on the manifold* $X$, which we shall denote by $\mathcal{D}'(X)$. Again with $E = X \times \mathbb{C}$, $k = \infty$ but $q = 1$ we have *distributional densities on* $X$, denoted by $\mathcal{D}_d'(X)$. If $E$ is some tensor bundle $\mathcal{T}_s^r(X)$ and $q = 0$ resp. $q = 1$ we use the term *tensor distribution* resp. *tensor distribution density* and denote the respective spaces by $\mathcal{D}'{}^r_s(X)$ resp. $\mathcal{D}_d'{}^r_s(X)$. Given $T \in \mathcal{D}'^{(k)}(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X}))$ and $u \in \Gamma_c^k(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$ we denote the action of $T$ on $u$ synonymously by $\langle T, u\rangle$ or $T(u)$.

Basic properties of inductive limit topologies yield the following criterion for a linear form on $\Gamma_c^k(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$ to be a distribution (cf. [117], 1.9.3).

**1.15 Proposition.** *For an element $T$ in the algebraic dual of $\Gamma_c^k(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$ the following conditions are equivalent*

- *(i) $T$ is a distribution of order $k$ taking values in $E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X})$*
- *(ii) $T|_{\Gamma_{c,K}^k(X, E^* \otimes \mathrm{Vol}^{1-q}(X))}$ is continuous for each $K \subset\subset X$.*
- *(iii) For every $K \subset\subset X$ there are numbers $s \in \mathbb{N}_0$, $s \le k$, $l, m \in \mathbb{N}$, $c_K > 0$ and indices $\alpha_1, \dots, \alpha_l$ such that $|T(u)| \le c_K \sup_i p_{s,m,\alpha_i}(u)$ for all $u \in \Gamma_{c,K}^k(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$.*
- *(iv) For every $K \subset\subset X$ there are numbers $s \in \mathbb{N}_0$, $s \le k$ and $c_K > 0$ such that $|T(u)| \le c_K p_{K,s}(u)$ for all $u \in \Gamma_{c,K}^k(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$.*

**1.16** Next we turn to the issue of *regular distributions* in $\mathcal{D}'(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X}))$. Denote by $\mathrm{tr}_{\mathrm{E}}$ the canonical vector bundle homomorphism $E \otimes E^* \to X \times \mathbb{C}$ induced by the pointwise dual action of $v^* \in E_p^*$ on $v \in E_p$ and by $\Pi_{q,q'}$ the product of $q$- and $q'$-densities. The vector bundle homomorphism

$$(\,.\,|\,.\,) := \mathrm{tr}_{\mathrm{E}} \otimes \Pi_{\mathrm{q},1-\mathrm{q}} : \mathrm{E} \otimes \mathrm{E}^* \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X}) \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}) \to (\mathrm{X} \times \mathbb{C}) \otimes \mathrm{Vol}\,(\mathrm{X}) = \mathrm{Vol}\,(\mathrm{X})$$

in turn allows us to assign to a pair $(f, u) \in \Gamma^0(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X})) \times \Gamma(\mathrm{X}, \mathrm{E}^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$ a density which may be integrated over $X$ provided it is compactly supported. Hence (by a slight abuse of notation) we define the action of $f \in \Gamma^0(X, E \otimes \mathrm{Vol}^{\mathrm{q}}(\mathrm{X}))$ on a section $u \in \Gamma_c(X, E^* \otimes \mathrm{Vol}^{1-\mathrm{q}}(\mathrm{X}))$ by

$$\langle f, u\rangle := \int_X (f|u)\,.$$

By the properties of the integral the map $T_f := (f|\,.\,)$ is linear and continuous, hence defines a distribution. Moreover the assignment $f \to T_f$ is linear and one to one. Summing up we have the following

**1.17 Proposition and Definition.** *The space of locally integrable sections of* $E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ *is (naturally) embedded into* $\mathcal{D}'(X, \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$ *by*

$$i : f \mapsto T_f : \quad \langle T_f, u \rangle := \int\limits_X (f \,|\, u) \qquad \forall u \in \Gamma_c^\infty(X, E^* \otimes \mathrm{Vol}^{\,1-\mathrm{q}}(\mathrm{X}))\,.$$

*A distribution* $T \in \mathcal{D}'(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$ *is called* regular *if there exists a locally integrable section* $f$ *of* $E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ *such that*

$$T = T_f\ .$$

We close this section stating a theorem that guarantees the "natural" localization properties for the distributional spaces allowing for a local description which is the topic of the next section.

**1.18 Theorem.** $\mathcal{D}'^{(k)}(V, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$ *is a fine sheaf of* $\mathcal{C}^\infty(X)$*-modules.*
A proof for spaces of currents (see 1.29 below) may be found in [53], 17.4.2. The general case can be proved analogously.

### 1.D. Local Description

**1.19** The aim of this section is to work out the details of the local description of distributions with values in $E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ enabling a simple operational approach to the geometric definition 1.14. We shall decompose distributions $T \in \mathcal{D}'(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$ into local families $(T_{\alpha\, i})_\alpha$ $(i = 1, \ldots, N)$, each $T_{\alpha\, i}$ being a distribution on an open set of $\mathbb{R}^n$, characterized by their transformation behavior. Moreover we show that locally distributions with values in $E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ are just sections in the bundle $E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})$ with distributional coefficients, i.e., for any chart $V_\alpha$, $\mathcal{D}'^{(k)}(V_\alpha, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})) \cong \mathcal{D}'^{(\mathrm{k})}(\mathrm{X}) \otimes \Gamma^\infty(\mathrm{V}_\alpha, \mathrm{E} \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$.
For simplicity we only treat the case $k = \infty$ and leave the obvious modifications for distributions of finite order to the reader.

**1.20** To begin with, recall the definition of $\Phi_* : \Gamma(E) \to \Gamma(F)$, assigning $\Phi \circ u \circ \varphi^{-1} \in \Gamma(F)$ to $u \in \Gamma(E)$ where $\Phi : E \to F$ is a vector bundle homomorphism along the diffeomorphism $\varphi : X \to Y$. Restricting $\Phi_*$ to sections with compact support we obtain a map (also denoted by $\Phi_*$)

$$\Phi_* : \Gamma_c(E) \to \Gamma_c(F)\,.$$

We have

**1.21 Lemma.** $\Phi_* : \Gamma_c(E) \to \Gamma_c(F)$ *is continuous w.r.t. the respective (LF)-topologies.*

**Proof.** Obviously, $\Phi_* : \Gamma_c(E) \to \Gamma_c(F)$ is linear and bijective with inverse $\Phi^* : v \mapsto \Phi^{-1} \circ v \circ \varphi$. Since the spaces $\Gamma_c(\ldots)$ are endowed with the respective inductive limit topologies, it suffices to prove continuity on the subspaces $\Gamma_{c,K}(E)$ resp. $\Gamma_{c,\varphi(K)}(F)$ with $K \subseteq\subseteq X$. But this is immediate since the generating sets of semi norms are respected by $\Phi$ and $\varphi$. □

If, in addition to the above, $\Phi$ is a diffeomorphism, $\Phi_*$ is an isomorphism of (LF)-spaces.

**1.22** Considering a vector bundle homomorphism $\Phi : E \to F$ again, the adjoint map of $\Phi_*$, that is

$$(\Phi_*)' : \Gamma_c(F)' \to \Gamma_c(E)'$$
$$\langle (\Phi_*)' T, u \rangle := \langle T, \Phi_* u \rangle = \langle T, \Phi \circ u \circ \varphi^{-1} \rangle$$

is weakly continuous and strongly continuous ( [87], corollary following 3.12, prop. 3). Moreover, $(\Phi_1 \circ \Phi_2)'_* = (\Phi_2)'_* \circ (\Phi_1)'_*$ and $\mathrm{id}'_* = \mathrm{id}$ (provided $\varphi = \mathrm{id}$). If in addition, $\Phi$ is a diffeomorphism, we set

$$\Phi^\wedge := ((\Phi^{-1})_*)' : \Gamma_c(E)' \to \Gamma_c(F)'$$
$$\langle \Phi^\wedge T, u \rangle := \langle T, \Phi_*^{-1} u \rangle = \langle T, \Phi^{-1} \circ u \circ \varphi \rangle\,.$$

Clearly, $\Phi^\wedge$ is an isomorphism with respect to the weak as well as the strong topologies. Note that $(\Phi_1 \circ \Phi_2)^\wedge = \Phi_1^\wedge \circ \Phi_2^\wedge$ and $\mathrm{id}^\wedge = \mathrm{id}$ (provided $\varphi = \mathrm{id}$, again).

**1.23** In particular, let $(V_\alpha, \Psi_\alpha)$ denote a chart for a vector bundle $(F, X, \pi)$. By 1.21, $(\Psi_\alpha)_* = (\Psi_\alpha^{-1})^* : \Gamma_c(V_\alpha, F) \to \Gamma_c(\psi_\alpha(V_\alpha), \mathbb{R}^N)$ is an isomorphism of (LF)-spaces. Setting $F = E^* \otimes \mathrm{Vol}^{1-q}(X)$, we obtain that

$$\begin{aligned} \Psi_\alpha^\wedge : \mathcal{D}'(V_\alpha, E \otimes \mathrm{Vol}^q(X)) \;&\to\; \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^N) \\ \langle \Psi_\alpha^\wedge(T), u\rangle \;&:=\; \langle T, \Psi_\alpha^*(u)\rangle \end{aligned} \tag{13}$$

is a topological isomorphism of the respective distributional spaces.
So to any distribution $T \in \mathcal{D}'(X, E \otimes \mathrm{Vol}^q(X))$, we may assign the family $(T_\alpha)_\alpha$ of the respective *local expressions* $T_\alpha := \Psi_\alpha^\wedge(T|_{V_\alpha}) \in \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^N)$ satisfying

$$\langle T, u\rangle \;=\; \langle T_\alpha, u_\alpha\rangle \;=\; \sum_{i=1}^N \langle T_{\alpha\, i}, u_\alpha^i\rangle$$

for $\operatorname{supp} u \in V_\alpha$ where $T_{\alpha\, i} \in \mathcal{D}'(\psi_\alpha(V_\alpha))$ $(i = 1, \dots, N)$ simply are distributions on an open subset of $\mathbb{R}^n$.
For $u \in \Gamma_c(\psi_\alpha(V_\alpha \cap V_\beta), \mathbb{R}^N)$ we have

$$\langle T_\alpha, u\rangle \;=\; \langle T, \Psi_\alpha^* u\rangle \;=\; \langle T, \Psi_\beta^* \circ (\Psi_\alpha \circ \Psi_\beta^{-1})^* u\rangle \;=\; \langle T_\beta, (\Psi_\alpha \circ \Psi_\beta^{-1})^* u\rangle \;=\; \langle (\Psi_\alpha \circ \Psi_\beta^{-1})^\wedge(T_\beta), u\rangle\,, \text{ so}$$

$$T_\alpha = (\Psi_\alpha \circ \Psi_\beta^{-1})^\wedge(T_\beta) \tag{14}$$

Note that (14) of course also holds for any $T$ in the *algebraic* dual $\Gamma_c^\infty(X, E^* \otimes \mathrm{Vol}^{1-q}(X))^*$ of $\Gamma_c^\infty(X, E^* \otimes \mathrm{Vol}^{1-q}(X))$.
Conversely, suppose we are given a family $(T_\alpha)_\alpha$ satisfying the transformation rule (14), where $T_\alpha \in \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^N)$. Then there exists a unique $T \in \mathcal{D}'(X, E \otimes \mathrm{Vol}^q(X))$ such that $\Psi_\alpha^\wedge(T|_{V_\alpha}) = T_\alpha$ for all $\alpha$. Indeed, setting $\widetilde{T}_\alpha = (\Psi_\alpha^{-1})^\wedge T_\alpha$, by 1.18 it is enough to prove that $\widetilde{T}_\alpha|_{V_\alpha \cap V_\beta} = \widetilde{T}_\beta|_{V_\alpha \cap V_\beta}$ for all $\alpha, \beta$. Now if $u \in \Gamma_c(V_\alpha \cap V_\beta, E^* \otimes \mathrm{Vol}^{1-q}(X))$ it follows that

$$\langle \tilde{T}_\alpha, u\rangle \;=\; \langle (\Psi_\alpha^{-1})^\wedge(T_\alpha), u\rangle \;\overset{(14)}{=}\; \langle (\Psi_\alpha^{-1})^\wedge(\Psi_\alpha \circ \Psi_\beta^{-1})^\wedge(T_\beta), u\rangle \;=\; \langle \tilde{T}_\beta, u\rangle \tag{15}$$

verifying our claim.

Summing up, we have proved the following

**1.24 Theorem.** *Distributions with values in $E \otimes \mathrm{Vol}^q(X)$ can be identified with families $(T_\alpha)_\alpha$ of distributions $T_\alpha \in \mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R}^N)$ satisfying the transformation law (14).*

**1.25** To obtain an even more explicit description of $E$-valued distributions of density character $q$ we consider the bilinear map

$$\begin{aligned} \beta_0 :\; D'(X) \times \Gamma(X, E \otimes \mathrm{Vol}^q(X)) &\to \mathcal{D}'(X, E \otimes \mathrm{Vol}^q(X)) \\ \langle \beta_0(T, z), u\rangle &:= \langle T, (z|u)\rangle\,, \end{aligned}$$

for $u \in \Gamma_c(X, E^* \otimes \mathrm{Vol}^{1-q}(X))$. By the usual algebraic techniques $\beta_0$ induces a linear map

$$\beta :\; \mathcal{D}'(X) \otimes_\infty \Gamma(X, E \otimes \mathrm{Vol}^q(X)) \to \mathcal{D}'(X, E \otimes \mathrm{Vol}^q(X))\,,$$

where $\otimes_\infty$ denotes the balanced tensor product over the module of $\mathcal{C}^\infty$-functions on $X$. Now we have the following

**1.26 Proposition.** *Locally $\beta$ is an isomorphism of $\mathcal{C}^\infty$-modules, i.e., for any chart $V_\alpha$ we have*

$$\mathcal{D}'(V_\alpha) \otimes_\infty \Gamma(V_\alpha, E \otimes \mathrm{Vol}^q(X)) \;\cong \mathcal{D}'(V_\alpha, E \otimes \mathrm{Vol}^q(X))\,.$$

**Proof.** To simplify notation we only treat the case $q = 0$ explicitly and leave the obvious modifications for general density characters to the reader.
Surjectivity: First note that any $u \in \Gamma_c(V_\alpha, E^* \otimes \mathrm{Vol}\,(\mathrm{X}))$ may be written in the form $u = \sum_{i=1}^N u_i \otimes \mu_i$ with $u_i := (\Psi_{E^*})^*_\alpha(e_i)$, where by $(\Psi_{E^*})_\alpha$ and $(e_j)_{j=1}^N$ we denote the bundle charts in $E^*$ and the standard basis of $\mathbb{R}^N$ respectively, and $\mu_i \in \Gamma_c(V_\alpha, \mathrm{Vol}\,(\mathrm{X}))$. Let $v_j \in \Gamma(V_\alpha, E)$ be dual to $u_j$, then we have

$$(v_j|u) \;=\; \sum_{i=1}^N (v_j|u_i \otimes \mu_i) \;=\; \sum_{i=1}^N \delta_{ij}\mu_i \;=\; \mu_j\,.$$

Now let $T \in \mathcal{D}'(V_\alpha, E)$. Defining $t_i \in \mathcal{D}'(V_\alpha)$ $(1 \le i \le N)$ by $\langle t_i, \mu\rangle := \langle T, u_i \otimes \mu\rangle$ we finally obtain

$$\langle T, u\rangle \;=\; \sum_{i=1}^N \langle T, u_i \otimes \mu_i\rangle \;=\; \sum_{i=1}^N \langle t_i, \mu_i\rangle \;=\; \sum_{i=1}^N \langle t_i, (v_i|u)\rangle \;=\; \sum_{i=1}^N \langle \beta_0(t_i, v_i), u\rangle\,.$$

Hence $\mathrm{span}(\mathrm{im}(\beta_0)) = \mathrm{im}(\beta) = \mathcal{D}'(\mathrm{V}_\alpha, \mathrm{E})$.
Injectivity: We denote the classes in the $\mathcal{C}^\infty(V_\alpha)$-balanced tensor product by square brackets. Let $[S] = [\sum_{i=1}^N T_i \otimes z_i] \in \mathcal{D}'(V_\alpha) \otimes_\infty \Gamma(V_\alpha, E)$. Expanding $z_i$ with respect to the natural basis $(v_j)_{j=1}^N$ as above, i.e., $z_i = \sum_{j=1}^N z_{ij} v_j$ we find

$$[S] \;=\; [\sum_{i=1}^N T_i \otimes \sum_{j=1}^N z_{ij} v_j] \;=\; [\sum_{i,j=1}^N z_{ij} T_i \otimes v_j]\,.$$

Assuming $\beta[S]$ to vanish we have to show that $t_j := \sum_{i=1}^N z_{ij} T_i = 0 \in \mathcal{D}'(V_\alpha)$. By assumption we have for all $u \in \Gamma_c(V_\alpha, E^* \otimes \mathrm{Vol}\,(\mathrm{X}))$

$$0 \;=\; \sum_{i=1}^N \langle T_i, (z_i|u)\rangle \;=\; \sum_{j=1}^N \langle t_j, (v_j|u)\rangle\,.$$

Choosing $u = u_k \otimes \mu_k$ with $u_k$ dual to $v_k$ and $\mu_k \in \Gamma_c(V_\alpha, \mathrm{Vol}\,(\mathrm{X}))$ we obtain $t_k = 0 \in \mathcal{D}'(X)$. Noting that $k$ and $\mu_k$ were arbitrary finishes the proof. □
We remark that since the above proof of surjectivity may be easily extended to the whole of $X$ using a partition of unity subordinate to the charts of a finite atlas $\beta$ is onto in general. Moreover the proof of injectivity also holds globally provided $E$ is trivial.

**1.27** It is a source of many important results in distribution theory on $\mathbb{R}^n$ that smooth regular objects are sequentially dense in $\mathcal{D}'$. The following result carries this property over to vector bundle distributions. We are going to use the following notation: Let $\langle G, H\rangle$ be a dual pairing of vector spaces and let $A \subseteq H$ $(\subseteq G^*)$. $\overline{A}^{wsc}$ denotes the weak sequential closure of $A$ in $G^*$, i.e., $\overline{A}^{wsc} = \{u \in G^* | \exists u_n \in A,\, \langle u_n, \varphi\rangle \to \langle u, \varphi\rangle\, \forall \varphi \in G\}$.

**1.28 Theorem.**

$$\overline{i(\Gamma^\infty(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X})))}^{wsc} = \mathcal{D}'(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))\,.$$

**Proof.** $\subseteq$: We will even show that $\mathcal{D}'(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$ is weakly sequentially complete: Let $(T_m)_{m \in \mathbb{N}}$ be a sequence in $\mathcal{D}'(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$ with $T_m(u) \to T(u)$ for all $u \in \Gamma_c^\infty(X, E^* \otimes \mathrm{Vol}^{\,1-\mathrm{q}}(\mathrm{X}))$. In order to show that $T \in \mathcal{D}'(X, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))$, by 1.24 and the remark following (14) it suffices to show that each $T_{\alpha\, i}$ is an element of $\mathcal{D}'(\psi_\alpha(V_\alpha), \mathbb{R})$. However, this is clear since $T_{\alpha\, i}$ is the weak limit of $T_{m\,\alpha\, i}$ which by themselves are in $\mathcal{D}'(\psi_\alpha(V_\alpha))$.
$\supseteq$: Choose an atlas such that each $V_\alpha$ is relatively compact and let $(\chi_\alpha)_\alpha$ be a partition of unity subordinate to $(V_\alpha)_\alpha$. First of all, it suffices to show that each $T|_{V_\alpha}$ is the weak limit of a sequence $i(f_m^\alpha)$ $(f_m^\alpha \in \Gamma^\infty(V_\alpha, E \otimes \mathrm{Vol}^{\,\mathrm{q}}(\mathrm{X}))))$. Indeed, suppose this is the case and let $u \in \Gamma_c^\infty(X, E^* \otimes \mathrm{Vol}^{\,1-\mathrm{q}}(\mathrm{X}))$. Then

$$\langle T, u\rangle = \sum_\alpha \langle T|_{V_\alpha}, \chi_\alpha u\rangle = \lim_m \sum_\alpha \int_{V_\alpha} (f_m^\alpha|\chi_\alpha u) = \lim_m \int \sum_\alpha (\chi_\alpha f_m^\alpha|u) = \lim_m \langle \iota(\sum_\alpha \chi_\alpha f_m^\alpha), u\rangle\,.$$

To finish the proof, for each $1 \le i \le N$ choose a sequence $(\tilde{f}^\alpha_{m\,i})_m$ of smooth functions in $\psi_\alpha(V_\alpha)$ converging to $T_{\alpha\, i}$ weakly in $\mathcal{D}'(\psi_\alpha(V_\alpha))$ and set $f_m^\alpha = \Psi_\alpha^*((\tilde{f}^\alpha_{m\,i})_{i=1}^N)$. □

**1.29 Example.** Currents.
As an important example of vector bundle distributions let us introduce the notion of *currents* due to de Rham ( [49]). To this end we first have to fix some terminology: For $1 \leq k \leq n$ we denote by $\Omega^k(X)$ the space of smooth sections of the vector bundle $\bigwedge^k T^*X$. Elements of $\Omega^k(X)$ are called differential forms of order $k$. We set $\Omega^0(X) = \mathcal{C}^\infty(X)$ and $\Omega(X) = \bigoplus_{k=0}^n \Omega^k(X)$. In de Rham's terminology, $k$-forms are called *even* k-forms in order to distinguish them from *odd* $k$-forms, whose definition (see below) may be motivated as follows: Any even $k$-form over some coordinate patch $(V_\alpha, \psi_\alpha)$ can be written as

$$\omega = \sum_{i_1,\dots,i_k} \omega^\alpha_{i_1\dots i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} ,$$

where $1 \leq i_j \leq n$, $\omega_{i_1\dots i_k} = 0$ unless $1 \leq i_1 < \dots < i_k \leq n$, and $\omega^\alpha_{\sigma(i_1)\dots\sigma(i_k)} = \mathrm{sign}(\sigma)\omega^\alpha_{\mathrm{i}_1\dots\mathrm{i}_\mathrm{k}}$ for any permutation $\sigma$. The components $\omega^\beta_{j_1\dots j_k}$ of $\omega$ with respect to another chart $(V_\beta, \psi_\beta)$ (with coordinates $y^i$) are related to the $\omega^\alpha_{i_1\dots i_k}$ by

$$\omega^\beta_{j_1\dots j_k} = \sum_{i_1,\dots,i_k} \omega^\alpha_{i_1\dots i_k} \frac{\partial x^{i_1}}{\partial y^{j_1}} \cdots \frac{\partial x^{i_k}}{\partial y^{j_k}} \tag{16}$$

This cocycle of transition functions uniquely characterizes $\bigwedge^k T^*X$. Similarly replacing (16) by

$$\omega^\beta_{j_1\dots j_k} = \mathrm{sign}(\det(D(\psi_{\alpha\beta}))) \sum_{i_1,\dots,i_k} \omega_{i_1,\dots,i_k} \frac{\partial x^{i_1}}{\partial y^{j_1}} \cdots \frac{\partial x^{i_k}}{\partial y^{j_k}} \tag{17}$$

characterizes the smooth sections (called *odd* $k$-forms) of a unique vector bundle which we denote by $\bigwedge^k_{\mathrm{o}} T^*X$. Equivalently, in the language of 1.9, we may say that $\bigwedge^k_{\mathrm{o}} T^*X = \mathrm{Or}(X) \otimes \bigwedge^k T^*X$. The space of odd $k$-forms is denoted by $\Omega^k_{\mathrm{o}}(X)$. Note that odd $n$-forms are exactly one-densities. Set

$$E = \bigoplus_{k=0}^{n} \bigwedge^{k} T^*(X) \oplus \bigoplus_{k=0}^{n} \bigwedge\nolimits_{\mathrm{o}}^{k} T^*(X) .$$

A *current* is an element of

$$(\Gamma^\infty_c(X, E))' = \bigoplus_{k=0}^{n} (\Gamma^\infty_c(X, \bigwedge^{k} T^*(X)))' \oplus \bigoplus_{k=0}^{n} (\Gamma^\infty_c(X, \bigwedge\nolimits_{\mathrm{o}}^{k} T^*(X)))' .$$

Elements of $(\Gamma^\infty_c(X, \bigwedge^{n-k} T^*(X)))'$ (resp. $(\Gamma^\infty_c(X, \bigwedge^{n-k}_o T^*(X)))'$) are called odd (resp. even) $k$-currents in [49], chap. III (unfortunately, in [53], 17.3.1 $n-k$ is exchanged with $k$ in this definition). In the terminology introduced in section 1.C, the space of currents is written as $\mathcal{D}'(X, E \otimes \mathrm{Vol}^1(\mathrm{X}))$. Odd forms and, dually, even currents are quite useful on non-oriented manifolds as a number of notions from tensor analysis (wedge product, pullback, ...) can easily be generalized to this setting. However, we will confine ourselves to a more systematic study of currents on orientable manifolds in the following section where even and odd forms resp. currents can be identified and where a much more refined machinery for generalizing geometric results is available.

**1.30** We close this section with some remarks on the interrelations of different approaches to distributions on manifolds thereby completing the discussion of sec. 1.A. Dieudonné ( [53], chap. XVII) defines distributions ($k$-currents) as continuous linear functionals on the space of compactly supported test *functions* ($n$-$k$-forms) hence as objects we have called distributional *densities* (with values in the respective bundle). Hörmander ( [88], chap. 6), on the other hand defines distributions on $X$ as families of distributions on $\psi_\alpha(V_\alpha)$ which transform according to the *distributional pullback* which is modelled as to coincide with the pullback of regular objects, i.e., functions. More precisely (cf. [88], 6.1.2), given a $\mathcal{C}^\infty$-map $f : V_1 \to V_2$ with $f'(x)$ onto for every $x$ the distributional pullback is the unique continuous linear map $f^* : \mathcal{D}'(V_2) \to \mathcal{D}'(V_1)$ such that $f^*u = u \circ f$ for every continuous function $u$. This, of course, accounts to burdening the test object with the determinant appearing in the transformation formula for multiple integrals, i.e., to our transformation law (14) in the case $F = \mathrm{Vol}\,(\mathrm{X})$. Hence in accordance with Hörmander we defined distributions on $X$ as (continuous linear) functionals on (test) one-densities and distributional densities as functionals on (test) functions.

### 1.E. Orientable Manifolds, Distributional Geometry

**1.31** In local distribution theory on $\mathbb{R}^n$ the main method of generalizing definitions and results from classical analysis to distributions consists in extending operations from regular objects to distributions by transposition. Let us exemplify this principle in a very simple situation, namely differentiation of distributions on $\mathbb{R}^n$: How do we know that setting $\langle u', \varphi\rangle = -\langle u, \varphi'\rangle$ is the "right" generalization of classical calculus?
The answer is of course that this definition agrees with classical differentiation on smooth functions by integration by parts. The second question that arises is if the above definition is the only one with this property. In fact, it is easy to see that if we want differentiation to be sequentially continuous and to agree with classical differentiation on smooth functions then indeed the above definition is the only one possible.
Thus if we want to pursue a similar path in extending classical differential geometry to distributional objects on manifolds we first of all need a class of smooth test objects densely contained in the distributional space and endowed with a sufficiently rich structure to allow for a continuous extension of classical concepts. For example, if we intend to extend Lie derivatives with respect to smooth vector fields $\xi$ then we need some kind of analogue to partial integration in order to be able to use transposition as above for defining $L_\xi$ on distributions. Now on an oriented manifold $X$, for $\omega \in \Omega^k(X)$, $\tau \in \Omega_c^{n-k}(X)$ we have

$$\int L_\xi(\omega \wedge \tau) = \int d\, i_\xi(\omega \wedge \tau) = 0$$

by Stokes' theorem, so $\int L_\xi \omega \wedge \tau = -\int \omega \wedge L_\xi \tau$, which is exactly what we need.
For this reason, in the remainder of this section we shall consider $X$ to be an oriented manifold with orientation induced by some fixed non-vanishing $\theta \in \Omega^n(X)$ and develop a *generalized tensor analysis for spaces of currents on* $X$. In doing this we will mainly follow Marsden's presentation in [138]. Henceforth we shall denote the space of *odd* $k$*-currents* on $X$ by $\Omega^k(X)'$ and also refer to it as the space of *generalized*[1] *or distributional* $k$*-forms* on $X$. Note that this notation (cf. Marsden [138]) does *not* imply that $\Omega^k(X)'$ is the dual space of $\Omega^k(X)$, rather it is the dual of $\Omega_c^{n-k}(X)$ (cf. 1.29 and 1.32 below), i.e., in the notation of 1.C we have

$$\Omega^k(X)' = \mathcal{D}'(X, \bigwedge^k T^*X \otimes \mathrm{Vol}\,(\mathrm{X}))\,.$$

**1.32 Regular objects.** The map

$$\begin{aligned} j : \Omega^k(X) &\to \Omega^k(X)' \\ j(\omega)(\tau) &= \int \omega \wedge \tau \end{aligned}$$

is a linear embedding. Indeed, $j(\Omega^k(X)) \subseteq \Omega^k(X)'$ as well as linearity of $j$ are clear and injectivity of $j$ follows immediately from the local description. In fact, $j$ also provides an embedding of the space of *locally integrable* $k$-forms into $\Omega^k(X)'$. (Here, a form $\omega$ is called locally integrable if for all $\tau \in \Omega_c^{n-k}(X)$ we have $\omega \wedge \tau = f_\tau \theta$ with $f_\tau$ locally integrable with respect to $\theta$.) *Distributions* on $X$ are elements of $\Omega^0(X)'$, i.e., continuous linear forms on the space of compactly supported $n$-forms on $X$. Since $X$ is oriented, we may canonically identify smooth $n$-forms and one-densities on $X$. Thus we have

$$\Omega^0(X)' \cong \mathcal{D}'(X)\,.$$

Moreover, since we have assumed the orientation to be induced by some fixed non-vanishing $\theta \in \Omega^n(X)$ we can identify

$$\Omega^0(X)' \cong \mathcal{D}'_d(X)\,,$$

the isomorphism being dependent on theta this time. Altogether, we have

---

[1]The terminology of "generalized quantities" is due to Marsden and we shall use it throughout this section for historical reasons. However, from chapter 3 on the term "generalized" will be reserved for generalized functions in the sense of Colombeau.

$$\mathcal{D}'(X) \cong \Omega^0(X)' \cong \mathcal{D}'_d(X)\,.$$

Therefore, whenever $X$ is an orientable manifold (which we assume throughout this section) let us agree to use the notations $\mathcal{D}'(X)$ and $\Omega^0(X)'$ synonymously. Furthermore, we introduce the following conventions:

$$\mathcal{I}^n_k = \{I = (i_1, \dots, i_k) \in \mathbb{N}^k :\ 1 \le i_1 < \dots < i_k \le n\}$$
$$dx^I := dx^{i_1} \wedge \dots \wedge dx^{i_k}$$

For $I \in \mathcal{I}^n_k$ we set $I^c = (j_1, \dots, j_{n-k})$ where $1 \le j_1 < \dots < j_{n-k} \le n$ and $I \cup I^c = \{1, \dots, n\}$. Then any $k$-form can locally be written as $\nu = \sum_{I \in \mathcal{I}^n_k} \nu_I dx^I$ [2]; moreover, define $S(I) \in \{-1, +1\}$ by $dx^I \wedge dx^{I^c} = S(I) d^n x = S(I) dx^1 \wedge \dots \wedge dx^n$. Denote by $(V_\alpha, \Psi^{(k)}_\alpha)_\alpha$ the vector bundle atlas of the bundle $\bigwedge^k T^*X$ corresponding to an atlas $(V_\alpha, \psi_\alpha)$ of the base. If $\omega \in \Omega^k(X)'$ then (according to the local description of sec. 1.D) $\omega_\alpha = (\Psi^{(n-k)}_\alpha)^\wedge(\omega|_{V_\alpha})$ corresponds to a compatible family $(\omega_{\alpha\, I})_{I \in \mathcal{I}^n_{n-k}}$ in $\mathcal{D}'(\psi_\alpha(V_\alpha))^N$ ($N = \binom{n}{k}$) such that for $\tau \in \Omega^{n-k}_c(V_\alpha)$, $\tau = \sum_{I \in \mathcal{I}^n_{n-k}} \tau_I dx^I$, we have $\langle \omega, \tau \rangle = \sum_{I \in \mathcal{I}^n_{n-k}} \langle \omega_{\alpha\, I}, \tau_I \circ \psi_\alpha^{-1} \rangle$. Also, for $\omega \in \Omega^k(X)$ with $\omega|_{V_\alpha} = \sum_{J \in \mathcal{I}^n_k} \omega_J dx^J$ we have

$$(j(\omega))_{\alpha\, I} = \left( \varphi \mapsto \int_{\psi_\alpha(V_\alpha)} s(I)(\omega_{I^c} \circ \psi_\alpha^{-1})(x)\varphi(x) d^n x \right) \qquad (I \in \mathcal{I}^n_{n-k}). \tag{18}$$

In fact, using (18), for $\tau \in \Omega^{n-k}_c(V_\alpha)$ as above we get

$$\langle j(\omega), \tau \rangle = \sum_{I \in \mathcal{I}^n_{n-k}} \langle j(\omega)_{\alpha\, I}, \tau_I \circ \psi_\alpha^{-1} \rangle = \sum_{I \in \mathcal{I}^n_{n-k}} \int_{\psi_\alpha(V_\alpha)} \omega_{I^c} \circ \psi_\alpha^{-1} dx^{I_c} \wedge \tau_I \circ \psi_\alpha^{-1} dx^I$$
$$= \int_{\psi_\alpha(V_\alpha)} (\omega \wedge \tau)_\alpha = \int_{V_\alpha} \omega \wedge \tau\,.$$

Thus $(j(\omega))_{\alpha\, I}$ is the regular distribution $s(I)\omega_{I^c} \circ \psi_\alpha^{-1}$.
The key to extending tensor analysis from $\Omega^k(X)$ to $\Omega^k(X)'$ is the following result which may be proved analogously to thm. 1.28.

**1.33 Theorem.**

$$\overline{j(\Omega^k(X))}^{wsc} = \Omega^k(X)'\,.$$

**1.34** We know from theorem 1.18 that $\Omega^k(X)'$ is a fine sheaf. In particular, the support of any generalized $k$-form is a meaningful concept. Moreover, for any $\omega \in \Omega^k(X)'$ we define its *singular support* by

$$\mathrm{singsupp}(\omega) := \mathrm{X} \setminus \bigcup \{\mathrm{U} \subseteq \mathrm{X} : \mathrm{U}\ \mathrm{open}, \omega|_{\mathrm{U}}\ \mathrm{is\ smooth}\}$$

Here $\omega$ is called smooth on $U$ if there exists some $\tau \in \Omega^k(X)$ with $j(\tau)|_U = \omega|_U$. $\omega \in \Omega^n(X)'$ is called *positive* $\omega \ge 0$, if $\omega(\varphi) \ge 0$ for all $\varphi \ge 0$. Similarly $\omega \in \mathcal{D}'(X)$ is called positive with respect to $\theta$ if $\omega(\varphi\theta) \ge 0$ for all $\varphi \ge 0$. We have the following result on positive distributional $n$-forms:

**1.35 Theorem.** *Let $\omega \in \Omega^n(X)'$, $\omega \ge 0$. Then there exists a unique positive Radon measure $\mu_\omega$ on $X$ such that $\omega(\varphi) = \int \varphi d\mu_\omega$ for all $\varphi \in \mathcal{D}(X)$.*

**Proof.** For $f \ge 0$ continuous with compact support let $\omega(f) = \sup\{\omega(\varphi) :\ \varphi \in \mathcal{D}(X),\ 0 \le \varphi \le f\}$. This defines a linear and positive extension of $\omega$, hence it uniquely determines a positive Radon measure. $\square$

---

[2] To simplify notation we here omit the upper index $\alpha$ (cf. 3.5).

**1.36** The operation of taking *exterior products* of differential forms is extended to distributional forms by transposition, i.e., for $\omega \in \Omega^k(X)'$, $\tau \in \Omega^l(X)$ we define $\omega \wedge \tau \in \Omega^{k+l}(X)'$ by

$$(\omega \wedge \tau)(\nu) = \omega(\tau \wedge \nu) \qquad (\nu \in \Omega_c^{n-k-l})$$
$$\tau \wedge \omega = (-1)^{kl}\omega \wedge \tau\,. \tag{19}$$

Moreover, since obviously $\omega \mapsto \omega \wedge \tau$, $\Omega^k(X)' \to \Omega^{k+l}(X)'$ is weakly sequentially continuous and since for $\omega \in \Omega^k(X)$ $j(\omega) \wedge \tau = j(\omega \wedge \tau)$, (19) is the unique continuous extension of the classical wedge product to $(\omega, \tau) \in \Omega^k(X)' \times \Omega^l(X)$. This constitutes the first example of the general strategy for extending classical concepts to generalized forms as discussed in 1.31 above. To simplify notations for this typical pattern, in the remainder of this section we agree to use the terms “continuous” and “weakly sequentially continuous” synonymously. Also, we will usually suppress the embedding $j : \Omega^k(X) \hookrightarrow \Omega^k(X)'$ and consider $\Omega^k(X)$ as a subspace of $\Omega^k(X)'$.

**1.37** The above notion provides us with the following more refined *local description* of distributional $k$-forms. Let $V_\alpha$ be a chart with local coordinates $x^1, \dots, x^n$. If $\omega \in \Omega^k(X)'$ then $\omega|_{V_\alpha}$ can uniquely be written in the form

$$\omega|_{V_\alpha} = \sum_{I \in \mathcal{I}_k^n} \omega_I^\alpha dx^I$$

with $\omega_I^\alpha \in \mathcal{D}'(V_\alpha) = \Omega^0(V_\alpha)'$. For $I \in \mathcal{I}_k^n$, $\omega_I^\alpha$ is given by

$$\omega_I^\alpha(\varphi d^n x) = s(I)\omega(\varphi dx^{I^c}) \qquad (\varphi \in \mathcal{D}(V_\alpha))\,. \tag{20}$$

Indeed, using (19) we compute ($J \in \mathcal{I}_{n-k}^n$)

$$(\sum_{I \in \mathcal{I}_k^n} \omega_I^\alpha dx^I)(\varphi dx^J) = \sum_{I \in \mathcal{I}_k^n} \omega_I^\alpha(\varphi dx^I \wedge dx^J) = \omega_{J^c}^\alpha(\varphi dx^{J^c} \wedge dx^J) = \omega_{J^c}^\alpha(s(J^c)\varphi d^n x) = \omega(\varphi dx^J)\,.$$

To determine the relationship between $\omega_{\alpha\, I}$ introduced in 1.32 and $\omega_I^\alpha$, note that for any $\varphi \in \mathcal{D}(V_\alpha)$ and any $J \in \mathcal{I}_{n-k}^n$ we have

$$\langle \omega_{\alpha\, J}, \varphi \circ \psi_\alpha^{-1} \rangle = \omega(\varphi dx^J) = \omega_{J^c}^\alpha(s(J^c)\varphi d^n x) = \langle (\Psi_\alpha^{(n)})^\wedge(s(J^c)\omega_{J^c}^\alpha), \varphi \circ \psi_\alpha^{-1} \rangle$$

so

$$\omega_{\alpha\, J} = s(J^c)(\Psi_\alpha^{(n)})^\wedge(\omega_{J^c}^\alpha) \qquad (J \in \mathcal{I}_{n-k}^n)\,.$$

Next we generalize the notion of a pullback by a diffeomorphism to the distributional level.

**1.38 Theorem.** *Let $f : X \to Y$ be an orientation preserving diffeomorphism. The map $f^* : \Omega^k(Y) \to \Omega^k(X)$ has a unique continuous extension $f^* : \Omega^k(Y)' \to \Omega^k(X)'$. For $\omega \in \Omega^k(Y)'$ and $\tau \in \Omega_c^{n-k}(X)$ we have*

$$(f^*\omega)(\tau) = \omega((f^{-1})^*\tau)\,. \tag{21}$$

*$f^*$ is bijective, $(f \circ g)^* = g^* \circ f^*$ and $(f^*)^{-1} = (f^{-1})^*$. Moreover, for $\omega \in \Omega^k(Y)'$ and $\tau \in \Omega_c^{n-k}(Y)$ we have*

$$f^*(\omega \wedge \tau) = f^*(\omega) \wedge f^*(\tau)\,. \tag{22}$$

**Proof.** Clearly (21) defines a continuous map, which for the moment we call $(f^*)'$ . Moreover, for $\omega \in \Omega^k(Y)$, $\tau \in \Omega_c^{n-k}(X)$ we have

$$((f^*)'(\omega))(\tau) = \omega((f^{-1})^*\tau) = \int \omega \wedge (f^{-1})^*\tau = \int f^*\omega \wedge \tau = (f^*\omega)(\tau)$$

by the change of variables formula (cf. [1], 2.6.7.), so $(f^*)'$ coincides with $f^*$ on $\Omega^k(X)$. $(f^*)'(\Omega^k(Y)') \subseteq \Omega^k(X)'$ as well as the statements on composition and inverses are immediate consequences of (21). Finally, (22) follows from the classical result by continuous extension. □

**1.39** Now we turn to the notion of *integration of generalized n-forms.* Let $(\chi_i)_i$ be a partition of unity with $\mathrm{supp}(\chi_\mathrm{i})$ compact for each $i$. For any compactly supported $\omega \in \Omega^n(X)'$ we define

$$\int \omega = \sum_i \omega(\chi_i)\,. \tag{23}$$

This definition is independent of the partition of unity chosen since:

$$\sum_i \omega(\chi_i) = \sum_{i,j} \omega(\chi_i \zeta_j) = \sum_j \omega(\zeta_j)\,.$$

Also, for $\omega$ smooth (23) coincides with the usual definition. If $\omega \in \Omega^n(X)'$ is not necessarily compactly supported and the sum in (23) converges independently of the partition of unity (where convergence to $+\infty$ is admissible), $\omega$ is called integrable. Positive generalized forms are always integrable and we have

$$\int \omega = \sum_i \omega(\chi_i) = \int d\mu_\omega$$

by monotone convergence.
Next, let us examine *change of variables* for integrals of generalized $n$-forms.

**1.40 Theorem.** *Let $f : X \to Y$ be an orientation-preserving diffeomorphism. If $\omega \in \Omega^n(Y)'$ is integrable then $\int f^*\omega = \int \omega$.*

**Proof.** Note that

$$\int f^*\omega = \sum_i f^*\omega(\chi_i) = \sum_i \omega((f^{-1})^*\chi_i) = \sum_i \omega(\chi_i \circ f^{-1})$$

from which the claim follows since $(\chi_i \circ f^{-1})_i$ is a partition of unity for $Y$. □

Integration of generalized $n$-forms violates our construction scheme based on continuous extension of classical operations since it is *not* continuous (as a simple counterexample, observe that $\delta_n$ on $\mathbb{R}$ converges weakly to 0 although $\int \delta_n = 1$ for all $n \in \mathbb{N}$). The example already indicates how to remedy this defect: Consider the space $\Omega^n_c(X)'$ of compactly supported generalized $n$-forms:

**1.41 Proposition.** *$\int$ is the unique map on $\Omega^n_c(X)'$ coinciding with the usual integral on $\Omega^n_c(X)$ and continuous in the following sense: If $\omega_m \to \omega_0$ such that there exists $K \subset\subset X$ with $\mathrm{supp}(\omega_\mathrm{m}) \subseteq \mathrm{K}$ for all $m \geq 0$ then $\int \omega_m \to \int \omega$.*

**Proof.** $\int$ is continuous in this sense since

$$\int \omega = \sum_i \omega(\chi_i) = \sum_i \lim_m \omega_m(\chi_i) = \lim_m \sum_i \omega_m(\chi_i) = \lim_m \int \omega_m$$

($i$ only ranges over the finite set of all $i$ with $K \cap \mathrm{supp}(\chi_\mathrm{i}) \neq \emptyset$). Finally, if $\omega \in \Omega^n_c(X)'$ choose some $\chi \in \mathcal{D}(X)$ with $\chi \equiv 1$ in a neighborhood of $\mathrm{supp}(\omega)$. By 1.33 there exists a sequence $(\omega_m)_m$ in $\Omega^n(X)$ converging to $\omega$. Then $\chi\omega_m \to \omega$ and $\mathrm{supp}(\chi\omega_\mathrm{m}) \subseteq \mathrm{supp}(\chi)$ for all $m$. □

**1.42 Theorem.** *The exterior derivative $d : \Omega^k(X) \to \Omega^{k+1}(X)$ has a unique continuous extension $d : \Omega^k(X)' \to \Omega^{k+1}(X)'$ given by*

$$d\omega(\tau) = (-1)^{k+1}\omega(d\tau) \qquad (\tau \in \Omega^{n-k-1}(X))\,, \tag{24}$$

*with the following properties*

*(i) $d$ is $\mathbb{R}$-linear and $d(\omega \wedge \tau) = d\omega \wedge \tau + (-1)^k \omega \wedge d\tau \qquad (\omega \in \Omega^k(X)',\ \tau \in \Omega^l(X))$.*

*(ii) $d \circ d = 0$.*

*(iii) If $\omega \in \Omega^{n-1}_c(X)'$ then $\int d\omega = 0$.*

*(iv) If $f : X \to Y$ is an orientation preserving diffeomorphism and $\omega \in \Omega^k(Y)'$ then $f^*(d\omega) = df^*(\omega)$.*

**Proof.** We temporarily denote the map defined by (24) by $d' : \Omega^k(X)' \to \Omega_c^{n-k}(X)^*$ (algebraic dual). Then $d'$ is obviously continuous and for any $\omega \in \Omega^k(X)$, $\tau \in \Omega_c^{n-k-1}(X)$ we have

$$d'\omega(\tau) = (-1)^{k+1} \int \omega \wedge d\tau = \int d\omega \wedge \tau - \int d(\omega \wedge \tau) = \int d\omega \wedge \tau = d\omega(\tau)$$

by Stokes' theorem. Also, $d'(\Omega^k(X)') \subseteq \Omega^{k+1}(X)'$ by 1.33: If $\omega_m \in \Omega^k(X)$ converge to $\omega$ then $d\omega_m = d'\omega_m \to d'\omega$. (i), (ii) and (iv) follow immediately by continuous extension and concerning (iii) we note that

$$\int d'\omega = \sum_i d'\omega(\chi_i) = \sum_i (-1)^n \omega(d\chi_i) = (-1)^n \omega(d(1)) = 0\,.$$

□

In order to derive a local description of $d$ over some chart $(\psi_\alpha, V_\alpha)$ with local coordinates $x^i$, we introduce the following notion of partial derivative for any $u \in \mathcal{D}'(V_\alpha)$: By $\frac{\partial u}{\partial x^i}$ we denote the element of $\mathcal{D}'(V_\alpha)$ satisfying

$$\frac{\partial u}{\partial x^i}(\varphi d^n x) = -u(\frac{\partial \varphi}{\partial x^i} d^n x) \qquad (\varphi \in \mathcal{D}(V_\alpha))\,.$$

For $u$ smooth this coincides with usual partial derivatives by partial integration. Moreover, setting $I = (1, \dots, i-1, i+1, \dots, n)$ we have for any $u \in \mathcal{D}'(V_\alpha)$

$$(-1)^{i+1} du(\varphi dx^I) = (-1)^i u(d(\varphi dx^I)) = (-1)^i u(\frac{\partial \varphi}{\partial x^i} dx^i \wedge dx^I) = -u(\frac{\partial \varphi}{\partial x^i} d^n x) = \frac{\partial u}{\partial x^i}(\varphi d^n x)$$

from which we conclude

$$du = \sum_i \frac{\partial u}{\partial x^i} \wedge dx^i \qquad (u \in \mathcal{D}'(V_\alpha))\,.$$

**1.43 Theorem.** *Let $\xi \in \mathfrak{X}(X)$. The Lie-derivative $L_\xi : \Omega^k(X) \to \Omega^k(X)$ has a unique continuous extension (again denoted by) $L_\xi : \Omega^k(X)' \to \Omega^k(X)'$. $L_\xi$ is given by*

$$(L_\xi \omega)(\tau) = -\omega(L_\xi \tau) \qquad (\omega \in \Omega^k(X)',\ \tau \in \Omega_c^{n-k}(X)) \tag{25}$$

*and has the following properties*

*(i) $L_\xi$ is $\mathbb{R}$-linear, and $L_\xi(\omega \wedge \tau) = L_\xi \omega \wedge \tau + \omega \wedge L_\xi \tau$ ($\omega \in \Omega^k(X)'$, $\tau \in \Omega^l(X)$).*

*(ii) $L_\xi d\omega = dL_\xi \omega$.*

*(iii) $L_{[\xi,\eta]} = L_\xi \circ L_\eta - L_\eta \circ L_\xi$.*

*(iv) For any orientation preserving diffeomorphism $f : X \to Y$, $f^*(L_\xi \omega) = L_{f^*\xi} f^* \omega$.*

*(v) If $\tau \in \Omega_c^{n-k}(X)$ then $\int L_\xi \omega \wedge \tau = -\int \omega \wedge L_\xi \tau$.*

*(vi) For any $V \subseteq X$ open, $\xi \in \mathfrak{X}(X)$ and $\omega \in \Omega^k(X)'$ we have $(L_\xi \omega)|_V = L_{\xi|_V}(\omega|_V)$.*

**Proof.** The map from $\Omega^k(X)'$ into $\Omega_c^{n-k}(X)^*$ defined by (25) is obviously continuous and coincides with the classical Lie derivative on $\Omega^k(X)$ since for $\omega \in \Omega^k(X)$, $\tau \in \Omega_c^{n-k}(X)$ we have

$$\int L_\xi \omega \wedge \tau = \int L_\xi(\omega \wedge \tau) - \int \omega \wedge L_\xi \tau = -\int \omega \wedge L_\xi \tau\,.$$

The last equality is due to Stokes' theorem as $L_\xi(\omega \wedge \tau) = \int d\ i_\xi(\omega \wedge \tau) = 0$. That $L_\xi(\Omega^k(X)') \subseteq \Omega^k(X)'$ follows by a similar argument as in 1.42. Now (i) – (iv) and (vi) follow by continuous extension (or directly by transposition) from the corresponding classical results . To prove (v), choose a sequence $(\omega_m)_m$ converging weakly to $\omega$. For $\omega_m$ instead of $\omega$ in (v) the desired equality has already been established. Since both integrands in (v) then converge in the sense of proposition 1.39, the claim follows from that result. □

Our next aim is to generalize the insertion operator $i_\xi : \Omega^k(X) \to \Omega^{k-1}(X)$ (cf. [1], 2.4.12).

**1.44 Theorem.** *For any $\xi \in \mathfrak{X}(X)$ the insertion operator $i_\xi$ has a unique continuous extension $i_\xi : \Omega^k(X)' \to \Omega^{k-1}(X)'$ given by*

$$i_\xi\omega(\tau) = (-1)^{k+1}\omega(i_\xi\tau) \qquad (\omega \in \Omega^k(X)',\, \tau \in \Omega_c^{n-k+1}(X)) \tag{26}$$

*with the following properties*

*(i)* $i_\xi$ *is* $\mathbb{R}$*-linear, and* $i_\xi(\omega \wedge \tau) = i_\xi\omega \wedge \tau + (-1)^k\omega \wedge i_\xi\tau$ $(\omega \in \Omega^k(X)',\, \tau \in \Omega_c^l(X))$

*(ii)* $i_\xi \circ i_\xi = 0$.

*(iii)* $L_\xi = i_\xi \circ d + d \circ i_\xi$.

*(iv)* $i_{[\xi,\eta]} = L_\xi \circ i_\eta - i_\eta \circ L_\xi$.

*(v)* $L_{f\xi} = fL_\xi + (df) \wedge i_\xi$ $(f \in \mathcal{C}^\infty(X))$.

*(vi) For any orientation preserving diffeomorphism* $f : X \to Y$, $f^*(i_\xi\omega) = i_{f^*\xi}f^*\omega$.

*(vii) For any* $V \subseteq X$ *open* $(i_\xi\omega)|_V = i_{\xi|_V}(\omega|_V)$.

**Proof.** (26) defines a continuous map from $\Omega^k(X)'$ into $\Omega_c^{n-k+1}(X)^*$ coinciding with $i_\xi$ on $\Omega^k(X)$ since for $\omega \in \Omega^k(X)$ and $\tau \in \Omega_c^{n-k+1}(X)$ we have (using [1], 2.4.13 (i))

$$i_\xi\omega \wedge \tau = (-1)^{k+1}\omega \wedge i_\xi\tau + i_\xi(\underbrace{\omega \wedge \tau}_{=0}).$$

$i_\xi(\Omega^k(X)') \subseteq \Omega^{k-1}(X)'$ is seen as in the preceding proofs. Finally, (i) – (vii) follow from the corresponding classical results by continuous extension or by direct computation, respectively. □

In order to derive a generalized Poincaré lemma we will make use of the following regularization result due to de Rham (for a proof see [49], ch. III, §15.)

**1.45 Theorem.** *For each $m \in \mathbb{N}$ there exist $\mathbb{R}$-linear maps*

$$\begin{aligned} &R_m : \Omega^k(X)' \to \Omega^k(X) \\ &A_m : \Omega^k(X)' \to \Omega^{k-1}(X)' \quad (k \geq 1)\,, \quad A_m(\Omega^0(X)') = \{0\} \end{aligned}$$

*with the following properties:*

*(i) For each* $\omega \in \Omega^k(X)'$ *we have* $R_m\omega - \omega = dA_m\omega + A_md\omega$.

*(ii)* $A_m(\Omega^k(X)) \subseteq \Omega^{k-1}(X)$.

*(iii)* $R_m\omega \to \omega$, $A_m\omega \to 0$ *uniformly on bounded sets.*

**Proof.** See [49], ch. III, §15. □

As for smooth forms, we call $\omega \in \Omega^k(X)'$ *closed* if $d\omega = 0$ and *exact* if there exists some $\tau \in \Omega^{k-1}(X)'$ with $\omega = d\tau$.

**1.46 Theorem.** *(Generalized Poincaré lemma)*

*(i) If* $\omega \in \Omega^k(X)'$ $(k \geq 1)$ *is closed then for each* $p \in X$ *there exists a neighborhood* $V_x$ *of* $x$ *and a* $\tau \in \Omega^{k-1}(X)'$ *such that*

$$\omega|_{V_x} = d\tau|_{V_x}$$

*(ii) Let* $X$ *be connected. If* $u \in \mathcal{D}'(X)$ *and* $du = 0$ *then* $u$ *is constant.*

**Proof.** (i) By 1.45 (i) we have $\omega = R_m\omega - dA_m\omega$, implying $dR_m\omega = 0$. Now $R_m\omega$ is smooth so by the classical Poincaré lemma there exists some $\varphi \in \Omega^{k-1}(X)$ whose exterior derivative coincides with $R_m\omega$ on some $V_x$. Taking $\tau = \varphi - A_m\omega$ establishes the result.
(ii) Again by 1.45 (i) we have $u = R_m u \in \mathcal{C}^\infty(X)$, so the result follows. □

Next we discuss the relation between *Lie derivatives* and *flows of vector fields.* For $\xi \in \mathfrak{X}(X)$ we denote by $(t,p) \to \mathrm{Fl}^\xi_t(p)$ the flow of $\xi$. $\xi$ is called *complete* if each integral curve of $\xi$ can be extended to have domain $(-\infty,\infty)$ or, equivalently, if the domain of $\mathrm{Fl}^\xi$ is $\mathbb{R} \times X$ ( [1], 2.1.12). The following result generalizes the classical flow theorem ( [1], 2.2.20) to Lie derivatives of generalized forms.

**1.47 Theorem.** *Let $\xi \in \mathfrak{X}(X)$ be complete. Then for each $\omega \in \Omega^k(X)'$ the map $t \to (\mathrm{Fl}^\xi_t)^*\omega$ is (weakly) differentiable and*

$$L_\xi((\mathrm{Fl}^\xi_t)^*\omega) = (\mathrm{Fl}^\xi_t)^* L_\xi\omega = \frac{d}{dt}(\mathrm{Fl}^\xi_t)^*\omega\,. \tag{27}$$

*Hence*

$$L_\xi\omega = 0 \Leftrightarrow \omega = (\mathrm{Fl}^\xi_t)^*\omega \ \forall t\,.$$

**Proof.** Let $a > 0$, $\tau \in \Omega^{n-k}_c(X)$ and set $g_m(t) = ((\mathrm{Fl}^\xi_t)^*(R_m(\omega)))(\tau) = \int (\mathrm{Fl}^\xi_t)^*(R_m(\omega)) \wedge \tau$, and $g(t) = ((\mathrm{Fl}^\xi_t)^*(\omega)))(\tau)$. Since $\mathrm{Fl}^\xi$ is smooth on $\mathbb{R} \times X$, the set $\{((\mathrm{Fl}^\xi_t)^{-1})^*\tau \,|\, t \in [-a,a]\}$ is bounded. Hence by 1.45 (iii) (plus a change of variables in the above integral) $g_m \to g$ uniformly on $[-a,a]$. Now

$$g_m'(t) = \int \frac{d}{dt}(\mathrm{Fl}^\xi_t)^*(R_m(\omega)) \wedge \tau = \begin{cases} \int L_\xi(\mathrm{Fl}^\xi_t)^*(R_m(\omega)) \wedge \tau &= (L_\xi(\mathrm{Fl}^\xi_t)^*(R_m(\omega)))(\tau) \\ \int (\mathrm{Fl}^\xi_t)^*(L_\xi R_m(\omega)) \wedge \tau &= ((\mathrm{Fl}^\xi_t)^*(L_\xi R_m(\omega)))(\tau) \end{cases} \tag{28}$$

by [1], 2.2.20. Again by 1.45 (iii), the right hand sides of (28) converge to $(L_\xi(\mathrm{Fl}^\xi_t)^*(\omega))(\tau)$ and $((\mathrm{Fl}^\xi_t)^*(L_\xi\omega))(\tau)$, respectively, uniformly on $[-a,a]$. Thus $g$ is differentiable on $(-a,a)$ and (27) holds. Since $a$ was arbitrary, the result follows. □
Hence $\xi$ can be viewed as inducing a one-parameter group on $\Omega^k(X)'$ with infinitesimal generator $L_\xi$. To conclude this section we shall describe the action of flows on Dirac distributions. For $p \in X$, by $\delta_p \in \Omega^n(X)'$ we denote the generalized $n$-form $\varphi \mapsto \varphi(p)$.

**1.48 Proposition.** *Let $\xi$ be a complete vector field on $X$. For any $p \in X$ we have $(\mathrm{Fl}^\xi_t)^*\delta_p = \delta_{\mathrm{Fl}^\xi_{-t}p}$ and the following conditions are equivalent:*

*(i) $\xi(p) = 0$ (i.e., $p$ is a critical point of $\xi$).*

*(ii) $(\mathrm{Fl}^\xi_t)^*\delta_p = \delta_p$ for all $t \in \mathbb{R}$.*

*(iii) $L_\xi\delta_p = 0$.*

**Proof.** First, note that $(\mathrm{Fl}^\xi_t)^*\delta_p(\varphi) = \delta_p(\varphi \circ \mathrm{Fl}^\xi_{-t}) = \varphi(\mathrm{Fl}^\xi_{-t}(p)) = \delta_{\mathrm{Fl}^\xi_{-t}(p)}(\varphi)$.
(i) ⇒ (ii) From (i) it follows that $\mathrm{Fl}^\xi_t(p) = p$ for all $t$. Thus $(\mathrm{Fl}^\xi_t)^*\delta_p = \delta_{\mathrm{Fl}^\xi_{-t}p} = \delta_p$.
(ii) ⇒ (iii) By (27), $L_\xi\delta_p = (\frac{d}{dt}(\mathrm{Fl}^\xi_t)^*\delta_p)(0) = 0$.
(iii) ⇒ (i) For any $\varphi \in \mathcal{D}(X)$ we obtain $0 = L_\xi\delta_p(\varphi) = -\delta_p(L_\xi\varphi) = -(L_\xi\varphi)(p) = -d\varphi(\xi(p))$, so $\xi(p) = 0$. □

**1.49** In order to carry on our generalization of classical tensor analysis we still have to introduce distribution valued tensor fields on $X$. To this end, in this section we are first going to consider a different description of generalized $k$-forms. We know from 1.36 that any $\omega \in \Omega^k(X)'$ can locally be written in the form

$$\omega|_{V_\alpha} = \sum_{I \in \mathcal{I}^n_k} \omega^\alpha_I dx^I\,.$$

As in the smooth case, $\omega|_{V_\alpha}$ thereby gives rise to a $\mathcal{C}^\infty(V_\alpha)$-$k$-linear alternating map

$$M_a(\omega|_{V_\alpha}) : \mathfrak{X}(V_\alpha)^k \to \mathcal{D}'(V_\alpha)$$
$$(\xi_1, \dots, \xi_k) \to \sum_{I \in \mathcal{I}^n_k} (dx^I(\xi_1, \dots, \xi_k))\omega^\alpha_I\,.$$

(the subscript $a$ stands for "alternating"). We are now going to show that *every* $\mathcal{C}^\infty(X)$-$k$-linear alternating map

$$\rho : \mathfrak{X}(X)^k \to \mathcal{D}'(X) \tag{29}$$

uniquely corresponds to some $\omega \in \Omega^k(X)'$ and vice versa. To begin with, any such $\rho$ is *local*, i.e., if $\xi_i|_V = 0$ for some $1 \le i \le k$ and some $V \subseteq X$ open then $\rho(\xi_1, \dots, \xi_k)|_V = 0$. To see this, suppose $i = 1$ and write $\xi_1 = \sum_j \xi_1^j \partial_{x^j}$ on a chart $V_\alpha \subset\subset V$. Choose $\varphi \in \mathcal{D}(V)$, $\varphi|_{V_\alpha} \equiv 1$. Then

$$\rho(\xi_1, \dots, \xi_k)|_{V_\alpha} = \varphi^2 \rho(\xi_1, \dots, \xi_k)|_{V_\alpha} = \sum_j \varphi \xi_1^j \rho(\varphi\partial_{x^j}, \dots, \xi^k) = 0\,,$$

so since $V_\alpha$ was arbitrary (and since $\mathcal{D}'$ is a sheaf) the claim follows. We can therefore define the restriction of $\rho$ to $V$ unambiguously by

$$\rho|_V : \mathfrak{X}(V)^k \to \mathcal{D}'(X)$$
$$\rho|_V(\xi_1, \dots, \xi_k) = \rho(\tilde\xi_1, \dots, \tilde\xi_k)|_V$$

where $\tilde\xi_1, \dots, \tilde\xi_k$ are arbitrary (smooth) extensions of $\xi_1, \dots, \xi_k$ to $X$. Now let

$$(\xi_1, \dots, \xi_k) = \sum_{i_1, \dots, i_k} \xi^{i_1, \dots, i_k}(\partial_{x^{i_1}}, \dots, \partial_{x^{i_k}})$$

be vector fields on $V_\alpha$. Then

$$\begin{aligned}\rho|_{V_\alpha}(\xi_1, \dots, \xi_k) &= \sum_{i_1, \dots, i_k} \xi^{i_1, \dots, i_k} \rho|_{V_\alpha}(\partial_{x^{i_1}}, \dots, \partial_{x^{i_k}}) \\ &= (\sum_{i_1, \dots, i_k} \rho|_{V_\alpha}(\partial_{x^{i_1}}, \dots, \partial_{x^{i_k}}) dx^{i_1} \otimes \ldots \otimes dx^{i_k})(\xi_1, \dots, \xi_k)\end{aligned} \tag{30}$$

and from this it follows easily by multilinear algebra that $\rho|_{V_\alpha}$ can uniquely be written as $\sum_{I \in \mathcal{I}^n_k} \rho^\alpha_I dx^I$ with $\rho^\alpha_I \in \mathcal{D}'(X)$. Because $\rho$ is globally defined it follows that these locally defined generalized $k$-forms form a coherent family, so since $\Omega^k(X)'$ is a sheaf they define a unique element $\omega_\rho \in \Omega^k(X)'$ corresponding to $\rho$. Conversely, let $\omega \in \Omega^k(X)'$ and set $\rho^\alpha_\omega = M_a(\omega|_{V_\alpha})$ as above. Let $(\chi_j)_j$ be a partition of unity subordinate to $(V_\alpha)_\alpha$ ($\mathrm{supp}\chi_j \subset\subset V_{\alpha_j}$) and for $\xi_1, \dots, \xi_k \in \mathfrak{X}(X)$ set

$$\rho_\omega(\xi_1, \dots, \xi_k) = \sum_{j_1, \dots, j_k} \rho^{\alpha_{j_1}}_\omega (\chi_{j_1}\xi_1, \dots, \chi_{j_k}\xi_k) \tag{31}$$

Obviously this definition is independent of the partition of unity chosen (and $\rho^{\alpha_{j_1}}_\omega$ could be replaced by any $\rho^{\alpha_{j_l}}_\omega$ ($1 \le l \le k$)). By construction, the linear maps $\rho \mapsto \omega_\rho$ and $\omega \mapsto \rho_\omega$ are inverse.

By 1.44 we actually already have a way of "inserting" smooth vector fields into generalized $k$-forms. In fact, both approaches give the same result due to

**1.50 Lemma.** *Let $\omega \in \Omega^k(X)'$, $\xi_1, \dots, \xi_k \in \mathfrak{X}(X)$. Then*

$$\rho_\omega(\xi_1, \dots, \xi_k) = i_{\xi_k} \dots i_{\xi_1}\omega$$

**Proof.** Again since $\mathcal{D}'(X)$ is a sheaf it suffices to establish this equality on a coordinate patch $V_\alpha$. Let $\omega|_{V_\alpha} = \sum_{I \in \mathcal{I}^n_k} \omega^\alpha_I dx^I$. Since $i_\xi(\tau) = 0$ for any $\tau \in \Omega^0(X)'$ we get (making use of 1.44 (i) and (vii))

$$i_{\xi_k} \dots i_{\xi_1}(\sum_{I \in \mathcal{I}^n_k} \omega^\alpha_I dx^I) = i_{\xi_k} \dots i_{\xi_2}(\sum_{I \in \mathcal{I}^n_k} \omega^\alpha_I i_{\xi_1} dx^I) = \ldots = (\sum_{I \in \mathcal{I}^n_k} \omega^\alpha_I i_{\xi_1} \dots i_{\xi_k} dx^I) = \sum_{I \in \mathcal{I}^n_k} (dx^I(\xi_1, \dots, \xi_k))\omega^\alpha_I \tag{32}$$

from which the claim follows by the definition of $\rho^\alpha_\omega$. □

**1.51** Let us now view the above identification from a different angle, again using continuous extension of classical concepts. To this end, let

$$L_a^k = \{\rho : \mathfrak{X}(X)^k \to \mathcal{C}^\infty(X),\ k\text{-linear alternating }\}$$
$$(L_a^k)' = \{\rho : \mathfrak{X}(X)^k \to \mathcal{D}'(X),\ k\text{-linear alternating }\}$$

Let $\rho \in L_a^k$. Then $(\xi_1, \dots, \xi_k) \to j(\rho(\xi_1, \dots, \xi_k))$ defines an element $\rho'$ of $(L_a^k)'$. $\rho \mapsto \rho'$ is an injection of $L_a^k$ into $(L_a^k)'$ and we will consider $L_a^k$ as a subspace of $(L_a^k)'$ by identifying $\rho'$ with $\rho$. We endow $(L_a^k)'$ with the topology of pointwise weak convergence, i.e., $\rho_m \to \rho$ iff $\rho_m(\xi_1, \dots, \xi_k) \to \rho(\xi_1, \dots, \xi_k)$ in $\mathcal{D}'(X)$ for all $\xi_1, \dots, \xi_k \in \mathfrak{X}(X)$. From classical tensor analysis we know that the map

$$F : \Omega^k(X) \to L_a^k$$
$$F(\omega)(\xi_1, \dots, \xi_k)(p) = \omega(p)(\xi_1(p), \dots, \xi_k(p))$$

is a linear isomorphism.

**1.52 Theorem.** *$F$ possesses a unique continuous extension $F' : \Omega^k(X)' \to (L_a^k)'$. $F'$ is a bijective homeomorphism given by*

$$F'(\omega)(\xi_1, \dots, \xi_k) = i_{\xi_k} \dots i_{\xi_1} \omega \,.$$

**Proof.** Clearly $F'(\omega) \in (L_a^k)'$ for each $\omega \in \Omega^k(X)'$ and $F'$ is continuous since for $\tau \in \Omega_c^n(X)$ we have $F'(\omega)(\xi_1, \dots, \xi_k)(\tau) = \omega(i_{\xi_k} \dots i_{\xi_1} \tau)$. Since any (compactly supported) $(n-k)$-form can be written as $i_{\xi_k} \dots i_{\xi_1} \tau$ for a suitable $\tau \in \Omega_c^n(X)$, this formula also shows injectivity of $F'$. Surjectivity of $F'$ has already been established above. To show that $F'$ coincides with $F$ on $\Omega^k(X)$ let $\omega \in \Omega^k(X)$. Then

$$j(F(\omega)(\xi_1, \dots, \xi_k)) = j(i_{\xi_k} \dots i_{\xi_1} \omega) = i_{\xi_k} \dots i_{\xi_1} j(\omega)$$

by 1.44, which yields the result. Finally, to show that $(F')^{-1}$ is continuous, suppose that $F'(\omega_m) \to 0$. Then for all $\xi_1, \dots, \xi_k \in \mathfrak{X}(X)$ it follows that $i_{\xi_k} \dots i_{\xi_1} \omega_m \to 0$. But then by (32) for each coordinate patch $V_\alpha$ and all $I \in \mathcal{I}_k^n$ we have $(\omega_m)_I^\alpha \to 0$ which is equivalent to $\omega_m \to 0$. □

Having established this result we may now simply write $\omega(\xi_1, \dots, \xi_k)$ for $i_{\xi_k} \dots i_{\xi_1} \omega$ ($\omega \in \Omega^k(X)'$), i.e., we may use the same notation as in the classical case. Moreover, by continuous extension we immediately derive the following result (for its classical counterpart, cf. [1], 2.4.15).

**1.53 Proposition.** *Let $\omega \in \Omega^k(X)'$, $\tau \in \Omega^l(X)$ and $\xi_i \in \mathfrak{X}(X)$. Then*

*(i)* $\omega \wedge \tau(\xi_1, \dots, \xi_{k+l}) = \sum_{\sigma \in \Upsilon_{k+l}} (\text{sign}\sigma) \omega(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) \tau(\xi_{\sigma(k+1)}, \dots, \xi_{\sigma(k+l)})$.

*(ii)* $(L_\xi \omega)(\xi_1, \dots, \xi_k) = L_\xi(\omega(\xi_1, \dots, \xi_k)) - \sum_{i=1}^k \omega(\xi_1, \dots, [\xi, \xi_i], \dots, \xi_k)$.

*(iii)* $d\omega(\xi_0, \dots, \xi_k) = \sum_{i=0}^k (-1)^i L_{\xi_i} \omega(\xi_0, \dots, \widehat{\xi_i}, \dots, \xi_k) + \sum_{0 \le i < j \le k} (-1)^{i+j} \omega([\xi_i, \xi_j], \xi_0, \dots, \widehat{\xi_j}, \dots, \xi_k)$.

**1.54** In complete analogy to the previous section we now define the set of *distributional tensor fields of type* $(r, s)$ by

$$\mathcal{T}_s^r(X)' = \{t : \mathfrak{X}^*(X)^r \times \mathfrak{X}(X)^s \to \mathcal{D}'(X),\ \mathcal{C}^\infty(X)\text{-multilinear}\}\,. \qquad (33)$$

(where $\mathfrak{X}^*(X) = \Omega^1(X) = \mathcal{T}_1^0(X)$). We equip $\mathcal{T}_s^r(X)'$ with the topology of pointwise weak convergence, i.e., $t_m \to t$ means $t_m(\omega_1, \dots, \omega_r, \xi_1, \dots, \xi_s)(\tau) \to t(\omega_1, \dots, \omega_r, \xi_1, \dots, \xi_s)(\tau)$ for all $\omega_i \in \mathfrak{X}^*(X)$, $\xi_j \in \mathfrak{X}(X)$, $\tau \in \Omega_c^n(X)$. The two most important special cases of (33) are $\mathfrak{X}(X)' := \mathcal{T}_0^1(X)'$, the space of generalized vector fields and $\mathfrak{X}^*(X)' := \mathcal{T}_1^0(X)' = \Omega^1(X)'$, the space of generalized one-forms (by 1.49). Let

$$(L^k)^* = \{t : \mathfrak{X}^*(X)^r \times \mathfrak{X}(X)^s \to \Omega_c^n(X)^*,\ \mathcal{C}^\infty(X)-\text{multilinear}\}\,.$$

(equipped with the topology of pointwise weak convergence) and define

$$G_s^r : \mathcal{T}_s^r(X) \to (L^k)^*$$
$$G_s^r(t)(\omega_1, \dots, \omega_r, \xi_1, \dots, \xi_s) = j(t(\omega_1, \dots, \omega_r, \xi_1, \dots, \xi_s))$$

The analogue of thm. 1.52 is given by

**1.55 Theorem.** *$G^r_s$ is injective and*

$$\overline{G^r_s(\mathcal{T}^r_s(X))}^{wsc} = \mathcal{T}^r_s(X)'\,. \tag{34}$$

Clearly any $t \in \mathcal{T}^r_s(X)'$ is local (just adapt the argument following (29) to the present context) and as in (30) it follows that on any $V_\alpha$ each $t \in \mathcal{T}^r_s(X)'$ can uniquely be written in the form

$$t|_{V_\alpha} = \sum_{\substack{i_1,\dots,i_r\\ j_1,\dots,j_s}} (t^\alpha)^{i_1,\dots,i_r}_{j_1,\dots,j_s} \partial_{x^{i_1}} \otimes \ldots \otimes \partial_{x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s} \tag{35}$$

with $(t^\alpha)^{i_1,\dots,i_r}_{j_1,\dots,j_s} \in \mathcal{D}'(V_\alpha)$. Hence we see from the local description of sec 1.D that on an oriented manifold $\mathcal{T}^r_s(V_\alpha)'$ may be identified with the spaces $\mathcal{D}'{}^r_s(V_\alpha)$ of tensor distributions defined in 1.14.
Since also $\mathcal{T}^r_s(V)'$ is a fine sheaf on $X$ we can again talk about support and singular support of generalized tensor fields. Moreover, it is now possible to extend classical tensor analysis to generalized tensor fields. However, we have to carefully avoid ill-defined products (e.g. in the following prop. only one factor may be generalized).

**1.56 Proposition.** *The map $\otimes : \mathcal{T}^r_s(X) \times \mathcal{T}^p_q(X) \to \mathcal{T}^{r+p}_{s+q}(X)$ possesses a unique continuous extension $\otimes : \mathcal{T}^r_s(X)' \times \mathcal{T}^p_q(X) \to \mathcal{T}^{r+p}_{s+q}(X)'$ given by*

$$t_1 \otimes t_2(\omega_1,\dots,\omega_{r+p},\xi_1,\dots,\xi_{s+q})(\tau) = t_1(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s)(t_2(\omega_{r+1},\dots,\omega_{r+p},\xi_{s+1},\dots,\xi_{s+q})\tau) \tag{36}$$

**Proof.** If $t_1 \in \mathcal{T}^r_s(X)$ and $t_2 \in \mathcal{T}^p_q(X)$ then

$$((t_1 \otimes t_2)(\omega_1,\dots,\omega_{r+p},\xi_1,\dots,\xi_{s+q}))(\tau) = \int t_1(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s) t_2(\omega_{r+1},\dots,\omega_{r+p},\xi_{s+1},\dots,\xi_{s+q})\tau \tag{37}$$

Thus since the map defined by (36) is obviously continuous and coincides with (37) in the smooth case, the claim follows. □

**1.57** By a similar strategy we shall now extend the action of a diffeomorphism $F : X \to Y$ to generalized tensor fields: To begin with, let $t \in \mathcal{T}^r_s(Y)$ and $\tau \in \Omega^n_c(X)$. Then

$$\begin{aligned}(F^*t)(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s)(\tau) &= \int F^*(t((TF^{-1})^*(\omega_1),\dots,TF(\xi_s)))\tau\\ &= \int t((TF^{-1})^*(\omega_1),\dots,TF(\xi_s)))(F^{-1})^*(\tau)\,.\end{aligned}$$

Hence the unique continuous extension of $F^*$, again denoted $F^* : \mathcal{T}^r_s(Y)' \to \mathcal{T}^r_s(X)'$ is given by

$$((F^*t)(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s))(\tau) = t((TF^{-1})^*(\omega_1),\dots,TF(\xi_s)))(F^{-1})^*(\tau)) \tag{38}$$

Next, let $t \in \mathcal{T}^r_s(X)$, $\xi \in \mathfrak{X}(X)$ and $\tau \in \Omega^n_c(X)$. Then

$$\begin{aligned}&(L_\xi t)(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s)(\tau) = \int (L_\xi t)(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s)\tau = \int L_\xi(t(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s))\tau\\ &-\sum_{i=1}^r \int t(\omega_1,\dots,L_\xi\omega_i,\dots,\omega_r,\xi_1,\dots,\xi_s)\tau - \sum_{j=1}^s \int t(\omega_1,\dots,\omega_r,\xi_1,\dots,L_\xi\xi_j,\dots,\xi_s)\tau\end{aligned}$$

Thus the unique continuous extension of $L_\xi$ to $\mathcal{T}^r_s(X)'$ is given by

$$\begin{aligned}&(L_\xi t)(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s)(\tau) = (L_\xi(t(\omega_1,\dots,\omega_r,\xi_1,\dots,\xi_s)))(\tau)\\ &-\sum_{i=1}^r t(\omega_1,\dots,L_\xi\omega_i,\dots,\omega_r,\xi_1,\dots,\xi_s)(\tau) - \sum_{j=1}^s \int t(\omega_1,\dots,\omega_r,\xi_1,\dots,L_\xi\xi_j,\dots,\xi_s)(\tau)\end{aligned} \tag{39}$$

Moreover, by continuous extension or direct calculation, the following properties hold:

$$\begin{aligned} L_\xi(t_1 \otimes t_2) &= (L_\xi t_1) \otimes t_2 + t_1 \otimes L_\xi t_2 \\ F^*(L_\xi t) &= L_{F^*\xi}(F^* t) \\ (F \circ G)^* &= G^* \circ F^* \end{aligned}$$

Finally, we have the following generalization of thm. 1.47:

**1.58 Theorem.** *Let $\xi \in \mathfrak{X}(X)$ be complete. Then for each $s \in \mathcal{T}^r_s(X)$ the map $t \to (\mathrm{Fl}^\xi_t)^* s$ is (weakly) differentiable and*

$$L_\xi((\mathrm{Fl}^\xi_t)^* s) = (\mathrm{Fl}^\xi_t)^* L_\xi s = \frac{d}{dt}(\mathrm{Fl}^\xi_t)^* s\,. \tag{40}$$

**1.59** So far we have only considered Lie derivatives with respect to smooth vector fields. We are now going to consider the case where the vector field itself is generalized. To begin with, we introduce the following spaces of derivations:

$$\begin{aligned} \mathrm{Der}(X) &= \{h : \mathcal{C}^\infty(X) \to \mathcal{C}^\infty(X),\ \mathbb{R}\text{-linear},\ h(fg) = h(f)g + fg(h)\} \\ \mathrm{Der}(X)^* &= \{h : \mathcal{C}^\infty(X) \to \Omega^n_c(X)^*,\ \mathbb{R}\text{-linear},\ h(fg) = h(f)g + fg(h)\} \end{aligned}$$

The map $k : \mathrm{Der}(X) \to \mathrm{Der}(X)^*$, $k(h)(f) = (\tau \mapsto \int h(f)\tau)$ provides an injection of $\mathrm{Der}(X)$ into $\mathrm{Der}(X)^*$ and we set $\mathrm{Der}(X)' = \overline{k(\mathrm{Der}(X))}^{wsc}$ ($h_m \to h$ in $\mathrm{Der}(X)^*$ means that $h_m(f)\tau \to h_m(f)\tau$ for all $f \in \mathcal{C}^\infty(X)$ and all $\tau \in \Omega^n_c(X)$). Elements of $\mathrm{Der}(X)'$ are called *generalized derivations on $X$*. By continuous extension it is clear that any $h \in \mathrm{Der}(X)'$ is a derivation from $\mathcal{C}^\infty(X)$ to $\mathcal{D}(X)'$, but not necessarily vice versa.

**1.60 Theorem.** *The map $L : \mathfrak{X}(X) \to \mathrm{Der}(X)$, $\xi \to L_\xi$ has a unique continuous extension (again denoted by) $L : \mathfrak{X}(X)' \to \mathrm{Der}(X)'$. $L$ is bijective, $L$ and $L^{-1}$ are continuous and $L_\xi(f) = \xi(df)$ for all $f \in \mathcal{C}^\infty(X)$ and all $\xi \in \mathfrak{X}(X)'$.*

**Proof.** We temporarily set $L' : \mathfrak{X}(X)' \to \mathrm{Der}(X)^*$, $L'_\xi(f) = \xi(df)$. Then obviously $L'(\mathfrak{X}(X)') \subseteq \mathrm{Der}(X)'$, $L'$ is continuous and coincides with $L$ on $\mathfrak{X}(X)$. Let $L'_\xi = 0$. Then $\xi(df) = 0$ for all $f \in \mathcal{C}^\infty(X)$. From the local representation of $\xi$ it follows that $\xi = 0$, hence $L'$ is injective. To show that $L'$ is surjective, let $h \in \mathrm{Der}(X)'$ and let $(h_m)_m$ be a sequence in $\mathrm{Der}(X)$ converging to $h$. There exists a sequence $(\xi_m)_m$ in $\mathfrak{X}(X)$ such that $h_m = L_{\xi_m}$ for each $m$. It follows that $\xi_m(df)$ converges for each $f \in \mathcal{C}^\infty(X)$. From the local representation of the $\xi_m$ we conclude that there exists some $\xi \in \mathfrak{X}(X)'$ such that $\xi_m \to \xi$. Now $L'_\xi(f) = \lim L'_{\xi_m}(f) = h(f)$ for all $f$, so $h = L'_\xi$ and the claim follows. Finally, $(L')^{-1}$ is continuous since $L'_{\xi_m} \to L'_\xi$ implies $\xi_m(df) \to \xi(df)$ for all $f$, hence $\xi_m \to \xi$. □

**1.61 Theorem.** *The Lie bracket $[\,,\,] : \mathfrak{X}(X) \times \mathfrak{X}(X) \to \mathfrak{X}(X)$ has a unique extension $[\,,\,] : \mathfrak{X}(X)' \times \mathfrak{X}(X) \to \mathfrak{X}(X)'$ which is continuous in the first variable. This extension is uniquely determined by*

$$L_{[\xi,\eta]} f = L_\xi(L_\eta f) - L_\eta(L_\xi f) \in \mathcal{D}'(X) \qquad (\xi \in \mathfrak{X}(X)',\ \eta \in \mathfrak{X}(X)) \tag{41}$$

*and satisfies the Jacobi identity*

$$[\xi_1, [\xi_2, \xi_3]] = [[\xi_1, \xi_2], \xi_3] + [\xi_2, [\xi_1, \xi_3]] \qquad (\xi_1 \in \mathfrak{X}(X)',\ \xi_2, \xi_3 \in \mathfrak{X}(X))\,.$$

**Proof.** Let $h(f) := L_{[\xi,\eta]} f$ as defined in (41). Then clearly $h$ is a derivation. In fact, $h \in \mathrm{Der}(X)'$: Choose a sequence $(\xi_m)_m$ in $\mathfrak{X}(X)$ converging to $\xi$. Since $L_\eta$ is continuous by thm. 1.43 it follows that $[L_{\xi_m}, L_\eta] \to h$. By thm. 1.60 there exists a unique generalized vector field $[\xi, \eta]$ with $h = L_{[\xi,\eta]}$. Furthermore, if $(\xi_m)_m$ is a sequence in $\mathfrak{X}(X)$ converging to some $\xi \in \mathfrak{X}(X)'$ then as above it follows that $L_{[\xi_m,\eta]} = [L_{\xi_m}, L_\eta]$ converges to $L_{[\xi,\eta]}$. But then by 1.60 $[\xi_m, \eta] \to [\xi, \eta]$, so the Lie bracket is continuous with respect to $\xi$. The Jacobi identity now follows by continuous extension. □

For $\xi \in \mathfrak{X}(X)'$, $\eta \in \mathfrak{X}(X)$, the local expression of $[\xi, \eta]$ with respect to coordinates $x^i$ is given by

$$[\xi,\eta]|_{V_\alpha} = \sum_{i,j}(\xi^i \frac{\partial \eta^j}{\partial x^i} - \eta^i \frac{\partial \xi^j}{\partial x^i})\partial_{x^j}$$

This again follows by continuous extension, as does the following: For any orientation preserving diffeomorphism $f : X \to Y$ we have

$$f^*[\xi,\eta] = [f^*\xi, f^*\eta] \qquad (\xi \in \mathfrak{X}(X)',\, \eta \in \mathfrak{X}(X))\,,$$

where $f^*\xi(g) = f^*(\xi((f^{-1})^*g))$ is the unique continuous extension of $f^*$ in the smooth case.

Analogously we obtain the following

**1.62 Theorem.** *The map $L : \mathfrak{X}(X) \times \mathcal{T}^r_s(X) \to \mathcal{T}^r_s(X)$, $(\xi, t) \to L_\xi t$ has a unique extension $L : \mathfrak{X}(X)' \times \mathcal{T}^r_s(X) \to \mathcal{T}^r_s(X)'$ continuous in $\xi$. $L$ is given by (39) (with $\xi \in \mathfrak{X}(X)'$).*

**1.63** (Stokes) Suppose $X$ is compact with boundary $\partial X$ and let $\omega \in \Omega^{n-1}_c(X)'$ have singular support $C \subset int(X)$. Then $d\omega$ is integrable on $X$ and

$$\int_X d\omega = \int_{\partial X} i^*\omega$$

(with $i : \partial X \hookrightarrow X$).

**Proof.** To show that $d\omega$ is integrable let $(\chi_i)_i$ be a partition of unity and let $\chi = \sum_{\mathrm{supp}(\chi_i)\cap C \neq 0} \chi_i$. Then

$$\sum d\omega(\chi_i) = d\omega(\chi) + \int (1-\chi)d\omega$$

For a second partition of unity $(\zeta_j)_j$ we have $\int(1-\chi)d\omega - \int(1-\zeta)d\omega = \int(\zeta-\chi)d\omega$ and $d\omega(\chi) - d\omega(\zeta) = \int(\chi-\zeta)d\omega$, so the claim follows. By 1.42 (iii) we have $\int d(\chi d\omega) = 0$, so $\int \chi d\omega = -\int (d\chi)\wedge\omega$. Moreover, $(1-\chi)d\omega = d((1-\chi)\omega) + (d\chi)\wedge\omega$. Both summands in this expression are smooth (note that $d\chi = 0$ on the singular support of $\omega$). Thus, finally, the classical theorem of Stokes gives

$$\int_X d\omega = \int_X d((1-\chi)\omega) = \int_{\partial X} i^*((1-\chi)\omega) = \int_{\partial X} i^*\omega$$

(since $\chi = 0$ on $\partial X$). □

**1.64** A *distributional semi-Riemannian metric* is an element $g$ of $\mathcal{T}^0_2(X)'$ which is symmetric and non-degenerate in the sense that $g(\omega_1,\omega_2) = 0$ for all $\omega_2 \in \mathfrak{X}(X)$ implies $\omega_1 = 0$. A *distributional connection* is a map $\nabla : \mathfrak{X}(X) \times \mathfrak{X}(X)' \to \mathfrak{X}(X)'$ satisfying (we write $\nabla_\xi\eta$ for $\nabla(\xi,\eta)$)

(i) $\nabla_\xi\eta$ is $\mathcal{C}^\infty(X)$-linear in $\xi$.

(ii) $\nabla_\xi\eta$ is $\mathbb{R}$-linear in $\eta$.

(ii) $\nabla_\xi(f\eta) = f\nabla_\xi\eta + \eta L_\xi f$.

However, the limitations of distributional geometry become apparent when trying to define generalized curvature: The natural choice, namely

$$R_{\xi\eta}\zeta = \nabla_{[\xi,\eta]}\zeta - [\nabla_\xi, \nabla_\eta]\zeta$$

is not defined in general as it would require multiplication of distributions. Only if the coefficients of $\nabla$ are in some kind of function algebra (e.g. a Sobolev space) one can make sense out of such expressions within classical distribution theory.

In the next chapter, after reviewing the problems of a multiplication of distributions in general, we shall discuss in detail the limitations of a distributional general relativity originating from the above mentioned problems. In section 4.C, however, we will introduce a more general framework—mainly due to J. F. Colombeau—which will avoid these shortcommings.

## 2. LIMITATIONS OF DISTRIBUTION THEORY

This chapter provides an overview of the limitations of the *linear* theory of distributions with special emphasis on its implications to applications in the genuinely *nonlinear* theory of general relativity.
In section 2.A we give an outline of the basic inconsistencies one encounters on trying to define a multiplication of distributions. Since this topic has been discussed in great detail and generality elsewhere we only give a brief summary and refer the interested reader to the literature, in particular to the books of Rosinger [165], part I, chaps. 2 and 3, and most notably of Oberguggenberger [144], chap. I.
Section 2.B is devoted to a study of applicability of linear distribution theory to general relativity. There we essentially follow a prominent paper by Geroch and Traschen [71] reaching the conclusion that sources of the gravitational field may at most be concentrated to submanifolds of codimension one in spacetime in order to be described consistently within distributions.

While the first section provides a strong intrinsically mathematical motivation for a "nonlinear theory of generalized functions," the latter one points out that such a theory is indispensable in the description of such physically interesting sources of the gravitational field like e.g. cosmic strings and point particles.

### 2.A. Multiplication of Distributions: Difficulties and Impossibility Results

Before focusing on the problems that arise on stepping only little over the borders of linear distribution theory we make some remarks on the strength and usefulness of the latter in its natural domain, the theory of linear PDEs. Since it is so tempting, however, we begin with a few historical remarks on the "prehistory" of distribution theory, the details of which may be found in the book of Lützen [134].

**2.1** Some physicists and also mathematicians were using "generalized functions ideas" thereby for a long time anticipating the later rigorous theory. The names to mention here are most of all Fourier, Kirchhoff and Heaviside. Dirac in 1926 [54] and later in his famous book [55] introduced the concept of the $\delta$-function which allowed him to draw an analogy between "discrete" and "continuous variables," thereby reaching a unified theory of quantum mechanics combining matrix mechanics and wave mechanics. Since these ideas were so beautiful and convincing the "Dirac-$\delta$" very soon became a widespread tool for physicists. However, the notation $\delta$ does not stand for "Dirac" but was originally chosen by Dirac to put emphasis on the analogy between $\delta(p-q)dp$ and the unity matrix which, in physics, commonly is written as the "Kronecker-$\delta$" $\delta_{pq}$ (cf. [134], p. 124). On the other hand, descriptions of the $\delta$-distribution as a limit of a series of (smooth) functions go back as far as 1822 and Fourier [64]. Also Kirchhoff [107] already in 1882 fully captured the concept of $\delta$ calculating the fundamental solution of the wave operator in $\mathbb{R}^{1+3}$.
A rigorous theory which first (implicitly) used distributions was given in 1932 by Bochner [26] while the first definition of distributions in the modern sense (as functionals) appeared in Sobolev's 1936 paper [170]. Hence (according to Lützen [134], p. 159ff.) he may be be called the *inventor* of distributions, while finally L. Schwartz *created the theory* of distributions in his classical monograph [167], first published in 1950.
Schwartz' theory rapidly was well received both by mathematicians and physicists who now could use "improper functions" in a well-defined sense. In mathematics, distribution theory was the essential tool to build an elaborate solution concept for linear PDEs. Most notably among the numerous contributions in that field is the famous theorem proved individually by Ehrenpreis [58] and Malgrange [135] guaranteeing the existence of a fundamental solution (in $\mathcal{D}'$) for every non zero constant coefficient linear partial differential operator. This result later on was refined (cf. [89], chap. X) by Hörmander [90] (existence in a local Besov space), Hörmander [91] and Łojasiewicz [127] (existence even in $\mathcal{S}'$), Treves [181] (smooth dependence on parameters), again Hörmander [92] (constructive proof) and even recently Ortner and Wagner [148] (containing the most explicit formula involving only one integral over $S^1$).

**2.2** However, the natural limitations of distribution theory became clear even in 1957 when Lewy [123] gave his famous example of a linear PDE with smooth coefficient functions which does not allow for a distributional solution.
There is quite a large number of occasions displaying the need for a concept of multiplication of distributions or even more general nonlinear operations. We give a short list and again refer to [144], I §1 for a more comprehensive treatment.

- Intrinsic problems in distribution theory such as restriction to submanifolds, no continuous calculation of convolutions via Fourier transform, etc.

- Nonlinear PDEs with singular coefficients or data, e.g., semilinear systems with rough initial data, delta waves in semilinear hyperbolic systems, shock waves in quasilinear nonconservative systems, equations without $\mathcal{D}'$-solutions, etc.
- Applications in mathematical physics, e.g., quantum field theory (the vast topic of renormalization), general relativity (ultrarelativistic limits, distributional geometries), etc.

**2.3** Let us now start looking at some examples showing the various difficulties one encounters when trying to define nonlinear operations of distributions. Generally speaking, one may divide the strategies to overcome them into two categories. The first—*intrinsic products*—deals with multiplying two arbitrary distributions to obtain a distribution again. This indeed is impossible in a sense to be made precise later.
The second one deals with differential algebras containing the space of distributions. Here the product of each pair of distributions is well-defined but consistency with classical operations together with the algebraic properties of commutativity and associativity becomes impossible.

**2.4 Example.** The product of $\delta$ and $H$.
We try to define the product of the $\delta$-distribution with the Heaviside step function $H$ by means of a regularization and passage to the $\mathcal{D}'$-limit. For this, let $\rho_\varepsilon$ be a *a strict delta net* (cf. [144], def. 7.1), i.e., a net of smooth functions satisfying $(0 < \varepsilon < 1)$

$$\begin{array}{ll} \text{(i)} & \operatorname{supp}(\rho_\varepsilon) \to 0 \quad (\varepsilon \to 0) \\ \text{(ii)} & \displaystyle\int \rho_\varepsilon(x)\,dx \to 1 \quad (\varepsilon \to 0) \\ \text{(iii)} & |\rho_\varepsilon|_{L^1} \text{ is uniformly bounded in } \varepsilon\ . \end{array} \tag{42}$$

Clearly $\rho_\varepsilon \to \delta$ in $\mathcal{D}'$. Similarly let $H_\varepsilon$ be smooth for all $\varepsilon$, converging almost everywhere to $H$ and bounded uniformly in $\varepsilon$. Evidently the limit $\lim_{\varepsilon\to 0} H_\varepsilon\, \rho_\varepsilon$ depends on the chosen regularizations. More precisely, if $\operatorname{supp}(\rho_\varepsilon) \subseteq (-\varepsilon, \varepsilon)$ and $H_\varepsilon \equiv 0$ on $(-\infty, +\varepsilon]$ for all $\varepsilon$ then $H_\varepsilon\rho_\varepsilon \to 0$. On the other hand, if $H_\varepsilon$ equals one on $[-\varepsilon, \infty)$ we immediately obtain $H_\varepsilon\rho_\varepsilon \to \delta$. However, there is the seemingly reasonable special choice $H_\varepsilon := H * \rho_\varepsilon$ leading to

$$(H * \rho_\varepsilon)\,\rho_\varepsilon = \frac{d}{dx}\,\frac{1}{2}\,(H * \rho_\varepsilon)^2 \;\to\; \frac{1}{2}\,\frac{d}{dx}\,H \;=\; \frac{1}{2}\,\delta\ , \tag{43}$$

which indeed is proposed by some authors. Accepting this for a moment we try to solve the simple ODE (for an almost humorous discussion of this equation see [79])

$$\frac{d}{dt}\,y(t) \;=\; \delta(t)\,y(t) \qquad y(-\infty) = 1\ .$$

Using the *ansatz* $y(t) = 1 + \alpha H(t)$ as well as (43) we find $\alpha\delta = (1 + \alpha/2)\delta$. Hence $\alpha = 2$ and

$$y(t) \;=\; 1 + 2\,H(t)\ . \tag{44}$$

Now, should we trust this solution? An alternative approach motivated by the requirement of stability against perturbations is as follows: First regularize the singular coefficient $\delta$, solve the equation and then pass to the limit. The regularized equation

$$\frac{d}{dt}\,y_\varepsilon(t) \;=\; \rho_\varepsilon(t)\,y_\varepsilon(t) \qquad y_\varepsilon(-\infty) = 1$$

is solved by $y_\varepsilon(t) = \exp(\int_{-\infty}^t \rho_\varepsilon(s)ds)$ which converges to

$$y_0(t) \;=\; 1 + (e-1)\,H(t)\ ,$$

telling us that our previous solution (44) based on formula (43) is not a good one. However, the discrepancy between solutions $y$ and $y_0$ may be resolved within the framework of measure differential equations [155,156]. Also physicists have succeeded in extracting unambiguous information from such situations (cf. [52]).
To indeed show the failure of any formula for $H\delta$ we consider the conservation law

$$u_t + (\frac{1}{m}u^m)_x \;=\; 0\ . \tag{45}$$

Using the classical weak solution concept we have to interpret it as follows: First take the product of $u$ in the algebra $L^\infty(\mathbb{R}^2)$, then take the weak derivative. One immediately finds the constant traveling wave solution

$$u(t,x) \;=\; H(x - \frac{1}{m}t)\ .$$

On the other hand interpreting (45) as

$$u_t + u^{m-1}u_x \;=\; 0 \tag{46}$$

and using $H^{m-1}(x-ct) = H(x-ct)$ as well as (43) we find the solution

$$u(t,x) \;=\; H(x - \frac{1}{2}t)$$

for whatever value of $m$. The reason behind this inconsistency is that transforming (45) to (46) we used the laws of a commutative differential algebra which are incompatible with the formulas $H^m = H$ and $H' \neq 0$ as we shall see in the next

**2.5 Example.** The powers of the Heaviside function.
We are going to show that given an associative and commutative multiplication on $\mathcal{D}'$ satisfying the Leibniz rule the formula $H^2 = H$ implies that $H$ is a constant, i.e., $\partial H = 0$. Indeed, $\partial(H^2) = (\partial H)H + H\partial H = 2H\partial H$ and $\partial(H^3) = 3H^2\partial H$ together with $H^2 = H (\Rightarrow H^3 = H)$ imply

$$2H\partial H \;=\; \partial H \;=\; 3H\partial H$$

Hence $H\partial H = 0$ and finally $\partial H = 0$.

**2.6 Example.** What is $\delta^2$?
Next we try to define the square of the $\delta$-distribution by means of a regularization. Take first a strict delta net $(\rho_\varepsilon)_\varepsilon$ (as defined in (42)) which we additionally require to be *real valued*. In order to check whether $\rho_\varepsilon^2$ converges in $\mathcal{D}'(\mathbb{R})$ take a test function $\varphi$ equal to unity in a neighborhood of $x = 0$. Then for $\varepsilon$ small

$$\int \rho_\varepsilon^2(x)\,\varphi(x)\,dx \;=\; \int \rho_\varepsilon^2(x)\,dx \tag{47}$$

showing that if $(\rho_\varepsilon)_\varepsilon$ converges weakly it is bounded in $L^2$, hence possesses a $L^2$-convergent subsequence. This in turn implies $\delta \in L^2$ which is wrong! Hence we have proved that squares of real valued strict delta nets never converge in $\mathcal{D}'$.
However, one may construct complex valued regularizations of $\delta$ whose squares converge in distributions to, e.g., $0$, $c\delta$, $c\delta + 1/(2\pi i)\delta'$, $c\delta + c'\delta'$, etc. (with $c$, $c'$ arbitrary complex constants). Similarly, given any complex number $z$ one may construct a net $(\rho_\varepsilon)_\varepsilon$ with $\rho_\varepsilon \to \delta$ and $\rho_\varepsilon^2 \to z\delta$ in distributions (cf. [144], ex 10.6). These facts strongly indicate that one should not try to define $\delta^2$ as an element of $\mathcal{D}'$.

**2.7 Example.** Non-associativity.
Since there is already a product on $\mathcal{D}' \times \mathcal{C}^\infty$ (defined by transposition) it seems only natural to require a product in $\mathcal{D}'$ to be an extension of it. However, this already destroys any chance of getting an *associative* operation as shown by the one line calculation

$$0 \;=\; (\delta(x)\,x)\,\mathrm{vp}\frac{1}{x} \;\neq\; \delta(x)\,(x\,\mathrm{vp}\frac{1}{x}) \;=\; \delta(x)\ ,$$

where $\mathrm{vp}\frac{1}{x}$ denotes the Cauchy principal value of $\frac{1}{x}$.

**2.8 Example.** Non-commutativity.
Motivated by ex. 2.7 we drop associativity from the assumptions of ex. 2.5. Defining $\eta = 2H - 1$, $\eta$ and $\eta^2 \in L^1_{\mathrm{loc}}$ and $\eta^2 = 1$ on $\mathbb{R} \setminus \{0\}$. Assuming the desired multiplication on $\mathcal{D}'$ to satisfy the Leibniz rule and $\eta^2 = 1$ we obtain

$$0 \;=\; (\eta^2)' \;=\; 2\eta\delta + 2\delta\eta\ .$$

Hence either $\delta\eta = 0 = \eta\delta$ or $\delta\eta = -\eta\delta$ which violates *commutativity*.

**2.9** Summing up, the above examples show that in order to define *any* multiplication on $\mathcal{D}'$ one has to pay the price of giving up at least some of the desired properties for such an operation—or to spell it out more drastically: *There is no way to define a "reasonable" product on $\mathcal{D}' \times \mathcal{D}'$.* The alternatives (following once again [144], chap. I.3) may be classified as follows

1. Regular intrinsic operations: This approach only deals with subspaces of $\mathcal{D}'$, where multiplication is defined classically, i.e., function algebras like $L^\infty_{\mathrm{loc}}$, $C^k$ or the Sobolev spaces $H^s(\mathbb{R}^n)$, where $s > n/2$. The prize to pay here is that a product obtained that way can not be extended to all of $\mathcal{D}'$.
2. Irregular intrinsic operations: This category comprises all approaches to assign a product to particular pairs of (possibly singular) distributions such as Fourier products, duality method, regularization, etc. The literature on this topic can only be described as vast. However, as may easily be seen, e.g., from example 2.6, this approach offers many degrees of freedom according to the compatibility conditions one is willing to take into account, i.e., to call "natural."
3. Extrinsic products and algebras containing $\mathcal{D}'$: Here the vector space of distributions is embedded into an algebra, hence a product of any pair of distributions is indeed defined. What becomes impossible here is—as already indicated in 2.3—full compatibility with classical operations.

**2.10** Since from chap. 4 on we are going to follow the latter approach our next aim is collect some natural requirements on an embedding of $\mathcal{D}'(\Omega)$ ($\Omega \subseteq \mathbb{R}^n$ open) into some associative and commutative algebra $(\mathcal{A}(\Omega), +, \circ)$. We then check whether such an algebra indeed may be constructed. We start with the following list

(i) $\mathcal{D}'$ is linearly embedded into $\mathcal{A}$ and $f(x) \equiv 1$ is the unity in $\mathcal{A}$.

(ii) There exist linear derivation operators $\partial_i : \mathcal{A} \to \mathcal{A}$ $(i = 1, \ldots, n)$ satisfying the Leibniz rule.

(iii) $\partial_i|_{\mathcal{D}'}$ is the usual partial derivative $(i = 1, \ldots, n)$.

(iv) $\circ|_{L^\infty_{\mathrm{loc}} \times L^\infty_{\mathrm{loc}}}$ is the usual pointwise product of functions.

However, from ex. 2.5 it follows immediately that in an associative and commutative algebra satisfying (i) and (ii) requirements (iii) and (iv) cannot hold simultaneously. Hence we try relaxing (iv) to

(iv') $\circ|_{\mathcal{C} \times \mathcal{C}}$ is the usual pointwise product.

This situation has already been analyzed by L. Schwartz [168] himself, as early as 1954, reaching the famous negative answer.

**2.11 Theorem.** (Impossibility result) *There exists no associative and commutative algebra* $(\mathcal{A}(\mathbb{R}), +, \circ)$ *satisfying (i)-(iii) and (iv').*

**Proof.** Suppose that $(\mathcal{A}(\mathbb{R}), +, \circ)$ satisfies the above properties and define, as usual $x_+(x) = \int_0^x H(t)dt$. Then (iv') implies $x_+ \circ x = x_+^2$ and $x \circ (x \log|x| - x) = x^2 \log|x| - x^2$. Thus

$$\begin{aligned} D^2(x_+) \circ x &= D^2(x_+ \circ x) - 2D(x_+) \circ D(x) - x_+ \circ D^2(x) = D^2(x_+^2) - 2D(x_+) = 0, \text{ and} \\ x \circ D^2(x \log|x| - x) &= D^2(x \circ (x \log|x| - x)) - 2D(x)D(x \log|x| - x) - D^2 x \circ (x \log|x| - x) \\ &= D^2(x^2 \log|x| - x^2) - 2D(x \log|x| - x) = D(2x \log|x| - x) - D(2x \log|x| - 2x) = D(x) = 1 \end{aligned}$$

Therefore,

$$D^2(x_+) = D^2(x_+) \circ (x \circ D^2(x \log|x| - x)) = (D^2(x_+) \circ x) \circ D^2(x \log|x| - x) = 0,$$

contradicting $D^2(x_+) = D(H) = \delta \neq 0$. $\square$

**2.12** By integrating the functions involved in the above proof we see that further weakening condition (iv') to $C^k$-functions won't do the trick either. It is actually this result that has led to the popular statement that "distributions cannot be multiplied." However, by taking a closer look at the proof of the impossibility result we see that the basic incompatibilities already arise between *differentiation* of $C^1$-functions, *multiplication* of continuous functions and the concept of a *$\delta$-function.* Hence we see that in order to actually construct

associative and commutative algebras containing $\mathcal{D}'$ we have to give up such seemingly harmless formulas as, e.g., $H^n = H$ or $x\delta = 0$.
Alternatively it's possible to give up associativity. Indeed König [108,109] constructed non-associative algebras satisfying conditions (i)-(iii) and (iv'). On the other hand, E. E. Rosinger [164–166] has undertaken fundamental investigations of the structure of algebras containing distributions creating a general theory characterizing these spaces.

**2.13** After all these discouraging insights it comes somewhat as a surprise that it is indeed possible to construct associative and commutative algebras $\mathcal{G}$ satisfying (i)-(iii) and the further weakened condition

(iv") $\circ|_{\mathcal{C}^\infty \times \mathcal{C}^\infty}$ is the usual pointwise product.

Such algebras have been introduced in the early 1980s by J. F. Colombeau [38,39,41] and will be our home from chap. 4 on. In view of the above discussions they have optimal consistency properties with respect to classical operations. Additionally, the embedding of $\mathcal{D}'$ into $\mathcal{G}$ is given explicitly using convolution (a feature which distinguishes $\mathcal{G}$ from the algebras introduced by Egorov [59,60]). So far, Colombeau algebras are the only known differential algebras providing all these favorable properties which makes them so special within Rosinger's general framework.

Before heading into the depths of Colombeau theory in chaps. 3 and 4 where we are going to set up a framework of nonlinear generalized functions suited to the needs of general relativity, the next section discusses in some detail the limitations of the linear theory of distributions in this context.

## 2.B. Implications for General Relativity

This section provides a discussion of the applicability of linear distribution theory in the context of general relativity. We shall see that the limitations discussed in detail above together with the nonlinearity of Einstein's equations put quite strong constraints on the nature of distributional sources of the gravitational field to be described within the classical (i.e., linear) theory. More precisely, we are going to set up a "maximal" distributional framework by finding the "largest possible" class of spacetime metrics which allow for a distributional formulation of the field equations. Moreover, we are going to define a notion of convergence in this class, ensuring the convergence also of the respective curvature quantities.
However, by a result of Geroch and Traschen [71] this maximal class of metrics only allows for a description of gravitational sources supported on a three dimensional submanifold of spacetime (thin shells). This seriously restricts its usefulness since physically interesting sources like cosmic strings (concentrated on two-dimensional submanifolds of spacetime) or even point particles (one-dimensional support) are strictly excluded. Speaking more clearly, the framework of linear distribution theory is too restrictive for general relativity. The way out provided by Colombeau's theory of nonlinear generalized functions which will be the topic of chaps. 4 and 5.
Our discussion widely parallels the one give in the seminar notes [172] which again is heavily based upon the original article by Geroch and Traschen [71].

**2.14** Our notations as far as general relativity is concerned are quite standard. We follow the conventions of Wald [187]; in particular we use the metric signature $(-+++)$ and abstract index notation . That is, we denote an $(r, s)$-tensor field by $T^{a_1\ldots a_r}_{b_1\ldots b_s} \in \mathcal{T}^r_s(X)$ while Greek indices, i.e., $T^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s}$, are used to denote its components with respect to a certain basis. Hence equations involving Latin indices are "true" tensor equations holding in any basis (cf. [187], 2.4).
Also the more mathematically oriented reader may wish to consult the monograph by O'Neill [147]. For the sake of $L^p$- and $H^s$-theory we refer to the standard book by Adams [2], chaps. II and III.

**2.15** Idealizations play an overall role in modelling physical situations; a particularly useful one is to replace smooth extended densities by "concentrated sources" whenever the density is confined to a "small region" in space and its internal structure is negligible (e.g. point charges in electrodynamics). On trying to describe this idealization mathematically one is led to distributions in a natural way.
In the case of, e.g., electrodynamics distribution theory in fact furnishes a consistent framework, i.e., provides the following two features: First since Maxwell equations are linear with respect to sources and fields they

make sense within distributions[3] and second it is guaranteed that (say smooth) charge densities close—in the sense of $\mathcal{D}'$-convergence—to, e.g., a point charge produce fields that are close to the Coulomb field. While the first property allows for a mathematically sound formulation it is precisely the latter one which renders the idealization physically sensible.
One would wish for a similar mathematical description of concentrated sources in the theory of general relativity. However, its field equations, i.e., Einstein's equations form a highly complicated system of *nonlinear* PDEs. More precisely, since the spacetime metric and its first derivatives enter nonlinearly, the field equations simply cannot be formulated for distributional metrics.

**2.16** Despite this conceptual obstacle spacetimes involving an energy-momentum tensor supported on a hypersurface of spacetime have long since been used in general relativity (see [121,122,45] and [97,146], as well as the references therein).
Consider a submanifold $S$ of codimension one dividing spacetime into a "lower" and "upper" part and let the metric be smooth up to and including $S$ from each of its sides but allow for a jump of its first derivatives across $S$. Writing out Einstein equations in terms of the extrinsic curvature of $S$ one finds junction conditions closely similar to the ones in electrodynamics (see e.g. [141], §21.23). More precisely, the jump of the extrinsic curvature is interpreted as the surface stress-energy of a surface layer located at $S$. In the case of $S$ being timelike this arrangement represents a thin shell of matter, while if $S$ is null it may be interpreted as a thin shell of radiation (see e.g. [103]). In [97] W. Israel has given the final formulation of this widely applied approach providing the practical advantage that no reference to any special coordinate system is required; the four-dimensional coordinates may be chosen freely hence adapted to possibly different symmetries in the upper and lower part of spacetime.
On the other hand Lichnerowicz [124], VI, 88 has given an alternative description using tensor distributions assuming the existence of an admissible continuous coordinate system across $S$. This formalism was used by Lichnerowicz [125,126] and Choquet-Bruhat [32] to study gravitational shock waves. They derived algebraic conditions on the metric across the shock (the "gravitational Rankine-Hugoniot conditions") as well as equations governing the propagation of the discontinuities. The respective formalisms of Israel and Lichnerowicz were shown to be equivalent in [136].

**2.17** The description of gravitational sources supported on two-dimensional submanifolds of spacetime, however, is more delicate. Israel [98] has given conditions under which a sensible treatment of the field of a "thin massive wire" is possible. He isolated a class of "simple line sources" which possess a linear energy-momentum tensor hence allow for a well-defined limit as the wire's radius shrinks to zero.
On the other hand Taub [177] has claimed to have generalized Lichnerowicz's formalism to include gravitational sources supported on submanifolds of arbitrary codimension in spacetime. However, he had to fix the ill-defined products by "multiplication rules" like, e.g., $H\delta = (1/2)\delta$ (cf. 2.4).
Parker [149] has presented a distributional geometry again using tensor distributions which allowed him to distributionally formulate Einstein's equations as long as the respective curvature quantities remained linear. Recently Garfinkle [69] has generalized the formalism of Geroch and Traschen (see below) to include a slightly more general class of metrics giving up the requirement of "limit consistency" as discussed in 2.15 above.
Finally the distributional curvature of a cosmic string has been rigorously computed by Clarke et al. [36] using Colombeau's nonlinear framework of generalized functions. This approach was later generalized to time dependent strings [188,189] and its diffeomorphism invariance (cf. also the discussions in chaps. 4 and 5) was explicitly shown in [185].

**2.18** We now begin our systematic approach by analyzing in detail the structure of the nonlinearities of the field equations to see how far we can get avoiding products ill-defined within classical distributions. More precisely, we are searching for a class of metrics allowing for a distributional formulation of the Einstein tensor in order to assign to the spacetime—via the field equations—a distributional energy-momentum tensor representing the "concentrated" source. Note that we have two contradicting demands on this—yet to find—class of metrics which essentially following Geroch and Traschen we shall call *gt-regular*[4]. On one hand these metrics should be "nice enough" to permit the distributional calculation of the curvature entities, while

[3] This feature—of course—is lost if one considers the coupled nonlinear Maxwell-Lorentz system. For the treatment of related problems in Colombeau's framework see [93,94]

[4] Geroch and Traschen [71] actually used the term "regular" which we reject due to its general flavor.

on the other hand they should be "bad enough" to have the Einstein tensor and hence the energy-momentum tensor concentrated to a submanifold of a high codimension in spacetime.
We set out by writing down the definition of the Riemann curvature tensor in coordinates, i.e.,

$$R_{abc}{}^{d} \;=\; 2\Gamma^{d}_{e[b}\Gamma^{e}_{a]c} + 2\partial_{[b}\Gamma^{d}_{a]c}\;, \tag{48}$$

where the square brackets are used to indicate antisymmetrization in the respective indices (i.e., $2A_{[ab]} := A_{ab} - A_{ba}$) and the Christoffel symbols (of the second kind) $\Gamma^a_{bc}$ are given by

$$\Gamma^{a}_{bc} \;=\; g^{ae}(\partial_{(b}g_{c)e} - \frac{1}{2}\partial_c g_{bc})\;. \tag{49}$$

Here the brackets stand for the symmetric part with respect to the affected indices (i.e., $2A_{(ab)} = A_{ab} + A_{ba}$) and $g_{ab}$ denotes the components of the spacetime metric.
Now we try to "save" these equations by putting just as much restrictions on the metric tensor as needed to allow for a distributional interpretation of the respective right hand sides. For the first term in (48) we obviously need $\Gamma^a_{bc}$ to be locally square integrable. Since $L^2_{\text{loc}} \subseteq L^1_{\text{loc}}$ this requirement actually also suffices to interpret the second term in (48) as weak derivative of the regular distribution $\Gamma^a_{bc}$. Furthermore from equation (49) we see that it is sufficient to demand $g^{ab}$, the inverse metric, to be bounded locally (Lebesgue) almost everywhere in order to produce locally square integrable $\Gamma^a_{bc}$ from locally square integrable first weak derivatives og $g_{ab}$. This finally motivates the following

**2.19 Definition.** *A symmetric tensor field $g_{ab}$ on a four-dimensional (smooth paracompact) manifold $M$*[5] *is called a* gt-regular metric *if*

*(i) the inverse metric $g^{ab}$ exists almost everywhere and $g_{ab}$ as well as $g^{ab}$ are locally bounded almost everywhere, and*

*(ii) the first weak derivatives of $g_{ab}$ exist and are locally square integrable.*

**2.20** Some remarks on this definition are in order. First it treats the metric and its inverse on equal footing, i.e., if $g_{ab}$ is gt-regular then $\partial_c g^{ab} \in L^2_{\text{loc}}$ ($g^{ab} = \text{cof}(g_{ab})/\det g_{ab}$, where $\text{cof}(g_{ab})$ denotes the cofactor of $g_{ab}$). Hence expressed in a more condensed language the above definition may be formulated as $g_{ab}$ and $g^{ab} \in L^\infty_{\text{loc}} \cap H^1_{\text{loc}}$.
For a detailed discussion of the gt-conditions in the context of axial and cylindrical symmetry and especially its interrelations to the notion of $\mathcal{C}^k$-regularity we refer to [188], chaps. 2.3-2.5.
Further we remark that it is hard to see how—in a general situation—one could find a wider class of metrics satisfying our needs, e.g., how the individual terms failing to be a distribution could conspire to yield a sum well-defined within $\mathcal{D}'$. Also we shall see in 2.22 (ii) that we have already exhausted the framework of general relativity in the sense that the Bianchi identities for general gt-regular metrics may not be formulated within $\mathcal{D}'$. Moreover, gt-regular metrics may not be used to pull up or down indices of general tensor distributions since the tensor product again would involve a multiplication of distributions. Note however, that gt-regular metrics still have a "function character," hence itself may not be concentrated on some lower dimensional submanifold. However, in algebraically special situations—in a preferred coordinate system—some of the curvature quantities may be defined even for non gt-regular metrics as we shall see in chap. 5 below.

**2.21** Our original motivation was to isolate a class of metrics that would allow for a distributional formulation of Einstein's equations. To see that this is actually possible for gt-regular metrics we have to convince ourself that we are able to build the Einstein tensor within distributions, i.e., that the tensor product of the metric with the Riemann tensor is well-defined. Expressing the term involving second derivatives of the metric as total derivative we may write

$$g^{ef}R_{abc}{}^{d} \;=\; 2g^{ef}\Gamma^{d}_{m[b}\Gamma^{m}_{a]c} + 2\partial_{[b}(g^{ef}\Gamma^{d}_{a]c}) - 2(\partial_{[b}g^{ef})\Gamma^{d}_{a]c}\;.$$

---

[5] We have reserved the letter $M$ for spacetime manifolds (in this case a somewhat generalized one) to distinguish it from a merely smooth (paracompact) manifold always denoted by $X$.

Now the first term involves a product $L^\infty_{\mathrm{loc}} \times L^1_{\mathrm{loc}}$ hence stays locally integrable. The second term involves a weak derivative of a $L^1_{\mathrm{loc}}$-tensor field hence may be interpreted as a distribution as well as the third term since a product of two locally square integrable fields is locally integrable.
Hence we have indeed succeeded in defining the "largest reasonable" class of spacetime metrics allowing for a distributional interpretation of the Einstein equations. Before investigating an appropriate notion of convergence for gt-regular metrics we discuss two of its properties.

**2.22 Remark.**

(i) We ask for assumptions on a tensor field in order for its *metric covariant derivative* ${}^g\nabla$ *with respect to a gt-regular metric* $g_{ab}$ to be well-defined. Denoting by $\xi^a$ a vector field we have

$$ {}^g\nabla_a \xi^b \;=\; \partial_a \xi^b + \Gamma^b_{ac}\xi^c \;, $$

showing that $\xi^a$ being locally square integrable is a sufficient condition. In particular, the metric covariant derivative of a gt-regular metric is well-defined and vanishes by virtue of (49). Also for a smooth one form $\mu_a$ both sides of the equation

$$ {}^g\nabla_{[a}\, {}^g\nabla_{b]}\mu_c \;=\; R_{abc}{}^d \mu_d \tag{50} $$

are well-defined. Moreover by (48) and (49) equation (50) indeed holds showing that we also could have taken it to define the distributional Riemann tensor.

(ii) The *Bianchi identities* generally may not be formulated for gt-regular metrics since their left hand side ${}^g\nabla_{[a}R_{bc]de}$ involves a product of the distributional coefficients of the Riemann tensor with the nonsmooth Christoffel symbols. We can take this fact as an indication that we have already overstretched the present framework.

**2.23** Our next aim is to set up an appropriate notion of *convergence for gt-regular metrics.* As already indicated in 2.15 we expect the Einstein tensors of a sequence of metrics approximating a gt-regular one to approximate the respective Einstein tensor of the gt-regular metric. We start by recalling the natural notion of weak convergence for a sequence of locally square integrable tensor fields $((\mu^{i_1\dots i_r}_{j_1\dots j_s})_n)_n$, i.e., convergence locally in square integral defined by

$$ (\mu^{i_1\dots i_r}_{j_1\dots j_s})_n \to 0 \quad (n\to\infty) \text{ iff } \int (\mu^{i_1\dots i_r}_{j_1\dots j_s})_n (\mu^{k_1\dots k_r}_{l_1\dots l_s})_n \, t^{j_1\dots j_s l_1\dots l_s}_{i_1\dots i_r k_1\dots k_r} \to 0 \quad \in \mathbb{C} $$

for all smooth compactly supported $(2s,2r)$-tensor densities $t^{i_1\dots i_{2s}}_{j_1\dots j_{2r}} \in \Gamma_0(T^{2r}_{2s}(M)\otimes \mathrm{Vol}\,(\mathrm{M}))$. Using this we have the following

**2.24 Theorem.** (Convergence of gt-regular metrics) *Let $g_{ab}$ and $((g_{ab})_n)_n$ be a gt-regular metric and a sequence of gt-regular metrics respectively and let*

*(i) $((g_{ab})_n)_n$ and $((g^{ab})_n)_n$ locally uniformly bounded, and*

*(ii) $(g_{ab})_n \to g_{ab}$, $(g^{ab})_n \to g^{ab}$, and $(\partial_a g_{bc})_n \to \partial_a g_{bc}$ locally in square integral.*

*Then $(R_{abc}{}^d)_n \to R_{abc}{}^d$ and hence $(G_{ab})_n \to G_{ab}$ in $\mathcal{D}'(M, T^1_3(M))$ respectively $\mathcal{D}'(M, T^0_2(M))$.*
*Also the space of gt-regular metrics is complete with respect to the notion of convergence defined by hypotheses (i) and (ii).*
*Moreover, let $g_{ab}$ be a continuous gt-regular metric then there exists a sequence of smooth metrics $((g_{ab})_n)_n$ converging to $g_{ab}$ in the sense of (i) and (ii).*

**Proof.** To prove the first statement fix a test density $t^{abc}{}_d \in \Gamma_c(M, T^3_1(M)\otimes \mathrm{Vol}\,(\mathrm{M}))$. Then

$$ \langle (R_{abc}{}^d)_n, t^{abc}{}_d\rangle \;=\; -2\int (\Gamma^d{}_{e[a})_n\,(\Gamma^e{}_{b]c})_n\, t^{abc}{}_d \;+\; 2\int (\Gamma^a{}_{c[b})_n\,\partial_{a]} t^{abc}{}_d \;. $$

Applying Schwarz' inequality to the second term on the right hand side we see that it suffices to show that $(\Gamma^a{}_{bc})_n \to \Gamma^a{}_{bc}$ locally in square integral. To maintain clarity of formulae we use the somewhat sloppy notation $f_n$ and $h_n$ for the components of $(g^{ab})_n$ and $(\partial_{(b}g_{c)d})_n - (1/2)(\partial_d g_{bc})_n$, respectively. Now for all test densities $t \in \Gamma_c(\mathrm{Vol}\,(\mathrm{M}))$

$$\begin{aligned}\int |f_n h_n - fh|^2\, t &= \int |f_n h_n - f_n h + f_n h - fh|^2\, t \\ &\leq \int |f_n|^2 |h_n - h|^2\, t + \int |h|^2 |f_n - f|^2\, t\end{aligned}$$

and the claim for the Riemann tensor follows from (i) and (ii). The same holds true for any product of the Riemann tensor and the metric by similar arguments.
To prove the second statement observe that by the completeness of $L^2_{\mathrm{loc}}$ all Cauchy sequences $((g_{ab})_n)_n$ and $((g^{ab})_n)_n$ converge to their respective limits $g_{ab}$ and $g^{ab}$ which by (i) are even locally bounded. Since every $L^2_{\mathrm{loc}}$ converging sequence possesses a subsequence which converges pointwise almost everywhere (cf. [2], 2.11) $g^{ab}$ is almost everywhere the inverse of $g_{ab}$. The respective statement on the weak derivative follows easily from the completeness of $H^1_{\mathrm{loc}}$ (cf. [2], 3.2).
Finally, given a continuous gt-regular metric $g_{ab}$ we construct its prospective approximation $((g_{ab})_n)_n$ by convolution with an approximative identity. Clearly $(g_{ab})_n$ are smooth and converge together with their first derivatives locally in square integral to $g_{ab}$ and $\partial_c g_{ab}$, respectively (cf. [2], 3.15). Since $g_{ab}$ is continuous $(g_{ab})_n$ converges locally even uniformly ( [2], 2.18), hence is locally uniformly bounded. For large $n$ the inverse metric $(g^{ab})_n$ exists and is locally uniformly bounded by the formula already used in 2.20 and converges locally in square integral to $g^{ab}$. □

**2.25** Summing up we have constructed a framework for treating spacetime metrics of low differentiability based on classical distribution theory which is both mathematically sound and physically reasonable (cf. 2.15). Before actually checking which class of gravitational sources may now be described we start out with the following heuristically considerations. Suppose $S$ is a submanifold in spacetime of dimension $d = (0, 1, 2, 3)$ and the metric $g_{ab}$ is smooth on $M \setminus S$ but some of its components diverging on $S$. Which order of divergence is allowed if $g_{ab}$ is to be gt-regular? Let $r$ be a typical distance from $S$ measured by, e.g., a Riemannian background metric $h_{ab}$ and suppose some components of $g_{ab}$ diverge by the rate of $r^{-s}$ for some positive number $s$. Hence the weak derivatives of $g_{ab}$ behave like $r^{-1-s}$ while the volume element is proportional to $r^{3-d}$. In order to render these derivatives locally square integrable we have to fulfill the following inequality

$$\begin{aligned}2(-s-1) + 3 - d > -1 &\text{ implying} \\ s < 1 - \frac{d}{2}\ .\end{aligned}$$

Hence we see that the components of gt-regular metrics grow at most by a rate of $r^{-1+d/2}$ on approaching a $d$-dimensional submanifold in spacetime. More precisely, they are allowed to be worsely behaved near a submanifold of larger codimension. However, this behavior is still not sufficient to catch, e.g., the notion of a string, as the next theorem shows.

**2.26 Theorem.** (Concentrated sources from gt-regular metrics) *Let $S$ be a submanifold of dimension $d = (0, 1, 2, 3)$ of a four-dimensional manifold $M$ and let $T^{i_1\cdots i_r}_{j_1\cdots j_s} \neq 0$ a tensor distribution satisfying*

*(i)* $\mathrm{supp}(\mathrm{T}^{\mathrm{i}_1\cdots\mathrm{i}_r}_{\mathrm{j}_1\cdots\mathrm{j}_s}) \subseteq \mathrm{S}$, *and*

*(ii) $T^{i_1\cdots i_r}_{j_1\cdots j_s}$ is the sum of a locally integrable tensor field and the weak derivative of a locally square integrable tensor field*

*(hence is of the form of the Riemann tensor of a gt-regular metric). Then* $d = 3$.

**Proof.** Without loss of generality let $T$ be a scalar. Hence by (ii) we may (in a sloppy fashion) write $T = f_1 + \partial f_2$, where $f_1$ resp. $f_2$ are locally resp. locally square integrable functions on $M$. Now fix a test density $t$ and an auxiliary Riemann metric $h_{ab}$ and define for all positive $\varepsilon$ a $h_{ab}$-$\varepsilon$ neighborhood $U_\varepsilon$ of $S$ and a function $h_\varepsilon$ satisfying (a) $h_\varepsilon$ smooth and non negative on $M$, (b) $h_\varepsilon|_{M\setminus U_\varepsilon} \equiv 1$, (c) $h_\varepsilon \equiv 0$ on a neighborhood of $S$ and (d) $|\partial h_\varepsilon|_{h_{ab}} \leq 2/\varepsilon$ on the support of $R$. Then we have

$$| \int_M (f_1 t - f_2 \partial t) h_\varepsilon | = |\langle T, t\, h_\varepsilon\rangle - \int_M t f_2 \partial h_\varepsilon| \leq \int_{U_\varepsilon} |t|\, |f_2 \partial h_\varepsilon| \quad \text{(b) and (c)}$$

$$\leq [\int_{U_\varepsilon} |t|\, f_2^2\,]^{1/2}\, [\int_{U_\varepsilon} |t|\, (\partial h_\varepsilon)^2\,]^{1/2} \leq [\int_{U_\varepsilon} |t|\, f_2^2\,]^{1/2}\, [\frac{4}{\varepsilon^2} \int_{U_\varepsilon} |t|\,]^{1/2} \quad \text{(d)} .$$

For $\varepsilon \to 0$ the left hand side of the first line above approaches $|\langle T, t\rangle|$, while the first term in the last line converges to zero since $f_2$ is locally square integrable. Hence the second term has to diverge and since it is proportional to $\frac{1}{\varepsilon}(\varepsilon^{4-d})^{1/2} = \varepsilon^{(2-d)/2}$ we conclude $d = 3$. □

**2.27 Example.** We now consider an explicit counterexample showing that it is indeed impossible—in a manner consistent within our present framework—to assign to a spacetime an energy-momentum tensor concentrated on a two-dimensional submanifold . We start with the following static cylindrically symmetric metric [72,84,71]

$$ds^2 \;=\; -dt^2 + dz^2 + dr^2 + \beta(r)^2 d\phi^2 \;, \tag{51}$$

where the coordinates $t, z$ are unrestricted, $r \in [0, \infty)$, and $\phi \in [0, 2\pi]$ (with the endpoints identified) and the function $\beta(r)$ is given by

$$\beta(r) \;=\; \begin{cases} \frac{l}{\gamma}\, \sin(\frac{\gamma r}{l}) & (r \leq l) \\ (r - l + \frac{l}{\gamma}\, \tan\gamma)\, \cos\gamma & (r > l)\, . \end{cases}$$

Here $l > 0$ and $\gamma \in (0, \pi/2]$ are constants. This metric is smooth everywhere but at $r = l$ where it is merely $\mathcal{C}^1$. Actually the second derivatives on either side exist but do not match continuously. In the interior region of the spacetime ($r < l$) the Einstein tensor has two nonvanishing components, i.e., $G_{zz} = -G_{tt} = \gamma^2/l^2$ hence an energy density (putting $8\pi G/c^2 = 1$) $\rho = -T^t{}_t = \gamma^2/l^2$. In the exterior region ($r > l$) the Riemann tensor vanishes via the transformation $R = r - l + (l/\gamma)\tan\gamma$ and we may cast the metric in the standard conical form

$$ds^2 \;=\; -dt^2 + dz^2 + dR^2 + R^2 \cos^2\gamma\, d\phi^2 \;, \tag{52}$$

with angular deficit $\Delta = 2\pi(1 - \cos\gamma)$. Hence the whole spacetime consists of a massive static cylinder ($r < l$) and an exterior cone. The mass density of the cylinder per unit length $\mu$ defined as the integral of the energy density over the two-surface $z, t$ =const. yields exactly the angular deficit of the exterior, i.e.,

$$\mu \;=\; \int_0^{2\pi} \int_0^l \frac{\gamma}{l}\, \sin\frac{\gamma r}{l}\, dr\, d\phi \;=\; 2\pi(1 - \cos\gamma) \;=\; \Delta \;. \tag{53}$$

To describe a string we consider the limit $l \to 0$. The exterior region now covers all of spacetime, i.e., the metric is given by (52); mass density per unit length and angular deficit do not depend on $l$, hence converge trivially. Note, however, that in the limit $\rho$ diverges and $\mu$ can not be computed. Nevertheless, it is tempting (and has indeed be done in the literature) to assign to the string the mass density

$$\mu_s \;=\; 2\pi(1 - \cos\gamma) \;. \tag{54}$$

Moreover, interpreting the situation as all the energy density to be concentrated to the axis $R = 0$ one even might assign to the string the mass density

$$\rho_s \;=-\lim T^t{}_t \;=\; 2\pi(1 - \cos\gamma)\, \delta^{(2)}(r) \tag{55}$$

Although this description of the situation might be physically desirable, it is by no way mathematically justified as we shall see in a moment. First, the metric (52) is not gt-regular; this statement needs some explanation. In principle the answer to the question whether or not a metric is gt-regular depends on the underlying differentiable structure; here it is most reasonable to take the differentiable structure as induced by Cartesian coordinates associated to the polar coordinates $R, \phi$ and the spacetime—of course—extended to the axis at $R = 0$. Then it is indeed easy to check that the weak first derivatives of the metric are not locally

square integrable. However, from the behavior of the Einstein tensor of (51) (in the limit) and thm. 2.26 it follows that there exists no differentiable structure rendering this metric gt-regular.
Second—and more explicitly—consider the metric

$$\bar{g}_{ab} = e^{2\lambda f(r/l)}\, g_{ab} \tag{56}$$

where $g_{ab}$ stands for the metric given by (51), $\lambda$ is a positive constant and $f \in \mathcal{D}(\mathbb{R})$, supp(f) $\subset\subset [1/2, 1]$ and $f \geq 0$. In the exterior the metrics (51) and (56) coincide, however in the interior we now have

$$\bar{G}_{tt} = \frac{\lambda^2 f^2(r/l)\sin(\gamma r/l) + 2\lambda f''(r/l)\sin(\gamma r/l) + 2\lambda f'(r/l)\gamma\cos(\gamma r/l) - \gamma^2\sin(\gamma r/l)}{l^2\,\sin(\gamma r/l)}\ ,$$

leading to a mass density per unit length

$$\bar{\mu} = 2\pi(1-\cos\gamma) - 2\pi\int\limits_0^1 \frac{e^{2\lambda f(x)}}{\gamma}\,[\lambda^2 f'^2(x)\sin(\gamma x)]\,dx\ . \tag{57}$$

Since all the terms in the integral are positive we have $\mu > \bar{\mu}$. Taking the limit $l \to 0$ for $\bar{g}_{ab}$ we again end up with the conic metric with angular deficit $\bar{\Delta} = \Delta = 2\pi(1-\cos\gamma)$ but following the procedure proposed above we would have to assign to it the energy density

$$\bar{\rho}_s = \bar{\mu}\,\delta^{(2)}(r) \neq 2\pi(1-\cos\gamma)\,\delta^{(2)}(r) = \rho_s\ .$$

Hence we see that assigning a distributional energy density to a string by virtue of a (naive) regularization procedure leads to real inconsistencies. The above example gains even more importance by the fact that the second regularization leading to the physically undesirable answer is not chosen particularly exotic. The metric (56) has the same symmetries as (51) and also fulfills the strong energy condition.
For a more general discussion of various regularizations of the cone metric also from the viewpoint of Colombeau's nonlinear theory of generalized functions we refer the reader to the thesis of J. Wilson [188], chaps. 5.5 and 5.6.

**2.28 Conclusion.** Summing up, the present chapter should make it pretty clear that the framework of classical *linear* distribution theory is too restrictive for general relativity. Staying strictly within the mathematically and physically consistent setting, one has to restrict oneself to a class of metrics that excludes physically interesting cases such as strings and point particles.
If one wants to describe more general gravitational sources—due to the non-linearity of the field equations—one is forced to go beyond the limits of classical distribution theory, thus encountering true conceptional problems. A consistent framework allowing for nonlinear operations on singular (e.g. distributional) objects is provided by Colombeau's algebras of generalized functions, which we shall introduce in the next chapter.

# 3. COLOMBEAU ALGEBRAS

This chapter provides a detailed introduction into the nonlinear theory of generalized functions due to J. F. Colombeau [38,39,41], i.e., Colombeau algebras. As already indicated in 2.13, these are associative and commutative algebras containing the space of distributions while at the same time providing maximal consistency properties with respect to classical operations (i.e., properties (i)-(iii) in 2.10 and (iv") in 2.13). Additionally, the embedding of distributions is given explicitly via convolution. These properties render Colombeau algebras a perfect framework for nonlinear analysis of singular, i.e., distributional data, which we will use in chap. 5 to distributionally describe the geometry of impulsive gravitational waves.
In our presentation we mainly focus on the so called "special" (or "simplified") version (in the sense of [144], p. 109) of the theory for the following two reasons. First—for the convenience of the reader—it is technically much easier to access while at the same time exposing all the main features of the "full" theory (as presented e.g. in [39]). Second, for the applications to be presented in chap. 5 the special algebra provides a perfectly suitable and flexible tool.

In sec. 3.A we motivate the main definitions and derive some of the basic properties of the Colombeau algebra $\mathcal{G}$ while sec. 3.B deals with the embedding of distributions into the algebra. The theory on (open sets of) $\mathbb{R}^n$ is further developed in sec. 3.C (point values of generalized functions), sec. 3.D (integration) and sec. 3.E (association and coupled calculus). Finally Colombeau algebras on manifolds as well as generalized sections in vector bundles will be treated in detail in chapter 4.
Our presentation widely parallels the ones given in [171], chap. 2 and [117], chap. 1, which by themselves are based upon [75]. For the "full" theory we refer the reader to [22], [144] and most of all to Colombeau's "*Elementary Introduction*" [39].

## 3.A. Definition and Basic Properties

**3.1** One key notion in constructing algebras containing distributions is *regularization by sequences (nets) of smooth functions*. Hence, to begin with, we recall the basic idea of the sequential approach to distribution theory (see [8]) which has been developed by Mikusinski [140] and Temple [179]. Setting $I = (0, 1]$ we define the set of weakly converging respectively weakly zero-nets by

$$\mathcal{V} := \{(u_\varepsilon)_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^n)^I : \exists u \in \mathcal{D}'(\mathbb{R}^n) \text{ with } u_\varepsilon \to u \text{ in } \mathcal{D}'\}$$
$$\mathcal{V}_0 := \{(u_\varepsilon)_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^n)^I : u_\varepsilon \to 0 \text{ in } \mathcal{D}'\}\ .$$

Now the space of distributions is isomorphic to the quotient space of the weakly converging modulo the weakly vanishing nets, i.e., we have the following

**3.2 Proposition.** *The map*

$$\begin{aligned} \Psi : \mathcal{V}/\mathcal{V}_0 &\to \mathcal{D}'(\mathbb{R}^n) \\ (u_\varepsilon)_{\varepsilon\in I} + \mathcal{V}_0 &\mapsto \lim_{\varepsilon\to 0} u_\varepsilon \end{aligned}$$

*is a linear isomorphism.*

**Proof.** Injectivity is obvious. To prove that $\Psi$ is onto choose a test function $\rho$ with unit integral. Then $u_\varepsilon(x) := \varepsilon^{-n} u(x) * \rho(x/\varepsilon)$ is smooth and converges weakly to $u \in \mathcal{D}'(\mathbb{R}^n)$. □

**3.3** Since we are interested in constructing an algebra containing $\mathcal{D}'$ the following observation is fundamental: The space $(\mathcal{C}^\infty(\mathbb{R}^n))^I$ endowed with componentwise operations, i.e., vector space operations, partial derivative and multiplication is a differential algebra.
Therefore proposition 3.2 suggests to proceed as follows. Choose some ideal $\mathcal{N} \subseteq \mathcal{V}_0$ of $(\mathcal{C}^\infty(\mathbb{R}^n))^I$ and embed distributions via convolution with a mollifier, more precisely

$$\begin{aligned} \mathcal{D}'(\mathbb{R}^n) &\hookrightarrow \mathcal{C}^\infty(\mathbb{R}^n)^I/\mathcal{N} \\ u &\mapsto (u * \rho_\varepsilon)_{\varepsilon\in I} + \mathcal{N}, \end{aligned} \tag{58}$$

where for the moment we choose $\rho \in \mathcal{D}(\mathbb{R}^n)$ such that $\int \rho = 1$ and set $\rho_\varepsilon(x) := (1/\varepsilon)\rho(x/\varepsilon)$.

We have hence constructed a linear embedding of $\mathcal{D}'$ into an associative, commutative differential algebra with $f \equiv 1$ mapped onto unity (condition (i) in 2.10). Furthermore the componentwise partial derivative clearly is linear, satisfies the Leibniz rule (condition (ii) in 2.10) and coincides with the usual weak derivative (condition (iii) in 2.10).
So it only remains to further arrange our setting in order to render $\mathcal{C}^\infty$ a faithful subalgebra (condition (iv") in 2.13). Hence we have to adjust the ideal $\mathcal{N}$ in such a way that for any pair of smooth functions $f, g$

$$((f * \rho_\varepsilon) \cdot (g * \rho_\varepsilon))_\varepsilon + \mathcal{N} = ((fg) * \rho_\varepsilon)_\varepsilon + \mathcal{N}$$

holds. Clearly setting $\mathcal{N} = \{0\}$ won't work. However, if for a moment we only consider the task of embedding $\mathcal{C}^\infty$ as a faithful subalgebra, the most natural way to do so is via the *"constant" embedding*

$$\begin{aligned} \sigma: \quad \mathcal{C}^\infty(\mathbb{R}^n) &\hookrightarrow \mathcal{C}^\infty(\mathbb{R}^n)^I/\mathcal{N} \\ f &\mapsto (f)_{\varepsilon\in I} + \mathcal{N}\ . \end{aligned}$$

On the other hand we have already assigned to $f$ the element $(f * \rho_\varepsilon)_{\varepsilon\in I} + \mathcal{N}$. Hence let us try to define $\mathcal{N}$ in such a way that $(f * \rho_\varepsilon)_\varepsilon$ and $(f)_\varepsilon$ may be identified modulo the ideal. To gain some analytical insight into this condition we expand the smooth function $f$ into a Taylor series (setting $n = 1$ for simplicity) to get for the difference

$$\begin{aligned} (f * \rho_\varepsilon - f)(x) &= \int (f(x-y) - f(x))\rho_\varepsilon(y)\,dy \\ &= \int \sum_{k=1}^{m} \frac{(-\varepsilon y)^k}{k!} f^{(k)}(x)\rho(y)\,dy + \int \frac{(-\varepsilon y)^{m+1}}{(m+1)!} f^{(m+1)}(x - \theta\varepsilon y)\rho(y)\,dy\ . \end{aligned} \tag{59}$$

Having in mind to estimate this expression we see that if we additionally require $\int \rho(x)x^k dx$ to vanish for all $k$ (this actually implies $\rho \notin \mathcal{D}$ as we shall see in 3.10 below) the above expression is bounded by some $C\varepsilon^{m+1}$ ($C$ a constant) uniformly on compact sets. Noting that $m$ was an arbitrary natural number and that the above procedure may be carried out analogously for any derivative of the left hand side a description of $\mathcal{N}$ may be read off: A sequence $(u_\varepsilon)_{\varepsilon\in I}$ should be *negligible* if $u_\varepsilon$ together with all its derivatives vanishes faster than any power of $\varepsilon$ uniformly on compact sets if $\varepsilon$ tends to zero.
Since $\mathcal{C}^\infty(\mathbb{R}^n)^I$ contains sequences of arbitrary growth in $1/\varepsilon$ (e.g. proportional to $\exp(1/\varepsilon)$) $\mathcal{N}$ is clearly not an ideal in it. However, if we restrict our space of representatives to consist of all sequences of *moderate growth* we finally are ready to give the definition of the Colombeau algebra.

**3.4 Definition.** *Let $\Omega$ be an open subset of $\mathbb{R}^n$. Set*

$$\begin{aligned} \mathcal{E}(\Omega) &:= (C^\infty(\Omega))^I \\ \mathcal{E}_M(\Omega) &:= \{(u_\varepsilon)_\varepsilon \in \mathcal{E}(\Omega) : \forall K \subset\subset \Omega, \forall \alpha \in \mathbb{N}_0^n\ \exists p \in \mathbb{N} \textit{ with } \sup_{x\in K} |\partial^\alpha u_\varepsilon(x)| = O(\varepsilon^{-p}) \textit{ as } \varepsilon \to 0\} \\ \mathcal{N}(\Omega) &:= \{(u_\varepsilon)_\varepsilon \in \mathcal{E}(\Omega) : \forall K \subset\subset \Omega, \forall \alpha \in \mathbb{N}_0^n\ \forall q \in \mathbb{N} : \sup_{x\in K} |\partial^\alpha u_\varepsilon(x)| = O(\varepsilon^{q}) \textit{ as } \varepsilon \to 0\}. \end{aligned}$$

*The Colombeau algebra on $\Omega$ is defined as the quotient space*

$$\mathcal{G}(\Omega) := \mathcal{E}_M(\Omega) \,/\, \mathcal{N}(\Omega)\,.$$

**3.5** Some remarks on this definition are in order. Obviously $\mathcal{E}_M(\Omega)$, the space of moderate sequences, endowed with componentwise operations is a differential algebra with $\mathcal{N}(\Omega)$ a differential (i.e., stable under differentiation) ideal in it. Hence $\mathcal{G}(\Omega)$ is an associative and commutative differential algebra with $\mathcal{C}^\infty(\Omega)$ a faithful subalgebra via the "constant" embedding $\sigma$.
We adopt the following notation. Elements of $\mathcal{G}$—henceforth called *generalized* [6] *functions*—will be denoted by capital letters, i.e., $U \in \mathcal{G}$, while for representatives we are going to use small letters, i.e., $(u_\varepsilon)_{\varepsilon\in I}$.

[6]From now on we shall use the term "generalized" in the sense of Colombeau only. Note that this is somehow in contrast to the terminology of chap. 1 where—for historical reasons—we have called distributional objects "generalized." However, form now on we refer to (and consider) distributional notions as "classical."

Furthermore we shall write $U = \mathrm{cl}[(u_\varepsilon)_\varepsilon] = (u_\varepsilon)_\varepsilon + \mathcal{N}(\Omega)$ to express that $U$ is the class of $(u_\varepsilon)_\varepsilon$. Depending on whether the functions $u_\varepsilon$ are real or complex valued we shall write $\mathcal{G}(\Omega,\mathbb{R})$ or $\mathcal{G}(\Omega,\mathbb{C})$. However, most of the time we are not interested in this distinction and treat both cases simultaneously by simply writing $\mathcal{G}(\Omega)$.

The following observation, first made by M. Grosser [76], substantially simplifies the characterization of negligible functions and will be repeatedly used in the sequel. In fact, for representatives $(u_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}_M$ it suffices to prove the $\mathcal{N}$-estimate for the zeroth derivative to guarantee that they are in the ideal.

**3.6 Proposition.** *Let* $(u_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}_M(\Omega)$ *and* $\forall K \subset\subset \Omega\ \forall q \in \mathbb{N}$

$$\sup_{x\in K} |u_\varepsilon(x)| \;=\; O(\varepsilon^q) \;\text{ as } \varepsilon \to 0\,.$$

*Then* $(u_\varepsilon)_{\varepsilon\in I} \in \mathcal{N}(\Omega)$.

**Proof.** It suffices to show that $\forall K \subset\subset \Omega\ \forall q \in \mathbb{N}\ \forall \alpha \in \mathbb{N}_0^n$ with $|\alpha| = 1$, $\sup_{x\in K} |\partial^\alpha u_\varepsilon(x)| = O(\varepsilon^q)$. The rest follows easily by induction. Let $K \subset\subset \Omega$ and choose $\tilde{K} \subset\subset \Omega$ such that $K \subseteq \tilde{K}^\circ$. Then we have $\sup_{x\in\tilde{K}} |u_\varepsilon(x)| = O(\varepsilon^{2q+p})$, where we choose $p$ in such a way that $\sup_{x\in\tilde{K}} |\partial^\alpha u_\varepsilon(x)| = O(\varepsilon^{-p})\ \forall |\alpha| = 2$. Let now $x \in K$, $|y| = 1$ and $\varepsilon$ so small that $\varepsilon^{p+q} \le \mathrm{dist}(K,\tilde{K})$. Then by Taylor expansion we have

$$u_\varepsilon(x + y\varepsilon^{p+q}) \;=\; u_\varepsilon(x) + \varepsilon^{p+q}(y\cdot\nabla u_\varepsilon)(x) + \frac{1}{2}\varepsilon^{2(p+q)}((y\cdot\nabla)^2 u_\varepsilon)(\theta)\,,$$

where $\theta \in \tilde{K}$. Hence for $x$ in $K$

$$(y\cdot\nabla u_\varepsilon)(x) = \varepsilon^{-(p+q)}\underbrace{(u_\varepsilon(x+y\varepsilon^{p+q}) - u_\varepsilon(x))}_{O(\varepsilon^{2q+p})} + \varepsilon^{p+q}\underbrace{((y\cdot\nabla)^2 u_\varepsilon)(\theta)}_{O(\varepsilon^{-p})} \;=\; O(\varepsilon^q)\,.$$

□

In $\mathcal{G}$ not only products and polynomials of generalized functions (like in *any* algebra) are well-defined but in fact a much wider class of nonlinear operations. Denoting by $\mathcal{O}_M$ the space of smooth functions which together with all their derivatives grow at most like some power of $||x||$ as $||x||$ tends to infinity and by $\mathbb{K}$ either $\mathbb{C}$ or $\mathbb{R}$ we have the following

**3.7 Proposition.** *Let* $U_1,\ldots,U_n \in \mathcal{G}(\Omega)$ *and* $F \in \mathcal{O}_M(\mathbb{K}^n)$ *then*

$$F(U_1,\ldots,U_n)(.) \;:=\; (F(u_{1\varepsilon}(.),\ldots,u_{n\varepsilon}(.)))_\varepsilon + \mathcal{N}(\Omega)$$

*is a well-defined element of* $\mathcal{G}(\Omega)$.

**Proof.** The $\mathcal{E}_M$-bounds follow from the fact that $F$ is polynomially bounded in all derivatives. To prove uniqueness choose some other representative $(\tilde{u}_{k\varepsilon})_\varepsilon$ of $U_k$ $(1 \le k \le n)$. Then by the mean value theorem, $|F(u_{1\varepsilon}(x),\ldots,u_{n\varepsilon}(x)) - F(\tilde{u}_{1\varepsilon}(x),\ldots,\tilde{u}_{n\varepsilon}(x))|$ can be estimated in terms of a sum of expressions involving derivatives of $F$ (composed with $\mathcal{E}_M$-functions, thus satisfying $\mathcal{E}_M$-bounds by the above), multiplied by terms of the form $u_{k\varepsilon} - \tilde{u}_{k\varepsilon}$ which belong to $\mathcal{N}$. Since $\mathcal{N}(\Omega)$ is an ideal, this gives the required $\mathcal{N}$-bounds. (For a more detailed proof see [39], p. 28.) □

*Composition of generalized functions*, however, is slightly more subtle. Let $\Omega \subseteq \mathbb{R}^n$ resp. $\Omega' \subseteq \mathbb{R}^m$ open and denote by $U$ the $\mathbb{R}^m$-valued generalized function $U = (U^1,\ldots,U^m)$ on $\Omega$ where each $U^i \in \mathcal{G}(\Omega)$. We call $U$ *valued in* $\Omega'$ if there exists a representative $(u_\varepsilon)_\varepsilon$ of $U$ satisfying

$$\forall K \subset\subset \Omega\, \exists K' \subset\subset \Omega',\ \eta > 0 \text{ such that } u_\varepsilon(K) = (u^1_\varepsilon(K),\ldots,u^m_\varepsilon(K)) \subseteq K' \quad \forall \varepsilon \le \eta\,. \tag{60}$$

Clearly every representative of $U$ has the same property. Now we may prove the following

**3.8 Proposition.** *(Composition of generalized functions) Let* $U = (U^1,\ldots,U^m) \in \mathcal{G}(\Omega)^m$ *be* $\Omega'$*-valued for some* $\Omega' \subseteq \mathbb{R}^m$ *open and* $V \in \mathcal{G}(\Omega')$*. Then the composition*

$$V \circ U \;:=\; \mathrm{cl}[(v_\varepsilon \circ u_\varepsilon)_\varepsilon] \tag{61}$$

*is a well-defined generalized function in* $\mathcal{G}(\Omega)$.

**Proof.** First to verify the $\mathcal{E}_M$-bounds for (61) recall that $\partial^\alpha(v_\varepsilon \circ u_\varepsilon)(x)$ may be written as a sum of terms of the form $(\partial^\beta v_\varepsilon)(u_\varepsilon(x))\partial^\gamma u_\varepsilon(x)$. Clearly $\partial^\gamma u_\varepsilon$ satisfies the desired estimate and so does the fist term by the assertion on $U$.
To show that (61) indeed is well-defined first choose another representative $(\tilde{u}_\varepsilon)_\varepsilon$ of $U$. By prop. 3.6 is suffices to estimate $v_\varepsilon \circ u_\varepsilon - v_\varepsilon \circ \tilde{u}_\varepsilon$. We have

$$|v_\varepsilon(u_\varepsilon(x)) - v_\varepsilon(\tilde{u}_\varepsilon(x))| \leq |Dv_\varepsilon((u_\varepsilon + \theta\tilde{u}_\varepsilon)(x))|\,|(u_\varepsilon - \tilde{u}_\varepsilon)(x)|\,,$$

with $0 \leq \theta \leq 1$. Since the second term on the right hand side is negligible and the first one again is moderate we are done. Finally choose another representative $\tilde{v}_\varepsilon$ of $V$. It is immediate that $|v_\varepsilon(u_\varepsilon(x)) - \tilde{v}_\varepsilon(u_\varepsilon(x))|$ satisfies the $\mathcal{N}$-bounds again by the assertion on $U$. $\square$

Let $U \in \mathcal{G}(\Omega)$ and $\Omega'$ an open subset of $\Omega$. We define the *restriction* $U|_{\Omega'}$ of $U$ to $\Omega'$ componentwise, that is $U|_{\Omega'} := \mathrm{cl}[(u_\varepsilon|_{\Omega'})_\varepsilon]$. We say that $U$ vanishes on $\Omega'$ if $U|_{\Omega'} = 0$ and define the *support* of $U$ as the complement of the union of all open subsets of $\Omega$ where $U$ vanishes, i.e., $\mathrm{supp(U)} := (\bigcup\{\Omega' \subset \Omega \text{ open} : \mathrm{U}|_{\Omega'} = 0\})^{\mathrm{c}}$. That this is indeed the complement of the largest open set where $U$ vanishes is one of the consequences of the following

**3.9 Theorem.** *$\mathcal{G}(\mathbb{R}^n)$ is a (fine) sheaf of differential algebras.*

**Proof.** (i) For $\Omega'' \subseteq \Omega' \subseteq \Omega$ open it is obvious that $(U|_{\Omega'})|_{\Omega''} = U_{\Omega''}$.
(ii) Let $(\Omega_\lambda)_{\lambda\in\Lambda}$ be an open covering of $\Omega$.
(a) We have to show that if $U, V \in \mathcal{G}(\Omega)$ satisfy $U|_{\Omega_\lambda} = V|_{\Omega_\lambda}$ for all $\lambda$ then $U = V$, that is $(u_\varepsilon - v_\varepsilon)_\varepsilon \in \mathcal{N}(\Omega)$ for some representatives $(u_\varepsilon)_\varepsilon$, $(v_\varepsilon)_\varepsilon$ of $U$ and $V$ respectively. Since every compact set $K \subseteq \Omega$ may be written as finite union of compact sets $K_j$ each contained in some say $\Omega_{\lambda_j}$ the $\mathcal{N}(\Omega)$-estimate is easily gained from the (finitely many) $\mathcal{N}(\Omega_{\lambda_j})$-estimates.
(b) Let $(U_\lambda)_\lambda$ be a *coherent family*, i.e., $U_\lambda \in \mathcal{G}(\Omega_\lambda)$ and for each $\emptyset \neq \Omega_\lambda \cap \Omega_\mu$ we have $U_\lambda|_{\Omega_\lambda\cap\Omega_\mu} = U_\mu|_{\Omega_\lambda\cap\Omega_\mu}$. Then we have to show that there exists a unique $U \in \mathcal{G}(\Omega)$ with $U|_{\Omega_\lambda} = U_\lambda$.
Uniqueness follows directly from (a). To show existence choose a smooth partition of unity $(\chi_j)_{j\in\mathbb{N}}$ subordinate to the open covering $(\Omega_\lambda)_{\lambda\in\Lambda}$ (i.e., $\chi_j \in \mathcal{D}(\Omega)$, $0 \leq \chi_j \leq 1$ and $\mathrm{supp}(\chi_\mathrm{j}) \subseteq \Omega_{\lambda_\mathrm{j}}$ for a suitable $\lambda_j$. Then $(\Omega_{\lambda_j})_{j\in\mathbb{N}}$ is a covering of $\Omega$; cf. e.g. [49], §2, thm. 1) and set

$$U \;:=\; \mathrm{cl}\,[(\sum_j \chi_j\, u_{\lambda_j\varepsilon})_\varepsilon]\;,$$

where $(u_{\lambda_j\varepsilon})_\varepsilon$ (to be precise $((u_{\lambda_j})_\varepsilon)_\varepsilon$ ) is a representative of $U_{\lambda_j}$ and we have set $u_{\lambda_j\varepsilon} = 0$ on $\Omega \setminus \Omega_{\lambda_j}$. To show that $U$ is moderate choose $K \subset\subset \Omega$. $K$ intersects only finitely many of the $\mathrm{supp}(\chi_\mathrm{j})$. For these, say $1 \leq j \leq N$, set $K_j = K \cap \mathrm{supp}(\chi_\mathrm{j})$. Now each $\chi_j u_{\lambda_j\epsilon}$ satisfies the $\mathcal{E}_M$-bounds on $K_j$. But since $\chi_j = 0$ outside $K_j$ the estimates even hold on all of $K$. Since we have to consider finitely many $j$'s only the claim follows.
It remains to show that $U|_{\Omega_\lambda} = U_\lambda$. Let $K$ be compact in $\Omega_\lambda$ and choose $M$ in such a way that on $K$ $\sum_{j=1}^M \chi_j(x) = 1$. For $x \in K$ we have

$$u_\varepsilon(x) - u_{\lambda\varepsilon}(x) = \sum_{j=1}^M \chi_j u_{\lambda_j\varepsilon}(x) - u_{\lambda\varepsilon}(x) = \sum_{j=1}^M \chi_j(x)\left(u_{\lambda_j\varepsilon}(x) - u_{\lambda\varepsilon}(x)\right)\,.$$

Hence it suffices to prove the $\mathcal{N}(\Omega_\lambda)$-estimate for each summand. For $1 \leq j \leq M$ choose a closed neighborhood $\Omega'_j \subseteq \Omega_{\lambda_j}$ of $\mathrm{supp}(\chi_\mathrm{j})$. For $x \in K \cap \Omega'_j \subset\subset \Omega_\lambda \cap \Omega_{\lambda_j}$ the uniform estimate for $u_{\lambda_j\varepsilon}(x) - u_{\lambda\varepsilon}(x)$ follows by assumption. Since $\chi_j$ is uniformly bounded it does not disturb the estimate. On the other hand, $\chi_j$ vanishes identically on $K \setminus \Omega'_j$. $\square$

### 3.B. Embedding of Distributions

While the construction of the previous section was carried out to define $\mathcal{G}$ as to set the stage for embedding $\mathcal{D}'$ into $\mathcal{G}$ via convolution with an appropriate mollifier, this section is devoted to a detailed study of the latter procedure.

**3.10** As pointed out already after eq. (59) we are bound to use a *mollifier* $\rho$ satisfying

$$\text{(i)} \quad \int \rho(x)\,dx \;=\; 1, \text{ and}$$
$$\text{(ii)} \quad \int \rho(x)\,x^\alpha\,dx \;=\; 0 \quad \forall|\alpha| \geq 1\ .$$

Such a $\rho$ indeed exists by the following argument. Take a Schwartz function equal to unity in a neighborhood of 0. Since all its derivatives vanish at the origin its Fourier transform—which again is in $\mathcal{S}$—satisfies condition (ii) above. By the same reasoning $\rho$ cannot be a test function; in this case its Fourier transform would have to be an entire function, hence to be equal to 1 identically. Now $f(x) = 1 \notin \mathcal{S}$ contradicts the fact that Fourier transform is an isomorphism on $\mathcal{S}$. From now on—unless otherwise stated—$\rho$ denotes a mollifier satisfying (i) and (ii) above and we set $\rho_\varepsilon(x) := \varepsilon^{-n}\rho(\frac{x}{\varepsilon})$, as usual.
The failure of $\mathrm{supp}(\rho)$ to be compact causes the following technical complication: In general its convolution with an arbitrary distribution is not defined forcing us to restrict the simple form of the embedding (58) to compactly supported distributions and use a sheaf theoretic construction to embed the rest of $\mathcal{D}'$.

In the sequel we are going to use the following main result on the local structure of distributions, a proof of which may be found, e.g., in [167], chap. III (theorems XXI and XXVI).

**3.11 Theorem.** *(i) Let $w \in \mathcal{D}'(\Omega)$ and let $\Omega'$ be a relatively compact open set with $\overline{\Omega'} \subseteq \Omega$. Then $w|_{\Omega'} = \partial^\alpha f|_{\Omega'}$ for some $f \in \mathcal{C}(\Omega')$ whose support is contained in an arbitrary neighborhood of $\overline{\Omega'}$ and some $\alpha \in \mathbb{N}_0^n$.*
*(ii) Every $w \in \mathcal{E}'(\Omega)$ is of the form $w = \sum_{|\alpha|\leq r} \partial^\alpha f_\alpha$, where $r \in \mathbb{N}$, $f_\alpha \in \mathcal{C}(\Omega)$ $(0 \leq |\alpha| \leq r)$ and the support of each $f_\alpha$ is contained in an arbitrary neighborhood $\Omega'$ of the support of $w$.*

We are now ready to prove the following properties of the embedding of distributions with compact support.

**3.12 Proposition.** *For any open $\Omega \subseteq \mathbb{R}^n$ the map*

$$\begin{aligned} \iota_0 : \quad & \mathcal{E}'(\Omega) \to \mathcal{G}(\Omega) \\ & w \mapsto ((w * \rho_\varepsilon)|_\Omega)_\varepsilon + \mathcal{N}(\Omega) \end{aligned} \tag{62}$$

*is a linear embedding coinciding with $\sigma$ on $\mathcal{D}(\Omega)$. Consequently, $\iota_0|_{\mathcal{D}(\Omega)}$ is an injective algebra homomorphism.*

Here and in any similar formula it goes without saying that we extend $w$ to all of $\mathbb{R}^n$ by setting it equal to 0 outside of $\Omega$.

**Proof.** To begin with we show that the image of $w \in \mathcal{E}'(\Omega)$ under $\iota_0$ is moderate. By theorem 3.11 is suffices to consider $w = \partial^\beta f$ with $f$ continuous and compactly supported. Then for $x$ in an arbitrary compact subset of $\Omega$ we have

$$\begin{aligned} \partial^\alpha(w * \rho_\varepsilon)(x) = f * \partial^{\alpha+\beta}\rho_\varepsilon(x) \;&=\; \int f(x-y)\partial^{\alpha+\beta}\rho_\varepsilon(y)\,dy \\ &= \int f(x-y)\,\varepsilon^{-(n+|\alpha|+|\beta|)}(\partial^{\alpha+\beta}\rho)(\frac{y}{\varepsilon})\,dy \\ &= \varepsilon^{-(|\alpha|+|\beta|)} \int f(x-y)\partial^{\alpha+\beta}\rho(y)\,dy \;=\; O(\varepsilon^{-(|\alpha|+|\beta|)})\,. \end{aligned}$$

To show that $\iota_0$ is one-to-one let $\iota_0(w) \in \mathcal{N}(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$. Since $w * \rho_\varepsilon \to 0$ uniformly on $\mathrm{supp}(\varphi)$ we have $0 = \lim_{\varepsilon\to 0}\langle w * \rho_\varepsilon, \varphi\rangle = \langle w, \varphi\rangle$.
Finally we show $\iota_0|_{\mathcal{D}(\Omega)} = \sigma|_{\mathcal{D}(\Omega)}$ by a Taylor series argument in combination with the vanishing moment condition 3.10 (ii) as already indicated in eq. (59). For any $f \in \mathcal{D}(\Omega)$ we have to show that $((f * \rho_\varepsilon)|_\Omega - f)_\varepsilon \in \mathcal{N}(\Omega)$. For $q$ an arbitrary natural number we have

$$
\begin{aligned}
(f * \rho_\varepsilon - f)(x) &= \int (f(x-y) - f(x))\rho_\varepsilon(y)\,dy \;=\; \int (f(x-\varepsilon y) - f(x))\rho(y)\,dy \\
&= \int [\sum_{k=1}^{q-1} \frac{1}{k!}((-\varepsilon y)\cdot\nabla)^k f)(x)]\rho(y)\,dy \;+\; \int [\frac{1}{q!}((-\varepsilon y)\cdot\nabla)^q f)(x-\theta\varepsilon y)]\rho(y)\,dy \\
&= 0 \;+\; \varepsilon^q \int \frac{1}{q!}((-y\cdot\nabla)^q f)(x-\theta\varepsilon y)\rho(y)\,dy \\
&= O(\varepsilon^q)\,,
\end{aligned}
$$

where for the last estimate we have to take into account that $f$ is bounded and $\rho$ is a Schwartz function. Note that the $\mathcal{N}$-estimate in this case holds even uniformly on $\Omega$. $\square$

The following technical result plays a key role in the sheaf theoretic construction of the embedding of $\mathcal{D}'$.

**3.13 Lemma.** *Let* $w \in \mathcal{E}'(\Omega)$ *then* $\mathrm{supp(w)} = \mathrm{supp}(\iota_0(\mathrm{w}))$.

**Proof.** We first prove the inclusion $\mathrm{supp}(\iota_0(\mathrm{w})) \subseteq \mathrm{supp(w)}$ by showing that $\iota_0(w)|_{\mathrm{supp(w)^c}} \in \mathcal{N}(\mathrm{supp(w)^c})$. Let $K \subset\subset \mathrm{supp(w)^c}$ then by theorem 3.11 we may write $w = \partial^\alpha f$ with $\alpha \in \mathbb{N}_0^n$ and $\mathrm{supp(f)} \subseteq \mathrm{\mathbb{R}^n \setminus K}$. For the representative $((f * \partial^\alpha \rho_\varepsilon)|_\Omega)_\varepsilon$ of $\iota_0(w)$ we have

$$
\begin{aligned}
f * \partial^\alpha \rho_\varepsilon(x) &= \int f(x-y)\partial^\alpha \rho_\varepsilon(y)\,dy \;=\; \int \varepsilon^{-|\alpha|} f(x-\varepsilon y)\partial^\alpha \rho(y)\,dy \\
&= \varepsilon^{-|\alpha|} \int\limits_{|y|<\frac{1}{\sqrt{\varepsilon}}} f(x-\varepsilon y)\partial^\alpha\rho(y)\,dy + \varepsilon^{-|\alpha|} \int\limits_{|y|\geq\frac{1}{\sqrt{\varepsilon}}} f(x-\varepsilon y)\partial^\alpha\rho(y)\,dy.
\end{aligned}
$$

For $\varepsilon$ small enough the first integral vanishes by the assumption on the support of $f$. The absolute value of the second integral is dominated by $\varepsilon^{-|\alpha|}\|f\|_\infty \int |\partial^\alpha \rho(y)|\,dy$. Since $\rho \in \mathcal{S}(\mathbb{R}^n)$, for any $q \in \mathbb{N}$ there exists some $C_q > 0$ with $|\partial^\alpha \rho(y)| \leq C_q(1+|y|)^{-2q-n-1}$. Thus we obtain the $\mathcal{N}$-estimate by

$$
\int\limits_{|y|\geq\frac{1}{\sqrt{\varepsilon}}} \varepsilon^{-|\alpha|}|\partial^\alpha\rho(y)|\,dy \leq C_q\,\varepsilon^{q-|\alpha|} \int (1+|y|)^{-n-1}\,dy = \widetilde{C}_q\,\varepsilon^{q-|\alpha|}.
$$

Conversely, to prove $\mathrm{supp(w)} \subseteq \mathrm{supp}(\iota_0(\mathrm{w}))$ let $x_0 \in \mathrm{supp(w)}$. For any $\eta > 0$ there exists some $\varphi \neq 0 \in \mathcal{D}(\mathbb{R}^n)$ with $\mathrm{supp}(\varphi) \subseteq \mathrm{B_\eta(x_0)}$ and $|\langle w, \varphi\rangle| = c > 0$. Since $w * \rho_\varepsilon \to w$ this implies $|\langle w * \rho_\varepsilon, \varphi\rangle| > \frac{c}{2}$ for $\varepsilon$ small. But then

$$
\iota_0(w)|_{B_\eta(x_0)} = \mathrm{cl}[((w*\rho_\varepsilon)|_{B_\eta(x_0)})_\varepsilon] \neq 0 \in \mathcal{G}(B_\eta(x_0)),
$$

so $x_0 \in \mathrm{supp}(\iota_0(\mathrm{w}))$. $\square$

**3.14** We are now prepared to start the sheaf theoretic construction of the embedding of $\mathcal{D}'$. Choose some open covering $(\Omega_\lambda)_{\lambda\in\Lambda}$ of $\Omega$ such that $\overline{\Omega}_\lambda \subset\subset \Omega$ and a family $(\psi_\lambda)_\lambda$ of test functions on $\Omega$ with $\psi_\lambda \equiv 1$ on an open neighborhood of $\overline{\Omega}_\lambda$. For each $\lambda \in \Lambda$ we define the mapping

$$
\begin{aligned}
\iota_\lambda : \mathcal{D}'(\Omega) &\to \mathcal{G}(\Omega_\lambda) \\
w &\mapsto \iota_\lambda(w) := (((\psi_\lambda w) * \rho_\varepsilon)|_{\Omega_\lambda})_\varepsilon + \mathcal{N}(\Omega_\lambda)
\end{aligned}
$$

The idea behind this definition is of course to use the cut-off functions $\psi_\lambda$ to render the convolution defined and then to use the embedding $\iota_0$ constructed previously.

**3.15 Proposition.** *Let* $w \in \mathcal{D}'(\Omega)$ *then* $(\iota_\lambda(w))_{\lambda\in\Lambda}$ *is a coherent family, i.e.,*

$$
\iota_\lambda(w)|_{\Omega_\lambda\cap\Omega_\mu} \;=\; \iota_\mu(w)|_{\Omega_\lambda\cap\Omega_\mu} \qquad \forall\lambda,\mu\in\Lambda\,.
$$

*Proof.* We have to show that $((((\psi_\lambda - \psi_\mu)w) * \rho_\varepsilon)|_{\Omega_\lambda\cap\Omega_\mu})_\varepsilon \in \mathcal{N}(\Omega_\lambda\cap\Omega_\mu)$. Set $v := (\psi_\lambda - \psi_\mu)w \in \mathcal{E}'(\Omega)$. Then $\Omega_\lambda \cap \Omega_\mu \subseteq \mathrm{supp(v)^c}$, so by 3.13 it follows that $\iota_0(v)|_{\Omega_\lambda\cap\Omega_\mu} \in \mathcal{N}(\Omega_\lambda\cap\Omega_\mu)$. $\square$

**3.16** By theorem 3.9 we are now guaranteed with the existence of a unique

$$\iota(w) \in \mathcal{G}(\Omega) \text{ with the property } \iota(w)|_{\Omega_\lambda} = \iota_\lambda(w) \quad \forall \lambda \in \Lambda\,. \tag{63}$$

Moreover from part (ii,b) of the proof of the theorem we even get the following explicit expression for $\iota(w)$: Let $(\chi_j)_{j\in\mathbb{N}}$ be a smooth partition of unity subordinate to $(\Omega_\lambda)_{\lambda\in\Lambda}$ then

$$\iota(w) = (\sum_{j=1}^{\infty} \chi_j((\psi_{\lambda_j} w) * \rho_\varepsilon))_{\varepsilon\in I} + \mathcal{N}(\Omega)\,. \tag{64}$$

Having collected all the necessary ingredients we are in the position to formulate and prove the following

**3.17 Theorem.** *For any open $\Omega \in \mathbb{R}^n$ the map $\iota : \mathcal{D}'(\Omega) \to \mathcal{G}(\Omega)$ is a linear embedding coinciding with $\iota_0$ on $\mathcal{E}'(\Omega)$. Moreover, we have*

$$\iota|_{\mathcal{C}^\infty(\Omega)} \;=\; \sigma$$

*turning $\mathcal{C}^\infty(\Omega)$ into a faithful subalgebra of $\mathcal{G}(\Omega)$.*

**Proof.** Linearity of $\iota$ is obvious. To prove that it is one-to-one we have to show that if $w \in \mathcal{D}'(\Omega)$ with $\iota(w) \in \mathcal{N}(\Omega)$ then $w = 0$. Choose some $\varphi \in \mathcal{D}(\Omega)$ and denote its support by $K$. Let $M \in \mathbb{N}$ so large that $K_j := K \cap \mathrm{supp}(\chi_j) = \emptyset \;\; \forall j \geq M$. Clearly $\sum_{j=1}^{M} \chi_j = 1$ on $K$ and we have

$$\langle w, \varphi\rangle = \langle w, \sum_{j=1}^{M} \chi_j \varphi\rangle = \sum_{j=1}^{M} \langle \chi_j \psi_{\lambda_j} w, \varphi\rangle$$

Since $\iota(w)|_{\Omega_{\lambda_j}} = \iota_{\lambda_j}(w)$ we have $\iota(w)_\varepsilon - (\psi_{\lambda_j} w) * \rho_\varepsilon \to 0$ uniformly on $K_j$. By assumption $\iota(w) \in \mathcal{N}(\Omega)$ hence $(\psi_{\lambda_j} w) * \rho_\varepsilon \to 0$ uniformly on $K_j$ which implies $\langle \chi_j \psi_{\lambda_j} w, \varphi\rangle = \lim \langle \chi_j \psi_{\lambda_j} w * \rho_\varepsilon, \varphi\rangle = 0$ verifying our claim.
To prove $\iota|_{\mathcal{E}'(\Omega)} = \iota_0$, by theorem 3.9 and (63) it suffices to show that for each $w \in \mathcal{D}'(\Omega)$ and for all $\lambda \in \Lambda$ we have

$$\iota_0(w)|_{\Omega_\lambda} = \iota(w)|_{\Omega_\lambda} = \iota_\lambda(w)\,.$$

We have $\iota_0(w)|_{\Omega_\lambda} - \iota_\lambda(w) = (((1-\psi_\lambda)w * \rho_\varepsilon)|_{\Omega_\lambda})_\varepsilon + \mathcal{N}(\Omega_\lambda) = \iota_0(v)|_{\Omega_\lambda}$ for $v := (1-\psi_\lambda)w \in \mathcal{E}'(\Omega)$. But $\Omega_\lambda \subseteq \mathrm{supp(v)}^{\mathrm{c}} = \mathrm{supp}(\iota_0(\mathrm{v}))^{\mathrm{c}}$ by 3.13.
To prove the last claim, again by theorem 3.9 we only have to show that for each smooth $f$ and for all $\lambda \in \Lambda$ $\sigma(f)|_{\Omega_\lambda} = \iota_\lambda(f)$. Let $K \subset\subset \Omega_\lambda$ and observe that $\psi_\lambda f \in \mathcal{D}'(\Omega)$ equals $f$ in an open neighborhood of $\overline{\Omega}_\lambda$. Now the desired $\mathcal{N}$-estimate on $K$ follows directly from 3.12. □

Finally we complete our program constructing associative and commutative algebras satisfying properties (i)-(iii) in 2.10 and (iv") in 2.13 by the following

**3.18 Theorem.** *Let $\alpha \in \mathbb{N}_0^n$ and $w \in \mathcal{D}'(\Omega)$ then*

$$\partial^\alpha(\iota(w)) \;=\; \iota(\partial^\alpha(w))\,.$$

**Proof.** As usual by theorem 3.9 we have to show that for all $\lambda$

$$(\partial^\alpha(\iota(w)))|_{\Omega_\lambda} \;=\; \iota(\partial^\alpha w)|_{\Omega_\lambda} \;=\; \iota_\lambda(\partial^\alpha w) \tag{65}$$

where the last equality follows directly from the definitions, i.e., from $(\partial^\alpha U)|_{\Omega'} = \partial^\alpha(U|_{\Omega'})$. A representative of the left hand side of (65) is given by

$$\partial^\alpha(((\psi_\lambda w) * \rho_\varepsilon)|_{\Omega_\lambda}) = ((\psi_\lambda \partial^\alpha w) * \rho_\varepsilon)|_{\Omega_\lambda} + \sum_{|\beta|=1}^{|\alpha|} c_\beta((\partial^\beta \psi_\lambda \partial^{\alpha-\beta} w) * \rho_\varepsilon)|_{\Omega_\lambda}$$

with certain constants $c_\beta$. The second term on the right hand side vanishes by the following argument: Take any $\beta$ with $1 \leq |\beta| \leq |\alpha|$ and set $v := \partial^\beta \psi_\lambda \partial^{\alpha-\beta} w \in \mathcal{E}'(\Omega)$. By 3.13 $\mathrm{supp}(\iota_0(\mathrm{v})) \subseteq \Omega_\lambda^{\mathrm{c}}$, so $\iota_0(v)|_{\Omega_\lambda} = cl[((\partial^\beta \psi_\lambda \partial^{\alpha-\beta} w) * \rho_\varepsilon)|_{\Omega_\lambda}] = 0$ in $\mathcal{G}(\Omega_\lambda)$. Hence the claim follows by the observation that a representative of the right hand side of (65) is also given by the expression $((\psi_\lambda \partial^\alpha w) * \rho_\varepsilon)|_{\Omega_\lambda}$. □

Once having reached our goal it is time to throw overboard the unnecessary technical accessories of our construction. We begin with the following observation

**3.19 Proposition.** *The embedding $\iota : \mathcal{D}'(\Omega) \hookrightarrow \mathcal{G}(\Omega)$ is independent of the particular choice of the covering $(\Omega_\lambda)_{\lambda\in\Lambda}$, the family of cut-off functions $(\psi_\lambda)_{\lambda\in\Lambda}$ and the partition of unity $(\chi_j)_{j\in\mathbb{N}}$.*

**Proof.** Independence of the choice of $(\chi_j)_j$ follows already from theorem 3.9. To prove the other statements choose another covering $(\Omega_{\lambda'})_{\lambda'\in\Lambda'}$ and another family $(\psi_{\lambda'})_{\lambda'\in\Lambda'}$ as above and denote the respective embedding with $\iota'$. Since $(\Omega_\lambda \cap \Omega_{\lambda'})_{(\lambda,\lambda')\in\Lambda\times\Lambda'}$ is an open covering of $\Omega$, again by the sheaf properties it suffices to show that for $w \in \mathcal{D}'(\Omega)$ and arbitrary $\lambda$, $\lambda'$

$$\iota(w)|_{\Omega_\lambda\cap\Omega_{\lambda'}} = \iota'(w)|_{\Omega_\lambda\cap\Omega_{\lambda'}}.$$

Note that $\iota(w)|_{\Omega_\lambda\cap\Omega_{\lambda'}} = \iota(w)|_{\Omega_\lambda}|_{\Omega_\lambda\cap\Omega_{\lambda'}} = \iota_\lambda(w)|_{\Omega_\lambda\cap\Omega_{\lambda'}}$, and mutatis mutandis for $\iota'$, so our claim follows from $\iota_\lambda|_{\Omega_\lambda\cap\Omega_{\lambda'}} = \iota'_{\lambda'}|_{\Omega_\lambda\cap\Omega_{\lambda'}}$ which is obvious from the proof of 3.15. $\square$

To stress more deeply the sheaf theoretic aspects of the whole construction we introduce the following notation. We write $\hat{\iota}$ for the entirety of all $\iota = \iota_\Omega : \mathcal{D}'(\Omega) \to \mathcal{G}(\Omega)$ where $\Omega$ is an arbitrary open subset of $\mathbb{R}^n$. $\hat{\iota}$ intuitively is the "lift" of the embeddings $\iota_\Omega$ to the level of the sheafs $\mathcal{D}$ and $\mathcal{G}$. The following proposition expresses the fact that "$\hat{\iota}$ commutes with restrictions."

**3.20 Proposition.** *$\widehat{\iota} : \mathcal{D}' \to \mathcal{G}$ is a sheaf morphism (in the category of complex vector spaces). More precisely let $\Omega_2 \subseteq \Omega_1 \subseteq \mathbb{R}^n$ open and $w \in \mathcal{D}'(\Omega_1)$ then we have*

$$\iota_{\Omega_1}(w)|_{\Omega_2} = \iota_{\Omega_2}(w|_{\Omega_2})$$

**Proof.** Let $\Omega_2 \subseteq \Omega_1 \subseteq \mathbb{R}^n$ open and choose $(\Omega_\lambda)_\lambda$ and $(\psi_\lambda)_\lambda$ with respect to $\Omega_2$. Now add some open sets to $(\Omega_\lambda)$ in order to get a suitable open cover of $\Omega_1$. Moreover, setting $\psi_\lambda = 0$ on $\Omega_1 \setminus \Omega_2$ we ensure that $\iota_{\Omega_1\lambda}$ can be defined using $\psi_\lambda$. For $w \in \mathcal{D}'(\Omega_1)$ and $\lambda \in \Lambda$ we have $(\iota_{\Omega_1}(w)\,|_{\Omega_2})\,|_{\Omega_\lambda} = \iota_{\Omega_1}(w)\,|_{\Omega_\lambda} = \iota_{\lambda\,\Omega_1}(w) = (\,(\psi_\lambda w) * \rho_\epsilon\,|_{\Omega_\lambda})_\epsilon + \mathcal{N}(\Omega_\lambda)$. On the other hand, $(\iota_{\Omega_2}(w\,|_{\Omega_2})\,)\,|_{\Omega_\lambda} = \iota_{\lambda\,\Omega_2}(w\,|_{\Omega_2}) = \iota_{\lambda\,\Omega_2}(w\,|_{\Omega_2}) = (\,(\psi_\lambda w\,|_{\Omega_2}\,) * \rho_\epsilon\,|_{\Omega_\lambda})_\epsilon + \mathcal{N}(\Omega_\lambda) = (\,(\psi_\lambda w) * \rho_\epsilon\,|_{\Omega_\lambda})_\epsilon + \mathcal{N}(\Omega_\lambda)$. The claim now follows once more from the sheaf properties of $\mathcal{G}$. $\square$

Finally we are able to formulate the main result of this section

**3.21 Theorem.** *There is a unique sheaf morphism (of complex vector spaces) $\widehat{\iota} : \mathcal{D}' \to \mathcal{G}$ which extends the canonical embedding $\widehat{\iota_0} : \mathcal{E}' \to \mathcal{G}$ (i.e., $\iota_\Omega|_{\mathcal{E}'(\Omega)} = \iota_{0\Omega}$ for each $\Omega \subseteq \mathbb{R}^n$ open). $\widehat{\iota}$ commutes with partial derivatives and its restriction to $\mathcal{C}^\infty$ is a sheaf morphism of algebras.*

**Proof.** Suppose that $\widehat{\kappa} : \mathcal{D}' \to \mathcal{G}$ is another sheaf morphism with this property and let $\Omega$ be an open subset of $\mathbb{R}^n$. Take some relatively compact open set $\Omega_1$ whose closure is contained in $\Omega$ and choose $\psi \in \mathcal{D}(\Omega)$ with $\psi|_{\Omega_1} = 1$. For any $w \in \mathcal{D}'(\Omega)$ we find

$$\kappa_\Omega(w)|_{\Omega_1} = \kappa_{\Omega_1}(w|_{\Omega_1}) = \kappa_{\Omega_1}(\psi w|_{\Omega_1}) = \kappa_\Omega(\psi w)|_{\Omega_1} = \iota_{0\Omega}(\psi w)|_{\Omega_1}$$

and, by the same reasoning: $\iota_\Omega(w)|_{\Omega_1} = \iota_{0\Omega}(\psi w)|_{\Omega_1}$. Hence $\kappa_\Omega(w)|_{\Omega_1} = \iota_\Omega(w)|_{\Omega_1}$. Since the $\Omega_1$ as above form an open covering of $\Omega$ we conclude that $\iota_\Omega = \kappa_\Omega$ for each $\Omega \subseteq \mathbb{R}^n$ open. Therefore, $\widehat{\iota} = \widehat{\kappa}$. The other statements follow from 3.18 and 3.17. $\square$

**3.22** Summing up, the above theorem tells us that if we choose to embed compactly supported distributions by convolution with a mollifier $\rho$ while at the same time demanding the localization properties of $\mathcal{D}'$ and $\mathcal{G}$ encoded in the sheaf structure, the construction carried out so far is both natural and unique. Indeed the extendibility of $\hat{\iota}_0$ is already a consequence of some general results from sheaf theory together with proposition 3.13. However, since we do not assume the reader to be familiar with any subtleties of this field we chose to carry out the construction along a "pedestrian road" and refer the reader interested in the details of the sheaf theoretic aspect of Colombeau algebras to [50].

**3.23** An important point to mention is that we were able to remove *all* the technical accessories of the construction, i.e., to show independence of the choice of the open covering, the partition of unity and the family of cut-off functions. On the other hand note that $\iota$ still depends on the choice of the mollifier

$\rho$. However, $\rho$ is not a technical accessory of the theory. It reflects a fundamental property of nonlinear modelling: *In general, nonlinear properties of a singular object depend on the regularization.* This opens a gate to a wide range of applications. Additional input on the regularization from, say, a physical model may enter the mathematical theory via this interface, leading to a sensible description of the problem at hand. The strength of the construction presented above is that it deals with the $\rho$-dependence in a consistent way. An even more aesthetical way of bookkeeping on this dependence is employed in the full version of Colombeau algebras (as presented in [39]). The key idea is to use the set of all (suitable) mollifiers as an index set for the generalized functions, i.e., a representative takes the form $(u_\rho(x))_\rho$ (see also section 4.D). We shall reenter the discussion on the features of Colombeau algebras in the context of nonlinear modelling in sec 3.E having at hand the full technical apparatus of the theory.

The exact form of the embedding $\iota$ is not too well suited for explicit calculations. Fortunately the simple convolution formula given in 3.12 may be used for a much wider class than just compactly supported distributions.

**3.24 Proposition.** *Let $f \in L^1_{loc}(\Omega)$ be polynomially bounded (i.e., $\exists\ C > 0,\ r \in \mathbb{N}$ such that $|f(x)| \leq C(1+|x|)^r$ almost everywhere). Then $\iota(w) = \mathrm{cl}[((f * \rho_\varepsilon)|_\Omega)_\varepsilon]$.*

**Proof.** Set $\widetilde{\iota}(w) = \mathrm{cl}[((f * \rho_\varepsilon)|_\Omega)_\varepsilon]$. Then

$$|f * \rho_\varepsilon(x)| = |\int f(x-\varepsilon y)\rho(y)\,dy| \leq C\int((1+|x-\varepsilon y|)^r)|\rho(y)|\,dy.$$

is bounded independently of $x \in K$ and $\varepsilon$, hence $\widetilde{\iota}(w) \in \mathcal{G}(\Omega)$. Using the same notation as before, we have to verify that $\iota_\lambda(f) = \mathrm{cl}[((f * \rho_\varepsilon)|_{\Omega_\lambda})_\varepsilon]\ \forall\lambda$. For $K \subseteq \Omega_\lambda$ compact and $x \in K$ we have

$$\begin{aligned}\iota_\lambda(f)(x) - f * \rho_\varepsilon(x) &= ((\psi_\lambda - 1)f) * \rho_\varepsilon(x)\\ &= \int\limits_{|y|\leq\frac{1}{\sqrt{\varepsilon}}}((\psi_\lambda-1)f)(x-\varepsilon y)\rho(y)\,dy + \int\limits_{|y|\geq\frac{1}{\sqrt{\varepsilon}}}(\psi_\lambda-1)(x-\varepsilon y)f(x-\varepsilon y)\rho(y)\,dy.\end{aligned}$$

For $\varepsilon$ small, the first integral vanishes. The absolute value of the second one can be estimated by $C\int_{|y|\geq\frac{1}{\sqrt{\varepsilon}}}(1+|x-\varepsilon y|)^r\rho(y)\,dy$. Now the proof is finished by an argument similar to the one used in the proof of 3.13. □

We close this section by presenting some simple

**3.25 Examples.**

(i) The $\delta$-distribution as an element of $\mathcal{G}(\mathbb{R})$.
By 3.12 we have

$$\iota(\delta) = (\rho_\varepsilon)_\varepsilon + \mathcal{N}(\Omega).$$

This formula reflects most clearly the influence of the mollifier. The $\delta$-distribution is smoothed out by the regularization sequence $\rho_\varepsilon$, hence its nonlinear properties are given by, i.e., depend on $\rho$.

(ii) The product $x\delta$.
Warned by 2.7 we expect $\iota(x)\iota(\delta) \neq 0$ in $\mathcal{G}(\mathbb{R})$ although $x\delta = 0$ in $\mathcal{D}'$. Choose some $x_0 \neq 0$ with $\rho(x_0) \neq 0$ and set $x = \varepsilon x_0$. Then

$$(\iota(x)\rho(x))_\varepsilon = (x\rho_\varepsilon(x))_\varepsilon = ((x_0\rho(x_0))_\varepsilon \neq 0$$

indeed implying $(x\rho_\varepsilon)_\varepsilon \notin \mathcal{N}(\mathbb{R})$.

(iii) What about the square of the $\delta$-distribution in $\mathcal{G}$?
No problem at all:

$$\iota(\delta)^2 = \mathrm{cl}[(\rho_\varepsilon^2)_\varepsilon],$$

hence diverging in the $L^\infty$-norm at a rate of $1/\varepsilon^2$ but still under control in $\mathcal{G}$.

(iv) The Heaviside function in $\mathcal{G}(\mathbb{R})$.
According to 3.24 we have

$$\iota(H) = \mathrm{cl}[(H * \rho_\varepsilon(x))_\varepsilon] = \mathrm{cl}[(\int\limits_{-\infty}^{x} \rho_\varepsilon(y)dy)_\varepsilon].$$

The same argument as in 2.5 may be employed to show that in $\mathcal{G}$, $\iota(H) \neq \iota(H)^2$.

### 3.C. Point values of Generalized Functions

In classical distribution theory a concept of point values has been introduced by Łojasiewicz [128] (see also [144], p. 67f). However, not every distribution admits a point value at arbitrary points. For example neither the Dirac-$\delta$ nor the Heaviside function can be assigned a point value at $x = 0$. All continuous functions have Łojasiewicz-point values but the converse does not hold. Moreover point values by no means characterize distributions in the way they characterize, say, continuous functions.
In the Colombeau algebra $\mathcal{G}$, point values arise in a very natural way, i.e., by inserting points into representatives. However, the objects arising that way are not elements in the field $\mathbb{K}$ but sequences of real (or complex) numbers representing generalized numbers. Our first aim is to give a precise formulation of this notion revealing that the set of generalized numbers is just the ring of constants in $\mathcal{G}$.
However, these point values do not characterize generalized functions. It was shown only recently by Oberguggenberger and Kunzinger [145] that generalized functions may be characterized by their generalized point values, i.e., generalized points inserted into representatives. We shall briefly discuss this result since it provides a powerful means to directly transfer many methods from classical analysis to Colombeau algebras.

**3.26 Definition.** *We set*

$$\mathcal{E} := \{(r_\varepsilon)_{\varepsilon\in I} \in \mathbb{K}^I : \exists\ p \in \mathbb{N}\ |r_\varepsilon| = O(\varepsilon^{-p})\ (\varepsilon \to 0)\}$$
$$\mathcal{N} := \{(r_\varepsilon)_{\varepsilon\in I} \in \mathbb{K}^I : \forall\ q \in \mathbb{N}\ |r_\varepsilon| = O(\varepsilon^{q})\ (\varepsilon \to 0)\}$$

*and define the* ring of generalized numbers *(with componentwise operations) by*

$$\mathcal{K} := \mathcal{E}/\mathcal{N}\,.$$

*In case $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) we set $\mathcal{K} = \mathcal{R}$ (resp. $\mathcal{K} = \mathcal{C}$).*

We shall mainly apply our notational conventions explained in 3.5 also in the case of $\mathcal{K}$. However, we denote generalized numbers by small letters, thus writing $r = \mathrm{cl}[(r_\varepsilon)_\varepsilon]$. That $\mathcal{K}$ is a ring but not a field is demonstrated by the following

**3.27 Example.** Set

$$r_\varepsilon := \begin{cases} 0 & \varepsilon = \frac{1}{n}\quad (n \in \mathbb{N}) \\ 1 & \text{otherwise} \end{cases}$$

Clearly $(r_\varepsilon)_\varepsilon \in \mathcal{E}$ but $(r_\varepsilon)_\varepsilon \notin \mathcal{N}$, implying $r \neq 0$ in $\mathcal{K}$. Now suppose there exists some generalized number $s = \mathrm{cl}[(s_\varepsilon)_\varepsilon]$ with $r \cdot s = 1$. Then for some sequence $(n_\varepsilon)_\varepsilon \in \mathcal{N}$ we have $r_\varepsilon s_\varepsilon + n_\varepsilon = 1$ for all $\varepsilon$. But this implies $n_\varepsilon = 1$ for $\varepsilon = \frac{1}{n}$, contradicting the $\mathcal{N}$-property.

**3.28** An Element $r \in \mathcal{R}$ is called *strictly nonzero* if there is some representative $(r_\varepsilon)_\varepsilon$ and $m \in \mathbb{N}$ such that $|r_\varepsilon| \geq \varepsilon^m$ for sufficiently small $\varepsilon$. A strictly nonzero generalized number $r$ is invertible; its inverse is given by $\mathrm{cl}[(1/r_\varepsilon)_\varepsilon]$.
Generalized *real* numbers can be endowed with some extra structure. For $r, s \in \mathcal{R}$ we define $r \leq s$ if there are representatives $(r_\varepsilon)_\varepsilon$ and $(s_\varepsilon)_\varepsilon$ satisfying $r_\varepsilon \leq s_\varepsilon$ for all $\varepsilon$. We now have

**3.29 Proposition.** $(\mathcal{R}, \leq)$ *is a partially ordered ring.*

**Proof.** Reflexivity is obvious. To show antisymmetry let $r \leq s$ and $s \leq r$. Then there are negligible sequences $(m_\varepsilon)_\varepsilon$ and $(n_\varepsilon)_\varepsilon$ such that for all $\varepsilon$, $r_\varepsilon \leq s_\varepsilon + m_\varepsilon$ and $s_\varepsilon \leq r_\varepsilon + n_\varepsilon$. Hence $|r_\varepsilon - s_\varepsilon| \leq \max(|m_\varepsilon|, |n_\varepsilon|)$ for all $\varepsilon$ implying $r_\varepsilon - s_\varepsilon \in \mathcal{N}$. For transitivity let $r \leq s \leq t$, then for all $\varepsilon$, $r_\varepsilon \leq s_\varepsilon + m_\varepsilon \leq t_\varepsilon + m_\varepsilon + n_\varepsilon$ for $m_\varepsilon, n_\varepsilon \in \mathcal{N}$, so $r \leq t$. □

**3.30** The field $\mathbb{K}$ is trivially embedded into $\mathcal{K}$ by

$$\sigma : \quad \mathbb{K} \to \mathcal{K}$$
$$r \mapsto (r)_\varepsilon + \mathbb{N}\,.$$

Moreover $\mathcal{K}$ is a subring of $\mathcal{G}(\Omega)$ for any open $\Omega \subseteq \mathbb{R}^n$ and—as already announced above—$\mathcal{K}$ is actually the ring of constants in any of the above algebras.

**3.31 Theorem.** *Let $\Omega \subseteq \mathbb{R}^n$ open and connected and $U \in \mathcal{G}(\Omega)$ then*

$$\nabla U = 0 \;\Leftrightarrow\; U \in \mathcal{K}$$

**Proof.** For the nontrivial implication let $(\nabla u_\varepsilon)_\varepsilon \in \mathcal{N}(\Omega)^n$ and—to begin with—let us assume $\Omega$ to be star-shaped. Then

$$|u_\varepsilon(x) - u_\varepsilon(x_0)| \;=\; |(x - x_0)\int\limits_0^1 \nabla u_\varepsilon(x_0 + \sigma(x - x_0))d\sigma| \;\leq\; C\varepsilon^q$$

for $q$ arbitrary, so $(u_\varepsilon)_\varepsilon - (u_\varepsilon(x_0))_\varepsilon \in \mathcal{N}$. If $\Omega$ is connected, any two points may be joined by a polygon, for which an analogous argument applies. □

In addition to 3.30 we have $\mathcal{N} = \mathcal{N}(\Omega) \cap \mathcal{K}$ rendering the following notion well-defined:

**3.32 Definition.** *For $U \in \mathcal{G}(\Omega)$ and $x_0 \in \Omega$, the* point value of $U$ at $x_0$ *is defined as* $\mathrm{cl}[(u_\varepsilon(x_0))_\varepsilon]$ *in $\mathcal{K}$.*

**3.33 Examples.**

(i) If $f$ is a smooth function on $\Omega$ and $x_0 \in \Omega$ then the point values of $f$ as a classical function and as a generalized function coincide in the following precise sense: $\iota(f)(x_0) = \sigma(f)(x_0) = \sigma(f(x_0))$ in $\mathcal{K}$.

(ii) For $f$ merely continuous, this is not true in general as can be seen from the generalized point value of $x_+ := H(x)x$ at $x_0 = 0$:

$$\iota(x_+)(x_0) = (x_+ * \rho_\varepsilon(x))_\varepsilon + \mathcal{N} = (\textstyle\int_0^\infty y\rho_\varepsilon(x_0 - y)\,dy)_\varepsilon + \mathcal{N}, \text{ so}$$
$$\iota(x_+)(0) = (\varepsilon \textstyle\int_0^\infty y\rho(-y)\,dy)_\varepsilon + \mathcal{N}.$$

Hence $\iota(x_+)(0)$ depends on the mollifier $\rho$ and in general will be non-vanishing in $\mathcal{R}$. On the other hand, the classical point value of $x_+$ at 0 is recovered in the limit $\varepsilon \to 0$. This is the first example for a general rule, as we shall see in 3.52 (iii).

(iii) Point values of $\delta$-distributions at 0.

(a) $\iota(\delta)(0) = (\rho_\varepsilon(0))_\varepsilon + \mathcal{N} = (\frac{1}{\varepsilon}\rho(0))_\varepsilon + \mathcal{N}$

(b) $\iota(\delta')(0) = (\rho'_\varepsilon(0))_\varepsilon + \mathcal{N} = (\frac{1}{\varepsilon^2}\rho'(0))_\varepsilon + \mathcal{N}$

(c) $(\iota(\delta))^2(0) = (\rho^2_\varepsilon(0))_\varepsilon + \mathcal{N} = (\frac{1}{\varepsilon^2}\rho(0))_\varepsilon + \mathcal{N}$

Thus point values of singular objects depend on the mollifier and display different degrees of divergence as measured by the parameter $\varepsilon$.

(iv) We know from 3.25 that $\iota(x)\iota(\delta) \neq 0$ in $\mathcal{G}(\mathbb{R})$. However, every point value of this generalized function vanishes since

$$x_0\rho_\varepsilon(x_0) = \frac{x_0}{\varepsilon}\rho(\frac{x_0}{\varepsilon}) \to 0$$

faster than any power of $\varepsilon$ since $\rho$ is a Schwartz function.

**3.34** From the last example above we explicitly see that a generalized function is not determined by its point values. Prescribing all point values of all derivatives will not help either by the following argument. Take some nonnegative $\varphi \in \mathcal{D}(\mathbb{R})$ with $\mathrm{supp}(\varphi) \subseteq [-1, 1]$ and $\int \varphi = 1$ and set $u_\varepsilon(x) = \varphi_\varepsilon(x - \varepsilon)$. Then $(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(\mathbb{R})$, so $U := \mathrm{cl}[(u_\varepsilon)_\varepsilon] \in \mathcal{G}(\mathbb{R})$. It is easily seen that every point value of every derivative of $U$ is 0 in $\mathcal{K}$. But clearly, $U \neq 0$ in $\mathcal{G}(\mathbb{R})$.
As already indicated at the beginning of this section this last result is not surprising from the point of view of distribution theory. However, viewing elements of $\mathcal{G}$ as direct generalizations of classical functions the question for an appropriate generalization of the classical point value-description arises naturally. It is one of the most important recent developments in the theory of algebras of generalized functions that such an extension is indeed possible (Oberguggenberger and Kunzinger [145]). Apart from reasons intrinsically to the theory of generalized functions the value of the following results must not be underestimated from the viewpoint of applications. It is the key ingredient to directly generalize a number of geometric aspects from classical analysis to the context of Colombeau algebras.

The basic idea is to generalize also the domain of definition of generalized functions.

**3.35 Definition.** *Let $\Omega$ be an open subset of $\mathbb{R}^n$. On $\Omega_M := \{(x_\varepsilon)_\varepsilon \in \Omega^I : \exists p > 0\ \exists\ \eta > 0\ |x_\varepsilon| \leq \varepsilon^{-p}\ (0 < \varepsilon < \eta)\}$ we introduce an equivalence relation by*

$$(x_\varepsilon)_\varepsilon \sim (y_\varepsilon)_\varepsilon \ \Leftrightarrow\ \forall q > 0\ \exists \eta > 0\ |x_\varepsilon - y_\varepsilon| \leq \varepsilon^q\ (0 < \varepsilon < \eta)$$

*and set $\widetilde{\Omega} := \Omega_M / \sim$. The set of* compactly supported points *is defined by*

$$\widetilde{\Omega}_c = \{\widetilde{x} \in \widetilde{\Omega} : \exists \text{ representative } (x_\varepsilon)_\varepsilon\ \exists K \subset\subset \Omega\ \exists \eta > 0 \text{ with } x_\varepsilon \in K \text{ for } 0 < \varepsilon < \eta\}$$

It is clear that if the $\widetilde{\Omega}_c$-property holds for one representative of $\widetilde{x} \in \widetilde{\Omega}$ then it holds for every representative. Also, for $\Omega = \mathbb{K}$ we have $\widetilde{\mathbb{K}} = \mathcal{K}$. Thus we have the canonical identification $\widetilde{\mathbb{K}^n} = \widetilde{\mathbb{K}}^n = \mathcal{K}^n$. For $\widetilde{\mathbb{K}}_c$ we write $\mathcal{K}_c$.

**3.36 Proposition and Definition.** Let $U \in \mathcal{G}(\Omega)$ and $\widetilde{x} \in \widetilde{\Omega}_c$. Then the *generalized point value of $U$ at* $\widetilde{x} = \mathrm{cl}[(x_\varepsilon)_\varepsilon]$, defined by

$$U(\widetilde{x}) := \mathrm{cl}[(u_\varepsilon(x_\varepsilon))_\varepsilon]\,,$$

is a well-defined element of $\mathcal{K}$.

**Proof.** If $\widetilde{x} = \mathrm{cl}[(x_\varepsilon)_\varepsilon] \in \widetilde{\Omega}_c$, there exists some $K \subset\subset \Omega$ such that $x_\varepsilon \in K$ for $\varepsilon$ small. Since $U \in \mathcal{G}(\Omega)$ it follows that $|u_\varepsilon(x_\varepsilon)| \leq \sup\limits_{x \in K} |u_\varepsilon(x)| \leq \varepsilon^{-p}$ for small $\varepsilon$, so $(u_\varepsilon(x_\varepsilon))_\varepsilon \in \mathcal{K}$. Next we show that $\widetilde{x} \sim \widetilde{y}$ implies $U(\widetilde{x}) \sim U(\widetilde{y})$. We have

$$|u_\varepsilon(x_\varepsilon) - u_\varepsilon(y_\varepsilon)| \leq |x_\varepsilon - y_\varepsilon| \int\limits_0^1 |\nabla u_\varepsilon(x_\varepsilon + \sigma(y_\varepsilon - x_\varepsilon))| d\sigma\,. \tag{66}$$

Since $x_\varepsilon + \sigma(y_\varepsilon - x_\varepsilon)$ remains within some compact subset of $\Omega$ for small $\varepsilon$ the second factor is bounded by some $\varepsilon^{-p}$. The first factor eventually is smaller than any $\varepsilon^q$ proving the claim. Finally, if $(w_\varepsilon)_\varepsilon \in \mathcal{N}(\Omega)$ then $(w_\varepsilon(x_\varepsilon))_\varepsilon \sim 0$ again because $x_\varepsilon$ stays within some compact set for $\varepsilon$ small. □

**3.37 Theorem.** *If $\Omega$ is an open subset of $\mathbb{R}^n$ then*

$$U = 0 \text{ in } \mathcal{G}(\Omega) \ \Leftrightarrow\ U(\widetilde{x}) = 0 \text{ in } \mathcal{K} \text{ for all } \widetilde{x} \in \widetilde{\Omega}_c.$$

**Proof.** The implication from left to right was already shown in the proof of 3.36. For the technical proof of the converse we refer the reader to [117], thm. 1.4.9 resp. [145], thm. 2.4. □

### 3.D. Integration of Generalized Functions

This section is devoted to another basic building block of the theory of generalized functions, namely integration. Technically this concept arises by transferring componentwise operations to the level of equivalence classes. The significance of integration in $\mathcal{G}$ is that it opens the possibility to describe distributions in a very elegant and explicit way within the generalized framework pointing directly in the direction of the fundamental concept of association which will be discussed in the next section.

**3.38 Definition.** *Let $M$ be a Lebesgue-measurable set with its closure compact and contained in $\Omega$. The* integral of the generalized function *$U \in \mathcal{G}(\Omega)$ is defined by*

$$\int\limits_M U(x)\,dx \;:=\; \left(\int\limits_M u_\varepsilon(x)\,dx\right)_\varepsilon + \mathcal{N}\,.$$

The integral of $U$ indeed is a well-defined element of $\mathcal{K}$ since on $\overline{M}$ we have $|u_\varepsilon| = O(\varepsilon^{-p})$ implying $\int_M u_\varepsilon(x)dx \in \mathcal{E}$. Moreover, if $u_\varepsilon \in \mathcal{N}(\Omega)$ the $\mathcal{N}$-estimate for $\int_M u_\varepsilon(x)dx \in \mathcal{E}$ follows directly.

Next we collect the basic properties of the integral in $\mathcal{G}$ which follow easily from the definition.

**3.39 Proposition.** *Let $M$, $M_1$, $M_2$ be Lebesgue-measurable sets with compact closures in $\Omega$, $U, V \in \mathcal{G}(\Omega)$ and $\alpha \in \mathcal{K}$. Then, denoting the Lebesgue-measure on $\mathbb{R}^n$ by $\lambda$, we have*

*(i)* $\int_M (U + \alpha V) = \int_M U + \alpha \int_M V$.

*(ii) If* $\lambda(M) = 0$ *then* $\int_M U = 0$.

*(iii) If* $\lambda(M_1 \cap M_2) = 0$ *then* $\int_{M_1 \cup M_2} U = \int_{M_1} U + \int_{M_2} U$.

*(iv) If* $\overline{M} \subseteq \Omega' \subseteq \Omega$ *then* $\int_M U = \int_M (U|_{\Omega'})$.

*(v) If* $f \in \mathcal{C}^\infty(\Omega)$ *then* $\int_M \iota(f) = \sigma(\int_M f)$ *in* $\mathcal{K}$.

□

Also the fundamental theorem of calculus in $\mathcal{G}$ is an easy consequence of our definitions.

**3.40 Theorem.** *Let $U \in \mathcal{G}(J)$ with $J$ some open interval of the real line. Setting*

$$I(U)(x) \;:=\; \left(\int\limits_{x_0}^{x} u_\varepsilon(s)\,ds\right)_\varepsilon + \mathcal{N} \tag{67}$$

*we have in $\mathcal{G}(J)$*

$$I(U)' \;=\; U \qquad \textit{and} \qquad U \;=\; U(x_0) + I(U')\,.$$

**Proof.** Clearly (67) defines an element in $\mathcal{G}(J)$ with derivative $U$. The result hence follows from 3.31. □

**3.41 Example.** $\int \delta$.
Let $K = [-a, a]$ $(a \neq 0)$ then

$$\int\limits_K \iota(\delta)(x)\,dx \;=\; \left(\int\limits_{-a}^{a} \rho_\varepsilon(x)\,dx\right)_\varepsilon + \mathcal{N} \;=\; 1 \quad \in \mathcal{C},$$

since for arbitrary positive $r$ we have $|1 - \int_{-a}^{a} \rho_\varepsilon(x)dx| \;=\; |\int_{x \geq a/\varepsilon} \rho(x)dx| \;\leq\; C\varepsilon^r$. Hence—in a sloppy notation—we have $\int_K \delta = 1$, as expected.

However, setting $K = \operatorname{supp}(\delta) = \{0\}$ in the above example we get $\int_K \delta = 0$. Actually the existence of generalized functions with support in isolated points forces us to employ the following technical definition for compactly supported generalized functions (instead of defining it just to be the integral over its support).

**3.42 Definition.** *Let $U \in \mathcal{G}(\Omega)$ compactly supported. We define the* integral of $U$ over $\Omega$ *by*

$$\int\limits_{\Omega} U(x)\,dx \;:=\; \int\limits_{K} U(x)\,dx\,,$$

*where $K$ is any compact subset of $\Omega$ containing the support of $U$ in its interior.*

Obviously the above definition does not depend on the choice of $K$. Next we look at some more examples

**3.43 Examples.**

(i) $\int \delta$ revisited. The above definition is sufficient to guarantee

$$\int\limits_{-\infty}^{\infty} \iota(\delta)(x)\,dx \;=\; 1\,.$$

(ii) $\int \delta^2$

$$\int\limits_{-\infty}^{\infty} \iota(\delta)^2(x)\,dx \;=\; \left(\int\limits_{-\infty}^{\infty} \rho_\varepsilon^2(x)\,dx\right)_\varepsilon + \mathcal{N} \;=\; \left(\frac{1}{\varepsilon}\int\limits_{-\infty}^{\infty} \rho^2(x)\,dx\right)_\varepsilon + \mathcal{N}$$

Hence $\int \delta^2$ will depend on the chosen mollifier.

Classical rules of calculus like integration by parts can also be "translated" to the generalized framework.

**3.44 Proposition.** *Let $U, V \in \mathcal{G}(\Omega)$ and at least one of them compactly supported. Then we have for any order of derivative $\alpha \in \mathbb{N}_0^n$*

$$\int_\Omega U(x)\partial^\alpha V(x)\,dx = (-1)^{|\alpha|}\int_\Omega \partial^\alpha U(x)V(x)\,dx\,.$$

**Proof.** It suffices to set $\partial^\alpha = \partial_{x_1}$. Without loss of generality assume $U$ to have compact support $K$ and choose some $\chi \in \mathcal{D}(\Omega)$ satisfying $\chi \equiv 1$ in some neighborhood of $K$ and $\operatorname{supp}(\chi) \subseteq \Omega' \subset \overline{\Omega'} \subset\subset \Omega$. Then $\sigma(\chi)U = U$ in $\mathcal{G}(\Omega)$, so we obtain a representative $(u_\varepsilon)_\varepsilon$ of $U$ with $\operatorname{supp}(u_\varepsilon) \subseteq \Omega'$ for each $\varepsilon$. Setting $L := \overline{\Omega'}$, a representative of $\int_\Omega U(x)\partial_{x_1}V(x)\,dx$ is given by

$$(\int_L u_\varepsilon(x)\partial_{x_1}v_\varepsilon(x)\,dx)_\varepsilon = (-\int_L \partial_{x_1}u_\varepsilon(x)v_\varepsilon(x)\,dx)_\varepsilon,$$

which is a representative of $-\int_\Omega \partial_{x_1}U(x)V(x)\,dx$ as well. □

As already indicated at the very beginning of this section we have the following natural description of distributions in the generalized functions framework.

**3.45 Theorem.** *If $w \in \mathcal{D}'(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$ then*

$$\int_\Omega \iota(w)(x)\iota(\varphi)(x)\,dx = \sigma(\langle w, \varphi\rangle) \text{ in } \mathcal{K}.$$

**Proof.** Choose some partition of unity $(\chi_j)_j$ subordinate to some open covering $(\Omega_\lambda)_\lambda$ of $\Omega$ and a family of cut-off functions $(\psi_\lambda)_\lambda$ as in 3.14. Further set $K = \operatorname{supp}(\varphi)$ and choose a compact neighborhood $L$ of $K$ in $\Omega$. If $M \in \mathbb{N}$ is such that $\sum_{j=1}^M \chi_j \equiv 1$ on $L$ then

$$\iota(w)\iota(\varphi) = \sum_{j=1}^{M} \iota(w)\iota(\chi_j\varphi) \quad \text{and} \quad \langle w, \varphi\rangle = \sum_{j=1}^{M} \langle w, \chi_j\varphi\rangle\,.$$

It is therefore enough to show that $\int_\Omega \iota(w)\iota(\chi_j\varphi) = \langle w, \chi_j\varphi\rangle$ for $1 \le j \le M$. This means that without loss of generality we may suppose $K \subset L \subset \Omega_{\lambda_j}$. Now $\iota_{\lambda_j}(w) = \iota(w)|_{\Omega_{\lambda_j}}$ and 3.39 (iv) imply

$$\int_\Omega \iota(w)(x)\iota(\varphi)(x)\,dx = (\int_L ((\psi_{\lambda_j} w) * \rho_\varepsilon)(x)\varphi(x)\,dx)_\varepsilon + \mathcal{N}\,.$$

Observing $\langle w, \varphi\rangle = \langle \psi_{\lambda_j} w, \varphi\rangle$ we get

$$\begin{aligned}
\int_L ((\psi_{\lambda_j} w) * \rho_\varepsilon)(x)\varphi(x)\,dx - \langle w, \varphi\rangle &= \langle \psi_{\lambda_j} w, \check{\rho}_\varepsilon * \varphi - \varphi\rangle \\
&= \langle \psi_{\lambda_j} w(x), \int \rho_\varepsilon(y)(\varphi(x+y) - \varphi(x))\,dy\rangle \\
&= \langle \psi_{\lambda_j} w(x), \int \rho(y)(\varphi(x+\varepsilon y) - \varphi(x))\,dy\rangle. \qquad (68)
\end{aligned}$$

By 3.11 (ii), it will suffice to consider $\psi_{\lambda_j} w = \partial^\alpha f$ for some $f \in \mathcal{C}(\Omega)$ with compact support. But then the last expression in (68) equals

$$(-1)^{|\alpha|} \int f(x) \int (\partial^\alpha \varphi(x+\varepsilon y) - \partial^\alpha \varphi(x))\rho(y)\,dy\,dx = O(\varepsilon^q)$$

for arbitrary $q > 0$ as can be seen from Taylor expansion analogous to 3.12. □

**3.46 Corollary.** *If $w \in \mathcal{D}'(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$ then*

$$\lim_{\varepsilon \to 0} \int \iota(w)_\varepsilon(x)\varphi(x)\,dx = \langle w, \varphi\rangle.$$

**Proof.** By 3.45 there exists some $[(n_\varepsilon)_\varepsilon]$ in $\mathcal{N}$ such that $\int \iota(w)_\varepsilon(x)\varphi(x)\,dx = \langle w, \varphi\rangle + n_\varepsilon$ for all $\varepsilon \in I$. □

3.46 already points into the direction of one of the main attractions of Colombeau algebras: the concept of association. The next section is devoted to a detailed study of this notion.

### 3.E. Nonlinear Modelling: Association and Coupled Calculus

So far we have constructed an associative and commutative differential algebra $\mathcal{G}$ of generalized functions containing distributions as a subspace and smooth functions as a faithful subalgebra, as well as providing maximal consistency properties with respect to classical operations (i.e., properties (i)-(iii) in 2.10 and (iv") in 2.13). Thereby—forced by L. Schwartz's impossibility result—we had to introduce a very narrow concept of equality forcing us to give up such well established distributional formulas like e.g. $x\delta = 0$ (cf. ex. 3.25 (ii)) and seemingly reasonable "multiplication rules" for distributions like $H^n = H$ or $H\delta = (1/2)\delta$. How can we recover (the remnants of) such relations—that after all have proved their usefulness in certain situations—in our rigorous framework where we are able to put them into the place they deserve?
This question is closely related to the fact that in general properties of singular objects in a nonlinear theory depend on the regularization, i.e., on our mollifier $\rho$, a fact which was already remarked in 3.23. To illustrate this more explicitly recall that $\rho$ totally determines the properties of $(\iota(\delta))_\varepsilon = \rho_\varepsilon$. However, there are many objects in $\mathcal{G}$ that, although being different from $\iota(\delta)$, converge to the $\delta$-distribution in the limit as $\varepsilon$ tends to zero. The same of course holds true—and may be most easily explained—for the Heaviside step function. Since in $\mathcal{D}'$ any power $H^n$ equals $H$, all these functions may be regarded as *equivalent in the linear theory.* However, by 3.33 (iv) we have $H^n \neq H$ in $\mathcal{G}(\mathbb{R})$; hence the powers of $H$ all display *different properties in a nonlinear context.*
This in turn has immediate consequences in applications, especially when modelling a, say, physical situation (c.f. also 5.14). Consider e.g. the example of 2.15, i.e., a charge density contained in a "small region" in space with negligible internal structure, which often is described adequately by a point charge. The latter, however, is a very highly idealized physical object, described by the $\delta$-distribution, a mathematical object with very little structure in the sense that it is completely banished to the linear regime; $\mathcal{D}'$ as a vector space is "blind" for any nonlinearities. To apply any nonlinear operations to the $\delta$-distribution our method of choice is to embed it into $\mathcal{G}$. This of course amounts to represent the physical object by a mathematical object with a much richer structure than the $\delta$-distribution. This "additional" structure (we might call it "nonlinear information")—in our case encoded into the mollifier $\rho$—has to be determined by a closer look at the physical object one wants to describe. This is what we mean by *nonlinear modelling.*
Summing up in $\mathcal{G}$ the whole so-called "nonlinear information" of the sequences $(u_\varepsilon)_\varepsilon$ (up to the ideal $\mathcal{N}$) is memorized; in the case of a distribution the "nonlinear information" of the regularization process is stored in $\iota(w)$. On the other hand in the classical $\mathcal{D}'$-picture one is only interested in the weak limit throwing away this additional information. Technically this is reflected by the fact that $\mathcal{N}$ is much smaller than $\mathcal{V}_0 \cap \mathcal{E}_M$.
In this section we are going to introduce a technical tool to identify in $\mathcal{G}$ nonlinearly distinct yet linearly equivalent objects; this is done by introducing an equivalence relation in $\mathcal{E}_M$ coarser than equality in $\mathcal{G}$.

**3.47 Definition.** *A generalized function $U \in \mathcal{G}(\Omega)$ is called* associated to 0 *, $U \approx 0$, if it has a representative $(u_\varepsilon)_\varepsilon$ belonging to $\mathcal{E}_M(\Omega) \cap \mathcal{V}_0$ (see 3.1), i.e.,*

$$\langle u_\varepsilon, \varphi \rangle \;\to\; 0 \qquad (\varphi \in \mathcal{D}(\Omega)). \tag{69}$$

Since for $(u_\varepsilon)_\varepsilon \in \mathcal{N}$, $u_\varepsilon \to 0$ uniformly on compact sets, $\mathcal{N}$ is a subset of $\mathcal{E}_M(\Omega) \cap \mathcal{V}_0$. Thus if $U$ belongs to the linear space $\mathcal{N}_\approx(\Omega) := \{U \in \mathcal{G}(\Omega) : U \approx 0\}$, every representative of $U$ satisfies (69). The equivalence relation

$$U \approx V :\Leftrightarrow U - V \approx 0 \tag{70}$$

($U, V \in \mathcal{G}(\Omega)$ are *associated to each other*) gives rise to the classes of the linear quotient space $\mathcal{G}(\Omega)/\mathcal{N}_\approx(\Omega)$. That $\approx$ is non-trivial on $\mathcal{G}(\Omega)$ (i.e., $\mathcal{N}_\approx \neq \{0\}$) will be clearly demonstrated by a number of examples below. However, $\mathcal{N}_\approx(\Omega)$ is not an ideal in $\mathcal{G}(\Omega)$, thus $\mathcal{G}(\Omega)/\mathcal{N}_\approx(\Omega)$ is not an algebra.
By 3.46, each association class contains at most one distribution $\iota(w)$ (resp. $w$). We define

**3.48 Definition.** *Let $U$ in $\mathcal{G}(\Omega)$ and $w \in \mathcal{D}'(\Omega)$. If $U \approx \iota(w)$ we say that $U$ admits $w$ as* associated distribution *and call $w$ the* distributional shadow *(or* macroscopic aspect*) of $U$. In that case we also write $U \approx w$.*

**3.49** The shadow of a generalized function—if it exists at all—is unique. Of course, there do exist $U \in \mathcal{G}$ not admitting any associated distribution. Such generalized functions may be called *vampires*[7]. Take for example the square of the $\delta$-distribution, then

$$\int \iota(\delta)^2(x)\varphi(x)\,dx = \frac{1}{\varepsilon^2}\int \rho^2(y)\varphi(\varepsilon y)\,dy \to \infty\,,$$

provided that $\varphi(0)$ is non-vanishing.

Association in $\mathcal{G}$ induces a equivalence relation on the ring of constants $\mathcal{K}$ again denoted by $\approx$. More precisely we have

**3.50 Definition.** *We call two generalized numbers $r, s$ associated to each other, $r \approx s$, if $r_\varepsilon - s_\varepsilon \to 0$ for one (hence any) pair of representatives $(r_\varepsilon)_\varepsilon$, $(s_\varepsilon)_\varepsilon$. If $\alpha \in \mathbb{K}$ and $r \approx \sigma(\alpha)$ then $r$ is said to admit $\alpha$ as* associated number *and $\alpha$ is called the* shadow *(or* macroscopic aspect*) of $r$.*

Let us now take a look at some typical examples

**3.51 Examples.**

(i) $x\delta$ revisited.
Here we add one more point to our by now quite broad knowledge on the product of $x$ with $\delta$. In ex. 3.25 (ii) we have already seen that $x\delta \neq 0$ in $\mathcal{G}(\mathbb{R})$ although—of course—$x\delta = 0$ in $\mathcal{D}'(\mathbb{R})$. Furthermore in ex. 3.33 (iv) we have proved that all point values of $x\iota(\delta)$ vanish. Now we show that $x\delta \approx 0$. Indeed, let $\varphi \in \mathcal{D}(\mathbb{R})$, then

$$\int x\rho_\varepsilon(x)\varphi(x)\,dx = \varepsilon\int y\rho(y)\varphi(\varepsilon y)\,dx \to 0\,.$$

(ii) The powers of the Heaviside function.
By 3.25 (iv) we know that although $H^n = H$ for all $n$ in $L_{\text{loc}} \subseteq \mathcal{D}'$, $H^n \neq H$ in $\mathcal{G}(\mathbb{R})$. As in the previous example we are going to recover the distributional result in the sense of association. Since by 3.24 $\iota(H)_\varepsilon = H * \rho_\varepsilon$ we have

$$\begin{aligned}
\int (\iota H)^n_\varepsilon(x)\varphi(x)\,dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{x/\varepsilon}\rho(y_1)\,dy_1\ldots\int_{-\infty}^{x/\varepsilon}\rho(y_n)\,dy_n\varphi(x)\,dx\\
&= \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\rho(y_1)\ldots\rho(y_n)\int_{\max\limits_{1\le i\le n}(\varepsilon y_i)}^{\infty}\varphi(x)\,dx\,dy_1\ldots dy_n\\
&\to \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\rho(y_1)\ldots\rho(y_n)\int_0^{\infty}\varphi(x)\,dx\,dy_1\ldots dy_n\\
&= \langle H, \varphi\rangle.
\end{aligned}$$

The above examples already anticipate some of the general properties of association given in the following

[7]It is well known that vampires do not have a mirror image. However, following [75] and [171] we assume that they—just like Peter Pan—do not even cast a shadow.

**3.52 Theorem.**

*(i) The product* $\mathcal{C}^\infty \times \mathcal{D}'$ *is compatible with association, i.e.,* $\iota(f)\iota(w) \approx \iota(fw)$ *for all* $f \in \mathcal{C}^\infty(\Omega), w \in \mathcal{D}'(\Omega)$.

*(ii) The product of continuous functions is compatible with association, i.e.,* $\iota(f)\iota(g) \approx \iota(fg)$ *for all* $f, g \in \mathcal{C}(\Omega)$.

*(iii) Point values of continuous functions are associated to their classical point values, i.e.,* $\iota(f)(x_0) \approx \sigma(f(x_0))$ *for all* $f \in \mathcal{C}(\Omega)$, $x_0 \in \Omega$.

*(iv) Composition of smooth functions is compatible with association, i.e., let* $g : \Omega_1 \to \Omega_2$ *be smooth and* $f \in \mathcal{C}(\Omega_2)$ *then* $\iota(f \circ g) \approx \iota(f) \circ g$.

*(v) Distributional pullback commutes with embedding in the sense of association, i.e., let* $g : \Omega_1 \to \Omega_2$ *be a diffeomorphism and* $w \in \mathcal{D}'(\Omega_2)$ *then* $\iota(g^*(w)) \approx \iota(w) \circ g$.

**Proof.** To prove (i) let $\varphi$ be a test function on $\Omega$ with $\mathrm{supp}(\varphi) = \mathrm{K}$. Now choose $(\Omega_\lambda)_\lambda$, $(\psi_\lambda)_\lambda$ and $(\chi_j)_j$ as in 3.14 and some compact neighborhood $L$ of $K$ in $\Omega$. If $M \in \mathbb{N}$ is such that $\sum_{j=1}^M \chi_j \equiv 1$ on $L$ then it is sufficient to show that

$$\lim_{\varepsilon \to 0} \int_L f(x)\iota(w)_\varepsilon(x)\varphi(x)\chi_j(x)\,dx = \lim_{\varepsilon \to 0} \int_L \iota(fw)_\varepsilon(x)\varphi(x)\chi_j(x)\,dx.$$

for $1 \le j \le M$. We further may assume with out loss of generality that $K \subseteq L \subseteq \Omega_{\lambda_j}$ for some $j$. Now

$$\int_L \iota(w)_\varepsilon(x)f(x)\varphi(x)\,dx = \int (\psi_{\lambda_j} w) * \rho_\varepsilon(x)f(x)\varphi(x)\,dx \;=\; \langle (\psi_{\lambda_j} w) * \rho_\varepsilon, f\varphi\rangle \;\to\; \langle \psi_{\lambda_j} w, f\varphi\rangle \;=\; \langle w, f\varphi\rangle\,.$$

On the other hand

$$\int_L \iota(fw)_\varepsilon(x)\varphi(x)\,dx = \int (\psi_{\lambda_j} fw) * \rho_\varepsilon(x)\varphi(x)\,dx \;=\; \langle (\psi_{\lambda_j} fw) * \rho_\varepsilon, \varphi\rangle \;\to\; \langle \psi_{\lambda_j} fw, \varphi\rangle \;=\; \langle w, f\varphi\rangle$$

proving our claim.
(ii) By the same arguments as in (i), for $\varphi \in \mathcal{D}(\Omega)$ with $\mathrm{supp}(\varphi) = K \subseteq L \subseteq \Omega_{\lambda_j}$ we have to show that

$$\lim_{\varepsilon \to 0} \int_L \iota(f)_\varepsilon(x)\iota(g)_\varepsilon(x)\varphi(x)\,dx = \lim_{\varepsilon \to 0} \int_L \iota(fg)_\varepsilon(x)\varphi(x)\,dx\,. \tag{71}$$

The left hand side of (71) is

$$\lim_{\varepsilon \to 0} \int ((\psi_{\lambda_j} f) * \rho_\varepsilon)(x)((\psi_{\lambda_j} g) * \rho_\varepsilon)(x)\varphi(x)\,dx = \int (\psi_{\lambda_j} f)(x)(\psi_{\lambda_j} g)(x)\varphi(x)\,dx \;=\; \int f(x)g(x)\varphi(x)\,dx\,,$$

where we have used Lebesgue's dominated convergence theorem and $\|(\psi_{\lambda_j} f) * \rho_\varepsilon\|_\infty \;\le\; \|\psi_{\lambda_j} f\|_\infty \|\rho_\varepsilon\|_1$. Analogously, the right hand side of (71) equals $\int f(x)g(x)\varphi(x)\,dx$.
(iii) There exists some $M \in \mathbb{N}$ such that

$$\lim_{\varepsilon \to 0} \iota(f)_\varepsilon(x_0) = \lim_{\varepsilon \to 0} \sum_{j=1}^M \chi_j(x_0)((\psi_{\lambda_j} f) * \rho_\varepsilon)(x_0) \;=\; \sum_{j=1}^M \chi_j(x_0)(\psi_{\lambda_j} f)(x_0) \;=\; \sum_{j=1}^M \chi_j(x_0) f(x_0) \;=\; f(x_0).$$

For the proof of (iv) and (v), observe that by the same reasoning as in (i) and (ii) it suffices to demonstrate the claims under the simplifying assumption that $f$ and $w$ are compactly supported.
(iv) For $\varphi \in \mathcal{D}(\Omega_1)$ we have

$$\int (f \circ g) * \rho_\varepsilon(x)\varphi(x)\,dx \;\to\; \int (f \circ g)(x)\varphi(x)\,dx$$

as well as

$$\int f * \rho_\varepsilon(g(x))\varphi(x)\,dx \;\to\; \int (f \circ g)(x)\varphi(x)\,dx$$

for $\varepsilon \to 0$, so the result follows.
(v) For $\varepsilon \to 0$ we have

$$\int \rho_\varepsilon * (g^*(w))(x)\varphi(x)\,dx \;\to\; \langle g^*(w), \varphi\rangle \;=\; \langle w(y), \varphi(g^{-1}(y))|\det(D(g^{-1})(y))|\rangle,$$

so the claim follows from

$$\int (\rho_\varepsilon * w)(g(x))\varphi(x)\,dx \;=\; \int (\rho_\varepsilon * w)(y)\varphi(g^{-1}(y))|\det(D(g^{-1})(y))|\,dy.$$

□

It is easily seen from the examples of this section that none of the above items will generally hold in the sense of equality in $\mathcal{G}$. Note that item (ii) above is the reconciliation of items (i)-(iii) and (iv') of 2.10 on the level of association. However, compatibility of association with differentiation and multiplication with smooth functions is guaranteed by the next

**3.53 Proposition.** *Let $U, V \in \mathcal{G}(\mathbb{R})$ with $U \approx V$, then*

*(i) $\partial^\alpha U \approx \partial^\alpha V$ for all orders of derivative $\alpha \in \mathbb{N}_0$.*

*(ii) $\iota(f)U \approx \iota(f)V$ for all smooth functions $f \in \mathcal{C}^\infty(\Omega)$.*

**Proof.** To prove (i) it suffices to show that $U \approx 0 \Rightarrow \partial^\alpha U \approx 0$ for all $\alpha \in \mathbb{N}_0$. This, however is an immediate consequence of

$$\int \partial^\alpha u_\varepsilon(x)\varphi(x)\,dx \;=\; (-1)^{|\alpha|}\int u_\varepsilon(x)\partial^\alpha\varphi(x)\,dx$$

for all test functions $\varphi$.
(ii) is proved by the observation that if $U \approx 0$ then by 3.17 we have

$$\lim_{\varepsilon\to 0}\int \iota(f)(x)\,u_\varepsilon(x)\,\varphi(x)\,dx \;=\; \lim_{\varepsilon\to 0}\int u_\varepsilon(x)\,(f\varphi)(x)\,dx\,.$$

□

In addition to the obvious compatibility of association with vector space operations proposition 3.53 gives the only basic differential algebraic operations compatible with $\approx$ as is clearly demonstrated by ex. 3.51 (ii) as well as the first of the examples below.

**3.54 Examples.**

(i) $x\delta \approx 0$ in $\mathcal{G}(\mathbb{R})$ by 3.51 (i), but $x\delta^2 \not\approx 0$. Indeed for $\varphi \in \mathcal{D}'(\mathbb{R})$ with $\varphi(0)$ non-vanishing we have

$$\int \iota(x)_\varepsilon\iota(\delta)^2_\varepsilon(x)\,dx \;=\; \frac{1}{\varepsilon^2}\int x\rho^2(\frac{x}{\varepsilon})\varphi(x)\,dx \;=\; \int y\rho^2(y)\varphi(\varepsilon y)\,dx \;\not\to\; 0\,.$$

(ii) $x\delta' \approx -\delta$, since from 3.51 (i) and the above prop. (i) we have $0 \approx (x\delta)' = \delta + x\delta'$.

(iii) The product of the Heaviside function with the Dirac-$\delta$.
From ex. 3.51 (ii) we have in particular $H^2 \approx H$ implying

$$H\delta \;\approx\; \frac{1}{2}\,\delta\,. \tag{72}$$

(iv) equation (72) above together with proposition 3.53 (i) implies $\delta^2 \approx (\frac{1}{2} - H)\delta'$.

**3.55** Despite its simplicity and rigor the above method of producing multiplication rules in the sense of association may not be well suited in all cases. The assignment of $(1/2)\delta$ to the product $H\delta$ in many cases will be too rigid to sensibly model a realistic situation. However, the technical apparatus of association immediately provides us with a solution to that problem. Note that $H\delta \approx (1/2)\delta$ just means that $\iota(H)\iota(\delta) \approx (1/2)\delta$, but there are many more generalized functions modelling a jump respectively a peak than just the canonical images of $H$ and $\delta$ under the embedding. Take e.g. $H^2$ which of course models a jump in a way different from the way $H$ does. From $H^3 \approx H$ one immediately derives

$$H^2\delta \;\approx\; \frac{1}{3}\,\delta\,.$$

Hence the statement that "in the Colombeau algebra Heaviside times $\delta$ is just $(1/2)\delta$" is—to say the least—misleading. On the contrary, association enables us to model Heaviside times $\delta$ in a large number of varieties. This possibility to model nonlinear properties of singular objects at an infinitesimal level may be regarded as a nonstandard aspect of the Colombeau algebra (see [144], §23 for details). This freedom to switch between a *macroscopic, linear description (association)* and a *microscopic, nonlinear one (equality in $\mathcal{G}$)* is best expressed in the term "coupled calculus."

**3.56** Association is not the only possible equivalence relation in $\mathcal{G}$ coinciding with equality in distributions. Two generalized functions $U, V$ are called *test equivalent* $U \sim V$ if

$$\int\limits_{\Omega} (U - V)(x)\,\varphi(x)\,dx \;=\; 0 \qquad \forall \varphi \in \mathcal{D}(\Omega)\,.$$

Obviously $\sim$ is a well-defined equivalence relation on $\mathcal{G}$. Moreover $U \sim V$ implies $U \approx V$ placing $\sim$ between equality in $\mathcal{G}$ and association. $\sim$ shares most of the properties of $\approx$ but is not compatible with the product of continuous functions, which is why we do not go into any further details here.
For many applications it is appropriate to introduce still different notions of association (see also chapter 4). One example is $L^\infty_{\mathrm{loc}}$-association defined as follows: $U$ is called $L^\infty_{\mathrm{loc}}$*-associated to* 0 if for one (hence any) representative $u_\varepsilon \to 0$ in $L^\infty_{\mathrm{loc}}$.
Depending on the practical situation at hand various concepts of association have proved to be a very flexible and useful tool. For a general investigation of association relations in $\mathcal{G}$ we refer the reader again to Rosinger [165].

# 4. COLOMBEAU ALGEBRAS ON MANIFOLDS

In this chapter we present a theory of generalized functions on manifolds as well as generalized sections of vector bundles providing a framework for a nonlinear distributional geometry. Since the main applications of Colombeau algebras so far have been in the field of nonlinear PDEs with singular data or coefficients a theory on open sets of $\mathbb{R}^n$ was a sufficient tool. Only recently applications in Lie group analysis (cf. [120]) and mainly general relativity (for an overview see [186]) have pointed out the need for a theory of generalized tensor fields on manifolds.

While the definition of Colombeau algebras on manifolds is a rather straightforward matter, the task of constructing a diffeomorphism-invariant embedding of distributions is a much harder problem giving rise to some controversy in the literature. In fact in our present setting, i.e., the special algebra, a *geometric* embedding of $\mathcal{D}'$ (i.e., independent of a chosen atlas) is not possible (although—of course—$\mathcal{D}'$ is injectively included in $\mathcal{G}$; [50]). The core of the problem is the (apparent) incompatibility of convolution (defining the embedding) and diffeomorphism invariance. There are, however, two ways to cope with this situation.

The first (radical) option is to dispense altogether with a canonical embedding of distributions and to use coupled calculus to make statements on the distributional level only. Such a theory is indeed practical and useful in applications as we shall see in the next chapter (cf. also [50,22]). The main idea is not to embed singular, i.e., distributional data into the algebra via some "canonical procedure" but to use a generalized function associated (probably in the sense of some relation of the kind discussed in 3.56) to the given object. Of course, this choice is not unique but has to be justified by additional input from, say, a physical model. Moreover, in some cases one might be able to prove a result (e.g. existence and uniqueness of the solutions to a differential equation) for a certain class of generalized data associated to a certain distribution (cf. chapter 5). Let us stress again that from the viewpoint of applications it is anything but clear that the "canonical" embedding via convolution really *is* canonical in the sense that it automatically models the singular object in the right way. (cf. also section 3.E on nonlinear modelling).

The second possibility—if one is not willing to follow the path indicated above—is to change the definition of the algebra in such a way as to allow for a canonical embedding of distributions (again given by convolution not with a single mollifier but with a suitable class of bounded paths $\Phi(\varepsilon, x) \in \mathcal{D}$). The main idea is to change the parameter set from $(0, 1]$ to a certain class of mollifiers $\mathcal{A}_0$ and to define generalized functions to be (classes of moderate modulo negligible) elements in the space of all $\mathcal{C}^\infty$-functions $R : \mathcal{A}_0 \times \mathbb{R}^n \to \mathbb{C}$ (cf. [99,61,76,77]). As may be seen immediately from the last notion the (technical) prize to pay for such a construction is the use of calculus in infinite dimensional spaces. Nevertheless very recent results have shown that in this approach it is indeed possible in a way to isolate the diffeomorphism invariant essence of the convolution based regularization process which then may be formulated intrinsically in geometrical terms [77].

In the first section below we are going to introduce definitions and basic properties of the (special version of the) algebra of generalized functions on a differentiable manifold. We also discuss in detail the problems arising in the context of a "canonical" embedding of distributions in this setting. In the following section we present (again: the special version of) the theory of generalized sections in arbitrary vector bundles, thereby further pursuing option one above. Specializing to tensor fields most results of classical tensor analysis carry over to the new setting which essentially consists of parameterized (componentwise) classical calculus. In sec. 4.C we finally set the stage for the applications to be presented in the last chapter by defining the curvature entities for a generalized metric.

On the other hand, the final section of the present chapter is devoted to an overview of the various variants of the theory of generalized functions on manifolds allowing for a canonical embedding of distributions. A fully detailed discussion of this topic which is still an area of current development, however, lies beyond the scope of this work and we are going to mainly refer to recent literature.

### 4.A. Generalized Functions on Manifolds

Throughout this chapter we shall use the same notational conventions on differential geometry as in chapter 1 (cf. also [53,1]). In particular, we always denote by $X$ a paracompact, smooth Hausdorff manifold of dimension $n$. To begin with, note that given a smooth function $f : \tilde{\Omega} \to \Omega$ between open subset of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, and a generalized function $U \in \mathcal{G}(\Omega)$, the pullback of $U$ by $f$

$$f^*(U) \;\equiv\; U \circ f \;:=\; \mathrm{cl}[(u_\varepsilon \circ f)_\varepsilon]$$

is a well-defined element of $\mathcal{G}(\tilde{\Omega})$. If in addition $g : \tilde{\tilde{\Omega}} \to \tilde{\Omega}$ is a smooth map as well we have $U \circ (f \circ g) = (U \circ f) \circ g$ in $\mathcal{G}(\Omega)$. These properties already enable us to define the algebra of generalized functions on $X$ by imposing the growth conditions in each chart (cf. [117], 1.9.18). However, we have chosen to give an equivalent intrinsic description of the space of moderate and negligible sequences, respectively (denoting the Lie derivative with respect to a vector field $\xi$ by $L_\xi$)

**4.1 Definition.** *Let $X$ be paracompact $\mathcal{C}^\infty$-manifold (and recall that $I = (0,1]$).*

$$\begin{aligned}
\mathcal{E}(X) &:= (\mathcal{C}^\infty(X))^I \\
\mathcal{E}_M(X) &:= \{(u_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}(X) : \forall K \subset\subset X,\, \forall k \in \mathbb{N}_0\, \exists N \in \mathbb{N}\, \forall \xi_1,\ldots,\xi_k \in \mathcal{T}^1_0(X) : \\
&\qquad \sup_{p\in K} |L_{\xi_1} \ldots L_{\xi_k}\, u_\varepsilon(p)| = O(\varepsilon^{-N}) \text{ as } \varepsilon \to 0\} \\
\mathcal{N}(X) &:= \{(u_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}(X) : \forall K \subset\subset X,\, \forall k,l \in \mathbb{N}_0\, \forall \xi_1,\ldots,\xi_k \in \mathcal{T}^1_0(X) : \\
&\qquad \sup_{p\in K} |L_{\xi_1} \ldots L_{\xi_k}\, u_\varepsilon(p)| = O(\varepsilon^{l})) \text{ as } \varepsilon \to 0\}
\end{aligned}$$

*The* Colombeau algebra of generalized functions on the manifold $X$ *is defined as the quotient*

$$\mathcal{G}(X) \;:=\; \mathcal{E}_M(X) \,/\, \mathcal{N}(X)\,.$$

**4.2** Again we call elements in $\mathcal{G}(X)$ *generalized functions* and denote them by capital letters. Also we write $U = \mathrm{cl}[(u_\varepsilon)_\varepsilon] = (u_\varepsilon)_\varepsilon + \mathcal{N}(X)$ for the class given by the representative $(u_\varepsilon)_\varepsilon$. Similar to the case of open sets in Euclidean space, $\mathcal{E}_M(X)$ is a differential algebra (w.r.t. the Lie derivative) with componentwise operations and $\mathcal{N}(X)$ is a differential ideal in it. The latter property allows us to define

**4.3 Definition.** *Let $U \in \mathcal{G}(X)$ and $\xi \in \mathcal{T}^1_0(X)$. The* Lie derivative of $U$ with respect to $\xi$ *is defined by*

$$L_\xi U \;:=\; \mathrm{cl}[(L_\xi\, u_\varepsilon)_\varepsilon]$$

This notion indeed is well-defined by linearity of the classical Lie derivative and the fact that $\mathcal{N}(X)$ is a differential ideal. More precisely, given two representatives $(u_\varepsilon)_\varepsilon$ and $(\tilde{u}_\varepsilon)_\varepsilon$ of $U$ then $\mathrm{cl}[(L_\xi u_\varepsilon)_\varepsilon] - \mathrm{cl}[(L_\xi \tilde{u}_\varepsilon)_\varepsilon] = \mathrm{cl}[(L_\xi(u_\varepsilon - \tilde{u}_\varepsilon))_\varepsilon] = 0 \in \mathcal{G}(X)$. Hence $\mathcal{G}(X)$ is a differential $\mathbb{K}$-algebra (w.r.t. the Lie derivative).
Since locally Lie derivatives are given by differential operators with smooth coefficients whose $L^\infty$-norm can be estimated on compacts sets we get (as already indicated above) the following localization of the respective notions of moderateness and negligibility.

**4.4 Proposition.** *Denote by $\mathcal{A} = \{(V_\alpha, \psi_\alpha) : \alpha \in A\}$ an atlas of $X$. Then*

$$\begin{aligned}
(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(X) &\Leftrightarrow (u_{\alpha\,\varepsilon})_\varepsilon := (u_\varepsilon \circ \psi_\alpha^{-1})_\varepsilon \in \mathcal{E}_M(\psi_\alpha(V_\alpha)) \quad \forall\alpha \text{ and} \\
(u_\varepsilon)_\varepsilon \in \mathcal{N}(X) &\Leftrightarrow (u_{\alpha\,\varepsilon})_\varepsilon \in \mathcal{N}(\psi_\alpha(V_\alpha)) \quad \forall\alpha\,.
\end{aligned}$$

Note that from proposition 3.6 it follows that in the above definition of negligibility, we also could have omitted the Lie derivatives provided the respective sequence $(u_\varepsilon)_\varepsilon$ was assumed to be moderate. It is now immediate that a generalized function $U$ on $X$ allows for the following local description via the assignment $\mathcal{G}(X) \ni U \mapsto (U_\alpha)_{\alpha\in A}$ with $U_\alpha := U \circ \psi_\alpha^{-1} \in \mathcal{G}(\psi_\alpha(V_\alpha))$. We call $U_\alpha$ the *local expression* of $U$ with respect to the chart $(V_\alpha, \psi_\alpha)$.

**4.5 Proposition.** *$\mathcal{G}(X)$ can be identified with the set of all families $(U_\alpha)_\alpha$ of generalized functions $U_\alpha \in \mathcal{G}(\psi_\alpha(V_\alpha))$ satisfying the following transformation law*

$$U_\alpha|_{\psi_\alpha(V_\alpha \cap V_\beta)} = U_\beta|_{\psi_\beta(V_\alpha \cap V_\beta)} \circ \psi_\beta \circ \psi_\alpha^{-1}$$

*for all $\alpha, \beta \in A$ with $V_\alpha \cap V_\beta \neq \emptyset$.*

As a straightforward generalization of theorem 3.9 we get

**4.6 Theorem.** *$\mathcal{G}(X)$ is a (fine) sheaf of $\mathbb{K}$-algebras.*

**4.7** Now we turn our attention to the task of *embedding distributions* into $\mathcal{G}(X)$. The "constant" embedding of smooth functions into the algebra is straightforward and we shall again denote it by $\sigma : \mathcal{C}^\infty(X) \hookrightarrow \mathcal{G}(X)$. Before turning to the interrelations between $\mathcal{D}'(X)$ and $\mathcal{G}(X)$ let us first clarify what we can expect at all from such an embedding. Lie derivatives will not commute with the embedding since locally they are given by operators with $\mathcal{C}^\infty$-coefficients and already on $\mathbb{R}^n$ multiplication with smooth functions does not commute with the embedding: More precisely, consider $X = \mathbb{R}$ and $\xi = x\partial_x$. We then have $\iota(L_\xi H) = \iota(x\delta) = 0$. On the other hand, $(L_\xi(\iota H))_\varepsilon = L_\xi(\iota H)_\varepsilon = x\partial_x(\iota H)_\varepsilon = x(\iota\delta)_\varepsilon = \iota(x)_\varepsilon \iota(\delta)_\varepsilon$. Hence $L_\xi(\iota H) = \iota(x)\iota(\delta) \neq 0$ (see 3.25(ii)). By the same reasoning, the $\mathcal{C}^\infty(X)$-module structure of $\mathcal{D}'$ will not be respected by the embedding. Note that the $\mathcal{C}^\infty$-module structure of $\mathcal{G}$ by construction is compatible with that of $\mathcal{C}^\infty(X)$ (via $\sigma$; cf. also [50]). Hence it is necessary to confine ourselves to construct an embedding rendering $\mathcal{C}^\infty(X)$ a faithful subalgebra, i.e., coinciding with $\sigma$ on $\mathcal{C}^\infty$.
Since it seems not feasible to construct something like a global $\mathcal{S}$-mollifier with vanishing moments, i.e., satisfying conditions 3.10 (ii) on the manifold, the method of choice would be to embed distributions chart-wise. (For the construction of such a type of mollifier on the tangent bundle, however, leading to an atlas depending embedding see [14].) This would also guarantee the prospective embedding to coincide with $\iota$ in the sense of 3.16 locally. Unfortunately chart-wise embedding of distributions does not define a generalized function on $X$ since embedding commutes with the action of a diffeomorphism on the level of association only (cf. 3.52 (iv)) but in general not on the level of equality in the algebra. To stress this important point more explicitly we give the following counterexample.

**4.8 Example.** Consider the diffeomorphism $\mu(x) = 2x$ on $\mathbb{R}$ and the distribution $w = \delta$. First we calculate the pullback of $\delta$ under $\mu$,

$$\langle \mu^*\delta, \varphi \rangle = \langle \delta, \varphi \circ \mu^{-1} |\det D\mu^{-1}| \rangle = \langle \delta(x), \frac{1}{2}\varphi(\frac{x}{2}) \rangle = \frac{1}{2}\varphi(0),$$

hence $\mu^*\delta = 1/2\,\delta$. Now we have

$$\begin{aligned} (\iota \circ \mu^*)\,\delta &= \iota(\tfrac{1}{2}\,\delta) = \tfrac{1}{2}\rho_\varepsilon \\ (\mu^* \circ \iota)\,\delta &= \mu^*\rho_\varepsilon = \rho_\varepsilon(2\,.)\,. \end{aligned}$$

From this we see that $(\iota \circ \mu^* - \mu^* \circ \iota)\delta = 1/2\rho_\varepsilon(x) - \rho_\varepsilon(2x)$ is not in the ideal $\mathcal{N}(\mathbb{R})$. However, it is evident that $(\iota \circ \mu^* - \mu^* \circ \iota)\delta \approx 0$.

**4.9** Slightly more general, let $\mu : \tilde{\Omega} \to \Omega$ be a diffeomorphism of open sets of $\mathbb{R}^n$. Then the above considerations may be illustrated by the following diagram which commutes only in the sense of association.

$$\begin{array}{ccccc} \mathcal{D}'(\Omega) & \xrightarrow{\mu^*} & \mathcal{D}'(\tilde{\Omega}) & & \\ \Big\downarrow \iota & & \Big\downarrow \iota & \quad \approx & \neq \\ \mathcal{G}(\Omega) & \xrightarrow{\mu^*} & \mathcal{G}(\tilde{\Omega}) & & \end{array}$$

From all of the above it is evident that a diffeomorphism invariant embedding may not be defined within the present framework (cf. also [50], the remarks after theorem 1). The deeper reason for this phenomenon is the incompatibility of embedding via convolution (with a single mollifier) and appropriate transformation behavior under diffeomorhisms. The first notion relies heavily on the linear structure of $\mathbb{R}^n$ which is not respected by the latter one. How this apparent road block may be circumvented by an embedding via convolution with paths of mollifiers will be the topic of sec. 4.D.

**4.10** In our present setting the existence of injective sheaf morphisms $\iota : \mathcal{D}'(X) \hookrightarrow \mathcal{G}(X)$ coinciding with $\sigma$ on $\mathcal{C}^\infty$ and satisfying $\langle w_\varepsilon, \varphi\rangle \to \langle w, \varphi\rangle$ for $\iota(w) = (w_\varepsilon)_\varepsilon$, $\varphi \in \mathcal{D}$ for all distributions $w$ has been proved by de Roever and Damsma [50] using de Rham-regularizations (1.45, [49], §15). Here, however, we shall confine ourselves with the following explicit construction of a global embedding (as has been mentioned previously, another non-geometric embedding using fiber integration in the tangent bundle $T(X)$ was proposed in [14]). Choose some atlas $\mathcal{A} = (V_\alpha, \psi_\alpha)_\alpha$ of the manifold, a subordinate partition of unity $(\chi_j)_j$ with $\mathrm{supp}(\chi_\mathrm{j}) \in \mathrm{V}_{\alpha_\mathrm{j}}$ and a family of cut-off functions $(\zeta_j)_j$ with each $\mathrm{supp}(\zeta_\mathrm{j}) \subset\subset \mathrm{V}_{\alpha_\mathrm{j}}$ and $\zeta_j \equiv 1$ on $\mathrm{supp}(\chi_\mathrm{j})$. Then we may define (denoting by $w_\alpha$ the local expression of the distribution $w$ with respect to the chart $(V_\alpha, \psi_\alpha)$, cf. 1.23)

$$\begin{aligned}\iota_\mathcal{A} &: \mathcal{D}'(X) \to \mathcal{G}(X)\\ &(w_\alpha)_\alpha \to \mathrm{cl}[(\sum_{j=1}^{\infty} \zeta_j \cdot (((\chi_j \circ \psi_{\alpha_j}^{-1}) w_{\alpha_j}) * \rho_\varepsilon) \circ \psi_{\alpha_j})_\varepsilon]\,.\end{aligned}$$

Despite providing a linear embedding coinciding with $\sigma$ on $\mathcal{C}^\infty(X)$ (for the rather technical proof see [117], 1.9.21) $\iota_\mathcal{A}$ clearly depends on all its ingredients, i.e., $\mathcal{A}$, $(\chi_j)_j$ and $(\zeta_j)_j$ hence is *non-geometric in an essential way.*

The option to be taken here—as already indicated in the introduction to this chapter—is to dispense with a "canonical" embedding of distributions altogether. A sufficient substitute for the latter concept allowing for statements on the distributional level will be given by coupled calculus to be introduced in a moment.

**4.11 Definition.** *A generalized function $U \in \mathcal{G}(X)$ is called* associated to 0 *, $U \approx 0$, if for all compactly supported one-densities $\mu$ (i.e., $\mu \in \Gamma_0^\infty(\mathrm{Vol}\,(\mathrm{X}))$ ) and one (hence any) representative $(u_\varepsilon)_\varepsilon$ of $U$*

$$\lim_{\varepsilon\to 0} \int_X u_\varepsilon\, \mu \;=\; 0\,.$$

Note that—at last—in the above definition it becomes clear that we regard generalized functions really as generalizations of *functions*, in the sense that we view them as regular objects in the space $\mathcal{D}'(X)$, hence acting on (compactly supported, smooth) one-densities. This remark also applies to generalized sections in vector bundles to be treated below. (For a detailed discussion of these matters see 1.A.) By the local description 4.5 we have

**4.12 Proposition.** $U \approx 0$ *in* $\mathcal{G}(X) \Leftrightarrow U_\alpha \approx 0$ *in* $\mathcal{G}(\psi_\alpha(V_\alpha)) \quad \forall\alpha$

**Proof.** Choose some $\mu \in \Gamma_0^\infty(\mathrm{Vol}\,(\mathrm{X}))$ with $\mathrm{supp}(\mu)$ contained in a single chart $V_\alpha$. Then for any representative $(u_\varepsilon)_\varepsilon$ of $U$ and all such $\mu$ we have

$$0 \;=\; \lim_{\varepsilon\to 0} \int_X u_\varepsilon\, \mu \;=\; \lim_{\varepsilon\to 0} \int_{\psi_\alpha(V_\alpha)} u_\varepsilon \circ \psi_\alpha^{-1}(x)\, \mu_\alpha(x)\, dx\,,$$

where $\mu_\alpha \in \mathcal{D}(\psi_\alpha(V_\alpha))$ is the component of $\mu$ with respect to the chart $(V_\alpha, \psi_\alpha)$ (cf.1.4). But this is obviously equivalent to

$$\int_{\psi_\alpha(V_\alpha)} u_{\alpha\,\varepsilon}(x) \varphi(x)\, dx \;\to 0$$

for all $\varphi \in \mathcal{D}(\Psi_\alpha(V_\alpha))$.

For the converse direction just note that any $\mu \in \Gamma_0^\infty(\mathrm{Vol}\,(\mathrm{X}))$ may be written as a finite sum of one-densities with their compact supports each contained in a single chart. □

As it is the case for open sets of $\mathbb{R}^n$

$$U \approx V \;:\Leftrightarrow\; U - V \approx 0$$

defines an equivalence relation in $\mathcal{G}(X)$ giving rise to a linear quotient of $\mathcal{G}(X)$. If $U \approx V$, $U$ and $V$ are called *associated to each other*. The next definition finally establishes the link from $\mathcal{G}(X)$ to the "distributional world."

**4.13 Definition.** *Let $U$ in $\mathcal{G}(X)$ and $w \in \mathcal{D}'(X)$. We say that $U$ admits $w$ as* associated distribution *and call $w$ the* distributional shadow *(or* macroscopic aspect*) of $U$ if for all $\mu \in \Gamma_c^\infty(\mathrm{Vol}\,(\mathrm{X}))$ and one (hence any) representative*

$$\lim_{\varepsilon\to 0}\int u_\varepsilon\,\mu \;=\; w(\mu)\,,$$

*where $w(\mu)$ denotes the distributional action of $w$ on $\mu$. In that case we use the notation $U \approx w$.*

Completely analogous to proposition 4.11 we obtain the following local characterization

**4.14 Proposition.** $U \approx w \;\Leftrightarrow\; U_\alpha \approx w_\alpha$ *in* $\mathcal{G}(\psi_\alpha(V_\alpha))\ \forall\alpha$

**4.15** Next we are going to discuss an analogue of theorem 3.52 clarifying consistency properties with classical products (in the sense of association). However, in the absence of an embedding $\iota$ we have to be slightly more cautious. For example the following generalization of 3.52 (i): "$U, V \in \mathcal{G}(X)$, $U \approx f \in \mathcal{C}^\infty$ and $V \approx w \in \mathcal{D}'(X) \Rightarrow UV \approx fw$" is wrong in general. To see this take $\rho \in \mathcal{D}(\mathbb{R})$ with $\int\rho = 1$. Then $\mathrm{cl}[(\rho(\frac{x}{\varepsilon})_\varepsilon] \approx 0$ and clearly $\mathrm{cl}[(\frac{1}{\varepsilon})\rho(\frac{x}{\varepsilon})_\varepsilon] \approx \delta$ but $\rho(\frac{x}{\varepsilon})\,(\frac{1}{\varepsilon})\rho(\frac{x}{\varepsilon}) \to \delta\int\rho^2$ in $\mathcal{D}'$. The reason for the validity of 3.52 (i), ultimately is that $f * \rho_\varepsilon \to f$ uniformly on compact sets already for a continuous function $f$ (cf. e.g. [2], 2.18), whereas $\rho(x/\varepsilon) \to 0$ only weakly. Therefore we introduce the following stronger equivalence relations on $\mathcal{G}(X)$ (cf. also 3.56).

**4.16 Definition.** *Let $U \in \mathcal{G}(X)$.*

*(i) $U$ is called $\mathcal{C}^k$*-associated *to $0$ $(0 \le k \le \infty)$, $U \approx_k 0$, if for all $l \le k$ (resp. for all $l \in \mathbb{N}_0$ in the case $k = \infty$) $\forall \xi_1, \dots, \xi_l \in \mathcal{T}_0^1(X)$ and one (hence any) representative $(u_\varepsilon)_\varepsilon$*

$$L_{\xi_1} \dots L_{\xi_l}\, u_\varepsilon \;\to\; 0 \;\; \textit{uniformly on compact sets.}$$

*(ii) We say that $U$ admits $f$ as $\mathcal{C}^k$*-associated function*, $U \approx_k f$, if for all $l \le k$ (resp. for all $l \in \mathbb{N}_0$ in the case $k = \infty$) $\forall \xi_1, \dots, \xi_l \in \mathcal{T}_0^1(X)$ and one (hence any) representative*

$$L_{\xi_1} \dots L_{\xi_l}\,(u_\varepsilon - f) \;\to\; 0 \; \textit{uniformly on compact sets.}$$

Clearly if $U$ is $\mathcal{C}^k$-associated to $f$ then $f \in \mathcal{C}^k(X)$. Moreover, if $U$ admits for a $\mathcal{C}^k$-associated function at all the latter is unique. Note also that the above notion of convergence may equivalently be expressed by saying that all $(u_{\alpha\,\varepsilon})_\varepsilon$ converge uniformly in all derivatives of order less or equal $k$ (resp. in all derivatives if $k = \infty$) on compact sets. We are now prepared to state the following

**4.17 Proposition.** *Let $U, V \in \mathcal{G}(X)$.*

*(i) If $V \approx w \in \mathcal{D}'(X)$, $f \in \mathcal{C}^\infty(X)$, and either (a) $U = \sigma(f)$ or (b) $U \approx_\infty f$, then $UV \approx fw$.*

*(ii) If $U \approx_k f$ and $V \approx_k g$ then $UV \approx_k fg$ $(f, g \in \mathcal{C}^k(X))$.*

**Proof.** (i)(a) is clear since $\int f v_\varepsilon \mu = v_\varepsilon(f\mu) \to w(f\mu)$ for all compactly supported one-densities $\mu$. To prove (i)(b) we use the fact that multiplication: $\mathcal{C}^\infty \times \mathcal{D}' \to \mathcal{D}'$ as a bilinear separately continuous map is jointly sequentially continuous since both factors are barreled ( [110], §42.2(3) and §40.1). (ii) follows from elementary analysis. □

Note that 4.17 (i)(a) is the reconciliation of the respective $\mathcal{C}^\infty$-module structures of $\mathcal{D}'$ and $\mathcal{G}$ on the level of association. Next we introduce the notion of integration of generalized functions.

**4.18 Definition.** *Let $U \in \mathcal{G}(X)$ and $\mu \in \Gamma^\infty(\mathrm{Vol}\,(\mathrm{X}))$. Then we define the* integral of $U$ with respect to $\mu$ *over the compact set $K \subset\subset X$ by*

$$\int_K U\mu \;=\; \mathrm{cl}[(\int_K u_\varepsilon \mu)_\varepsilon]\,.$$

We close this section by showing that the Lie derivative respects associated distributions.

**4.19 Proposition.** *Let $X$ be orientable and $U \approx w$. Then $L_\xi U \approx L_\xi w$.*

Note that in order to be able to define a Lie derivative of the test objects we have to assume $X$ to be orientable (cf. 1.31). Moreover, Stokes theorem is used in the following

**Proof.** Let $\nu \in \Omega^n_c(X)$ then

$$\int (L_\xi u_\varepsilon)\nu = -\int u_\varepsilon(L_\xi \nu) \to -w(L_\xi \nu) = L_\xi w(\nu)$$

□

### 4.B. Generalized Sections of Vector Bundles

In this section after defining generalized sections in vector bundles we discuss tensor fields, introduce the generalized Lie derivative and prove some consistency results (in the sense of association) with respect to classical tensor analysis.
We again use the same notations as in chapter 1 (see, in particular 1.4). We denote by $(E, X, \pi)$ a vector bundle with base space $X$ and by $(V_\alpha, \Psi_\alpha)_\alpha$ a vector bundle atlas. Let the dimension of the fibers be $N$ and denote by $\Psi^i_\alpha$ $(i = 1, \dots, N)$ the $i$-th component of the vector part of the bundle chart $(V_\alpha, \Psi_\alpha)$ whereas $\psi_\alpha$ denotes the corresponding chart on the base manifold. The space of $\mathcal{C}^k$-sections is denoted by $\Gamma^k(X, E)$ where, for simplicity, we drop the superscript in the case $k = \infty$. Finally $s^i_\alpha := \Psi^i_\alpha \circ s \circ \psi_\alpha^{-1}$ $(i = 1, \dots, N)$ denotes the $i$-th component of the section $s$ with respect to the chart $(V_\alpha, \Psi_\alpha)$.

**4.20 Definition.** *Let $(E, X, \pi)$ be a vector bundle, and again $I = (0, 1]$.*

$$\begin{aligned}
\mathcal{E}(X,E) &:= (\Gamma(X,E))^I\\
\mathcal{E}_M(X,E) &:= \{(s_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}(X,E) :\ \forall\alpha, \forall i = 1,\dots,N :\ (s^i_{\alpha\,\varepsilon})_\varepsilon := (\Psi^i_\alpha \circ s_\varepsilon \circ \psi_\alpha^{-1})_\varepsilon \in \mathcal{E}_M(\psi_\alpha(V_\alpha))\}\\
\mathcal{N}(X,E) &:= \{(s_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}(X,E) :\ \forall\alpha, \forall i = 1,\dots,N :\ (s^i_{\alpha\,\varepsilon})_\varepsilon \in \mathcal{N}(\psi_\alpha(V_\alpha))\}
\end{aligned}$$

First note that although the composition $f \circ U$ of a generalized function $U$ with a smooth function $f$ generally need not be moderate (cf. 3.7) the notions of moderateness and negligibility as defined above are preserved under the change of bundle charts due to the (fiberwise) linearity of the transition functions. In order to define generalized sections of the bundle $(E, X, \pi)$ we need the following

**4.21 Proposition.** *With operations defined componentwise (i.e., for each $\varepsilon$), $\mathcal{E}_M(X,E)$ is a $\mathcal{G}(X)$-module with $\mathcal{N}(X,E)$ a submodule in it.*

**Proof.** We need to establish the following statements (a) $(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(X)$, $(s_\varepsilon)_\varepsilon \in \mathcal{E}_M(X,E) \Rightarrow (u_\varepsilon s_\varepsilon)_\varepsilon \in \mathcal{E}_M(X,E)$, (b) $(u_\varepsilon)_\varepsilon \in \mathcal{N}(X)$, $(s_\varepsilon)_\varepsilon \in \mathcal{E}_M(X,E) \Rightarrow (u_\varepsilon s_\varepsilon)_\varepsilon \in \mathcal{N}(X,E)$ and (c) $(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(X)$, $(s_\varepsilon)_\varepsilon \in \mathcal{N}(X,E) \Rightarrow (u_\varepsilon s_\varepsilon)_\varepsilon \in \mathcal{N}(X,E)$, which easily follow from the local description 4.5 and the definitions above. □

Now we are in the position to define.

**4.22 Definition.** *The $\mathcal{G}(X)$-module of generalized sections of $(E, X, \pi)$ is defined as the quotient*

$$\mathcal{G}(X,E) \;:=\; \mathcal{E}_M(X,E)\,/\,\mathcal{N}(X,E)\,.$$

**4.23** As usual we denote the generalized objects by capital letters, e.g., $S = \mathrm{cl}[(s_\varepsilon)_\varepsilon]$. By the very definition of $\mathcal{G}(X,E)$ we may describe a generalized section $S$ by a family $(S_\alpha)_\alpha = ((S^i_\alpha)_\alpha)^N_{i=1}$, where $S_\alpha$ is called the *local expression* of $S$. Its *components* $S^i_\alpha := \Psi^i_\alpha \circ S \circ \psi_\alpha^{-1} \in \mathcal{G}(\psi_\alpha(V_\alpha))$ $(i = 1, \dots, N)$ satisfy

$$S^i_\alpha(x) \;=\; (\boldsymbol{\psi}_{\alpha\beta})^i_j(x)\, S^j_\beta(\psi_\beta \circ \psi_\alpha^{-1}(x)) \tag{73}$$

for all $p = \psi_\alpha^{-1}(x) \in V_\alpha \cap V_\beta$, where $\boldsymbol{\psi}_{\alpha\beta}$ denotes the transition functions of the bundle. Hence formally generalized sections of $(E, X, \pi)$ are locally simply given by "ordinary" sections with generalized "coefficients." Note that this is not the case in the framework of Vickers and Wilson [184] based on the "full" version of the algebra.

By a straightforward generalization of theorem 3.9 we have.

**4.24 Theorem.** *$\mathcal{G}(X,E)$ is a (fine) sheaf of $\mathcal{G}(X)$-modules.*

**4.25** As before smooth sections may be embedded into $\mathcal{G}(X,E)$ by the "constant" embedding now denoted by $\Sigma$, i.e., $\Sigma(s) = \mathrm{cl}[(s)_\varepsilon]$. In this way $\mathcal{G}(X,E)$ also becomes a $\mathcal{C}^\infty(X)$-module and the two respective module structures on the space of generalized sections are compatible in the sense of the following commutative diagram.

$$\begin{array}{ccc}
\mathcal{C}^\infty(X)\times\Gamma(X,E) & \xrightarrow{\sigma\times\Sigma} & \mathcal{G}(X)\times\mathcal{G}(X,E)\\
\big\downarrow\cdot & & \big\downarrow\cdot\\
\Gamma(X,E) & \xrightarrow{\Sigma} & \mathcal{G}(X)
\end{array}$$

Analogously to the earlier cases we set up coupled calculus.

**4.26 Definition.**

*(i) A generalized section $S\in\mathcal{G}(X,E)$ is called* associated to 0*, $S\approx 0$, if for all $\mu\in\Gamma_0(X,E^*\otimes\mathrm{Vol}\,(\mathrm{X}))$ and one (hence any) representative $(s_\varepsilon)_\varepsilon$ of $S$*

$$\lim_{\varepsilon\to 0}\int_X (s_\varepsilon|\mu) \;=\; 0\,.$$

*(ii) Let $S$ in $\mathcal{G}(X,E)$ and $w\in\mathcal{D}'(X,E)$. We say that $S$ admits $w$ as* associated distribution (with values in $E$) *and call $w$ the* distributional shadow *(or* macroscopic aspect*) of $S$ if for all $\mu\in\Gamma_c(X,E^*\otimes\mathrm{Vol}\,(\mathrm{X}))$ and one (hence any) representative*

$$\lim_{\varepsilon\to 0}\int_X (s_\varepsilon|\mu) \;=\; w(\mu)\,,$$

*where $w(\mu)$ denotes the distributional action of $w$ on $\mu$. In that case we use the notation $S\approx w$.*

Again $S\approx T :\Leftrightarrow S-T\approx 0$ defines an equivalence relation giving rise to a linear quotient of $\mathcal{G}(X,E)$. If $S\approx T$ we call $S$ and $T$ *associated to each other*. Analogously to 4.12 and 4.14, by straightforward localization we get the following

**4.27 Proposition.**

*(i) $S\approx 0$ in $\mathcal{G}(X,E)\Leftrightarrow S^i_\alpha\approx 0$ in $\mathcal{G}(\psi_\alpha(V_\alpha))$ $\forall\alpha, i=1,\dots,N$*

*(ii) $S\approx w$ in $\mathcal{G}(X,E)\Leftrightarrow S^i_\alpha\approx w^i_\alpha$ in $\mathcal{G}(\psi_\alpha(V_\alpha))$ $\forall\alpha, i=1,\dots,N$*

**4.28 Definition.** *Let $S\in\mathcal{G}(X,E)$*

*(i) $S$ is called $\mathcal{C}^k$-associated to 0 $(0\le k\le\infty)$, $S\approx_k 0$, if for one (hence any) representative $(s_\varepsilon)_\varepsilon$ and $\forall\alpha$, $i=1,\dots,N$ $s^i_{\alpha\,\varepsilon}\to 0$ uniformly on compact sets in all derivatives of order less or (if $k<\infty$) equal to $k$.*

*(ii) We say that $S$ allows $t\in\Gamma^k(X,E)$ as a $\mathcal{C}^k$-associated section, $S\approx_k t$, if for one (hence any) representative $(s_\varepsilon)_\varepsilon$ and $\forall\alpha$, $i=1,\dots,N$ $s^i_{\alpha\,\varepsilon}\to t^i_\alpha$ uniformly on compact sets in all derivatives of order less or (if $k<\infty$) equal to $k$.*

As is the case with $\mathcal{G}(X)$ the different $\mathcal{C}^\infty$-module structures of $\mathcal{D}'(X,E)$ and $\mathcal{G}(X,E)$, respectively, may be reconciled at the level of association.

**4.29 Proposition.** *Let $U\in\mathcal{G}(X)$ and $S\in\mathcal{G}(X,E)$.*

*(i) If $U\approx w\in\mathcal{D}'(X)$, $s\in\Gamma(X,E)$ and either (a) $S=\Sigma(s)$ or (b) $S\approx_\infty s$, then $U\,S\approx ws$ in $\mathcal{G}(X,E)$.*

*(ii) If $S\approx s\in\mathcal{D}'(X,E)$, $f\in\mathcal{C}^\infty(X)$ and either (a) $U=\sigma(f)$ or (b) $U\approx_\infty f$, then $U\,S\approx fs$ in $\mathcal{G}(X,E)$.*

*(iii) If $U\approx_k f$ and $S\approx_k s$ then $U\,S\approx_k fs$ in $\mathcal{G}(X,E)$ $(f\in\mathcal{C}^k(X),\ s\in\Gamma^k(X,E))$.*

Instead of proving the above results which is done by a straightforward (componentwise) application of 4.17 we illustrate them by the following commutative diagrams. Note that the horizontal arrows only have subspaces of the respective source spaces as their domains.

$$\begin{tikzcd}
\mathcal{D}'(X) \times \Gamma(X,E) \arrow[d, "\cdot"] & \mathcal{G}(X) \times \mathcal{G}(X,E) \arrow[l, "\approx\times\approx_\infty"'] \arrow[d, "\cdot"] \\
\mathcal{D}'(X,E) & \mathcal{G}(X,E) \arrow[l, "\approx"']
\end{tikzcd}$$

$$\begin{tikzcd}
\mathcal{C}^\infty(X) \times D'(X,E) \arrow[d, "\cdot"] & \mathcal{G}(X) \times \mathcal{G}(X,E) \arrow[l, "\approx_\infty\times\approx"'] \arrow[d, "\cdot"] \\
\mathcal{D}'(X,E) & \mathcal{G}(X,E) \arrow[l, "\approx"']
\end{tikzcd}$$

$$\begin{tikzcd}
\mathcal{C}^k(X) \times \Gamma^k(X,E) \arrow[d, "\cdot"] & \mathcal{G}(X) \times \mathcal{G}(X,E) \arrow[l, "\approx_k\times\approx_k"'] \arrow[d, "\cdot"] \\
\Gamma^k(X,E) & \mathcal{G}(X,E) \arrow[l, "\approx_k"']
\end{tikzcd}$$

**4.30** In the case where $(E, X, \pi)$ is some tensor bundle $T^r_s(X)$ over the manifold $X$ we shall use the notation $\mathcal{G}^r_s(X)$ for $\mathcal{G}(X, T^r_s(X))$ and similarly for $\mathcal{E}$, $\mathcal{E}_M$ and $\mathcal{N}$. First we prove that generalized tensor fields are just $\mathcal{G}(X)$-multilinear maps on generalized vector respectively covector fields. This result, in particular, will be very useful to carry over many of the results of classical tensor analysis into our setting.

**4.31 Theorem.** *The space of $\mathcal{G}(X)$-multilinear mappings $L_{\mathcal{G}(X)}(\mathcal{G}^0_1(X)^r, \mathcal{G}^1_0(X)^s; \mathcal{G}(X))$ is $\mathcal{G}(X)$-linearly isomorphic to $\mathcal{G}^r_s(X)$.*

To simplify the notation we will set $r = 1 = s$ in the proof which needs the following technical lemma.

**4.32 Lemma.** *Let $T \in L_{\mathcal{G}(X)}(\mathcal{G}^0_1(X), \mathcal{G}^0_1(X); \mathcal{G}(X))$, $A \in \mathcal{G}^0_1(X)$ and $\Xi \in \mathcal{G}^1_0(X)$ with $\Xi|_U = 0$ for some open $U \subseteq X$. Then $T(A, \Xi)|_U = 0$.*

**Proof.** Since $U$ can be written as the union of a collection of open sets $(U_p)_{p\in U}$ such that each $\overline{U}_p \subseteq V_\alpha$ for some chart $V_\alpha$ and due to the sheaf property of $\mathcal{G}(X)$ we may assume without loss of generality that $\overline{U} \subseteq V_\alpha$ and write $\Xi|_{V_\alpha} = \Xi^i\partial_i$ with $\Xi^i \in \mathcal{G}(V_\alpha)$ vanishing on $U$. Let now $f$ be a bump function on $\overline{U}$ (i.e., $f \in \mathcal{D}(V_\alpha)$, $f|_{\overline{U}} = 1$) then

$$\begin{aligned}
T(A,\Xi)|_U &= f^2|_U\, T(A,\Xi)|_U \;=\; f^2\, T(A,\Xi)|_U \\
&= T(A, f\Xi^i f\partial_i)|_U \;=\; f\Xi^i\, T(A, f\partial_i)|_U \;=\; f\Xi^i|_U\, T(A, f\partial_i)|_U \;=\; 0\,,
\end{aligned}$$

where we did not distinguish notationally between $f$ and $\sigma(f)$. □

**Proof of the theorem.** Let $T = \mathrm{cl}[(t_\varepsilon)_\varepsilon] \in \mathcal{G}^1_1(X)$, $A = \mathrm{cl}[(a_\varepsilon)_\varepsilon] \in \mathcal{G}^0_1(X)$ and $\Xi = \mathrm{cl}[(\xi_\varepsilon)_\varepsilon] \in \mathcal{G}^1_0(X)$. Using classical contraction we define componentwise the following map

$$\tilde{T} :\ (a_\varepsilon, \xi_\varepsilon) \mapsto f_\varepsilon := t_\varepsilon(a_\varepsilon, \xi_\varepsilon)\,.$$

Using the local description it is easy to see that $F = \mathrm{cl}[(f_\varepsilon)_\varepsilon] \in \mathcal{G}(X)$, $\tilde{T} : \mathcal{G}^0_1(X) \times \mathcal{G}^1_0(X) \to \mathcal{G}(X)$ is well-defined and $\mathcal{G}(X)$-bilinear, hence $\tilde{T} \in L_{\mathcal{G}(X)}(\mathcal{G}^0_1(X), \mathcal{G}^1_0(X); \mathcal{G}(X))$. Moreover, the assignment $T \mapsto \tilde{T}$ is also $\mathcal{G}(X)$-linear, so it only remains to show that the latter is an isomorphism.
To prove injectivity assume $\tilde{T} = 0$, that is $(t_\varepsilon(a_\varepsilon, \xi_\varepsilon))_\varepsilon \in \mathcal{N}(X)$ for all $A = \mathrm{cl}[(a_\varepsilon)_\varepsilon] \in \mathcal{G}^0_1(X)$ and all $\Xi = \mathrm{cl}[(\xi_\varepsilon)_\varepsilon] \in \mathcal{G}^1_0(X)$. To show that $T = 0 \in \mathcal{G}^1_1(X)$ it suffices to work locally. Choose $A = \Sigma(dx^i)$ and $\Xi = \Sigma(\partial_j)$ for some $1 \le i, j \le n$. Then $\mathcal{N}(X) \ni f_\varepsilon|_{V_\alpha} = {t_\alpha}^{\,i}_{j\,\varepsilon}$. Since $i, j$ were arbitrary we are done.

To show surjectivity choose $\tilde{T} \in L_{\mathcal{G}(X)}(\mathcal{G}^0_1(X), \mathcal{G}^1_0(X); \mathcal{G}(X))$. Similar as in the proof of the lemma we cover $X$ by a collection $(U_p)_{p\in X}$ such that each $\overline{U}_p \subseteq V_\alpha$ for some $\alpha$. Then on $U_p$ we may define $(1 \le i, j \le n)$

$$T_{\alpha p}{}^i{}_j \;=\; \tilde{T}(g dx^i, g\partial_j) \circ \psi_\alpha^{-1} \;\in \mathcal{G}(\psi_\alpha(U_p))\,,$$

where $g$ is a bump function on $\overline{U}_p$, i.e., $g \in \mathcal{D}(V_\alpha)$, $g|_{\overline{U}_p} \equiv 1$. By the above lemma this definition is independent of the choice of $g$. Moreover, since $\tilde{T}$, $gdx^i$ and $g\partial_j$ are globally defined the $(T_{\alpha p})_{p\in X}$ form a coherent family. Hence by the sheaf property of $\mathcal{G}^1_1(X)$ there exists a unique $T \in \mathcal{G}^1_1(X)$ with $T|_{U_p} = T_{\alpha p}$. □

**4.33** From now on given a generalized tensor field $T \in \mathcal{G}^r_s(X)$ we shall call the $n^{r+s}$ generalized functions on $V_\alpha$ defined by

$$T^{\alpha}{}^{i_1\dots i_r}_{j_1\dots j_s} \;:=\; T|_{V_\alpha}(dx^{i_1}, \dots, dx^{i_r}, \partial_{j_1}, \dots, \partial_{j_s})$$

its *components* with respect to the chart $(V_\alpha, \psi_\alpha)$. We shall also use abstract index notation (cf. 2.14, resp. [187], pp. 23ff) whenever convenient and write $T^{a_1\dots a_r}_{b_1\dots b_s} \in \mathcal{G}^r_s(X)$. To clearly distinguish between the notions of abstract and concrete indices we reserve the letters $a, b, c, d, e, f$ for the previous one and $i, j, k, l, \dots$ for the latter one. Hence we shall denote the components of $\Xi^a \in \mathcal{G}^1_0(X)$, $A_a \in \mathcal{G}^0_1(X)$ and $G_{ab} \in \mathcal{G}^0_2(X)$ by $\Xi^i$, $A_i$ and $G_{ij}$ respectively. Similarly the components of a representative $(t^{a_1\dots a_r}_{b_1\dots b_s}{}_\varepsilon)_\varepsilon \in \mathcal{E}_M{}^r_s(X)$ of $T^{a_1\dots a_r}_{b_1\dots b_s} \in \mathcal{G}^r_s(X)$ will be denoted by $(t^{\alpha}{}^{i_1\dots i_r}_{j_1\dots j_s}{}_\varepsilon)_\varepsilon$.

The spaces of moderate respectively negligible sequences of tensor fields may be characterized invariantly by the Lie derivative (similar to the scalar case treated in 4.4).

**4.34 Proposition.**

$$\begin{aligned}\mathcal{E}_M{}^r_s(X) = \{&(t_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}^r_s(X) : \forall K \subset\subset X,\, \forall k \in \mathbb{N}_0\, \exists N \in \mathbb{N}\, \forall \xi_1, \dots, \xi_k \in \mathcal{T}^1_0(X) :\\ &\sup_{p\in K} ||L_{\xi_1} \dots L_{\xi_k} t_\varepsilon(p)|| = O(\varepsilon^{-N}) \text{ as } \varepsilon \to 0\}\\ \mathcal{N}^r_s(X) = \{&(t_\varepsilon)_{\varepsilon\in I} \in \mathcal{E}^r_s(X) : \forall K \subset\subset X,\, \forall k, l \in \mathbb{N}_0\, \forall \xi_1, \dots, \xi_k \in \mathcal{T}^1_0(X) :\\ &\sup_{p\in K} ||L_{\xi_1} \dots L_{\xi_k} t_\varepsilon(p)|| = O(\varepsilon^{l}) \text{ as } \varepsilon \to 0\}\end{aligned}$$

*where $||.||$ denotes the norm induced on $\mathcal{T}^r_s(X)$ by any Riemannian metric on $X$.*

**4.35 Definition.** *Let $S \in \mathcal{G}^r_s(X)$ and $T \in \mathcal{G}^{r'}_{s'}(X)$. We define the* tensor product *$S \otimes T \in \mathcal{G}^{r+r'}_{s+s'}(X)$ of $S$ and $T$ by*

$$S \otimes T \;:=\; \mathrm{cl}[(s_\varepsilon \otimes t_\varepsilon)_\varepsilon]$$

Using again the local description it is easily checked that the tensor product is well defined. Moreover it is $\mathcal{G}(X)$-bilinear, associative and by a straightforward generalization of proposition 4.17 displays the following consistency properties with respect to the classical resp. distributional (cf. 1.56) tensor product.

**4.36 Proposition.** *Let $S \in \mathcal{G}^r_s(X)$ and $T \in \mathcal{G}^{r'}_{s'}(X)$.*

*(i) If $T \approx w \in \mathcal{D}'^{r'}_{s'}(X)$, $s \in \mathcal{T}^r_s(X)$ and either (a) $S = \Sigma(s)$ or (b) $S \approx_\infty s$ (or vice versa) then $S \otimes T \approx s \otimes w$ in $\mathcal{G}^{r+r'}_{s+s'}(X)$.*

*(ii) if $S \approx_k s$ and $T \approx_k t$ then $S \otimes T \approx_k s \otimes t$ in $\mathcal{G}^{r+r'}_{s+s'}(X)$ $(s \in \Gamma^k(X, T^r_s(X))$, $t \in \Gamma^k(X, T^{r'}_{s'}(X)))$.*

We may now easily generalize the following notions of classical tensor calculus.

**4.37 Definition.**

*(i) Let $T^{a_1\dots a_r}_{b_1\dots b_s} \in \mathcal{G}^r_s(X)$. We define the contraction of $T^{a_1\dots a_r}_{b_1\dots b_s}$ with respect to the affected indices by*

$$T^{a_1\dots i\dots a_r}_{b_1\dots i\dots b_s} := \mathrm{cl}[(t^{a_1\dots i\dots a_r}_{b_1\dots i\dots b_s}\,{}_\varepsilon)_\varepsilon] \in \mathcal{G}^{r-1}_{s-1}(X)\,.$$

*(ii) For any smooth vector field $\xi$ on $X$ we define the Lie derivative of $T \in \mathcal{G}^r_s(X)$ by*

$$L_\xi T \;:=\; \mathrm{cl}[(L_\xi t_\varepsilon)_\varepsilon]\,.$$

*(iii) Finally, we define the universal generalized tensor algebra over $X$ by*

$$\hat{G}(X) \;:=\; \bigoplus_{r,s} \mathcal{G}^r_s(X)\,.$$

The Lie derivative again displays the following consistency property with respect to its distributional (cf. 1.57) counterpart

**4.38 Proposition.** *Let $X$ be orientable and $T \approx t$ in $\mathcal{G}(X,E)$. Then $L_\xi T \approx L_\xi t$*

Remember that we had to assume $X$ to be orientable to define the Lie derivative of distributional tensor fields. Next we introduce the generalized Lie derivative, i.e., the Lie derivative with respect to a generalized vector field.

**4.39 Definition.** *Let $\Xi \in \mathcal{G}^1_0(X)$ and $T \in \mathcal{G}^r_s(X)$. We define the* generalized Lie derivative *of $T$ with respect to $\Xi$ by*

$$L_\Xi(T) \;:=\; \mathrm{cl}[(L_{\xi_\varepsilon}(t_\varepsilon))_\varepsilon]\,.$$

*In case $U \in \mathcal{G}(X)$ we also use the notation $\Xi(U)$ for $L_\Xi U$.*

The well-definedness of $L_\Xi(T)$ again is an easy consequence of the local description. Again literally all classical properties of the Lie derivative carry over since they hold componentwise. In particular, for generalized vector fields $\Xi, H$ we have $L_\Xi H = [\Xi, H] := \mathrm{cl}[\,([\xi_\varepsilon, \eta_\varepsilon])_\varepsilon]$ and for all generalized functions $U$ we have: $[U\Xi, H] = U[\Xi, H] - H(U)\Xi$. Also we immediately get the following consistency properties (cf. 1.62)

**4.40 Proposition.** *Let $\Xi \in \mathcal{G}^1_0(X)$ and $T \in \mathcal{G}^r_s(X)$*

*(i) If $\Xi = \Sigma(\xi)$ for some $\xi \in \mathcal{T}^1_0(X)$ then $L_\Xi(T) = L_\xi(T)$.*

*(ii) If $\Xi \approx_\infty \xi \in \mathcal{T}^1_0(X)$ and $T \approx t \in \mathcal{D}'^r_s(X)$ or conversely, if $\Xi \approx \xi \in \mathcal{D}'^1_0(X)$ and $T \approx_\infty t \in \mathcal{T}^r_s(X)$ then $L_\Xi(T) \approx L_\xi t$.*

*(iii) If $\Xi \approx_k \xi$ and $T \approx_{k+1} t$ then $L_\Xi(U) \approx_k L_\xi t$ ($\xi \in \Gamma^k(X, TX)$, $t \in \Gamma^{k+1}(X, T^r_s(X))$).*

In particular, for a generalized vector field $\Xi$ the map $L_\Xi \equiv \Xi : \mathcal{G}(X) \to \mathcal{G}(X)$ is clearly $\mathbb{R}$-linear (in fact even $\mathcal{R}$-linear) and obeys the Leibniz rule, hence is a derivation on $\mathcal{G}(X)$. Moreover any derivation on the algebra of generalized function arises this way.

**4.41 Theorem.** *$\mathcal{G}^1_0(X)$ is ($\mathbb{R}$-linearly) isomorphic to $Der(\mathcal{G}(X))$.*

**Proof.** Is suffices to show that for any derivation $\theta$ on $\mathcal{G}(X)$ we may construct a generalized vector field $\Xi$ such that $\theta(U) = \Xi(U)$ for all $U \in \mathcal{G}(X)$. We start by showing that $\theta$ is a local operator, i.e., that $U = 0$ on $V (\subseteq X)$ open implies $\theta(U)|_V = 0$. To this end chose $p \in V$ and a function $f \in \mathcal{D}(V)$ equal to 1 on a (smaller) neighborhood $W$ of $p$. Then $U = (1-f)U$ and

$$\theta(U)|_W \;=\; \theta(1-f)U|_W + (1-f)\theta(U)|_W \;=\; 0 \,\in \mathcal{G}(W)$$

By the sheaf properties, $\theta(U)|_V = 0$ on $V$. Now let $(V_\alpha, \psi_\alpha)$ be a chart in $X$, $x = \psi_\alpha(p)$ and $U \in \mathcal{G}(X)$. Then for $y$ in a neighborhood of $x$

$$\begin{aligned}(U \circ \psi_\alpha^{-1})(y) &= (U \circ \psi_\alpha^{-1})(x) + \int_0^1 \frac{d}{dt}(U \circ \psi_\alpha^{-1})(x + t(y-x))\, dt \\ &= (U \circ \psi_\alpha^{-1})(x) + \sum_{i=1}^n (y^i - x^i) \int_0^1 D_i(U \circ \psi_\alpha^{-1})(x + t(y-x))\, dt\,.\end{aligned}$$

Hence in a neighborhood of $p$ $(q = \psi_\alpha^{-1}(y))$

$$U(q) \;=\; U(p) + \sum_{i=1}^n (\psi_\alpha^i(q) - \psi_\alpha^i(p))\, g_i(q)\,,$$

where $g_i$ is given by the integral above whence, in particular, $g_i(p) = \frac{\partial}{\partial x^i}(U \circ \psi_\alpha^{-1})|_x$. Consequently

$$(\theta(U))(p) \;=\; \sum_{i=1}^n \frac{\partial}{\partial x^i}(U \circ \psi_\alpha^{-1})|_x\, \theta(\psi_\alpha^i)(p)$$

and we define $\Xi$ locally to be given by $\Xi_\alpha^i = \theta(\psi_\alpha^i)$. It is easily checked that this indeed defines a consistent family in the sense of (73). □

### 4.C. A Generalized Curvature Framework

In this section we are going to set up what may be called a "generalized curvature framework." That is, we define the notions of a generalized metric and a generalized connection and prove a "Fundamental Theorem of Riemannian Geometry" in this setting. Furthermore we define the generalized Riemannian curvature tensor, Ricci and Einstein tensor for a generalized metric and give some consistency results with respect to the classical resp. distributional theory. Here we assume all sections to be real valued.

**4.42** We start with some motivating remarks concerning the definition of a generalized metric. Since we are interested in handling "singular" metrics in general relativity we want a generalized metric to be a sequence of classical (smooth) metrics that converge to a distributional one. On the other hand we certainly want a generalized metric to be invertible in the generalized sense. Unfortunately, given a sequence of smooth metrics we cannot infer its invertibility in $\mathcal{G}^0_2$ solely from the fact that for any fixed $\varepsilon$ the determinant of the metric is nonvanishing, not even if it additionally converges weakly to some distributional (non-degenerate) metric. The basic reason of course is that generalized numbers do not constitute a field but merely a ring; not every nonzero element may be inverted. A simple counterexample was already given in 3.27 showing that a generalized number, even if not associated to 0, will not be invertible in general. A sufficient condition for invertibility, however, was given in 3.28 by the notion of a generalized number to be *strictly nonzero.* This finally motivates our choice of definition.

**4.43 Definition.**

(i) *A generalized* $(0,2)$ *tensor field* $\hat{G} \in \mathcal{G}^2_0(X)$[8] *is called a* generalized (Pseudo-Riemannian) metric *if it has a representative* $(\hat{g}_\varepsilon)_\varepsilon$ *satisfying*

(a) $\hat{g}_\varepsilon$ *is a smooth (Pseudo-Riemannian) metric for all* $\varepsilon$*, and*

(b) $(\det \hat{g}_\varepsilon)_\varepsilon$ *is strictly nonzero on compact sets, i.e.,* $\forall K \subset\subset X\ \exists m \in \mathbb{N}$*:* $\inf_{p\in K} |\det \hat{g}_\varepsilon(p)| \geq \varepsilon^m$.

(ii) *We call a manifold* $M$ *furnished with a generalized Pseudo-Riemannian metric* $\hat{G}$ generalized Pseudo-Riemannian manifold *or* generalized spacetime *and denote it by* $(M,\hat{G})$ *or merely by* $M$*. The action of the metric on a pair of generalized vector fields will be denoted by* $\hat{G}(\Xi, H)$ *and* $\langle \Xi, H\rangle$*, equivalently (if there is no danger of misinterpreting* $\langle\,,\,\rangle$ *as distributional action).*

A sufficient condition for a sequence of classical (smooth) metrics to constitute a representative of a generalized one—besides being moderate, of course—is to be zero-associated (i.e., to converge locally uniformly) to a classical (then necessarily continuous metric). We shall quickly convince ourselves that definition (i) above indeed guarantees the invertibility of the metric.

**4.44 Lemma.** *Let* $(M,\hat{G})$ *be a generalized Pseudo-Riemannian manifold then the inverse metric* $\hat{G}^{-1} \in \mathcal{G}^2_0(M)$ *exists and is given by* $\mathrm{cl}[(\hat{g}_\varepsilon^{-1})_\varepsilon]$*, where* $(\hat{g}_\varepsilon)_\varepsilon$ *is a representant of* $\hat{G}$ *according to definition 4.43 (i).*

**Proof.** Let $(\hat{g}_\varepsilon)_\varepsilon$ be like in 4.43 (i). Then in a chart for fixed $\varepsilon$ we may define (in usual notation) $\hat{g}^{ij}{}_\varepsilon$ to be the pointwise inverse of $\hat{g}_{ij\,\varepsilon}$ which is obviously a smooth $(2,0)$-tensor field. It remains to show that $(\hat{g}^{ji}{}_\varepsilon)_\varepsilon$ is moderate. By the cofactor formula of matrix inversion we have $\hat{g}^{ij}{}_\varepsilon = \mathrm{cof}(\hat{\mathrm{g}}_{\mathrm{ij}\,\varepsilon})/(\det \hat{\mathrm{g}}_{\mathrm{ij}\,\varepsilon})$ and the only trouble arises from the denominator. But for all $K \subset\subset M$ (which again we may assume to be contained in a single chart) we have from condition (b) on $\hat{G}$

$$\sup_{p\in K} \Big|\frac{1}{(\det \hat{g}_{ij\,\varepsilon})(p)}\Big| \;=\; \frac{1}{\inf\limits_{p\in K} |(\det \hat{g}_{ij\,\varepsilon})(p)|} \;=\; O(\varepsilon^{-m})\,.$$

The Lie derivatives of $\hat{g}^{ij}{}_\varepsilon$ may be estimated similarly. □

---

[8] Here we deviate from our convention of denoting generalized functions respectively tensor fields merely by capital letters since it would have led into a notational catastrophe in case of the Christoffel symbols, the Riemann and Ricci tensor etc. All generalized quantities in the context of curvature will be denoted by the usual (capital) letter plus a hat, the only exception being the Einstein tensor for obvious reasons (cf. definition 4.52 (iv) below).

From now on we denote the inverse metric (using abstract indices) by $\hat{G}^{ab}$, its components by $\hat{G}^{ij}$ and the components of a representative by $\hat{g}^{ij}{}_{\varepsilon}$.
Since taking the determinant is a polynomial (hence *nonlinear)* operation we cannot expect association to be compatible with inverting a metric. However, the analogous statement for $k$-association holds by an application of proposition 4.17 (ii).

**4.45 Proposition.** *Let $\hat{G}_{ab}$ a generalized metric and $G_{ab} \approx_k g_{ab}$, where $g_{ab}$ is a classical $\mathcal{C}^k$-Pseudo-Riemannian metric then $\hat{G}^{ab} \approx_k g^{ab}$.*

Next we state and prove some technical results needed in the sequel.

**4.46 Lemma.** *Let $(M, \hat{G})$ be a generalized Pseudo-Riemannian manifold.*

*(i) $\hat{G}$ is non-degenerate in the following sense: $\hat{G}(\Xi, H) = 0\ \forall H \in \mathcal{G}^1_0(M) \Rightarrow \Xi = 0$.*

*(ii) $\hat{G}$ is symmetric, i.e., $\hat{G}(\Xi, H) = \hat{G}(H, \Xi)\ \forall \Xi, H \in \mathcal{G}^1_0(M)$.*

*(iii) $\hat{G}$ induces a $\mathcal{G}(X)$-linear isomorphism $\mathcal{G}^1_0(M) \to \mathcal{G}^0_1(M)$ by $\Xi \mapsto \hat{G}(\Xi, .)$.*

**Proof.** (i) We have to show that for $\hat{G} = \mathrm{cl}[(\hat{g}_\varepsilon)_\varepsilon]$, $\Xi = \mathrm{cl}[(\xi_\varepsilon)_\varepsilon]$ $(\hat{g}_\varepsilon(\xi_\varepsilon, \eta_\varepsilon))_\varepsilon \in \mathcal{N}(M)$ for all $H = \mathrm{cl}[(\eta_\varepsilon)_\varepsilon]$ implies $(\xi_\varepsilon)_\varepsilon \in \mathcal{N}^1_0(M)$. It suffices to establish the claim locally, i.e.,

$$\big(\hat{g}_{ij\ \varepsilon}\xi^i_\varepsilon \eta^j_\varepsilon\big)_\varepsilon \in \mathcal{N}(V_\alpha) \ \Rightarrow\ (\xi^i{}_\varepsilon)_\varepsilon \in \mathcal{N}(V_\alpha) \quad i = 1 \ldots n\,.$$

Setting $\eta^j{}_\varepsilon = \sum_l \hat{g}^{jl}{}_\varepsilon\, \xi^l_\varepsilon$ gives $(\sum_{ij} \hat{g}_{ij\ \varepsilon}\xi^i_\varepsilon \eta^j_\varepsilon) = (\xi^i{}_\varepsilon)^2$ which finishes the proof, since every sum of positive negligible functions is negligible iff every term is.
(ii) is easily shown using the (symmetric) representative of definition 4.43 (i).
(iii) Due to the $\mathcal{G}(X)$-bilinearity of the metric $\Xi^* := \hat{G}(\Xi, .)$ is indeed a one form and the assignment $\Xi \mapsto \Xi^*$ is $\mathcal{G}(X)$-linear. Moreover, injectivity of the map follows directly from the non-degeneracy of $\hat{G}$ in the sense of (i) above. It remains to show that the assignment is onto. Locally we may write $\Xi^* = \Xi^*_i dx^i$ and define a generalized vector field by $V = \hat{G}^{jk}\, \Xi^*_j\, \partial_k$. Then

$$\langle V, \partial_l \rangle \ =\ \hat{G}^{jk}\, \Xi^*_k\, \langle \partial_j, \partial_l \rangle \ =\ \Xi^*_l\,.$$

□

**4.47** The isomorphism in (iii) above—as in the classical context—extends naturally to generalized tensor fields of higher types. Hence from now on we shall use the common conventions on upper and lower indices also in the context of generalized tensor fields. In particular, identifying a vector field $\Xi^a \in \mathcal{G}^1_0(M)$ with its metrically equivalent one-form $\Xi_a$ we denote its contravariant respectively covariant indices by $\Xi^i$ and $\Xi_i$. A similar convention will apply to representatives.

Now we introduce the generalized version of a connection and its Christoffel symbols.

**4.48 Definition.**

*(i) A* generalized connection *$\hat{D}$ on a manifold $X$ is a map $\mathcal{G}^1_0(X) \times \mathcal{G}^1_0(X) \to \mathcal{G}^1_0(X)$ satisfying*

*(D1) $\hat{D}_\Xi H$ is $\mathcal{R}$-linear in $H$.*

*(D2) $\hat{D}_\Xi H$ is $\mathcal{G}(X)$-linear in $\Xi$.*

*(D3) $\hat{D}_\Xi(UH) = U\,\hat{D}_\Xi H + \Xi(U)H$ for all $U \in \mathcal{G}(X)$.*

*(ii) Let $(V_\alpha, \psi_\alpha)$ be a chart on $X$ with coordinates $x^i$. We define the* generalized Christoffel symbols *for this chart to be the $n^3$ functions $\hat{\Gamma}^k_{ij} \in \mathcal{G}(V_\alpha)$ given by*

$$\hat{D}_{\partial_i}\partial_j \ =\ \sum_k \hat{\Gamma}^k_{ij}\,\partial_k \quad 1 \le i, j \le n\,.$$

We are now already in the position to prove the "Fundamental Theorem of Riemannian Geometry" in our setting.

**4.49 Theorem.** *Let $(M, \hat{G})$ be a generalized Pseudo-Riemannian manifold. Then there exists a unique generalized connection $\hat{D}$ such that*

*(D4)* $[\Xi, H] = \hat{D}_\Xi H - \hat{D}_H \Xi$ *and*

*(D5)* $\Xi\langle H, Z\rangle = \langle \hat{D}_\Xi H, Z)\rangle + \langle H, \hat{D}_\Xi Z\rangle$

*hold for all* $\Xi$, $H$, $Z$ *in* $\mathcal{G}^1_0(M)$. $\hat{D}$ *is called* generalized Levi-Civita connection *of $M$ and characterized by the—so called—Koszul formula*

$$2\langle \hat{D}_\Xi H, Z\rangle \;=\; \Xi\langle H, Z\rangle + H\langle Z, \Xi\rangle - Z\langle \Xi, H\rangle - \langle \Xi, [H, Z]\rangle + \langle H, [Z, \Xi]\rangle + \langle Z, [\Xi, H]\rangle\,. \tag{74}$$

**Proof.** Assume $\hat{D}$ to be a generalized connection additionally satisfying (D4) and (D5). As in the classical proof (see e.g. [147], §3, theorem 11) using the latter two properties one shows that equation (74) is just an identity. Hence $\hat{D}$ satisfies (74) and by the non-degeneracy of $\hat{G}$ (more precisely by the injectivity of the map in lemma 4.46 (iii)) is unique.
To show existence define $F(\Xi, H, Z)$ to be one half of the the right hand side of (74). Then for fixed $H, Z$ the function $\Xi \mapsto F(\Xi, H, Z)$ is $\mathcal{G}(X)$-linear, hence defines a generalized one-form (theorem 4.31). Again by lemma 4.46 (iii) there exists a unique generalized vector field associated with it which we may call $\hat{D}_\Xi H$. Now it is easy to derive (D1)-(D5) along the lines of the classical proof just using the bilinearity of $\hat{G}$ and the standard properties of the Lie bracket (cf. the remark after definition 4.39). □

As in the classical case from the torsion-free condition (i.e., (D4)) we immediately infer the symmetry of the Christoffel symbols of the Levi-Civita connection in the lower pair of indices. Moreover, from (D3) and the Koszul formula (74) we derive (again analogously to the classical case) the following

**4.50 Proposition.** *Given a chart as in definition 4.48 (ii) we have for the generalized Levi-Civita connection $\hat{D}$ of $(M, \hat{G})$ and any vector field $\Xi \in \mathcal{G}^1_0(X)$*

$$\hat{D}_{\partial_i}(\Xi^j \partial_j) \;=\; \left(\frac{\partial\, \Xi^k}{\partial x^i} + \hat{\Gamma}^k_{ij}\, \Xi^j\right) \partial_k\,.$$

*Moreover, the generalized Christoffel symbols are given by*

$$\hat{\Gamma}^k_{ij} \;=\; \frac{1}{2}\, \hat{G}^{km} \left(\frac{\partial \hat{G}_{jm}}{\partial x^i} + \frac{\partial \hat{G}_{im}}{\partial x^j} - \frac{\partial \hat{G}_{ij}}{\partial x^m}\right).$$

Now we see that we could equivalently have introduced the generalized Christoffel symbols of a generalized metric by just demanding the classical formula on the level of representatives. However, we have chosen to follow the more elegant way introducing them via a generalized connection.
To state the next result concerning consistency properties of generalized connections resp. generalized Christoffel symbols with respect to their classical counterparts we need to define the action of a classical (smooth) connection $D$ on generalized vector fields $\Xi$, $H$. This is easily done by setting

$$D_\Xi H \;:=\; \mathrm{cl}[(D_{\xi_\varepsilon}\eta_\varepsilon)_\varepsilon]\,.$$

From the local formulae in the above proposition we infer

**4.51 Proposition.** *Let $(M, \hat{G})$ be a generalized Pseudo-Riemannian manifold.*

*(i) If $\hat{G}_{ab} = \Sigma(g_{ab})$ where $g_{ab}$ is a classical smooth metric then we have, in any chart, $\hat{\Gamma}^i_{jk} = \Sigma(\Gamma^i_{jk})$ (with $\Gamma^i_{jk}$ denoting the Christoffel Symbols of $g_{ab}$). Hence for all $H \in \mathcal{G}^1_0(M)$*

$$\hat{D}_\Xi H \;=\; D_\Xi H\,.$$

*(ii) If $\hat{G}_{ab} \approx_\infty g_{ab}$, $g_{ab}$ a classical smooth metric, $\Xi$, $H \in \mathcal{G}^1_0(M)$ and $\Xi \approx_\infty \xi \in \mathcal{T}^1_0(M)$, $H \approx \eta \in \mathcal{D}'^1_0(M)$, or vice versa, i.e., $\Xi \approx \xi \in \mathcal{D}'^1_0(M)$, $H \approx_\infty \eta \in \mathcal{T}^1_0(M)$, then*

$$\hat{D}_\Xi H \approx D_\xi \eta .$$

*(iii) Let $\hat{G}_{ab} \approx_k g_{ab}$, $g_{ab}$ a classical $\mathcal{C}^k$-metric, then, in any chart, $\hat{\Gamma}^i_{jk} \approx_{k-1} \Gamma^i_{jk}$. If in addition $\Xi$, $H \in \mathcal{G}^1_0(M)$, $\Xi \approx_{k-1} \xi \in \Gamma^{k-1}(M, TM)$ and $H \approx_k \eta \in \Gamma^k(M, TM)$ then*

$$\hat{D}_\Xi H \approx_{k-1} D_\xi \eta .$$

Next we introduce the generalized Riemann, Ricci, scalar and Einstein curvature from an invariant point of view. It is then clear that all the classical formulae will hold on the level of representatives, i.e., all the symmetry properties of the respective classical tensor fields carry over to our setting. Moreover, the Bianchi identities hold in the generalized sense.

**4.52 Definition.** *Let $(M, \hat{G})$ be a generalized Pseudo-Riemannian manifold with Levi-Civita connection $\hat{D}$.*

*(i) We define the* generalized Riemannian curvature tensor $\hat{R}_{abc}{}^d \in \mathcal{G}^1_3(M)$ *by*

$$\hat{R}_{\Xi,H} Z := \hat{D}_{[\Xi,H]} Z - [\hat{D}_\Xi, \hat{D}_H] Z .$$

*(ii) We define the* generalized Ricci curvature tensor $\hat{R}_{ab} \in \mathcal{G}^0_2(M)$ *by the usual contraction of the generalized Riemann tensor*

$$\hat{R}_{ab} := \hat{R}^c{}_{abc} .$$

*(iii) The* generalized curvature scalar (or Ricci scalar) $\hat{R} \in \mathcal{G}(M)$ *is defined by the usual contraction of the generalized Ricci tensor*

$$\hat{R} := \hat{R}^a{}_a .$$

*(iv) Finally we define the generalized Einstein tensor $\hat{\mathfrak{G}}_{ab} \in \mathcal{G}^0_2(M)$ by*

$$\hat{\mathfrak{G}}_{ab} := \hat{R}_{ab} - \frac{1}{2} \hat{G}_{ab} .$$

From the coordinate formulae of the above defined fields we infer

**4.53 Theorem.** *Let $(M, \hat{G})$ a generalized Pseudo-Riemannian manifold with $\hat{G}_{ab} \approx_k g_{ab}$ then the all the generalized curvature quantities defined above will be $\mathcal{C}^{k-2}$-associated to their classical counterparts.*

In particular, if $\hat{G}_{ab}$ is $\mathcal{C}^2$-associated to a vacuum solution of Einstein equations then we have for the generalized Ricci tensor

$$\hat{R}_{ab} \approx_0 R_{ab} = 0 .$$

Hence $\hat{R}_{ab}$ satisfies the vacuum Einstein equations in the sense of 0-approximation (cf. the remarks after proposition 18 in the revised version of [184]).

**4.54** Finally, we may sum up the contents of this entire section by saying that Colombeau's generalized functions indeed provide a framework suited to the needs of general relativity. We have defined the notion of a generalized metric which allows for the description of a wide class of classical "singular" metrics. All the curvature quantities of a generalized metric are defined within our setting allowing for a mathematically rigorous formulation of the field equations.
Given a classical metric of low differentiability—which of course need not to be gt-regular (cf. sec. 2.B)—our framework may be applied in the following way: First one has to transfer the singular metric into a generalized one. This may be done by some "canonical" smoothing or by some other physically motivated

regularization (see also the section on nonlinear modelling 3.E). Of course the question of diffeomorphism invariance of such a procedure has to be carefully investigated. Then, once the generalized setting has been entered, all curvature quantities may be calculated simply using componentwise classical calculus. All classical concepts literally carry over to the new framework and one may treat e.g. the Ricci tensor, geodesics, geodesic deviation, etc. in the algebra. Finally one may use coupled calculus to return to the distributional (or $\mathcal{C}^k$-)level for the purpose of interpretation.
This program has been carried out for the cone-metric (in the full Colombeau algebra) by the Clarke, Vickers and Wilson [36,185,189]. They were able to rigorously assign to the (not gt-regular) metric a distributional curvature and (via the field equations) the heuristically expected energy-momentum tensor.
In the next chapter we are going to use the above developed framework to distributionally describe the geometry of impulsive gravitational waves. But before—in the last section of the present chapter—we summarize the current research in the field of diffeomorphism invariant Colombeau algebras.

### 4.D. Diffeomorphism invariant Colombeau algebras

The algebra $\mathcal{G}(\Omega)$ (with $\Omega$ open in $\mathbb{R}^n$) as presented in chapter 3 above has been introduced in order to allow for an unrestricted application of a large variety of nonlinear operations to spaces of distributions. Using sheaf theoretic methods in sec. 3.B we have constructed an embedding of $\mathcal{D}'(\Omega)$ into $\mathcal{G}(\Omega)$ that is unique and canonical, once a fixed mollifier $\rho$ is chosen. However, passing from $\mathbb{R}^n$ to smooth manifolds in sec. 4.A we rejected this sort of embedding for it is not diffeomorphism invariant, still constructing a nonlinear generalized framework well-suited to applications especially to the ones presented in the final chapter below. This section, on the other hand, is devoted to an overview of algebras of generalized functions indeed allowing for a diffeomorphism invariant embedding of distributions, which—in a sloppy way—we address as *diffeomorphism invariant Colombeau algebras.* To this end we have to go back to the theory on open sets of Euclidean space and introduce Colombeau's "elementary" algebra $\mathcal{G}^e(\Omega)$ (as presented in [39]; hence the name), which, to be sure, is the standard version among those being independent from the choice of a particular approximation of the delta-distribution. We refer to algebras of that type as *full* algebras as opposed to the *special* algebras treated so far.
Further we pave our way towards the diffeomorphism invariant algebras of [61] and [76], passing by the algebras of Colombeau and Meril [44] as well as that of Jelínek [99] on which the final construction is based upon. We conclude by reviewing the global formulation of the (standard) diffeomorphism invariant algebra $\mathcal{G}^d$ presented in [77] best suited to applications in general relativity.

**4.55** As pointed out several times, in the special version of Colombeau's algebra on open sets of $\mathbb{R}^n$ (as presented in chap. 3) the embedding of distributions depends on a particular choice of a mollifier (see sec. 3.B). However, we have also repeatedly seen (cf. sec. 3.E) that this dependence of nonlinear properties of singular objects on the chosen regularization is unavoidable as it expresses a fundamental principle of nonlinear modelling. Moreover, all of the results we have treated so far are concerned with those properties of distributions in $\mathcal{G}$ that hold for *each* such embedding, irrespective of the particular form of $\rho$. Nevertheless, there is a way of handling these dependencies by modifying (in fact: enlarging) the algebra $\mathcal{G}$ to the so-called "elementary" algebra $\mathcal{G}^e(\Omega)$ of [39]. This process will lead to a canonical embedding of $\mathcal{D}'(\Omega)$ into $\mathcal{G}^e(\Omega)$, while all the properties of $\mathcal{G}(\Omega)$ we have studied so far can be preserved. The prize one has to pay for such improvements is an increased load of technical apparatus that sometimes—especially in applications—tends to obstruct the view of what is really going on. However, since in the present context we are interested in $\mathcal{G}^e$ as a pioneer for the diffeomorphism invariant algebras, we give a brief outline of the construction and refer to the literature (e.g. [39], chaps. 1,2, [144], chap. III and [117], chap.1.7) for further details.

As was already indicated in the discussion at the beginning of this chapter the key idea will be a substitution of the index set $(0, 1]$ by a suitable class of mollifiers. This time, the mollifiers will be chosen from $\mathcal{D}$, which facilitates many proofs while necessitating the definition of the following grading on the parameter set (sometimes called *kernel spaces*) in order to secure an analogue of 3.10 (i) and (ii) (both properties cannot hold simultaneously for elements of $\mathcal{D}$, see 3.10).

**4.56 Definition.**

$$\mathcal{A}_0(\mathbb{R}^n) := \{\varphi \in \mathcal{D}(\mathbb{R}^n) \, : \, \int \varphi(x)\, dx = 1\}$$
$$\mathcal{A}_q(\mathbb{R}^n) := \{\varphi \in \mathcal{A}_0(\mathbb{R}^n) \, : \, \int \varphi(x) x^\alpha \, dx = 0\,,\, 1 \le |\alpha| \le q\} \;\; (q \in \mathbb{N})$$

It is easy to show (cf. [144], §9) that $\mathcal{A}_q \neq \emptyset$ for all $q$, $\mathcal{A}_{q+1} \subseteq \mathcal{A}_q$ and $\bigcap_{q\in\mathbb{N}} \mathcal{A}_q = \emptyset$. Next we define the basic space as well as the subspaces of moderate resp. negligible elements in the "elementary" setting of [39], yet following the more straight forward presentation of [76]. Again we write $\varphi_\varepsilon(x) := (1/\varepsilon^n)\varphi(x/\varepsilon)$.

**4.57 Definition.** *Let $\Omega \subseteq \mathbb{R}^n$ open.*

$$U(\Omega) := \{(\varphi, x) \in \mathcal{A}_0(\mathbb{R}^n) \times \Omega \, : \, \mathrm{supp}(\varphi) \subseteq \Omega - \mathrm{x}\}$$
$$\mathcal{E}^e(\Omega) := \{R : U(\Omega) \to \mathbb{C} \, : \, \forall (\varphi, x) \in U(\Omega) \textit{ the map } y \mapsto R(\varphi, y) \textit{ is smooth in a neighbourhood of } x\,\}$$
$$\mathcal{E}^e_M(\Omega) := \{R \in \mathcal{E}(\Omega) : \forall K \subset\subset \Omega \;\forall \alpha \in \mathbb{N}_0^n \;\exists p \in \mathbb{N}_0 \;\forall \varphi \in \mathcal{A}_p(\mathbb{R}^n) \;\sup_{x\in K} |\partial^\alpha R(\varphi_\varepsilon, x)| = O(\varepsilon^{-p}) \textit{ as } \varepsilon \to 0\}$$
$$\mathcal{N}^e(\Omega) := \{R \in \mathcal{E}(\Omega) : \forall K \subset\subset \Omega \;\forall \alpha \in \mathbb{N}_0^n \;\forall p \in \mathbb{N}_0 \;\exists q \;\forall \varphi \in \mathcal{A}_q(\mathbb{R}^n) : \sup_{x\in K} |\partial^\alpha R(\varphi_\varepsilon, x)| = O(\varepsilon^{p}) \textit{ as } \varepsilon \to 0\}$$

*The* "elementary" Colombeau Algebra *of generalized functions on $\Omega$ is defined as the quotient space*

$$\mathcal{G}^e(\Omega) := \mathcal{E}^e_M(\Omega) \,/\, \mathcal{N}^e(\Omega)\,.$$

As in the special version of the theory $\mathcal{E}^e_M$ is a differential algebra with operations defined componentwise (now for fixed $\varphi \in \mathcal{A}_0$) and $\mathcal{N}^e$ is a differential ideal in it. Hence $\mathcal{G}^e$ indeed is a differential algebra and, as in the previous cases, even a fine sheaf of differential algebras. Note that although the above defined index sets $\mathcal{A}_q$ depend on the choice of a basis of $\mathbb{R}^n$ there is a way of defining $\mathcal{G}^e(\Omega)$ avoiding such a choice ( [9], see also [22]) which we have left aside to avoid additional technical complications.
Analogously to section 3.B we define the "constant" embedding of smooth functions by $\sigma(f)(\varphi, x) = f(x)$ as well as the embedding $\iota$ for distributions by convolution with mollifiers. For reasons of consistency with later constructions, however, we replace $w * \varphi$ by $w * \check{\varphi}$ (i.e., $w * \check{\varphi}(x) = \langle w, \varphi(.-x)\rangle$) from now on. This also reveals the motivation behind the definition of $U(\Omega)$ above. The latter consist of pairs $(\varphi, x)$ of test functions and points, where the test function $\varphi$ has a favorable support w.r.t the point $x$, in the sense that convolution of $\check{\varphi}$ with any $w \in \mathcal{D}'(\Omega)$ is defined for all $x \in \Omega$. Indeed it is now easy to prove the following

**4.58 Proposition.** *The map*

$$\iota : \mathcal{D}'(\Omega) \to \mathcal{G}^e(\Omega)$$
$$w \to \mathrm{cl}[(w * \check{\varphi})_{\varphi \in \mathcal{A}_0}]$$

*is linear embedding coinciding with $\sigma$ on $\mathcal{C}^\infty(\Omega)$.*

Thus the advantage of $\mathcal{G}^e(\Omega)$ over $\mathcal{G}(\Omega)$ is that we can embed the whole of $\mathcal{D}'(\Omega)$ using the convolution formula; no sheaf theory is needed.
We again remark that all the concepts like e.g. point values, integration, coupled calculus etc. introduced in the special setting in chap. 3 carry over to the "elementary" setting with—more or less—obvious modifications. In fact, most monographs on Colombeau algebras first introduce the "elementary" setting before "specializing" to the special setting.
However, we do not go into any further details here but continue our journey towards diffeomorphism invariance by discussing the central issue of the action of a diffeomorphism on a Colombeau generalized function, since already in the present setting the basic problem may be seen clearly.

**4.59** Let $\mu : \tilde{\Omega} \to \Omega$ denote a diffeomorphism of open sets of $\mathbb{R}^n$. We try to tailor the definition of the action $R \mapsto R \circ \mu$ of a diffeomorphism in such a way to guarantee that $\iota(f) \circ \mu = \iota(f \circ \mu)$ at least for continuous functions $f$. We have for $\tilde{x} \in \tilde{\Omega}$

$$\iota(f)(\varphi,\mu(\tilde{x})) \;=\; (f*\check{\varphi})(\mu(\tilde{x})) \;=\; \int f(\xi+\mu(\tilde{x}))\varphi(\xi)\,d\xi\,,$$

whereas on the other hand

$$\begin{aligned}\iota(f\circ\mu)(\varphi,\tilde{x}) = f\circ\mu*\check{\varphi}(\tilde{x}) \;&=\; \int f(\mu(\eta+\tilde{x}))\varphi(\eta)\,d\eta\\ &= \int f(\xi+\mu(\tilde{x}))\varphi(\mu^{-1}(\xi+\mu(\tilde{x}))-\tilde{x})|\det D\mu^{-1}(\xi+\mu(\tilde{x}))|\,d\eta\,,\end{aligned}$$

where we have substituted $\mu(\eta+\tilde{x}) = \xi+\mu(\tilde{x})$. Now we see that the diffeomorphism cannot act in the simple form $\iota(f)(\varphi,\mu(\tilde{x}))$ but *has to act on $\varphi$ also!* Consequently the "elementary" algebra where the $\varphi$'s are treated as parameters only has no chance of being diffeomorphism invariant.
On the other hand, the above considerations clearly lead the way to the definition of the action of a diffeomorphism on a Colombeau function $R(\varphi,x)$; if we wish to set it up in a functorial way, i.e., $(\hat{\mu}R)(\tilde{\varphi},\tilde{x}) = R(\bar{\mu}(\tilde{\varphi},\tilde{x}))$ we have no choice but to define

**4.60 Definition.** *Let $\mu:\tilde{\Omega}\to\Omega$ a diffeomorphism of open subsets of $\mathbb{R}^n$. We define*

$$\begin{aligned}&\bar{\mu}:U(\tilde{\Omega})\to U(\Omega)\\ &\quad(\tilde{\varphi},\tilde{x})\mapsto\bar{\mu}(\tilde{\varphi},\tilde{x}) := \big(\tilde{\varphi}(\mu^{-1}(.+\mu(\tilde{x}))-\tilde{x})\cdot|\det D\mu^{-1}(.+\mu(\tilde{x}))|,\mu(\tilde{x})\big)\;.\end{aligned}\tag{75}$$

**4.61** Colombeau and Meril in their paper [44] (using earlier ideas of [38]) made the first decisive steps to incorporate formula (75) into the construction of a Colombeau algebra which they claimed to be diffeomorphism invariant. Before discussing their work in some more detail let us introduce some terminology (cf. [61], sec. 9 ) which eases the understanding of the definitions to be given below.
Every *full* Colombeau algebra is constructed as a quotient space of moderate modulo negligible sequences (nets) of (smooth) functions $R$ belonging to some *basic space* usually denoted by $\mathcal{E}$ (plus some superscript to distinguish the algebras to be constructed). The respective properties of moderateness and negligibility are then defined by inserting *scaled test objects* (e.g. $\varphi_\varepsilon$ with $\varphi\in\mathcal{A}_0$ resp. $\mathcal{A}_q$ in definition 4.57 above) into $R$ and analyzing the asymptotic behavior of the latter on these "paths" as the scaling parameter $\varepsilon$ tends to zero (and consequently $\varphi_\varepsilon\to\delta$ weakly); we shall refer to the respective processes as *testing for moderateness resp. negligibility.* In this terminology diffeomorphism invariance of a Colombeau algebra is ensured by diffeomorphism invariance of the respective tests, of course including diffeomorphism invariance of the respective class of (scaled) test objects.
As opposed to this testing procedure the elements of the algebra themselves *do not depend in any way from $\varepsilon$* (quite contrary to the special setting). We regard this distinction as fundamental clarifying several misinterpretations in the literature and call it following [61] the policy of *"separating definitions from testing."*

Also in [61], sec. 3 there was given a blueprint collecting all the definitions and theorems necessary for the construction of a (full) Colombeau algebra. In the following we shall use this collection as a guiding line in discussing the various variants of the algebra proposed in the literature beginning with the one of Colombeau and Meril [44].

**4.62** There are basically three modifications introduced by the authors of [44] distinguishing their construction, which—following [76]—we call $\mathcal{G}^1$ from $\mathcal{G}^e$, namely:

(i) Smooth dependence of $R$ on $\varphi$ in place of arbitrary dependence.

(ii) Dependence of test objects on $\varepsilon$, i.e., bounded paths $\varepsilon\mapsto\phi(\varepsilon)\in\mathcal{D}(\Omega)$.

(iii) Asymptotically vanishing moments (see below) of test objects as compared to the stronger condition $\phi(\varepsilon)\in\mathcal{A}_q(\mathbb{R}^n)$ for all $\varepsilon$ (which is the naive analog of $\varphi\in\mathcal{A}_q(\mathbb{R}^n)$ in the case of $\mathcal{G}^e$).

We briefly comment on these items (for a full discussion see [76]). First, (i) is necessary to guarantee smoothness of $\hat{\mu}R$ with respect to $\tilde{x}$ (cf. transformation (75)). However, the technical prize to pay here is the use of calculus in infinite dimensional spaces. Since these matters—although a central issue in the constructions—clearly lie beyond the scope of the present work we do not go into any further detail but remark that Colombeau and Meril in particular used the concept of Silva-differentiability [37]. However,

instead of giving the proofs they rather "invited the reader to admit"( [44], p. 362) the respective smoothness properties.
Change (ii) together with (iii) obviously was introduced to obtain a diffeomorphism invariant analog of the vanishing moment condition of definition 4.56. More precisely, following [44] we define $\widetilde{\mathcal{A}}_q(\mathbb{R}^n)$ to be the set of all smooth, bounded paths $\varepsilon \mapsto \phi(\varepsilon)$ satisfying

$$\int \phi(\varepsilon)(\xi)\, d\xi = 1 \quad \forall \varepsilon \in (0,1] \text{ and}$$
$$\int x^\alpha \phi(\varepsilon)(\xi)\, d\xi = O(\varepsilon^q) \quad \forall \alpha \in \mathbb{N}_0^n \text{ with } 1 < |\alpha| \le q\,.$$

It may now be shown ( [44], §3) that these moment conditions indeed are invariant under the action of a diffeomorphism as defined in 4.60.

**4.63** Colombeau and Meril chose their basic space to be $\mathcal{E}^1(\Omega) := \{R:\ \widetilde{\mathcal{A}}_0(\mathbb{R}^n)\times\Omega \to \mathbb{C}^{(0,1]}\}$. Note that this definition clearly mixes definition of basic objects and testing procedure, hence is not in accordance with the policy of "separating definitions from testing" as propagated above. Moreover, their definition of the objects constituting the Colombeau algebra was not unambiguous. However, following the interpretations of [99] and [76], the testing process in [44] is defined by inserting test objects of the form $S_\varepsilon\phi(\varepsilon) := (1/\varepsilon)^n\varphi(\varepsilon)(./\varepsilon)$ into the first slot of $R$. More precisely,

$$\mathcal{E}^1_M(\Omega) := \{R \in \mathcal{E}^1(\Omega):\ \forall K \subset\subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \exists p \in \mathbb{N}\ \forall \phi \in \widetilde{\mathcal{A}}_p\ \sup_{x\in K} |\partial^\alpha R(S_\varepsilon\phi(\varepsilon), x)| = O(\varepsilon^{-p}) \text{ as } (\varepsilon \to 0)\}$$
$$\mathcal{N}^1(\Omega) := \{R \in \mathcal{E}^1(\Omega):\ \forall K \subset\subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \forall p \in \mathbb{N}\ \exists q\ \forall \phi \in \widetilde{\mathcal{A}}_q\ \sup_{x\in K} |\partial^\alpha R(S_\varepsilon\phi(\varepsilon), x)| = O(\varepsilon^{p})\}$$

Finally the *Colombeau-Meril algebra* on $\Omega$ is defined as the quotient space

$$\mathcal{G}^1(\Omega) := \mathcal{E}^1_M(\Omega)\,/\,\mathcal{N}^1(\Omega)\,.$$

Using these definitions, all the main properties of $\mathcal{G}^e(\Omega)$ carry over to $\mathcal{G}^1(\Omega)$, with almost identical proofs. Indeed, boundedness of the paths $\phi(\varepsilon)$ in $\mathcal{D}(\mathbb{R}^n)$ assures similar estimates as in the case of single functions $\varphi$.

**4.64** Unfortunately, in addition to the ambiguities mentioned above (cf. also [61], sec. 9) the class of test objects as defined by Colombeau and Meril still is not preserved under the action of a diffeomorphism. Nevertheless, despite these defects (which, apparently, went unnoticed by nearly all workers in the field) their construction was quoted and used many times [117], [46], [14], [189], [185], [142], [184], [47], [186], [17], [118]. It was only in 1998 that J. Jelínek in [99] pointed out the error in [44] by giving a (rather simple) counterexample which we shall discuss in a moment. In the same paper, he presented another version of the theory which avoided (some of) the shortcomings of [44] and has to be considered as the second decisive step towards a diffeomorphism invariant version of a Colombeau algebra.

**4.65** Taking a closer look on the nature of test objects as used by Colombeau and Meril, from (75) we see that the action of a diffeomorphism $\mu$ introduces an additional $x$-dependence in the first slot of $R$, which was not there initially. This in turn may be exploited by giving an example of a function in $\mathcal{E}^1_M$ which is constant in $x$ (hence the estimates of the derivatives follow trivially) but whose $\mu$-transform depending on $x$ fails to be moderate, with the derivatives producing factors $e^{1/\varepsilon}$. More precisely, set $R(\phi, x) := \exp(i\exp(\int |\phi(\xi)|^2\, d\xi))$. Then according to 4.60 we have

$$\hat{\mu}R(S_\varepsilon\tilde{\phi}, \tilde{x}) = R(\bar{\mu}(S_\varepsilon\tilde{\phi}, \tilde{x})) = \exp(i\exp(\int |\tilde{\phi}(\frac{\mu^{-1}(\varepsilon\xi + \mu\tilde{x}) - \tilde{x}}{\varepsilon})\det D\mu^{-1}(\varepsilon\xi + \mu\tilde{x})|^2\, d\xi))\,.$$

One now easily verifies that $R \in \mathcal{E}^1_M$, yet $\hat{\mu}R \neq \mathcal{E}^1_M$.

**4.66** We next discuss in some detail the algebra proposed by J. Jelínek in [99] which—again following [76]—we call $\mathcal{G}^d$ [9]. Analogously to the previous construction we start by listing the main features distinguishing $\mathcal{G}^d$ from $\mathcal{G}^1$.

---

[9] "d" obviously stands for diffeomorphism invariant. In fact Jelínek's construction comes very close to diffeomorphism invariance although the last gaps were only closed in [61], however, without the necessity to change the main definitions.

(i) (Smooth) dependence of test objects also on $x \in \Omega$.

(ii) In testing for moderateness test objects may take arbitrary values in $\mathcal{A}_0(\mathbb{R}^n)$, independently of any moment condition.

While in the light of Jelínek's counterexample (i) is compelling there seems to be no apparent necessity for (ii). Apparently (ii) widens the range of test objects which in turn reduces $\mathcal{E}^d_M$ resp. $\mathcal{N}^d$ in size. Yet is has to be admitted that by this reduction no generalized function of interest, neither for the further development of the theory nor in applications is lost. However, M. Grosser in [76] raised the question whether one may construct another diffeomorphism invariant Colombeau algebra $\mathcal{G}^2$ omitting (ii) above, which he could answer in the positive; This algebra can be viewed as the minimal modification of Colombeau's and Meril's algebra $\mathcal{G}^1$ reaching the goal of diffeomorphism invariance. However, leaving aside this construction we focus on the algebra $\mathcal{G}^d$ of J. Jelínek which we regard as the standard diffeomorphism invariant algebra.

**4.67** While Colombeau in his "Elementary Introduction" [39] chose to embed distributions via convolution with a mollifier, i.e., (cf. also 4.58 above)

$$\iota^C(w)(\varphi, x) := \langle w, \varphi(. - x)\rangle\,, \tag{76}$$

Jelínek (following in fact earlier ideas of Colombeau presented in [38]) decided to embed distributions by simply letting them act on the test function, i.e.,

$$\iota^J(w)(\varphi, x) := \langle w, \varphi\rangle\,. \tag{77}$$

Since both embeddings are simply related by a translation, i.e., $\iota^C = T^*\iota^J$ with

$$\begin{aligned} T : \mathcal{D}(\mathbb{R}^n) \times \mathbb{R}^n &\to \mathcal{D}(\mathbb{R}^n) \times \mathbb{R}^n \\ (\varphi, x) &\mapsto (T_x\varphi, x) := (\varphi(. - x), x)\,, \end{aligned}$$

they give rise to equivalent descriptions of virtually every (full) Colombeau algebra, which—once again following [61]—we call *C- resp. J-formalism.* The authors of [61] felt the need to set up a translation formalism allowing to change from one setting to the other at any place of the construction. It was used in turn to clarify the subtle questions of infinite-dimensional calculus arising at the various steps of the construction. Before giving the actual definitions of $\mathcal{G}^d$ we make one further comment on the issue of calculus. Jelínek uses [190] as main reference while the presentation of [61,76] and [77] is based upon the more convenient calculus of [113]. The basic idea of the latter is that a map $f : E \to F$ between locally convex spaces is smooth if it transports smooth curves in $E$ to smooth curves in $F$, where the notion of smooth curves is straightforward (via limits of difference quotients). This notion of smoothness in general is weaker than Silva-differentiability but coincides with the latter on all the spaces used in the construction of Colombeau algebras. Moreover, it displays the following decisive advantage in applications to partial differential equation: If one is to construct a generalized solution to a nonlinear singular equation this is done componentwise (cf. also sec. 5), i.e., for fixed $\varphi$. Smoothness of the respective solution in $\varphi$ is then guaranteed already by classical theorems on smooth dependence of solutions on parameters.

**4.68** We now give a brief description of Jelínek's algebra $\mathcal{G}^d$, however contrary to the original presentation (of course) using the $C$-formalism for its better familiarity. For a comparison of the respective features of the two formalisms we refer to the table in [61], sec. 5. Forced by the choice of the embedding (76) we define the basic space to be [10]

$$\mathcal{E}^C(\Omega) := \mathcal{C}^\infty(U(\Omega))\,.$$

Partial derivatives on $\mathcal{E}^C(\Omega)$—which will become the derivatives in the algebra—in the C-formalism are simply defined by

$$D^C_i : \mathcal{E}^C(\Omega) \to \mathcal{E}^C(\Omega) \qquad D^C_i = \partial_i\,. \tag{78}$$

Recall that test objects have to depend on $\varepsilon$ and $x$, in particular are chosen to be smooth, bounded paths $\phi : (0, 1] \times \Omega \to \mathcal{A}_0(\mathbb{R}^n)$ (resp. $\mathcal{A}_q(\mathbb{R}^n)$). Denoting their space by $\mathcal{C}^\infty_b((0, 1] \times \Omega, \mathcal{A}_0(\mathbb{R}^n))$ we are able to formulate the tests for moderateness and negligibility.

---

[10] Here the superscript $C$ on $\mathcal{E}$ indicates the use of the C-formalism; there is no danger of misunderstanding since we distinguish between C- and J-formalism in case of the algebra $\mathcal{G}^d$ only.

**4.69 Definition.**

*(i)* $R \in \mathcal{E}^C(\Omega)$ *is called* moderate *( [99], 8) if*
$\forall K \subset\subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \exists N \in \mathbb{N}\ \forall \phi \in \mathcal{C}_b^\infty((0,1] \times \Omega, \mathcal{A}_0(\mathbb{R}^n))$ :

$$\sup_{x \in K} |\partial^\alpha (R(S_\varepsilon \phi(\varepsilon, x), x))| = O(\varepsilon^{-N}) \text{ as } \varepsilon \to 0\,.$$

*The set of all moderate elements* $R \in \mathcal{E}^C(\Omega)$ *will be denoted by* $\mathcal{E}_M^C(\Omega)$*.*

*(ii)* $R \in \mathcal{E}_M^C(\Omega)$ *is called* negligible *( [99], 18 (3°)) if*
$\forall K \subset\subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \forall n \in \mathbb{N}\ \exists q \in \mathbb{N}\ \forall \phi \in \mathcal{C}_b^\infty((0,1] \times \Omega, \mathcal{A}_q(\mathbb{R}^n))$ :

$$\sup_{x \in K} |\partial^\alpha (R(S_\varepsilon \phi(\varepsilon, x), x))| = O(\varepsilon^{n}) \text{ as } \varepsilon \to 0\,.$$

*The set of all negligible elements* $R \in \mathcal{E}^C(\Omega)$ *will be denoted by* $\mathcal{N}^C(\Omega)$*.*

**4.70** The key ingredients in proving diffeomorphism invariance as well as stability with respect to derivatives (i.e., that the $x$-derivative of a moderate resp. negligible function again is moderate resp. negligible; this becomes a peculiar issue due to the additional $x$-dependence of $\phi$) are several equivalent formulations of the tests given above.
To settle the question of stability w.r.t. differentiation Jelínek introduced an alternate, yet equivalent form of tests involving differentials of $R$ with respect to the test function-slot denoted by $d_1$. ( [99], thm. 17, resp. thm. 18, (2°) ⇔ (3°)). We only formulate the respective test for moderateness (the case of negligibility being analog) and refer to the original for the ingenious proofs. We presume that the author was completely aware of the role theorems 17 and 18 had to play in this respect yet for some reasons he decided not to address this issue.

**4.71 Theorem.** *$R \in \mathcal{E}(\Omega)$ is a member of $\mathcal{E}_M(\Omega)$ if and only if the following condition is satisfied:*

$$\forall K \subset\subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \forall k \in \mathbb{N}_0\ \exists N \in \mathbb{N}\ \forall B\,(\textit{bounded}) \subseteq \mathcal{D}(\mathbb{R}^n):$$

$$\partial^\alpha \mathrm{d}_1^k (R \circ S^{(\varepsilon)})(\varphi, x)(\psi_1, \dots, \psi_k) = O(\varepsilon^{-N}) \qquad (\varepsilon \to 0)$$

*uniformly for* $x \in K$, $\varphi \in B \cap \mathcal{A}_0(\mathbb{R}^n)$, $\psi_1, \dots, \psi_k \in B \cap \mathcal{A}_{00}(\mathbb{R}^n)$*.*

Here $\mathcal{A}_{00}(\mathbb{R}^n) = \{\varphi \in \mathcal{D}(\mathbb{R}^n) : \int \varphi = 0\}$ denotes the tangent space of $\mathcal{A}_0$ and the operator $S^{(\varepsilon)}$ is derived from $S_\varepsilon$ by

$$\begin{aligned} S^{(\varepsilon)} : \mathcal{D}(\mathbb{R}^n) \times \mathbb{R}^n &\to \mathcal{D}(\mathbb{R}^n) \times \mathbb{R}^n \\ (\varphi, x) &\mapsto (S_\varepsilon \varphi, x) = (\frac{1}{\varepsilon}\, \varphi(\frac{\cdot}{\varepsilon}), x)\,. \end{aligned}$$

Finally to prove that $\partial_i R$ is moderate if $R$ was, observe that $\partial_i(R \circ S^{(\varepsilon)}) = (\partial_i R) \circ S^{(\varepsilon)}$. Then the claim follows from

$$\partial^\alpha \mathrm{d}_1^k((\partial_i R) \circ S^{(\varepsilon)}) = \partial^\alpha \mathrm{d}_1^k \partial_i (R \circ S^{(\varepsilon)}) = \partial^{\alpha + e_i} \mathrm{d}_1^k (R \circ S^{(\varepsilon)})\,.$$

**4.72** We now turn to the central issue of diffeomorphism invariance. First we present a heuristical calculation which clearly shows which path has to be pursued. Suppose we want to prove moderateness of $\hat{\mu} R$. Given $\tilde{\phi} \in \mathcal{C}^\infty((0,1] \times \tilde{\Omega}, \mathcal{A}_0(\mathbb{R}^n))$ then we would have to estimate

$$\hat{\mu}(R)(S_\varepsilon \tilde{\phi}(\varepsilon, \tilde{x}), \tilde{x}) \;=\; R(\bar{\mu}(S_\varepsilon \tilde{\phi}(\varepsilon, \tilde{x}), \tilde{x})) \;=\; R(\bar{\mu} S^{(\varepsilon)}(\tilde{\phi}(\varepsilon, \tilde{x}), \tilde{x})) \;=\; R(S^{(\varepsilon)}(S^{(\varepsilon)})^{-1} \bar{\mu} S^{(\varepsilon)}(\tilde{\phi}(\varepsilon, \tilde{x}), \tilde{x}))\,.$$

Hence we would need $R$ to pass a test for moderateness w.r.t. test objects of the form ($\mathrm{pr}_1$ denoting the projection to the fist component)

$$\phi(\varepsilon, x) \;=\; \mathrm{pr}_1 (S^{(\varepsilon)})^{-1} \bar{\mu} S^{(\varepsilon)}(\tilde{\phi}(\varepsilon, \tilde{x})) \;=\; \tilde{\phi}(\varepsilon, \mu^{-1} x) \left( \frac{\mu^{-1}(\varepsilon \xi + x) - \mu^{-1} x}{\varepsilon} \right) \cdot |\det D\mu^{-1}(\varepsilon \xi + x)|\,.$$

But unfortunately $\phi(\varepsilon,x) \notin \mathcal{C}^\infty((0,1]\times\Omega, \mathcal{A}_0(\mathbb{R}^n))$ since it is only defined if $\xi \in \frac{\Omega-x}{\varepsilon}$, whereas we want $\xi \mapsto \phi(\varepsilon,x)(\xi)$ to be a test function on the whole of $\mathbb{R}^n$.
However, $\phi(\varepsilon,x)$ belongs to a class of test objects providing an apparently weaker, yet, as it finally turns out, equivalent test. More precisely, from [61], thm. 10.5 we have that $R \in \mathcal{E}^C(\Omega)$ is moderate if and only if it fulfills the following condition (Z) [11]

$$\forall K \subset\subset \Omega\ \forall \alpha \in \mathbb{N}_0^n\ \exists N \in \mathbb{N}\ \forall \phi : D \to \mathcal{A}_0(\mathbb{R}^n))\ (D, \phi \text{ as described below})$$

$$\exists C>0\ \exists \eta>0\ \forall \varepsilon\,(0<\varepsilon<\eta)\ \forall x \in K:\ (\varepsilon,x)\in D \text{ and } |\partial^\alpha(R(S_\varepsilon\phi(\varepsilon,x),x))| \leq C\varepsilon^{-N}\,,$$

where $D \subseteq (0,1]\times\Omega$ and for $D,\varphi$ the following holds: For each $L \subset\subset \Omega$ there exists $\varepsilon_0$ and a subset $U$ of $D$ which is open in $(0,1]\times\Omega$ such that

(1) $(0,\varepsilon_0]\times L \subseteq U(\subseteq D)$ and $\phi$ is smooth on $U$, and

(2) for all $\beta \in \mathbb{N}_0^n$, $\{\partial^\beta\phi(\varepsilon,x) \mid 0<\varepsilon\leq\varepsilon_0,\ x\in L\}$ is bounded in $\mathcal{D}(\mathbb{R}^n)$.

Now diffeomorphism invariance of the notion of moderateness is established by the following

**4.73 Theorem.** *( [61], thm. 7.14) Let $\mu : \tilde{\Omega} \to \Omega$ be a diffeomorphism and $\tilde{\phi} \in \mathcal{C}_b^\infty((0,1]\times\tilde{\Omega}, \mathcal{A}_0(\mathbb{R}^n))$. Define $D(\subseteq (0,1]\times\Omega)$ by $D := \{(\varepsilon,x)\in(0,1]\times\Omega \mid (\tilde{\phi}(\varepsilon,\mu^{-1}x),\mu^{-1}x) \in U_\varepsilon(\tilde{\Omega})\}$. For $(\varepsilon,x)\in D$, set*

$$\phi(\varepsilon,x)(\xi) := \tilde{\phi}(\varepsilon,\mu^{-1}x)\left(\frac{\mu^{-1}(\varepsilon\xi+x)-\mu^{-1}x}{\varepsilon}\right)\cdot|\det D\mu^{-1}(\varepsilon\xi+x)|.$$

*Then $\phi$ satisfies the requirements specified for test objects in condition (Z).*

In some more detail assume $R$ to be moderate. We show that $\hat{\mu}R$ passes the test used in definition 4.69 (i). Indeed given $\tilde{K} \subset\subset \tilde{\Omega}$, $\alpha \in \mathbb{N}_0^n$ and $\tilde{\phi} \in \mathcal{C}_b^\infty((0,1]\times\tilde{\Omega},\mathcal{A}_0(\mathbb{R}^n))$ define $\phi$ as in the preceding theorem. Then by the chain rule

$$\partial_{\tilde{x}}^\alpha\left((\hat{\mu}R)(S_\varepsilon\tilde{\phi}(\varepsilon,\tilde{x}),\tilde{x})\right) = \partial_{\tilde{x}}^\alpha(R(S_\varepsilon\phi(\varepsilon,\mu\tilde{x}),\mu\tilde{x})) = \sum_{|\beta|\leq|\alpha|}\partial_x^\beta(R(S_\varepsilon\phi(\varepsilon,x),x))\Big|_{x=\mu\tilde{x}}\cdot g_\beta(\tilde{x})$$

where each $g_\beta$ is bounded on $\tilde{K}$. Since $R$ satisfies condition (Z) the claim follows.

**4.74** However, matters become more complicated in the case of negligibility. First note that the resp. test objects take values in $\mathcal{A}_q$ $(q>0)$ which is not a diffeomorphism invariant property. The way out is provided by the re-introduction of asymptotically vanishing moments (cf. 4.62) into the theory by building up (another) equivalent test using this notion. Indeed Jelínek in [99], 18 (4°) has formulated such a condition which, however, unfortunately is *not equivalent* to the notion of negligibility as defined above. Moreover, his condition is so strong that even $\iota(x^2)\neq(\iota(x))^2$, hence the property of $\iota$ being an algebra homomorphism on $\mathcal{C}^\infty$ is lost (for more details see [61], 7.7). However, also in [61], sec. 7 this flaw was removed, namely by demanding also *all derivatives* of the test objects to have asymptotically vanishing moments. More precisely we say that a test object $\phi \in \mathcal{C}_b^\infty((0,1]\times\Omega,\mathcal{A}_0(\Omega))$ is of type $[A_l^\infty]_{K,q}$ if on (a given) $K \subset\subset \Omega$

$$\forall \beta \in \mathbb{N}_0^n\ 1\leq|\beta|\leq q\ \forall\gamma\in\mathbb{N}_0^n\ \sup_{x\in K}|\int \xi^\beta\partial^\gamma\phi(\varepsilon,x)(\xi)\,dx| = O(\varepsilon^q)\,.$$

Then we have

[11] Condition (Z) is one of 6 tests proved to be equivalent in [61], theorem 10.5 called Theorem A–Z there. Note, however, that this neither indicates that the authors of [61] originally intended to give 26 equivalent tests, nor that (Z) for some mysterious reason was considered to be the ultimative condition; rather it was invented during a train ride returning from a workshop in Novi Sad to Vienna and “Z” simply stands for “Zug” which is the German word for train.

**4.75 Theorem.** *$R \in \mathcal{E}^C(\Omega)$ is negligible if and only if*

$$\forall K \subset\subset \Omega\, \forall \alpha \in \mathbb{N}_0^n\ \forall N \in \mathbb{N}\ \exists q \textit{ such that } \forall \phi \textit{ of type } [A_l^\infty]_{K,q} \sup_{x\in K} |\partial^\alpha R(S_\varepsilon \phi(\varepsilon, x), x)| = O(\varepsilon^N)\,.$$

diffeomorphism invariance of the notion of negligibility is now established by the above theorem in conjunction with an analog of theorem 4.73, as well as an analog of condition (Z) above. Finally we define our main object of desire.

**4.76 Definition.** *The* diffeomorphism invariant Colombeau algebra on $\Omega$ *(in the C-formalism) is defined as the quotient*

$$\mathcal{G}^C(\Omega) := \mathcal{E}^C_M(\Omega) \,/\, \mathcal{N}^C(\Omega)\,.$$

**4.77** Summing up we have constructed a differential algebra $\mathcal{G}^C(\Omega)$ in a diffeomorphism invariant way, in particular allowing for a diffeomorphism invariant embedding of distributions. Moreover, $\mathcal{G}^C(\Omega)$ (as usual) is a fine sheaf of differential algebras ( [61], sec. 8).
We finally turn to the issue of commutativity of the embedding with partial derivatives in the algebra. This will guarantee that $\mathcal{G}^C(\Omega)$ indeed possesses all the favorable properties (i.e., (i)-(iii) in 2.10 and (iv") in 2.13) of a Colombeau algebra. To this end it is useful to change to the J-formalism. Recall from (78) that derivatives in the C-formalism are just given by partial derivatives. Using the translation formalism of [61] we derive that in the J-formalism, i.e., on $\mathcal{E}^J(\Omega) := (T^{-1})^*(\mathcal{E}^C(\Omega))$ we have

$$D_i^J = (T^{-1})^* \circ \partial_i \circ T^* \text{ i.e., } (D_i^J R)(\varphi, x) = -(d_1 R(\varphi, x))(\partial_i \varphi) + (\partial_i R)(\varphi, x)\,. \tag{79}$$

We now see immediately that if $F \in \mathcal{D}'(\Omega)$ then $\iota^J(F)$ (cf. (77)) is independent of $x$ hence the second term in (79) vanishes. Moreover, since $\iota^J(F)$ is linear in $\varphi$, $-d_1(\iota^J(F))(\varphi, x)(\partial_i\varphi) = \langle F, -\partial_i \varphi\rangle$ which is exactly the $\iota^J$-image of $\partial_i F$.

**4.78** We close the discussion on diffeomorphism invariant Colombeau algebras on opens sets of $\mathbb{R}^n$ with a few remarks on [76]. In this article M. Grosser has investigated all (reasonable) classes of possible test objects constituting tests for moderateness respectively negligibility leading, in turn, to a classification of the respective algebras. He found that in the range between the "elementary" algebra $\mathcal{G}^e$ and the diffeomorphism invariant algebra $\mathcal{G}^d$ (which, by the way is not injectively included in $\mathcal{G}^e$ via the canonical homomorphism as one could have conjectured!) there is one further construction leading to a diffeomorphism invariant differential[12] algebra, namely $\mathcal{G}^2$, which—for reasons already mentioned in 4.66 above—may called the "corrected" Colombeau—Meril algebra. However, for the reasons discussed in 4.66 above and following [76] we regard $\mathcal{G}^d$ as the standard diffeomorphism invariant Colombeau algebra on (open sets of) $\mathbb{R}^n$.
Finally we remark that from the viewpoint of certain applications it might also be of interest to construct Colombeau algebras possessing only weaker invariance properties than $\mathcal{G}^d$ does. So one could construct e.g. an algebra invariant only under (orthochronous, proper) Lorentz-transformations well-suited to applications in special relativity; in that case the class of test objects would have to be restricted to the images of constant paths under the Lorentz group.

**4.79** To end this section (and chapter) we turn to the issue of (full) Colombeau algebras on manifolds. Clearly the above constructed local diffeomorphism invariant theory lends itself to a formulation on a smooth manifold $X$. In this respect the J-formalism is clearly superior since—contrary to the C-formalism—its basic definitions involve neither scaling nor translation operations, which, a priori, are meaningless on a manifold. However, recent applications to general relativity (e.g. [36], [185]) have clearly underscored the need for a theory of algebras of generalized functions on manifolds enjoying two additional features: First, it should be *geometric* in the sense that its basic objects should be defined intrinsically on the manifold itself. Second, $\mathcal{G}(X)$ should be a *differential* algebra with Lie derivatives commuting with the embedding of $\mathcal{D}'(X)$. The

[12] There is still another diffeomorphism invariant algebra within that range, yet it is still questionable whether it is a *differential* algebra at all. However, it is not a Colombeau-type algebra since built from Jelinek's condition (4°) the product of smooth functions is not preserved.

construction of such an algebra, there called $\hat{\mathcal{G}}(X)$, indeed was carried out in [77] (for a brief overview see [78]), where also a number of localization results were proved allowing one to make full use of the well-developed local theory. We shall give a short outline of this construction since from the viewpoint of general relativity—besides the special algebra presented in the sections above—it is the most interesting version of the theory.

**4.80** The key notion allowing for a global formulation is that of *smoothing kernels* replacing the *scaled* test objects of the local theory. Definition 4.81(i) below may in a sense be viewed as the diffeomorphism invariant "essence" of the process of regularization via convolution and linear scaling on $\mathbb{R}^n$ while 4.81(ii) is the invariant formulation of the interplay between $x$- and $y$-differentiation in the local context.
We set $\hat{\mathcal{A}}_0(X) := \{\omega \in \Omega^n_c(X) \ : \ \int \omega = 1\}$ and define the basic space of the forthcoming global Colombeau algebra to be

$$\hat{\mathcal{E}}(X) = C^\infty(\hat{\mathcal{A}}_0(X) \times X)\,.$$

Smooth functions are again embedded "constantly," i.e.,

$$\begin{aligned}\sigma : C^\infty(X) &\to \hat{\mathcal{E}}(X)\\ f &\mapsto ((\omega, x) \mapsto f(x))\,,\end{aligned}$$

whereas we embed distributions by

$$\begin{aligned}\iota : \mathcal{D}'(X) &\to \hat{\mathcal{E}}(X)\\ u &\mapsto ((\omega, x) \mapsto \langle u, \omega\rangle)\,.\end{aligned}$$

Next we introduce the following notation of Lie derivatives: For any smooth $f : X \to \Omega^n_c(X)$ we set

$$L'_\xi f(p)(q) \;:=\; L_\xi(p \mapsto f(p)(q)) \;=\; \lim_{t\to 0} \frac{1}{t}(f(\mathrm{Fl}^\xi_t(p)) - f(p))(q)\,. \tag{80}$$

Here $\mathrm{Fl}^\xi_t$ denotes the flow of $\xi \in \mathcal{T}^1_0(X)$ and convergence is understood pointwise (i.e., fiberwise in $\Lambda^n T^* X$). On the other hand, we will denote the Lie derivative of $f(p)$ with respect to $q$ simply by $L_\xi f(p)$, i.e.,

$$L_\xi f(p)(q) \;:=\; L_\xi(q \mapsto f(p)(q)) \;=\; \frac{d}{dt}|_0 (\mathrm{Fl}^\xi_t)^*(f(p))(q)\,.$$

Now we are ready to introduce the space of smoothing kernels by the following

**4.81 Definition.** $\Phi \in C^\infty((0,1] \times X, \hat{\mathcal{A}}_0(X)) \subseteq C^\infty((0,1] \times X \times X, \Lambda^n T^* X)$ *is called a smoothing kernel if it satisfies the following conditions*

*(i)* $\forall K \subset\subset X\ \exists\, \varepsilon_0,\ C > 0\ \forall p \in K\ \forall \varepsilon \le \varepsilon_0$*:* $\mathrm{supp}(\Phi(\varepsilon, \mathrm{p})) \subseteq \mathrm{B}_{\varepsilon\mathrm{C}}(\mathrm{p})$*, and*

*(ii)* $\forall K \subset\subset X\ \forall k, l \in \mathbb{N}_0\ \forall \xi_1, \ldots, \xi_k, \eta_1, \ldots, \eta_l \in \mathcal{T}^1_0(X)$

$$\sup_{\substack{p\in K\\ q\in X}} \|L_{\eta_1} \ldots L_{\eta_l}(L'_{\xi_1} + L_{\xi_1}) \ldots (L'_{\xi_k} + L_{\xi_k})\Phi(\varepsilon, p)(q)\| = O(\varepsilon^{-(n+l)})$$

*The space of smoothing kernels on $X$ is denoted by $\tilde{\mathcal{A}}_0(X)$.*

In (i) the radius of the ball $B_{\varepsilon C}(p)$ has to be measured with respect to the Riemannian distance induced by any Riemannian metric $h$ on $X$. Similarly in (ii), $\|\,.\,\|$ denotes the norm induced on $\Omega^n_c(X)$ by any Riemannian metric on $X$.
Next we introduce the analog of the grading of the space $\tilde{\mathcal{A}}_0$ which in fact is modelled with a view towards reproducing the main technical ingredient for proving that the embedding of distributions into $\hat{\mathcal{G}}$ coincides with the "constant" embedding on $\mathcal{C}^\infty$.

**4.82 Definition.** *For all $m \in \mathbb{N}$ we denote by $\tilde{\mathcal{A}}_m(X)$ the set of all $\Phi \in \tilde{\mathcal{A}}_0(X)$ such that $\forall f \in \mathcal{C}^\infty(X)$ and $\forall K \subset\subset X$*

$$\sup_{p\in K} |f(p) - \int_X f(q)\Phi(\varepsilon, p)(q)| = O(\varepsilon^{m+1})$$

Finally we define the global tests for moderateness resp. negligibility by

**4.83 Definition.**

*(i)* $R \in \hat{\mathcal{E}}(X)$ *is called* moderate *if* $\forall K \subset\subset X\ \forall k \in \mathbb{N}_0\ \exists N \in \mathbb{N}\ \forall\ \xi_1, \dots, \xi_k \in \mathcal{T}_0^1(X)\ \forall\ \Phi \in \tilde{\mathcal{A}}_0(X)$

$$\sup_{p \in K} |L_{\xi_1} \dots L_{\xi_k}(R(\Phi(\varepsilon, p), p))| = O(\varepsilon^{-N}) .$$

*The subset of moderate elements of* $\hat{\mathcal{E}}(X)$ *is denoted by* $\hat{\mathcal{E}}_M(X)$.

*(ii)* $R \in \hat{\mathcal{E}}_M(X)$ *is called* negligible *if* $\forall K \subset\subset X\ \forall k, l \in \mathbb{N}_0\ \exists m \in \mathbb{N}\ \forall\ \xi_1, \dots, \xi_k \in \mathcal{T}_0^1(X)\ \forall \Phi \in \tilde{\mathcal{A}}_m(X)$

$$\sup_{p \in K} |L_{\xi_1} \dots L_{\xi_k}(R(\Phi(\varepsilon, p), p))| = O(\varepsilon^{l}) .$$

*The set of negligible elements of* $\hat{\mathcal{E}}(X)$ *will be denoted by* $\hat{\mathcal{N}}(X)$.

**4.84** For the further development of the construction, in particular to prove stability with respect to derivatives (cf. 4.70), it is necessary to use the host of local results made available by localization to be discussed in a moment. To this end writing again $(V_\alpha, \psi_\alpha)$ for a chart in $X$ note that

$$(\psi_\alpha^* \times \psi_\alpha^{-1})(\hat{\mathcal{A}}_0(\psi_\alpha(V_\alpha)) \times \psi_\alpha(V_\alpha)) \subseteq \hat{\mathcal{A}}_0(X) \times X .$$

and locally we have $\mathcal{E}^J(\psi_\alpha(V_\alpha)) \cong \hat{\mathcal{E}}(\psi_\alpha(V_\alpha))$. Furthermore, since the map $\psi_\alpha^* \times \psi_\alpha^{-1} : \Omega_c^n(\psi_\alpha(V_\alpha)) \times \psi_\alpha(V_\alpha) \to \Omega_c^n(V_\alpha) \times V_\alpha$ is smooth, for any $R \in \hat{\mathcal{E}}(X)$ its local representation

$$(\psi_\alpha^{-1})^* R := R \circ (\psi_\alpha^* \times \psi_\alpha^{-1})$$

may be view as an element of $\mathcal{E}^J(\psi_\alpha(V_\alpha))$. Now we are in a position to state the following theorem which is proved (using a quite involved notion of "transport" of test objects) in [77], sec. 4.

**4.85 Theorem.**

*(i) (Localization of moderateness) Let* $R \in \hat{\mathcal{E}}(X)$. *Then*

$$R \in \hat{\mathcal{E}}_M(X) \Leftrightarrow (\psi_\alpha^{-1})^* R \in \mathcal{E}_M^J(\psi_\alpha(V_\alpha))\ \forall \alpha.$$

*(ii) (Localization of negligibility) Let* $R \in \hat{\mathcal{E}}(X)$. *Then*

$$R \in \hat{\mathcal{N}}(X) \Leftrightarrow (\psi_\alpha^{-1})^* R \in \mathcal{N}^J(\psi_\alpha(V_\alpha)) \quad \forall \alpha.$$

**4.86** Finally we may define the *"global" Colombeau algebra on* $X$ by the quotient

$$\hat{\mathcal{G}}(X) := \hat{\mathcal{E}}_M(X) \,/\, \hat{\mathcal{N}}(X) .$$

As in all the previous cases, $\hat{\mathcal{G}}(X)$ is a fine sheaf of differential algebras with the Lie derivative with respect to some smooth vector field $\xi$ defined by

$$(\hat{L}_\xi R)(\omega, p) := -d_1 R(\omega, p)(L_\xi \omega) + L_\xi(R(\omega, \,.\,))|_p . \tag{81}$$

Formula (81) ensures the commutativity of Lie derivatives with the embedding by an simple argument analogous to the one given in 4.77, hence $\hat{\mathcal{G}}(X)$ possesses all the four magic properties (i)-(iii) in 2.10 and (iv") in 2.13 (with partial derivatives replaced by Lie derivatives).

# 5. DISTRIBUTIONAL DESCRIPTION OF IMPULSIVE GRAVITATIONAL WAVES

This final section of the present work is devoted to a detailed study of the geometry of impulsive gravitational pp-waves from a distributional point of view. In particular, using the distributional form of the metric tensor we study the geodesics as well as the geodesic deviation in these spacetimes. Our main tool to overcome the problems stemming from the nonlinear operations to be performed on the singular, i.e., distributional metric (cf. the discussions in chapter 2) will be the nonlinear theory of generalized functions as presented above (chapters 3 and 4).
In the literature impulsive pp-waves have also been frequently described by a continuous metric tensor. We analyze from a rigorous point of view the intuitively clear but formally ill-defined transformation first given explicitly by R. Penrose [152] (in the special case of a plane wave with constant linear polarization, see below) relating the distributional to the continuous form of the metric.

In section 5.A, after reviewing the class of gravitational pp-waves we introduce their *impulsive* form and discuss their physical relevance as well as their most basic properties. We treat the geodesic equation for these spacetimes in a distributional setting in section 5.B. In particular, we are going to prove existence and uniqueness of solutions to these nonlinear singular ODEs in the algebra $\mathcal{G}$ and calculate their associated distributions. In section 5.C we study the geodesic deviation equation for impulsive pp-waves. Although linear in the components of the Jacobi field it involves terms formally proportional even to the square of the Dirac $\delta$-distribution. Nevertheless we again succeed proving existence and uniqueness of generalized solutions as well as calculating physical sensible distributional shadows. In section 5.D we study in detail the distributionally ill-defined transformation between the continuous and the distributional form of the metric unmasking it as the distributional shadow of a perfectly well behaved *generalized coordinate transformation.* Finally in section 5.E we discuss our results and give an outlook to future research.
Most of the material presented here has also been published in references [174,118,119], as well as [175,176] by the author, in part together with M. Kunzinger.

## 5.A. Impulsive pp-Waves

This section consists of two parts: The first one is concerned with a review of the class of pp-wave spacetimes, while in the second one we introduce impulsive pp-waves to be described in detail in the following sections.

**5.1** The class of *plane fronted gravitational waves with parallel rays* or, for short, *pp-waves* was first considered by Brinkmann [30] already in 1923 and rediscovered subsequently by several authors among them Rosen [163], (in the special case of plane linearly polarized waves, see below), Robinson in 1956 (cf. [57], p.88), Hély [82] and Peres [154]. Robinson was the first to discover their physical significance and derive some of their properties.
Many reviews have appeared over the years, e.g. in [101], chap. 4, [57], §2-5, [192], [111], 21.5 and [73], chap. 4, so that here it is only necessary to recall the most important properties of these solutions and to refer to existing literature for details.

**5.2** Ehlers and Kundt in [57], §2-5.1 have collected six equivalent characterizations of pp-waves. However, it is most common to define pp-waves as spacetimes admitting a *covariantly constant null vector field* $k^a$. It is possible to physically interpret such a field as the rays of gravitational (or other null) waves. The defining condition immediately implies (among others properties) the vanishing of the optical scalars (see [187], chap. 9.2) *shear, twist and expansion.* Hence, $k^a$ is tangent to a non-expanding, shear-free and twist-free null geodesic congruence. From the twist-free property the existence of 2-surfaces orthogonal to $k^a$ follows, which may be interpreted as wave surfaces.
The Petrov type (cf. [187], Chap 13.2) of pp-waves is either $N$, with the multiple principal null direction given by $k^a$, or conformally flat. Recall that a vacuum solution is called null field if its Weyl tensor is of type $N$ everywhere; physically such fields are interpreted as pure radiation fields.
Within the wider class of spacetimes admitting a non-expanding, shear- and twist-free null congruence (Kundt's class, cf. [111], chap. 27), resp. null fields of that type (plane waves; called so in analogy to the electromagnetic case [57], 2-4.7) pp-waves are characterized by their vanishing rotation [57], 2-5.1. (Recall that the rotation may be defined in an invariant way exactly for such fields [57], 2-2.6.) On the other hand [25] all vacuum type-N (which implies vanishing shear [115]), non-twisting solutions with possibly non-vanishing

cosmological constant belong either to the (non-expanding) Kundt class (with pp-waves one of their six natural subclasses) or the expanding Robinson-Trautman class (cf. [111], chap. 24).

**5.3** Introducing a null coordinate $u$ by the condition $\partial_a u = k_a$ the metric of a pp-wave may be written in the form

$$ds^2 \;=\; H(u,x,y)\,du^2 - du\,dv + dx^2 + dy^2, \tag{82}$$

where $H$ called the *wave profile* is an arbitrary (smooth) function of its arguments, namely $v$, another null coordinate with $\partial_a v \partial^a u \;=\; 0$ and $x, y$ Cartesian coordinates spanning the wave surfaces. Moreover by [57], theorem 2-5.5 pp-waves are characterized by this form of the metric. Often (82) is referred to as the *Brinkmann form* of the pp-wave metric. It is clearly of Kerr-Schild type (cf. [111], 28.1) and its Ricci tensor is given by

$$R_{ab} \;=\; (H_{xx} + H_{yy})\,k_a k_b\,. \tag{83}$$

Hence in the vacuum case the metric (82) is most conveniently written ( [57], theorem 2-5.6)

$$ds^2 \;= 2\,\mathrm{Re(F)du^2 - du\,dv} + |\mathrm{d}\zeta|^2\,, \tag{84}$$

using the complex coordinate $\zeta = (1/\sqrt{2})\,(x+iy)$, $H$ =Re($F$) and $F(\zeta,u)$ a holomorphic function of $\zeta$ but otherwise arbitrary (smooth).
In general the group generated by $k^a$ is the maximal group of isometries but larger groups exist for special cases. Ehlers and Kundt [57], 2-5.4, 2-5.6 have completely classified the symmetries of pp-waves according to various choices of the (profile) function $F$.

**5.4** Of particular interest is the special case of *plane gravitational waves* first considered by Baldwin and Jefferey [18] in 1926 and later treated among others by Bondi [27], Takeno [178], Bondi, Pirani and Robinson [28] , Kundt [114], Jordan, Ehlers and Kundt [100] and again Bondi and Pirani [29].
A pp-wave is called plane or *homogeneous* if $F''' = 0$, where $'$ denotes derivatives with respect to $\zeta$. In this case the amplitude (cf. [57], 2-5.2) is constant on the wave surfaces and the metric may be written in the form

$$ds^2 \;=\; (h_{11}x^2 + 2h_{12}xy + h_{22}y^2)du^2 - du\,dv + dx^2 + dy^2\,, \tag{85}$$

with the $h_{ij}$'s depending on $u$ only. This metric describes a vacuum spacetime—in that case a pure gravitational wave—if $h_{22} = -h_{11}$, respectively a pure electromagnetic, conformally flat spacetime if $h_{22} = h_{11}$ and $h_{12} = 0$. In the vacuum case a plane wave is said to have *constant linear polarization* (cf. [57], 2-5.2) if in addition also $h_{12}$ is proportional to $h_{11}$. Then ( [73], 4.2) the metric function $H$ may be written as $H \;=\; h(u)(\cos\alpha(x^2-y^2) + 2\sin\alpha xy)$ with $h$ an arbitrary (smooth) function and $\alpha$ a constant called *polarization.* Furthermore it is then possible to rotate the coordinates such that $h_{12}$, respectively $\alpha$ vanishes, implying

$$ds^2 \;=\; h(u)(x^2-y^2)\,du^2 - du\,dv + dx^2 + dy^2\,. \tag{86}$$

For many purposes it is convenient to transform the line element (85) to the following form due to *Rosen* [163]

$$ds^2 \;=\; -du\,dV + (a^2+e^2)\,dX^2 + (ab+ce)dX\,dY + (b^2+c^2)\,dY^2\,, \tag{87}$$

where $x \;=\; aX + bY$, $y \;=\; eX + cY$ and $v \;=\; V + (1/2)(aa' + ee')X^2 + (1/2)(ba' + ab' + ec' + ce')XY + (1/2)(bb' + cc')Y^2$ with $a$, $b$, $c$ and $e$ all functions of $u$ constrained by the following set of equations (again see [73], 4.2):

$$\begin{aligned} a'' + h_{11}a + h_{12}e &= 0 & \qquad b'' + h_{11}b + h_{12}c &= 0\\ c'' + h_{12}b + h_{22}c &= 0 & \qquad e'' + h_{12}a + h_{22}e &= 0\\ ba' - ab' - ec' + ce' &= 0 \end{aligned}$$

In the case of constant linear polarization $b$ and $e$ may taken to be zero. Relative motion of test particles in plane waves has been studied in [28] and [29] (using yet another coordinate system) revealing a strong focusing effect. In particular for plane waves with constant linear polarization all particles will collide after finite time no matter how far they were apart initially. The occurrence of such caustics is associated with a coordinate singularity in the Rosen form (87) of the metric.

**5.5** A pp-wave of the form $H = \rho(u)f(x,y)$ is called a *sandwich pp-wave* [28] if $\rho$ is non-vanishing only in some finite region $u_0 \leq u \leq u_1$ of spacetime. The gravitational field then is confined to that region, with flat space in "front" ($u \leq u_0$), resp. "behind" ($u \leq u_1$) the wave.
Yurtsever in [191] has proved that in fact any gravitational wave spacetime that is flat before the arrival of the wave and returns to flatness after the wave has passed is necessarily a pp-wave.

**5.6** In [150], chap. 4 R. Penrose introduced *impulsive pp-waves* as idealization (*impulsive limit*) of sandwich waves of infinitely short duration (say $u_0,\, u_1 \to 0$) but still producing a nontrivial effect in the sense that $\rho$ equals the Dirac-$\delta$, i.e., the metric taking the form

$$ds^2 \;=\; f(x,y)\,\delta(u)\,du^2 - du\,dv + dx^2 + dy^2\,. \tag{88}$$

This spacetime is hence flat everywhere except for the null hyperplane $u = 0$ where a $\delta$-like impulse is located.
Of course such a spacetime, due to the appearance of a distribution in the metric, is not "really allowable" in the sense of classical general relativity. Also (written in this form) the metric is not even gt-regular (see 2.19) hence lies outside the "largest reasonable" class of distributionally tractable spacetimes (see chap. 2). On the other hand this special situation is "tame" in the sense that one may calculate the Ricci tensor within classical distribution theory since it is linear in the metric coefficients (cf. (83) above). Hence the Einstein vacuum equations may be formulated giving $\Delta f = 0$ on the waves surface $u = 0$.

**5.7** R. Penrose has also given a vivid description of impulsive pp-waves by his *"scissors and paste" approach* [150–152]. Ordinary Minkowski space $ds^2 = 2dudv + dx^2 d\zeta d\bar{\zeta}$ (again using the complex coordinate $\zeta = (1/\sqrt{2})(x+iy)$) is divided into two halves $M^+$ $(u > 0)$, resp. $M^-$ $(u < 0)$ by removal of the null hyperplane $u = 0$. Then the two halves are joined together with a "warp." More precisely, a null hyperplane $K$ is inserted as the common boundary of $M^-$ resp. $M^+$ in such a way that $M^- \cup K$ and $M^+ \cup K$ each have normal Minkowski structure but the two halves are joined in a way not compatible with a four-dimensional Minkowski structure at $K$, although the 3-metric induced on $K$ by either of the halves coincide. Using coordinates $u^\mp, v^\mp, \zeta^\mp$ in $M^\mp$ the entire manifold hence is defined by

$$\begin{aligned}
M^- \cup K:&\quad ds^2 \;=\; -du^-\,dv^- \;+\; 2\,d\zeta^-\,d\bar{\zeta}^- &&(u^- \leq 0,\; v^-, \zeta^- \text{ unrestricted})\\
M^+ \cup K:&\quad ds^2 \;=\; -du^+\,dv^+ \;+\; 2\,d\zeta^+\,d\bar{\zeta}^+ &&(u^+ \geq 0,\; v^+, \zeta^+ \text{ unrestricted})\\
K:&\quad u^- = u^+ = 0, \quad \zeta^- = \zeta^+, \quad v^- = v^+ + f(\xi,\bar{\xi})\,.
\end{aligned}$$

The equations in the last line above are often referred to as *Penrose junction conditions*. For a detailed discussion of the intrinsic geometries of null hypersurfaces in this context see [152].

**5.8** Physically impulsive pp-waves arise—besides as the impulsive limit of sandwich waves— in the *ultrarelativistic limit of black hole spacetimes* of the Kerr-Newman family as was first shown by Aichelburg and Sexl [3] already in 1971 for the Schwarzschild case. The technique employed consists essentially in applying a boost (relative to a static observer) to the spacetime metric (in isotropic coordinates) and taking the limit of the boost velocity $v$ to the speed of light $c \equiv 1$ (in a specific way). First the mass parameter has to be rescaled according to $m = p\sqrt{1-v^2} = p/\gamma$, where $p$ denotes the constant momentum and $\gamma$ the well-known "gamma-factor" of special relativity. Physically this choice is motivated by the need to keep the total energy of the "particle" finite in the limit, which is done by rescaling its rest mass. Second one has to employ a sensible coordinate transformation which becomes singular in the limit to end up with the following metric

$$ds^2 \;=\; 8\,p\,\delta(u)\ln\sqrt{x^2+y^2}\,du^2 - du\,dv + dx^2 + dy^2\,, \tag{89}$$

which clearly is of impulsive pp-wave-type (88) with $f = 8p\ln\sqrt{x^2+y^2}$. The field equations for that metric imply (denoting by $\mu$ the energy density and by $\delta^{(2)}$ the two-dimensional $\delta$-distribution)

$$\mu = \Delta f = 32\pi p\delta^{(2)}(x,y)\,.$$

Hence, physically this spacetime geometry is interpreted as a gravitational shock wave [13] produced by a

[13]Note that in Penrose's terminology [152] the term gravitational *shock* wave is reserved for waves of the form $H = \theta(u)f(x,y)$ ($\theta$ denoting the Heaviside function) whereas following the majority of the literature on ultrarelativistic limits (see below) we use the terms *shock*- resp. *impulsive* waves synonymously.

massless particle moving at the speed of light, with its world line given by $x = y = u = 0$ in Minkowski space.

**5.9** The procedure of Aichelburg and Sexl has been generalized by various authors. The ultrarelativistic limit of the Reissner-Nordstrøm metric was derived by Loustó and Sánchez in [131], whereas the Kerr case was treated by a number of authors, among them again Loustó and Sánchez [130,133], Ferrari and Pendenza [63] and Hayashi and Samura [81]. Even different sources like cosmic strings, domain walls and monopoles have been boosted to gain ultrarelativistic spacetimes of impulsive pp-wave form [132] which in turn have been used to describe quantum scattering processes of highly energetic particles (see [182,183] for an overview). The basic idea is that if a particle is moving at ultrarelativistic speed (at Planckian energies) its gravitational field modelled by a gravitational shock-wave has to be taken into account. Hence as an approximation to a (not yet existing) theory of quantum gravity one studies particle [52] or string scattering [51] in a background given by such ultrarelativistic limit spacetimes.
However, the mathematical formalism (i.e., linear distribution theory) employed to derive these results has clearly been pushed to (or even over) its limits. In the case of the ultrarelativistic Reissner-Nordstrøm solution e.g. the following physically paradox situation arises [131]: In oder to have the metric converging to a distributional limit as $v$ tends to 1 the charge parameter $e$ has to be rescaled according to $e = p_e/\sqrt{\gamma}$ ($p_e$ constant) which in turn implies that the electromagnetic field vanishes in the ultrarelativistic limit. However, the electromagnetic energy-momentum tensor does not vanish, i.e., $T_{uu} \to \frac{3p_e^2}{16(x^2+y^2)^{3/2}}\,\delta(u)$. This situation has been throughly analyzed from a mathematical point of view using non-linear generalized functions in [171] and [173] displaying that the electromagnetic field is given by a generalized tensor field associated to 0 (see section 3.E) with its "square" associated to a distribution proportional to $\delta$. This example clearly shows that if one is willing to deal with distributional idealizations of physical situations in a nonlinear context seemingly paradox situations have to be taken into account. Or—turning the argument—in such situations a more refined modelling also capturing the properties of singular objects which are relevant in the nonlinear context is needed (cf. the discussion at the beginning of sec. 3.E).

**5.10** An alternative approach to boosting Kerr-Newman black holes was developed by Balasin and Nachbagauer. Using the Kerr-Schild form of the metric a distributional energy-momentum tensor supported in the singularity is assigned to the black hole spacetime [10,11]. Then this energy-momentum tensor is subjected to a boost and the limit $v \to 1$, subsequently [12,13,16]. The ultrarelativistic metric may then be derived using Einstein's equations. However, the assignment of a distributional energy-momentum tensor to a singular spacetime depends on the choice of the underlying differentiable structure at the singularity (cf. also the highly "coordinate dependent" results in [105]).

**5.11** A generalization of Penrose's "scissors and paste" method to non-flat backgrounds which may be used as an alternative way to derive the Aichelburg-Sexl geometry (89) as well as more general gravitational shock waves was introduced by Dray and t' Hooft in [56]. Starting with a vacuum solution of Einstein's equations of the form

$$ds^2 \;=\; 2\,A(u,v)\,dudv + g(u,v)\,h_{ij}(x^l)dx^i dx^j$$

(with $A$ and $g$ smooth functions and $h_{ij}$ a two-dimensional (smooth) Riemannian metric) they asked under which conditions one may introduce a shift in the $v$-coordinate at $u = 0$, i.e., replacing $v$ by $v + f(x^i)$ for $u > 0$, such that the resulting metric satisfies Einstein equations with a photon at the origin of the wave surface, i.e., an energy-momentum tensor $T_{uu} = 4p\delta^{(2)}(x,y)\delta(u)$. The answer is given by the following set of equations which have to be fulfilled at the shock hypersurface $u = 0$

$$\begin{gathered} A_{,v} \;= 0 \;=\; g_{,v} \\ \frac{A}{g}\Delta f - \frac{g_{,uv}}{g}\,f = -32\pi\,p\,A^2\delta^{(2)}(x,y)\,. \end{gathered}$$

Choosing a flat background, i.e., $A = -1/2$, $g = 1$ and $h_{ij} = \delta_{ij}$, one ends up precisely with metric (89).
Using this method Dray and t' Hooft derived the spherical shock wave due to a massless particle moving at the speed of light along the horizon of a Schwarzschild black hole which was used to study the influence of matter, falling into the black hole on its Hawking-radiation. These ideas lie at the heart of t' Hooft's S-matrix approach to quantum gravity [180,48].
Dray and t' Hooft's approach to shock wave geometries has been generalized to include a non-vanishing cosmological constant as well as Maxwell fields again by Loustó and Sánchez in [129].

**5.12** The Aichelburg-Sexl procedure has been further generalized to include a non-vanishing cosmological constant by Hotta and Tanaka. In [95] they boosted the Schwarzschild-de Sitter solution to obtain a spherical gravitational wave generated by a pair of null particles in a de Sitter background. A similar solution was also given in an anti-de Sitter background. On the other hand solutions describing impulsive waves in de Sitter resp. anti-de Sitter spacetime generated by null particles with an arbitrary multipole structure have been found by Podolský and Griffiths [157]. As pointed out in [158] null monopole solutions can thus be regarded as the limits of static monopole solutions boosted to their ultrarelativistic limit. The analogous statement holds in Minkowski background for multipole particles also. In other words (cf. [158]), impulsive waves in flat space with arbitrary multipole structure arise as ultrarelativistic limits of solutions of the Weyl family (cf. [111], sec. 18.1). However, since no explicit exact solutions describing static sources of any multipole structure in de Sitter resp. anti-de Sitter space are known the question for an analogous relation in the case of non-vanishing cosmological constant remains unresolved.
The entire class of non-expanding impulsive gravitational waves (with possibly non-vanishing cosmological constant) has been throughly analyzed in [159]. It was shown that the only non-trivial impulsive waves in Minkowski background are in fact impulsive pp-waves.

Finally, the symmetries of impulsive pp-waves have been analyzed by Aichelburg and Balasin in a series of papers [4–6]. Working out a distributional symmetry classification analogous to the generic one given in [57], 2-5.4, 2-5.6 they found a richer structure of the symmetry algebra in the impulsive case. Aichelburg and Balasin in [7] have also studied the global structure of impulsive pp-waves.

### 5.B. The Geodesic Equation for Impulsive pp-Waves

In this section we study in detail the geodesic equation for impulsive gravitational pp-waves using the distributional form of the metric. This choice is mainly motivated by the fact that physically, i.e., in the ultrarelativistic limit 5.8, 5.9 (which is a delicate construction; cf. [171,173] and the remarks in 5.9), as well as simply modelled by the impulsive limit of a sandwich wave (cf. 5.6) the metric arises precisely that way. Thereby we follow the approaches of [62] and [15]. The interrelations between the continuous and the distributional form of the metric are discussed in detail in section 5.D below.
First we analyze the system of equations from a distributional point of view to discover that it is not well-defined within the framework of L. Schwartz' linear theory. Then we proceed transferring the system into our setting of Colombeau's generalized functions using a very natural and general regularization of the $\delta$-distribution. We give a short outline how—in general—this framework is used to treat singular, i.e., distributional, nonlinear differential equations before proving existence and uniqueness of generalized solutions of the system under consideration. Finally, to make sensible physical predictions we calculate the associated distributions of the unique $\mathcal{G}$-solutions, which fit perfectly well into the heuristical picture and, moreover, are completely independent of the chosen regularization of the singular metric coefficient.

**5.13** We start with an impulsive pp-wave metric of the form

$$ds^2 \;=\; f(x,y)\,\delta(u)\,du^2 - du\,dv + \sum_{i=1}^{2}(dx^i)^2 \tag{90}$$

where $x^i = (x,y)$ and we suppose $f$ to be smooth for the moment (for comments on non-smooth $f$ see 5.25). We apply a first analysis of the geodesic equation strictly in a linear distributional setting. It is straightforward to derive the Christoffel symbols in these coordinates (which cover the whole manifold). The non-vanishing components are

$$\Gamma^v_{uu} \;=\; -f\;\dot\delta\,, \quad \Gamma^i_{uu} = -\frac{1}{2}\,\partial_i f\;\delta\,, \quad \Gamma^v_{ui} \;=\; \Gamma^v_{iu} \;=\; -\partial_i f\;\delta\,,$$

where we have denoted the partial derivatives of $f$ by $\partial_i f$ and the derivative of the $\delta$-distribution by $\dot\delta$. Hence we get the geodesic equations

$$\begin{aligned} u'' &= 0 \\ v'' &= f\,\dot\delta + 2\partial_i f\,x^i\,u'\,\delta \\ x^{i\,\prime\prime} &= \frac{1}{2}\partial_i\,f\,u'^{\,2}\delta\,, \end{aligned}$$

where $'$ denotes the derivative with respect to an affine parameter and summation over $i$ is understood. We may use the first equation to introduce $u$ as a new affine parameter (thereby excluding only trivial geodesics parallel to the shock hypersurface) to get

$$\begin{aligned}\ddot{v}(u) &= f(x^j(u))\,\dot{\delta}(u) \,+\, 2\,\partial_i f(x^j(u))\,\dot{x}^i\,(u)\delta(u)\\ \ddot{x}^i\,(u) &= \frac{1}{2}\,\partial_i\,f(x^j(u))\;\delta(u)\;,\end{aligned} \tag{91}$$

where $\dot{}$ again denotes the derivative with respect to $u$ and we have inserted all the dependences explicitly. Equations (91) seem to form a system of three coupled, nonlinear ODEs of second order in the vector space $\mathcal{D}'$ of distributions. For $u \neq 0$ all the right hand sides vanish, which is clear from the form of the metric tensor, and we expect the geodesics to be broken, possibly refracted, straight lines.
However, if we take a closer look at system (91) we see immediately that the first equation is not well-defined within classical distributions as the terms $\dot{x}^i\,\delta$ involve the ill-defined product $H\delta$ [14]. In some more detail, if we integrate the last two equations using the (distributional) identity $f(x^i(u))\,\delta(u) = f(x^i(0))\,\delta(u)$ we get

$$x^i(u) \;=\; \text{initial values } \;+\frac{1}{2}\,\partial_i\,f(x^i(0))\,u_+\;, \tag{92}$$

where we have denoted the kink function $u\,H(u)$ by $u_+$. Now we see that $x^i$ merely is continuous, hence—strictly speaking—cannot be derived the way envisaged above. If we still try to go on by brute force—as obviously intended in [62]—things become even worse: Inserting (92) into the first equation (91) we see that the term $\dot{x}^i(u)\,\delta(u)$ gives rise to the ill-defined product $H\delta$. The ambiguities connected with this very "product" have been analyzed in great detail in 2.4. Here it is time for us to stop heuristics and come back to a mathematically rigorous treatment.

**5.14** We first briefly outline how to apply the generalized functions framework to handle *singular nonlinear differential equations.* The *first step* consists of transferring the equations into the generalized functions framework which, generally speaking, amounts to some kind of regularization of the singular objects. Depending on the context at hand one may either use the "canonical" embedding or some other appropriate method (cf. the discussions in sections 3.E and 4.A). The *second step* is to solve the equation in the $\mathcal{G}$-setting, that is, to prove existence and uniqueness of a generalized solution and/or to construct an explicit solution. This is naturally done componentwise, i.e., for fixed $\varepsilon$ in a classical (smooth) setting. The *third and final step* consists in calculating the associated distribution (if they exist, which has to be regarded as a special property of the equation at hand) of the $\mathcal{G}$-solution. These may be viewed as "distributional solution" to the original problem and serve for the purpose of (physical) interpretation.
However, we have used quotation marks above to stress the fact that these "distributional solutions" *cannot be said to obey the original distributional equations,* for in general, certain terms will not be defined if one tries to insert the "distributional solutions" into the equations. For this reason it is impossible to set up a distributional solution concept for nonlinear equations. This argument also affects the reliability of "naive regularization procedures," consisting merely in solving the regularized equation in a smooth setting and then passing to the weak limit (see e.g. [174] and the discussion in [118], p. 1480).
The decisive advantage of our generalized functions approach, as compared to a "naive regularization procedure," is that it does provide a *sensible solution concept* for such equations. Moreover, it may be regarded as an additional (stability) property of an equation to allow for generalized solutions, hence—from the viewpoint of applications—strengthens the reliability of "distributional solutions" derived via association from generalized ones.

**5.15** Let us now come to the first task, i.e., transporting system (91) into the setting of the Colombeau algebra $\mathcal{G}(\mathbb{R})^3$. We have to regularize the coefficient $\delta$ in the first equation, i.e., to replace it by some element of $\mathcal{G}(\mathbb{R})$. Remember that we discarded the "canonical" embedding already in section 4.A due to its lack of diffeomorphism invariance; hence we have to choose some generalized function associated to $\delta$. Since we aim at the most general result as well as diffeomorphism invariance (as discussed in 5.24 below) we are interested in putting as few restrictions on the explicit shape of the regularization as possible. For our purpose the

[14] We return to our convention to denote the Heaviside function by $H$, which—of course—has to be distinguished from the wave profile of pp-waves which for historical reasons was denoted also by $H$ in section 5.A.

"largest reasonable" class of smooth regularization of $\delta$ is given by so-called *strict delta nets* (cf. 2.4), i.e., nets $(\rho_\varepsilon)_{\varepsilon\in(0,1]} \in (\mathcal{C}^\infty(\mathbb{R}))^{(0,1]}$ satisfying

$$
\begin{aligned}
&\text{(a)} \quad \mathrm{supp}(\rho_\varepsilon) \to \{0\} \quad (\varepsilon \to 0)\ ,\\
&\text{(b)} \quad \int \rho_\varepsilon(x)\, dx \to 1 \quad (\varepsilon \to 0) \ \text{ and}\\
&\text{(c)} \quad \exists \eta > 0\ \exists C \geq 0 : \int |\rho_\varepsilon(x)|\, dx \leq C\ \forall \varepsilon \in (0,\eta]
\end{aligned}
$$

(Note that since $\mathcal{D}$ is dense in $L^1$ practically even (approximations to) discontinuous regularizations (e.g., boxes) are included.) Obviously any such net converges to $\delta$ in distributions as $\varepsilon$ tends to 0, hence its class is associated to $\delta$ in $\mathcal{G}(\mathbb{R})$. To simplify notation it is often convenient to replace (a) by

$$\mathrm{supp}(\rho_\varepsilon) \subseteq [-C'\varepsilon, C'\varepsilon] \qquad \forall \varepsilon \in (0,1)\,,$$

with $C'$ denoting a constant. In fact in the (nevertheless quite technical) proofs below we shall even choose $C' = 1$. However, this is just a technical easement since we only use the fact that the "support numbers" of $\rho_\varepsilon$ defined by $l(\rho_\varepsilon) := \inf\{x :\ \rho_\varepsilon(x) \neq 0\}$ and $L(\rho_\varepsilon) := \sup\{x :\ \rho_\varepsilon(x) \neq 0\}$ vanish in the limit as $\varepsilon \to 0$. This finally motivates the following

**5.16 Definition.** *We call an element $D$ of $\mathcal{G}(\mathbb{R})$ a* generalized delta function *if it admits a strict delta net as a representative, i.e., $D = \mathrm{cl}[(\rho_\varepsilon)_\varepsilon]$ with $\rho_\varepsilon$ satisfying conditions (a), (b) and (c) above.*

**5.17** Note that the above regularization of the $\delta$-distribution besides providing the key for a successful mathematical treatment of the geodesic equation has a sensible physical interpretation in terms of the impulsive limit of sandwich waves (cf. 5.6). Also substituting $\delta$ by a generalized delta function in the metric (90) turns it into a generalized metric in the sense of sec. 4.C (the determinant not depending on $\varepsilon$ at all). Then the regularized geodesic equations (see (93) below) are just the generalized geodesic equations for the generalized metric. We shall return to this point of view in sec. 5.D.

**5.18** Replacing the Dirac-$\delta$ in the system (91) by a generalized delta function $D$ we may write out explicitly the geodesic equations for impulsive pp-waves in the generalized functions setting. Denoting the generalized functions corresponding to $x^i$ and $v$ by the respective capital letters we get the *generalized geodesic equations*

$$
\begin{aligned}
\ddot{V}(u) &= f(X^j(u))\,\dot{D}(u) + 2\,\partial_i f(X^j(u))\,\dot{X}^i\,(u)\,D(u)\\
\ddot{X}^i(u) &= \frac{1}{2}\,\partial_i f(X^j(u))\,D(u)\,.
\end{aligned}
\tag{93}
$$

Next we shall be concerned with the question of initial conditions. We have to prescribe the initial positions and velocities at a certain value $u_0$ of $u$ which is playing the role of a (retarded) time. It is physically reasonable to prescribe initial values "long before" the shock at $u = 0$; so let us choose $u_0 = -1$. Note, however, that choosing $u_0 = 0$ would mean to "start at the shock" and one has to expect to (and indeed does) end up with a regularization dependent result in this case. Hence we choose our initial conditions as follows

$$
\begin{aligned}
V(-1) = v_0 \qquad X^i(-1) &= x_0^i\\
\dot{V}(-1) = \dot{v}_0 \qquad \dot{X}^i(-1) &= \dot{x}_0^i\,,
\end{aligned}
\tag{94}
$$

where $v_0$, $\dot{v}_0$, $x^i$ and $\dot{x}^i$ are real numbers. Note that our mathematical formalism actually is even capable of treating initial conditions in $\mathcal{R}$, which in this case, however, seems to have no sensible physical interpretation.

**5.19** A main problem one has to face in our approach is that, since generalized solutions are constructed componentwise, i.e., from componentwise classical (smooth) solutions to *nonlinear* equations, we have to ensure that the interval of existence does not shrink to a point in the limit $\varepsilon \to 0$. In proving general theorems one usually puts strong restrictions on the right hand side of the equations to guarantee this behavior (cf. [83], [117], chap. 1.10.1). However, putting restrictions on $f$ would seriously affect physical applicability of our considerations, and so we are forced to make use of the special form of the equations. Indeed we have the following technical result which guarantees global existence of solutions for small $\varepsilon$ and also ensures that the composition of the smooth function $f$ with the generalized solution $X^j$ is moderate. (cf. 3.7)

**5.20 Lemma.** *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R} \to \mathbb{R}^n$ smooth and $(\rho_\varepsilon)_\varepsilon$ a net of smooth functions satisfying conditions (a) and (c) as above. For any $x_0, \dot{x}_0 \in \mathbb{R}^n$ and any $\varepsilon \in (0,1]$ consider the system*

$$\begin{aligned} \ddot{x}_\varepsilon(t) &= g(x_\varepsilon(t))\rho_\varepsilon(t) + h(t) \\ x_\varepsilon(-1) &= x_0 \\ \dot{x}_\varepsilon(-1) &= \dot{x}_0 \,. \end{aligned} \tag{95}$$

*Let $b > 0$, $M = \int_{-1}^{1}\int_{-1}^{s} h(r)\,dr ds$, $I = \{x \in \mathbb{R}^n : |x - x_0| \leq b + |\dot{x}_0| + M\}$ and*

$$\alpha = \min\left\{ \frac{b}{C\|g\|_{L^\infty(I)} + |\dot{x}_0|},\, \frac{1}{2LC},\, 1 \right\},$$

*with L a Lipschitz constant for $g$ on $I$. Then (95) has a unique solution $x_\varepsilon$ on $J_\varepsilon = [-1, \alpha - \varepsilon]$. Consequently, for $\varepsilon$ sufficiently small $x_\varepsilon$ is globally defined and both $x_\varepsilon$ and $\dot{x}_\varepsilon$ are bounded, uniformly in $\varepsilon$, on compact sets.*

**Proof.** We shall work on the (nonempty, closed) subset $X_\varepsilon := \{x_\varepsilon \in \mathcal{C}(J_\varepsilon) : |\, x_\varepsilon(t) - x_0 \,| \leq b + |\dot{x}_0| + M\,\}$ of the Banach space of continuous functions on the interval $J_\varepsilon$. Since no first order derivative of $x_\varepsilon$ enters the right hand side of equation (95) we may define an integration operator $A$ by

$$[A\,x_\varepsilon]\,(t) \;:=\; x_0 + \dot{x}_0(t+1) + \int_{-\varepsilon}^{t}\int_{-\varepsilon}^{s} g(x_\varepsilon(r))\,\rho_\varepsilon(r)\,dr\,ds + \int_{-1}^{t}\int_{-1}^{s} h(r)\,dr\,ds \;.$$

By construction $Ax_\varepsilon \in \mathcal{C}(J_\varepsilon)$, but we even have $Ax_\varepsilon \in X_\varepsilon$ since

$$\begin{aligned} |\ [A\,x_\varepsilon]\,(t) - x_0 \ | \ &\leq |\dot{x}_0|(t+1) + \int_{-\varepsilon}^{t}\int_{-\varepsilon}^{s} |\ g(x_\varepsilon(r))\,|\,|\ \rho_\varepsilon(r)\ |\ dr\,ds + \int_{-1}^{t}\int_{-1}^{s} |h(r)|\,dr\,ds \\ &\leq |\dot{x}_0| + \alpha(|\dot{x}_0| + \|g\|_{\infty,I}\|\rho\|_1) + M \ \leq\ |\dot{x}_0| + b + M \end{aligned}$$

By a similar estimate one shows that $A$, in fact, is a contraction on $X_\varepsilon$. Hence by the fixed point theorem we have a unique solution $x_\varepsilon \in X_\varepsilon$.
Since $\alpha$ is independent of $\varepsilon$ we have for small $\varepsilon$ that: $-\varepsilon + \alpha \geq \varepsilon$. But at $t = \varepsilon$ the nonlinearity at the right hand side of the differential equation (95) has already been "turned off".
Furthermore, since for all $\varepsilon$ the solution lies in $X_\varepsilon$, $x_\varepsilon$ is bounded uniformly in $\varepsilon$ for say $t \leq \alpha/2$ (for $\varepsilon$ small enough). For larger values of $t$, $x_\varepsilon$ essentially grows only linearly with $\dot{x}_\varepsilon$ which is bounded by $\dot{x}_0 + \|g\|_{\infty,I}\|\rho\|_1$. Hence for small $\varepsilon$ the solution $x_\varepsilon(t)$ is defined on the whole real line and is bounded uniformly in $\varepsilon$ on every compact set. □

Now we are ready to prove the following

**5.21 Theorem.** *(Existence and uniqueness of generalized geodesics) Let $D \in \mathcal{G}(\mathbb{R})$ be a generalized delta function, $f \in \mathcal{C}^\infty(\mathbb{R}^2)$ and let $v_0$, $\dot{v}_0$, $x_0^i$, $\dot{x}_0^i \in \mathbb{R}$ $(i = 1, 2)$. The initial value problem*

$$\begin{aligned} \ddot{V}(u) \ &= f(X^j(u))\,\dot{D}(u) \ + \ 2\,\partial_i f(X^j(u))\,\dot{X}^i\,(u)\,D(u) \\ \ddot{X}^i(u) &= \frac{1}{2}\,\partial_i f(X^j(u))\,D(u) \\ V(-1) = v_0 \qquad X^i(-1) \ &= \ x_0^i \\ \dot{V}(-1) = \dot{v}_0 \qquad \dot{X}^i(-1) \ &= \ \dot{x}_0^i \end{aligned} \tag{96}$$

*has a unique, locally bounded solution $(V, X^1, X^2) \in \mathcal{G}(\mathbb{R})^3$.*

**Proof.** First we prove *existence*. Componentwise we have to deal with the equations

$$\begin{aligned} \ddot{v}_\varepsilon(u) &= f(x_\varepsilon^j(u))\,\dot{\rho}_\varepsilon(u) \ + \ 2\,\partial_i\,f(x_\varepsilon^j(u))\ \dot{x}_\varepsilon^i(u)\,\rho_\varepsilon(u) \ , \\ \ddot{x}_\varepsilon^i(u) &= \frac{1}{2}\,\partial_i\,f(x_\varepsilon^j(u))\,\rho_\varepsilon(u) \end{aligned} \tag{97}$$

$$\begin{aligned} v_\varepsilon(-1) \ = \ v_0, \quad x_\varepsilon^i(-1) \ = \ x_0^i \ , \\ \dot{v}_\varepsilon(-1) \ = \ \dot{v}_0, \quad \dot{x}_\varepsilon^i(-1) \ = \ \dot{x}_0^i \ . \end{aligned} \tag{98}$$

According to lemma 5.20 (setting $h \equiv 0$), the second line of (97) for fixed $\varepsilon$ has a unique globally defined solution $x^i_\varepsilon$ with the specified initial values. Inserting this into the first line and integrating we also obtain a solution $v_\varepsilon$. From the boundedness properties of $x^i_\varepsilon$ established in lemma 5.20 and the fact that $\rho_\varepsilon \in \mathcal{E}_M(\mathbb{R})$ it follows easily by induction that $x^i_\varepsilon$ and $v_\varepsilon$ are moderate as well. Hence their respective classes in $\mathcal{G}(\mathbb{R})$ define solutions to (96).

*Uniqueness.* Suppose that $\tilde{V} = \mathrm{cl}[(\tilde{v}_\varepsilon)_\varepsilon]$ and $\tilde{X}^i = \mathrm{cl}[(\tilde{x}^i_\varepsilon)_\varepsilon]$ are locally bounded solutions of (97) as well. On the level of representatives this means that there exist $M = \mathrm{cl}[(m_\varepsilon)_\varepsilon]$, $N^i = \mathrm{cl}[(n^i_\varepsilon)_\varepsilon] \in \mathcal{N}(\mathbb{R})$ and $n_{x^i}$, $n_{\dot{x}^i}$, $n_v$, $n_{\dot{v}} \in \mathcal{N}$ with

$$\begin{aligned}
\ddot{\tilde{v}}_\varepsilon(u) &= f(\tilde{x}^j_\varepsilon(u))\,\dot{\rho}_\varepsilon(u) \;+\; 2\,\partial_i\, f(\tilde{x}^j_\varepsilon(u))\;\dot{\tilde{x}}^i{}_\varepsilon(u)\,\rho_\varepsilon(u) + m_\varepsilon(u)\ ,\\
\ddot{\tilde{x}}^i{}_\varepsilon(u) &= \frac{1}{2}\,\partial_i\, f(\tilde{x}^j_\varepsilon(u))\,\rho_\varepsilon(u) + n^i_\varepsilon(u)\\
\tilde{v}_\varepsilon(-1) \;=\; v_0 + n_v, &\qquad \tilde{x}^i_\varepsilon(-1) \;=\; x^i_0 + n_{x^i}\ ,\\
\dot{\tilde{v}}_\varepsilon(-1) \;=\; \dot{v}_0 + n_{\dot{v}}, &\qquad \dot{\tilde{x}}^i{}_\varepsilon(-1) \;=\; \dot{x}^i_0 + n_{\dot{x}^i}\ .
\end{aligned} \tag{99}$$

We have to show that $(v_\varepsilon - \tilde{v}_\varepsilon)_\varepsilon$ and $(x^i_\varepsilon - \tilde{x}^i_\varepsilon)_\varepsilon$ belong to the ideal $\mathcal{N}(\mathbb{R})$. Since $N^i \in \mathcal{N}(\mathbb{R})$ it follows that for $\varepsilon$ sufficiently small $n^i_\varepsilon$ is bounded on compact sets, uniformly in $\varepsilon$. Thus by lemma 5.20 (with $h$ equal to this constant) the same holds true for $\tilde{x}^i_\varepsilon$ and its first derivative. From (99) we conclude

$$\begin{aligned}
(x^i_\varepsilon - \tilde{x}^i_\varepsilon)(u) = &-n_{x^i\,\varepsilon} - (u+1)n_{\dot{x}^i\,\varepsilon}\\
&\frac{1}{2}\int\limits_{-1}^{u}\int\limits_{-1}^{s} \rho_\varepsilon(r)[\partial_i f(x^j_\varepsilon(r)) - \partial_i f(\tilde{x}^j_\varepsilon(r))\,]drds \;-\; \int\limits_{-1}^{u}\int\limits_{-1}^{s} n_\varepsilon(r)\,drds
\end{aligned}$$

Hence $\forall T>0\ \forall q \in \mathbb{N}_0\ \exists C>0\ \exists \eta>0\ \forall \varepsilon \in (0,\eta)\ \forall u \in [-T,T]$:

$$|(x^i_\varepsilon - \tilde{x}^i_\varepsilon)(u)| \le C\varepsilon^q + \frac{1}{2}\int\limits_{-1}^{u}\int\limits_{-r}^{u}\int\limits_{0}^{1} |\nabla\partial_i f(\sigma x^i_\varepsilon(r) + (1-\sigma)\tilde{x}^i_\varepsilon(r)|d\sigma|(x^i_\varepsilon - \tilde{x}^i_\varepsilon)(r)|\,|\rho_\varepsilon(r)|dsdr$$

By the boundedness properties of $x^i$ and $\tilde{x}^i$ and by condition (c) on $\rho_\varepsilon$ an application of Gronwall's lemma to the above inequality yields the $\mathcal{N}$-estimates for $(x^i_\varepsilon - \tilde{x}^i_\varepsilon)$. Thanks to proposition 3.6 we are done. □

Note that in the above proof we only have used conditions (a) and (c) in the definition of strict delta nets. On the other hand condition (b) will be the essential ingredient to establish the following

**5.22 Theorem.** *(Associated distributions for the generalized geodesics) The unique solution $(V, X^i)$ of the geodesic equation (96) satisfies the following association relations:*

$$\begin{aligned}
X^i &\approx x^i_0 + \dot{x}^i_0(1+u) + \frac{1}{2}\partial_i f(x^j_0 + \dot{x}^j_0)u_+\\
V &\approx v_0 + \dot{v}_0(1+u) + f(x^j_0 + \dot{x}^j_0)H(u) + \partial_i f(x^j_0 + \dot{x}^i_0)\left(\dot{x}^i_0 + \frac{1}{4}\partial_i f(x^j_0 + \dot{x}^j_0)\right)u_+\ ,
\end{aligned} \tag{100}$$

*where $H$ and $u_+$ denote the Heaviside and kink function, respectively. The first line even holds in the sense of 0-association.*

**Proof.** Representatives of the unique $\mathcal{G}$ solution are implicitly given by

$$\begin{aligned}
x^i_\varepsilon(u) &= x^i_0 + \dot{x}^i_0\,(1+u) + (H * H * \ddot{x}^i_\varepsilon)\,(u)\\
v_\varepsilon(u) &= v_0 + \dot{v}_0\,(1+u) + (H * H * \ddot{v}_\varepsilon)\,(u)\ ,
\end{aligned} \tag{101}$$

where "$*$" denotes convolution. Since distributions supported in an acute cone form an (associative) convolution algebra (where, in particular, convolution is a separately continuous operation [87], §4.9 proposition 8) it suffices to calculate the limits of the right hand sides of (97); the distributional limits of the solutions (101) are then computed simply by integration.

We begin with the latter two equations of system (97) and choose a test function $\psi$. We have

$$
\begin{aligned}
\langle \ddot{x}_\varepsilon - \frac{1}{2}\partial_i f(x_0^j + \dot{x}_0^j)\delta, \psi\rangle &= \frac{1}{2}\int\limits_{-\varepsilon}^{\varepsilon} \partial_i f(x_\varepsilon^i(u))\rho_\varepsilon(u)\psi(u)\, du - \frac{1}{2}\partial_i f(x_0^j + \dot{x}_0^j)\psi(0) \\
&= \frac{1}{2}\int\limits_{-\varepsilon}^{\varepsilon} \Big[\partial_i f(x_\varepsilon^i(u))\psi(u) - \partial_i f(x_0^j + \dot{x}_0^j)\psi(0)\Big]\rho_\varepsilon(u)\, du \\
&\quad +[\int\limits_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(u)\, du - 1]\frac{1}{2}\partial_i f(x_0^j + \dot{x}_0^j)\psi(0)
\end{aligned}
$$

The last term above vanishes in the limit $\varepsilon \to 0$ due to condition (b), while the absolute value of the first term may estimated by

$$
\frac{1}{2}\sup_{|u|\le 1} |\partial_i f(x_\varepsilon^i(\varepsilon u))\psi(\varepsilon u) - \partial_i f(x_0^j + \dot{x}_0^j)\psi(0)| \int\limits_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(u)|\, du
$$

which also vanishes in the limit due to condition (c) and the fact $x_\varepsilon^j(\varepsilon u) \to x_0^j + \dot{x}_0^j$ uniformly. To see this we estimate on an arbitrary $K \subset\subset \mathbb{R}$

$$
\sup_{u\in K} |x_\varepsilon^j(\varepsilon u) - x_0^i + \dot{x}_0^i| \le \sup_{u\in K}[\, |\dot{x}_0^i \varepsilon u| + \frac{1}{2}\int\limits_{-\varepsilon}^{\varepsilon u}\int\limits_{-\varepsilon}^{s} |\partial_i f(x_\varepsilon^i(r))|\, |\rho_\varepsilon(r)|\, drds\,] \tag{102}
$$

which vanishes in the limit by the boundedness properties of $x^i$ established in lemma 5.20 and, again by condition (c).
Since $\dot{x}_\varepsilon^i$ is bounded on compact sets, uniformly in $\varepsilon$, it follows that the family $\{x_\varepsilon^i : \varepsilon \in (0, \varepsilon_0]\}$ (with $\varepsilon_0$ small enough) is locally equicontinuous. Hence Ascoli's theorem implies locally uniform convergence.
Finally concerning $V$ we have to calculate the limit of

$$
\ddot{v}_\varepsilon(u) = [f(x_\varepsilon^i(u))\rho_\varepsilon(u)]^{\cdot} + \partial_i f(x_\varepsilon^i(u))\dot{x}_\varepsilon^i(u)\rho_\varepsilon(u)\,.
$$

The first summand converges to $f(x_0^i + \dot{x}_0^i)\dot{\delta}$ by an argument similar to the one above. For the second summand we have

$$
\partial_i f(x_\varepsilon^i(u))\dot{x}_\varepsilon^i(u)\rho_\varepsilon(u) = \underbrace{\partial_i f(x_\varepsilon^i(u))\rho_\varepsilon(u)\dot{x}_0^i}_{(*)} + \frac{1}{2}\partial_i\partial_i f(x_\varepsilon^i(u))\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} \partial_i f(x_\varepsilon^i(s))\rho_\varepsilon(s)ds
$$

and $(*) \to \partial_i f(x_0^i + \dot{x}_0^i)\dot{x}_0^i\delta$ again by similar arguments. Finally, we write out

$$
\begin{aligned}
&\int\limits_{-\varepsilon}^{\varepsilon} \psi(u)\partial_i f(x_\varepsilon^j(u))\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} \partial_i f(x_\varepsilon^j(s))\rho_\varepsilon(s)\, dsdu - \frac{1}{2}\partial_i f(x_0^i + \dot{x}_0^i)^2\psi(0) \\
&= \int\limits_{-\varepsilon}^{\varepsilon} \psi(u)\partial_i f(x_\varepsilon^j(u))\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} [\partial_i f(x_\varepsilon^j(s)) - \partial_i f(x_0^i + \dot{x}_0^i)]\rho_\varepsilon(s)\, dsdu \\
&\quad + \int\limits_{-\varepsilon}^{\varepsilon} \psi(u)[\partial_i f(x_\varepsilon^j(u)) - \partial_i f(x_0^i + \dot{x}_0^i)]\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} \partial_i f(x_0^i + \dot{x}_0^i)\rho_\varepsilon(s)\, dsdu \\
&\quad + \int\limits_{-\varepsilon}^{\varepsilon} (\psi(u) - \psi(0))\partial_i f(x_0^i + \dot{x}_0^i)\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} \partial_i f(x_0^i + \dot{x}_0^i)\rho_\varepsilon(s)\, dsdu \\
&\quad +\psi(0)\partial_i f(x_0^i + \dot{x}_0^i)^2[\int\limits_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(u)\int\limits_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(s)\, dsdu - \frac{1}{2}]\,.
\end{aligned} \tag{103}
$$

Now we may estimate the absolute value of the above by

$$
\begin{aligned}
| \quad | &\le \underbrace{\sup_{|s|\le\varepsilon} |\partial_i f(x^j_\varepsilon(u)) - \partial_i f(x^i_0 + \dot x^i_0)|}_{\to 0} \underbrace{\int\limits_{-\varepsilon}^{\varepsilon} |\psi(u)|\, |\partial_i f(x^j_\varepsilon(u))|\, |\rho_\varepsilon(u)| \int\limits_{-\varepsilon}^{u} |\rho_\varepsilon(s)|\, ds du}_{\text{bounded}} \\
&+ \underbrace{\sup_{|u|\le\varepsilon} |\partial_i f(x^j_\varepsilon(u)) - \partial_i f(x^i_0 + \dot x^i_0)|}_{\to 0} \underbrace{\int\limits_{-\varepsilon}^{\varepsilon} |\psi(u)|\, |\rho_\varepsilon(u)|\, |\partial_i f(x^i_0 + \dot x^i_0)| \int\limits_{-\varepsilon}^{u} |\rho_\varepsilon(u)|\, ds du}_{\text{bounded}} \\
&+ \underbrace{\sup_{|u|\le\varepsilon} |\psi(u) - \psi(0)|}_{\to 0} \underbrace{\int\limits_{-\varepsilon}^{\varepsilon} |\partial_i f(x^i_0 + \dot x^i_0)|^2 |\rho_\varepsilon(u)| \int\limits_{-\varepsilon}^{u} |\rho_\varepsilon(s)|\, ds du}_{\text{bounded}} \\
&+ |\psi(0)|\, |\partial_i f(x^i_0 + \dot x^i_0)^2|\, \underbrace{[\frac{1}{2}(\int\limits_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(s)|\, ds)^2 - \frac{1}{2}]}_{\to 0}
\end{aligned}
\tag{104}
$$

Hence the claim follows. □

**5.23** We finally come to the last point of our program outlined in 5.14, namely interpretation of our results. From theorem 5.22 we conclude that viewed distributionally, the geodesics are given by refracted, broken straight lines, as suggested by the form of the metric (90). The constant coefficients of the step resp. kink function are given by the values of the function $f$ and its first derivatives at the shock hypersurface which may be traced back to the initial values.
Of course equations (100) coincide with the earlier results [62,15]. However from the point of view of our approach the (deeper) reason why here the "rule" $H\delta = (1/2)\,\delta$ used by [62] (which in fact coincides with the "determination of the point value" $H(0) = 1/2$ used by [15]) leads to a physically reasonable result is the following: The metric involves *only one* singular object and hence the $\delta$'s as well as the $H$'s appearing in the geodesic equations share the same root; namely the $\delta$-shaped wave profile. Hence, when regularizing the equations both factors of the ill-defined product naturally involve the same regularization which immediately leads to the (regularization-independent) result $\rho_\varepsilon \int \rho_\varepsilon \to (1/2)\,\delta$. This point is even more strengthened if one views equations (96) as the geodesic equations of the the generalized spacetime (in the sense of sec. 4.C). Moreover, note that the regularization independence of our results has the following important physical consequence: In the impulsive limit the geodesics are totally independent of the particular shape of the sandwich wave. Hence the impulsive wave "totally forgets its seed" which is in accordance with the results in [160].

**5.24** Finally we turn to the central issue of diffeomorphism invariance. Although we have chosen to perform our analysis in a special coordinate system, namely, in a geometrically preferred one, easing the actual calculations as well as aiding intuition our results are actually independent of this choice. The key fact is that we have modelled the singular metric coefficient by the very general class of generalized delta functions. Since this class is manifestly diffeomorphism invariant, modelling the singular metric in any other coordinate system (to be sure: a diffeomorphically related one, as opposed to the continuous *form* of the metric) just amounts to the choice of a different generalized delta function in the original system. Moreover, due to the fact that our results, especially the distributional shadows of the generalized solutions, do not depend on the regularization within this class and since the action of a diffeomorphism is compatible with association (resp. since association is a well-defined concept on manifolds at all, cf. 3.52 (v) and 4.26) they are in fact diffeomorphism invariant.
We thus conclude that the geodesic equation can be treated consistently in a nonlinear distributional setting, leading to a regularization-independent distributional result. This, of course, is only possible due to the relatively mild character of the singular terms which allows for a distributional shadow of the generalized solutions at all. However, this is not obvious (cf. for example [173]), but should rather strengthen our trust in impulsive waves as reasonable solutions of Einstein's equations. Furthermore, we shall see in a moment that even in the considerably more complicated case of the Jacobi equation our strategy can be applied successfully.

**5.25** To conclude this section we remark that unique solvability of the geodesic (and also the geodesic deviation equation which is the topic of the next section) for (90) is not confined to the case where the profile function $f$ is smooth. Indeed, it turns out that for a large class of generalized profile functions (those that are not "too singular") theorems 5.21 and 5.22 (resp. 5.29 and 5.30 below) retain their validity. More precisely, we have to demand that $f$ belongs to the algebra of *tempered* generalized functions (see [39], chap. 4) to make sure that the composition $f(X)$ is well-defined and that $\nabla\nabla f$ is of $L^\infty$-log-type(cf. [83], definition 2.3) to ensure existence and uniqueness of solutions to (96) and (106). However, to include many physically interesting examples (cf. [132]) one has to cut out the world line of the ultrarelativistic particle, i.e., the $v$-axis from the domain of definition (cf. [173]).

#### 5.C. Geodesic Deviation for Impulsive pp-Waves

Our next goal is an analysis of the Jacobi equation for impulsive pp-waves in the framework of algebras of generalized functions again following the blueprint in 5.14.

**5.26** As in references [174] and [118] to keep formulas more transparent we make some simplifying assumptions concerning geometry (namely axisymmetry) and initial conditions. More precisely, we restrict the function $f$ of the transverse coordinates $x^i = (x, y)$ in the metric tensor (90) to depend on the two-radius $r = \sqrt{x^2+y^2}$ only and work entirely within the $y = 0$-hypersurface (initial conditions $x_0^2 \equiv y_0 = 0 = \dot{x}_0^2 \equiv \dot{y}_0$). Furthermore we take initial values $v_0 = 0 = \dot{x}_0$.

**5.27** Like in the previous section we start with a purely distributional analysis of the equations at hand. With the assumptions on the geometry made above the distributional geodesics, i.e, the shadows of the unique $\mathcal{G}$ solutions to the geodesic equation given in theorem 5.22 simplify to

$$\begin{aligned}
V &\approx \dot{v}_0\,(1+u) + f(x_0)\,H(u) + \frac{1}{4}\,f'(x_0)^2\,u_+ \ ,\\
X^1 &\approx x_0 + \frac{1}{2}\,f'(x_0)\,u_+ \ ,\\
X^2 &\approx 0\ ,
\end{aligned}$$

where $f'$ denotes the derivative with respect to the single variable $r = x$. Hence we have to deal with a geodesic tangent vector of the form (recall that differentiation commutes with association)

$$T^a(u) = \begin{pmatrix} 1 \\ \dot{v}_0 + f(x_0)\,\delta(u) + \frac{1}{4}\,f'(x_0)^2\,H(u) \\ \frac{1}{2}\,f'(x_0)\,H(u) \\ 0 \end{pmatrix} ,$$

where in the following we are going to use the abbreviations $A := f(x_0)\,\delta(u) + (1/4)f'(x_0)^2\,H(u)$ and $B := (1/2)f'(x_0)\,H(u)$ for its components.

Next we compute the explicit form of the Jacobi equation $\frac{D^2N^a}{du^2} = -R^{\;a}_{b\;cd}T^bT^dN^c$ for a vector field $N^a(u) = (N^u(u), N^v(u), N^x(u), N^y(u))$ over the geodesic. After some (tedious) calculations we end up with the following form of the system

$$\begin{aligned}
\ddot{N}^u &= 0\ ,\\
\ddot{N}^v &= 2[N^x f'\delta]^{\cdot} - N^x f'\dot{\delta} + [N^u f\delta]^{\cdot\cdot} - N^u f'' B^2\delta - N^u f'\dot{B}\delta\ ,\\
\ddot{N}^x &= [\dot{N}^u f' + \frac{1}{2}\,N^x f'']\,\delta + \frac{1}{2}\,f' N^u\dot{\delta}\ ,\\
\ddot{N}^y &= 0\ ,
\end{aligned} \tag{105}$$

where we have suppressed the dependence on the parameter $u$ and the variable $x$. Equations (105) form a system of four coupled ODEs linear in the components of the vector field $N^a$ but nonlinear in the derivatives of the metric. From the fact that $B$ involves the step function we see immediately that (in the second equation) we again have to deal with distributionally ill-defined expressions, but now of even worse type than before. Indeed the term $\dot{B}\delta$ is proportional to the "square" of the Dirac $\delta$-distribution, and the term

$B^2\delta$ involves an expression "$H^2\delta$"[15]. Note, however, that the critical terms arise from the second covariant derivative where some of the Christoffel symbols get multiplied, yet not from the Riemann tensor whose components are just proportional to the $\delta$-distribution.

**5.28** We now transfer system (105) into our generalized functions setting by again replacing every $\delta$-distribution by a generalized delta function and $x$ by the unique solution to system (96) additionally satisfying the simplifications on geometry and initial conditions as discussed in 5.26. Again this procedure may also be interpreted as writing out the generalized Jacobi equation for the generalized metric (cf. 5.17). Finally, we choose (classical) initial positions and velocities at $u = -1$. Existence and uniqueness of solutions to the initial value problem in the Colombeau algebra is established in the following result where, for the sake of brevity, we denote the $\mathcal{G}$-functions corresponding to $N^a$ again by $N^a$.

**5.29 Theorem.** *(Existence and uniqueness of the generalized Jacobi field) Let $D \in \mathcal{G}(\mathbb{R})$ be a generalized delta function, $f \in \mathcal{C}^\infty(\mathbb{R})$, $n^a$, $\dot{n}^a \in \mathbb{R}^4$ and let $X$ denote the (unique) solution to system (96) with initial conditions and simplifications as discussed above. The initial value problem*

$$\begin{aligned}
\ddot{N}^v &= 2[N^x f'(X)D]^{\cdot} - N^x f'(X)\dot{D} + [N^u f(X)D]^{\cdot\cdot} - \\
&\quad -N^u f''(X)\dot{X}^2 D - N^u f'(X)\ddot{X} D \\
\ddot{N}^x &= [\dot{N}^u f'(X) + \frac{1}{2} N^x f''(X)]D + \frac{1}{2} f'(X) N^u \dot{D} \qquad (106)\\
\ddot{N}^y &= \ddot{N}^u \ = \ 0 \\
N^a(-1) = n^a &\qquad \dot{N}^a(-1) \ = \ \dot{n}^a \qquad (107)
\end{aligned}$$

*has a unique solution $N^a \in \mathcal{G}(\mathbb{R})^4$.*

**Proof.** We start proving *existence*. Since the equations are linear in the components of the deviation field we are provided with globally defined solutions on the level of representatives.
The last two equations of (106) are actually trivial and so is the first one once we know that its right hand side belongs to $\mathcal{G}(\mathbb{R})$, since then it is of the form $\ddot{N}(u) = G(u)$, where $G = \mathrm{cl}[(g_\varepsilon)_\varepsilon]$ is moderate and $\mathrm{supp}(g_\varepsilon) \subseteq [\varepsilon, \varepsilon]$. Hence the $\mathcal{E}_M$-bounds follow from

$$|N_\varepsilon(u)| \ \leq \ |N_0| + |\dot{N}_0(1+u)| + \int_{-\varepsilon}^{t}\int_{-\varepsilon}^{\varepsilon} |g_\varepsilon(r)|\, drds\,.$$

So we are left with the equation for $N^x$ which is of the form $\ddot{N}(u) = f''(X(u))D(u)N(u) + G(u)$ with $G$ in $\mathcal{G}(\mathbb{R})$. Using the boundedness properties of $X$ established in lemma 5.20 as well as condition (c) of the strict delta net the $\mathcal{E}_M$-bounds for $N^x$ follow from Gronwall's lemma.
*Uniqueness* is established along the same lines again using Gronwall-type arguments. □

Note, that in the above proof we have again only used properties (a) and (c) of the generalized delta function $D$.
Our next task is to calculate associated distributions for the solution of the Jacobi equation (where we shall make use also of (b)). To maintain clarity of formulae we shall make further simplifying assumptions, this time on the initial conditions, i.e.,

$$\begin{aligned}
N^a(-1) &= (0,0,0,0) \\
\dot{N}^a(-1) &= (a,b,0,0)\ . \qquad (108)
\end{aligned}$$

Then we have the following

[15] Note that if "$H^2\delta$" is treated by the "multiplication rules" $H\delta = 1/2\delta$ and $H^2 = H$ we encounter the following non-associativity (see also 2.7) $(HH)\delta = H\delta = 1/2\delta \neq 1/4\delta = H(1/2)\delta = H(H\delta)$.

**5.30 Theorem.** *(Associated distributions for the generalized Jacobi field) The unique solution of the geodesic deviation equation (106) with initial conditions (108) satisfies the following association relations*

$$N^x \approx \frac{1}{2} a f'(x_0)(u_+ + H(u)) \tag{109}$$

$$N^v \approx b(1+u) + a[f(x_0)\delta(u) + \frac{1}{4} f'(x_0)^2 (H(u) + u_+)] \,. \tag{110}$$

**Proof.** As in the case of theorem 5.22 it suffices to calculate the limits of the right hand sides of (106) and then to perform integration. The last two equations are actually trivial and we start with the equation for $\ddot{N}^x_\varepsilon$. For that purpose we write

$$\ddot{N}^x_\varepsilon(u) \;=\; I_\varepsilon(u)\rho_\varepsilon(u) + II_\varepsilon(u)\dot{\rho}_\varepsilon(u) + II_\varepsilon(u)u\dot{\rho}_\varepsilon(u)\,,$$

where we have used the abbreviations $I_\varepsilon(u) = af'(x_\varepsilon(u)) + \frac{1}{2}N^x_\varepsilon(u)f''(x_\varepsilon(u))$ and $II_\varepsilon(u) = \frac{1}{2}af'(x_\varepsilon(u))$. We start by calculating the limits term by term. First by partial integration ($\psi$ again a test function)

$$\langle II_\varepsilon\dot{\rho}_\varepsilon, \psi\rangle = \underbrace{-\frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(u)\dot{\psi}(u) f'(x_\varepsilon(u))\, du}_{A_\varepsilon}$$
$$\underbrace{-\frac{1}{4}\, a \int_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(u)\psi(u) f''(x_\varepsilon(u)) \int_{-\varepsilon}^{u} f'(x_\varepsilon(s))\rho_\varepsilon(s)\, ds\, dt}_{B_\varepsilon}$$

and $A_\varepsilon \to -(1/2)af'(x_0)\dot{\psi}(0)$ by an argument analogous to the one used in equations (103) resp. (104) in the proof of theorem 5.22. Similarly, by adding and subtracting the obvious terms one finds that $B_\varepsilon \to -(1/8)af'(x_0)f''(x_0)\psi(0)$ hence

$$II_\varepsilon\dot{\rho} \to \frac{1}{2}af'(x_0)\dot{\delta} - \frac{1}{8}af'(x_0)f''(x_0)\delta\,. \tag{111}$$

Again by a similar argument we conclude that

$$II_\varepsilon(u)u\dot{\rho}_\varepsilon(u) \;\to\; -\frac{1}{2}af'(x_0)\delta(u)\,. \tag{112}$$

So only the difficult term $I_\varepsilon\rho_\varepsilon$, involving the unknown function $N_\varepsilon$ itself, remains. More precisely we have

$$I_\varepsilon(u)\rho_\varepsilon(u) \;=\; \underbrace{af'(x_\varepsilon(u))\rho_\varepsilon(u)}_{\to af'(x_0)\delta(u)} + \frac{1}{2}N^x_\varepsilon(u)f''(x_\varepsilon(u))\rho_\varepsilon(u)$$

and we start by inserting $H * H * \ddot{N}^x_\varepsilon$ for $N^x_\varepsilon$ to get

$$\frac{1}{2}N^x_\varepsilon(u)f''(x_\varepsilon(u))\rho_\varepsilon(u) = \underbrace{\frac{1}{2}f''(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\rho_\varepsilon(r)\, dr\, ds}_{C_\varepsilon}$$
$$\underbrace{+\frac{1}{4}af''(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} N^x_\varepsilon(r)f''(x_\varepsilon(r))\rho_\varepsilon(r)\, dr\, ds}_{D_\varepsilon}$$
$$\underbrace{+\frac{1}{4}af''(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot{\rho}_\varepsilon(r)\, dr\, ds}_{E_\varepsilon}$$

$$\underbrace{+\frac{1}{4}af''(x_\varepsilon(u))\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot\rho_\varepsilon(r)\,r\,dr\,ds}_{F_\varepsilon}\;.$$

$C_\varepsilon \to 0$ since we have

$$|\langle C_\varepsilon,\psi\rangle| \;\le\; \frac{1}{2}\,a||\psi||_\infty \sup_{|u|\le\varepsilon}|f'(x_\varepsilon(u))|^2\underbrace{\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(u)|\,du}_{\text{bounded}}\underbrace{\int\limits_{-\varepsilon}^{\varepsilon}\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(r)|\,dr\,ds}_{\to 0}\;.$$

Also one finds $E_\varepsilon \to (1/8)af'(x_0)f''(x_0)\delta$ by again adding and subtracting the obvious terms and $F_\varepsilon \to 0$ by the estimate

$$|\langle F_\varepsilon,\psi\rangle| \;\le\; \frac{1}{4}\,a||\psi||_\infty \sup_{|u|\le\varepsilon}|f''(x_\varepsilon(u))|^2\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(u)|\,du[\underbrace{\int\limits_{-\varepsilon}^{\varepsilon}|s|\,|\rho_\varepsilon(s)|\,ds}_{\to 0}-\underbrace{\int\limits_{-\varepsilon}^{\varepsilon}\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(r)|\,dr\,ds}_{\to 0}]\,.$$

Finally, to prove that $D_\varepsilon$ also vanishes in the limit we first show that $N^x_\varepsilon$ is bounded on compact sets (note that we cannot expect the limit of $N^x_\varepsilon$ to be continuous). Denoting constants by $C$, $C'$, and so on, we find

$$\begin{aligned}
|N^x_\varepsilon(u)| &\le C + C'\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}|N^x_\varepsilon(r)|\,|\rho_\varepsilon(r)|\,dr\,ds\\
&\le C + C'\int\limits_{-\varepsilon}^{u}\int\limits_{r}^{u}|N^x_\varepsilon(r)|\,|\rho_\varepsilon(r)|\,ds\,dr\\
&\le C + C''(u)\int\limits_{-\varepsilon}^{u}|N^x_\varepsilon(r)|\,|\rho_\varepsilon(r)|\,dr\,.
\end{aligned}$$

Hence by Gronwall's lemma $|N^x_\varepsilon(u)| \le C\exp[C''\int\limits_{-\varepsilon}^{u}|\rho_\varepsilon(r)|dr]$ which we use to conclude that

$$|\langle D_\varepsilon,\psi\rangle| \;\le\; C'''\sup_{|u|\le\varepsilon}|f''(x_\varepsilon(u))|^2\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(u)|\,du\underbrace{\int\limits_{-\varepsilon}^{\varepsilon}\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(r)|\,dr\,ds}_{\to 0}\;.$$

Collecting things together we have proved that $I_\varepsilon\rho_\varepsilon \to af'(x_0)[(1/8)f''(x_0)+1]\,\delta$, hence using (111) and (112)

$$N^x \;\approx\; \frac{1}{2}\,af'(x_0)(u_+ + H(u))$$

as stated.

We now turn to the equation for $\ddot N^v_\varepsilon$ in (106) and treat each of the five terms on the right hand side separately. We begin with the *first* one:

$$\begin{aligned}
&\langle (N^x_\varepsilon(u)f'(x_\varepsilon(u))\rho_\varepsilon(u))\dot{}\,,\psi\rangle\\
&= \underbrace{-\int\limits_{-\varepsilon}^{\varepsilon}f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}[af'(x_\varepsilon(r))+\frac{1}{2}N^x_\varepsilon(r)f''(x_\varepsilon(r))]\,\rho_\varepsilon(r)\,dr\,ds\,du}_{|\quad|\le C\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(u)|\,du\int\limits_{-\varepsilon}^{\varepsilon}\int\limits_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(r)|\,dr\,ds\to 0}
\end{aligned}$$

$$-\underbrace{\frac{1}{2}\,a\int\limits_{-\varepsilon}^{\varepsilon} f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot\rho_\varepsilon(r)\,r\,dr\,ds\,du}_{|\quad|\le C\,\varepsilon[\int_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(u)|\,du]^2\to 0}$$

$$-\underbrace{\frac{1}{2}\,a\int\limits_{-\varepsilon}^{\varepsilon} f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot\rho_\varepsilon(r)\,dr\,ds\,du}_{A_\varepsilon}\,,$$

and $A_\varepsilon \to -(1/4)af'(x_0)^2\dot\psi(0)$ by inserting and subtracting again the obvious terms. Hence the first term converges to

$$\frac{1}{2}\,af'(x_0)^2\dot\delta\,. \tag{113}$$

We proceed by showing that the *third term* on the right hand side of the first equation of (106) tends to

$$af(x_0)\ddot\delta\,. \tag{114}$$

Indeed,

$$\begin{aligned}\langle(a(1+u)f(x_\varepsilon(u))\rho_\varepsilon(u))^{\cdot\cdot},\psi\rangle &= a\underbrace{\int\limits_{-\varepsilon}^{\varepsilon} f(x_\varepsilon(u))\rho_\varepsilon(u)\ddot\psi(u)\,du}_{\to f(x_0)\ddot\psi(0)}\\ &\quad + a\int\limits_{-\varepsilon}^{\varepsilon} f(x_\varepsilon(u))\rho_\varepsilon(u)\ddot\psi(u)\,u\,du\end{aligned}$$

and the absolute value of the last expression may be estimated by $a\sup_{|u|\le\varepsilon}|uf(x_\varepsilon(u))\ddot\psi(u)|\int_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon(u)|du$, which vanishes in the limit.
Next for the *fourth term* we have

$$\begin{aligned}&\langle -a(1+u)f''(x_\varepsilon(u))(H*\ddot x_\varepsilon)^2(u)\rho_\varepsilon(u),\psi(u)\rangle\\ &= \underbrace{-\frac{1}{2}\,a\int\limits_{-\varepsilon}^{\varepsilon} f''(x_\varepsilon(u))\psi(u)\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} f'(x_\varepsilon(s))\rho_\varepsilon(s)\int\limits_{-\varepsilon}^{s} f'(x_\varepsilon(r))\rho_\varepsilon(r)\,dr\,ds\,du}_{A_\varepsilon}\\ &\quad\underbrace{-\frac{1}{2}\,a\int\limits_{-\varepsilon}^{\varepsilon} uf''(x_\varepsilon(u))\psi(u)\rho_\varepsilon(u)\int\limits_{-\varepsilon}^{u} f'(x_\varepsilon(s))\rho_\varepsilon(s)\int\limits_{-\varepsilon}^{s} f'(x_\varepsilon(r))\rho_\varepsilon(r)\,dr\,ds\,du}_{|\quad|\le C\sup_{|u|\le\varepsilon}|u|[\int_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon|]^3\to 0}\,,\end{aligned}$$

and $A_\varepsilon \to -(1/12)af'(x_0)^2f''(x_0)\psi(0)$ again by inserting the obvious terms and integrating by parts several times. Hence the whole forth term converges to

$$-\frac{1}{12}\,af'(x_0)^2f''(x_0)\delta \tag{115}$$

So we are left with the two most complicated terms namely the second and the fifth one. For the *second one* we get by integration by parts

$$\langle -N^x_\varepsilon(u) f'(x_\varepsilon(u))\dot\rho_\varepsilon(u), \psi(u)\rangle = \underbrace{\int_{-\varepsilon}^{\varepsilon} N^x_\varepsilon(u) f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u)\, du}_{A_\varepsilon}$$

$$+ \underbrace{\int_{-\varepsilon}^{\varepsilon} N^x_\varepsilon(u) f''(x_\varepsilon(u))\dot x_\varepsilon(u)\psi(u)\rho_\varepsilon(u)\, du}_{B_\varepsilon} + \underbrace{\int_{-\varepsilon}^{\varepsilon} \dot N^x_\varepsilon(u) f'(x_\varepsilon(u))\psi(u)\rho_\varepsilon(u)\, du}_{C_\varepsilon}\,.$$

Inserting for $N^x_\varepsilon$ we find

$$A_\varepsilon = \int_{-\varepsilon}^{\varepsilon} f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u) \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} [a f'(x_\varepsilon(r)) + \frac{1}{2} N^x_\varepsilon(r) f''(x_\varepsilon(r))]\rho_\varepsilon(r)\, dr\, ds\, du$$
$$+\frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u) \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} f'(x_\varepsilon(r))(1+r)\dot\rho_\varepsilon(r)\, dr\, ds\, du\,,$$

hence (again denoting constants by $C$, $C'$) by the boundedness of $N^x_\varepsilon$

$$|A_\varepsilon| \le \underbrace{C \int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(u)| du \int_{-\varepsilon}^{\varepsilon}\int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(r)|\, dr\, ds}_{\to 0} + \underbrace{C' \sup_{|u|\le\varepsilon} |u|\, [\int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(u)| du\,]^2}_{\to 0}$$
$$\underbrace{+\frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} f'(x_\varepsilon(u))\dot\psi(u)\rho_\varepsilon(u) \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot\rho_\varepsilon(r)\, dr\, ds\, du}_{\to \frac{1}{4} a f'(x_0)^2 \dot\psi(0)}\,,$$

where the last convergence once again is shown by inserting and subtracting the obvious terms. Summing up $A_\varepsilon \to -(1/4) a f'(x_0)^2 \psi(0)$. Similarly by inserting for $N^x_\varepsilon$ and using its boundedness we get for $B_\varepsilon =$

$$\underbrace{\frac{1}{2}\int_{-\varepsilon}^{\varepsilon} \psi(u) f''(x_\varepsilon(u))\rho_\varepsilon(u) \int_{-\varepsilon}^{u} f'(x_\varepsilon(s)\rho_\varepsilon(s)\, ds \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} [a f'(x_\varepsilon(r)) + \frac{1}{2} N^x_\varepsilon(r) f''(x_\varepsilon(r))\rho_\varepsilon(r)]\, dr\, ds\, du}_{|\quad |\le C\, [\int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(u)|\, du\,]^2 \int_{-\varepsilon}^{\varepsilon}\int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(r)|\, dr\, ds \to 0}$$

$$\underbrace{\frac{1}{4}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f''(x_\varepsilon(u))\rho_\varepsilon(u) \int_{-\varepsilon}^{u} f'(x_\varepsilon(s))\rho_\varepsilon(s)\, ds \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot\rho_\varepsilon(r)\, r\, dr\, ds\, du}_{|\quad |\le C' \sup_{|u|\le\varepsilon} |u| [\int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon(u)|\, du]^3 \to 0}$$

$$+\frac{1}{4}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f''(x_\varepsilon(u))\rho_\varepsilon(u) \int_{-\varepsilon}^{u} f'(x_\varepsilon(s))\rho_\varepsilon(s)\, ds \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\dot\rho_\varepsilon(r)\, dr\, ds\, du\,.$$

The last term of the above equation in the limit equals to

$$\frac{1}{4}\, a\, \psi(0) f''(x_0) f'(x_0)^2 \underbrace{\lim_{\varepsilon\to 0} [\int_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(u) \int_{-\varepsilon}^{u} \rho_\varepsilon(s) \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{r} \dot\rho_\varepsilon(r)\, dr\, ds\, du]}_{\to 1/3}\,,$$

hence $B_\varepsilon \to (1/12) a f''(x_0) f'(x_0)^2 \psi(0)$. The most troublesome term is $C_\varepsilon$. Again by inserting for $N^x_\varepsilon$ we split it up into three terms according to

$$
\begin{aligned}
C_\varepsilon = {} & \underbrace{\frac{1}{2}\,a\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}f'(x_\varepsilon(s))\dot\rho_\varepsilon(s)\,ds\,du}_{C1_\varepsilon} \\
& \underbrace{+\frac{1}{2}\,a\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}f'(x_\varepsilon(s))\dot\rho_\varepsilon(s)\,s\,ds\,du}_{C2_\varepsilon} \\
& \underbrace{+\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}[af'(x_\varepsilon(s))+\frac{1}{2}N^x_\varepsilon(s)f''(x_\varepsilon(s))]\rho_\varepsilon(s)\,ds\,du}_{C3_\varepsilon}\;.
\end{aligned}
$$

$C2_\varepsilon$ is diverging whereas $C1_\varepsilon$ converges to a limit depending on the the concrete shape of $\rho_\varepsilon$ and we leave these terms for later treatment. Now we calculate the limit of $C3_\varepsilon$.

$$
\begin{aligned}
C3_\varepsilon = {} & \underbrace{a\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}f'(x_\varepsilon(s))\rho_\varepsilon(s)\,ds\,du}_{\to\frac{1}{2}af'(x_0)^2\psi(0)} \\
& \underbrace{+\frac{1}{2}\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}N^x_\varepsilon(s)f''(X_\varepsilon(s))\rho_\varepsilon(s)\,ds\,du}_{(*)}
\end{aligned}
$$

The limit of $(*)$ may be calculated by inserting one more time the implicit solution for $N^x_\varepsilon$. More precisely, $(*) =$

$$
\begin{aligned}
& \underbrace{\frac{1}{2}\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}f''(x_\varepsilon(s))\rho_\varepsilon(s)\int_{-\varepsilon}^{s}\int_{-\varepsilon}^{r}[af'(x_\varepsilon(t))+\frac{1}{2}N^x_\varepsilon(t)f''(x_\varepsilon(t))]\rho_\varepsilon(t)\,dt\,dr\,ds\,du}_{|\quad|\le C\,[\int_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon|]^2\int_{-\varepsilon}^{\varepsilon}\int_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon|\to 0} \\
& \underbrace{+\frac{1}{4}\,a\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}f''(x_\varepsilon(s))\rho_\varepsilon(s)\int_{-\varepsilon}^{s}\int_{-\varepsilon}^{r}f'(x_\varepsilon(t))\dot\rho_\varepsilon(t)\,t\,dt\,dr\,ds\,du}_{|\quad|\le C'\sup_{|t|\le\varepsilon}|t|\,[\int_{-\varepsilon}^{\varepsilon}|\rho_\varepsilon|]^3\to 0} \\
& \underbrace{+\frac{1}{4}\,a\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))\rho_\varepsilon(u)\int_{-\varepsilon}^{u}f''(x_\varepsilon(s))\rho_\varepsilon(s)\int_{-\varepsilon}^{s}\int_{-\varepsilon}^{r}f'(x_\varepsilon(t))\dot\rho_\varepsilon(t)\,dt\,dr\,ds\,du}_{\to\frac{1}{4}af'(x_0)^2f''(x_0)\psi(0)\,\lim\limits_{\varepsilon\to 0}\underbrace{[\int_{-\varepsilon}^{\varepsilon}\rho_\varepsilon(u)\int_{-\varepsilon}^{u}\rho_\varepsilon(s)\int_{-\varepsilon}^{s}\rho_\varepsilon(r)\,dr\,ds\,du]}_{1/6}}\,,
\end{aligned}
$$

hence $C3_\varepsilon \to af'(x_0)^2[(1/2)+(1/24)f''(x_0)]\psi(0)$.

On the other hand the *fifth term* in the first equation in (106) may be split up according to

$$
\langle -a(1+u)f'(x_\varepsilon(u))\ddot x_\varepsilon(u)\rho_\varepsilon(u),\psi(u)\rangle = \underbrace{-\frac{1}{2}\,a\int_{-\varepsilon}^{\varepsilon}\psi(u)f'(x_\varepsilon(u))^2\rho_\varepsilon(u)^2\,du}_{D_\varepsilon}
$$

$$\underbrace{-\frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f'(x_\varepsilon(u))^2 \rho_\varepsilon(u)^2\, u\, du}_{E_\varepsilon}\ .$$

We now calculate the limit of $D_\varepsilon + C1_\varepsilon$:

$$\begin{aligned}
D_\varepsilon + C1_\varepsilon &= \frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f'(x_\varepsilon(u))\rho_\varepsilon(u)\, \underbrace{[\int_{-\varepsilon}^{u} f'(x_\varepsilon(s))\dot\rho_\varepsilon(s)\, ds - f'(x_\varepsilon(u))\rho_\varepsilon(u)]}_{f'(x_\varepsilon(u))\rho_\varepsilon(u) - \int_{-\varepsilon}^{u} f''(x_\varepsilon(s))\dot x_\varepsilon(s)\rho_\varepsilon(s)\, ds}\, du \\
&= -\frac{1}{4}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f'(x_\varepsilon(u))\rho_\varepsilon(u) \int_{-\varepsilon}^{u} f''(x_\varepsilon(s))\rho_\varepsilon(s) \int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\rho_\varepsilon(r)\, dr\, ds\, du \\
&\to -\frac{1}{4}\, a f'(x_0)^2 f''(x_0)\psi(0)\, \underbrace{\lim_{\varepsilon\to 0} \int_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(u) \int_{-\varepsilon}^{u} \rho_\varepsilon(s) \int_{-\varepsilon}^{s} \rho_\varepsilon(r)\, dr\, ds\, du}_{\to 1/6}
\end{aligned}$$

Hence $D_\varepsilon + C1_\varepsilon \to -(1/24) a f'(x_0)^2 f''(x_0)\psi(0)$.
Finally,

$$\begin{aligned}
C2_\varepsilon + E_\varepsilon &= \frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f'(x_\varepsilon(u))\rho_\varepsilon(u)\ \underbrace{[\int_{-\varepsilon}^{u} f'(x_\varepsilon(s))\dot\rho_\varepsilon(s)\, s\, ds - f'(x_\varepsilon(u))\rho_\varepsilon(u)\, u]}_{f'(x_\varepsilon(u))\rho_\varepsilon(u) u - \int_{-\varepsilon}^{u} [f'(x_\varepsilon(s)) + f''(x_\varepsilon(s))\dot x_\varepsilon(s) s]\rho_\varepsilon(s)\, ds}\, du \\
&= \underbrace{-\frac{1}{4}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f'(x_\varepsilon(u))\rho_\varepsilon(u) \int_{-\varepsilon}^{u} f''(x_\varepsilon(s))\rho_\varepsilon(s)\, s \int_{-\varepsilon}^{s} f'(x_\varepsilon(r))\rho_\varepsilon(r)\, dr\, ds\, du}_{|\quad | \leq C \sup_{|s|\leq\varepsilon} |s|\, [\int_{-\varepsilon}^{\varepsilon} |\rho_\varepsilon|\, ]^3 \to 0} \\
&\quad - \underbrace{-\frac{1}{2}\, a \int_{-\varepsilon}^{\varepsilon} \psi(u) f'(x_\varepsilon(u))\rho_\varepsilon(u) \int_{-\varepsilon}^{u} f'(x_\varepsilon(s))\rho_\varepsilon(s)\, ds\, du}_{\to -\frac{1}{4} a f(x_0)^2 \psi(0)}\,,
\end{aligned}$$

hence $C2_\varepsilon + E_\varepsilon \to -(1/4) a f'(x_0)^2 \psi(0)$ and we have proved that the sum of the second and the fifth term on the right of the first equation of (106) converges to

$$a f'(x_0)^2 [\frac{1}{4} + \frac{1}{12} f''(x_0)]\delta - \frac{1}{4}\, a f'(x_0)\dot\delta\,. \tag{116}$$

Combining the results of (113), (114), (115) and (116) we have proved

$$\ddot N^x_\varepsilon \to a f(x_0)\ddot\delta + \frac{1}{4} a f'(x_0)^2(\dot\delta + \delta)\,,$$

thereby establishing the claim. □

**5.31** Hence, viewed distributionally, the Jacobi field suffers a kink, a jump and a $\delta$-like pulse in the $v$-direction as well as a kink and jump in the $x$-direction overlapping the linear, flat space behavior. These effects can be understood heuristically from the corresponding behavior of the geodesics, given by equation (100). The constant factor $a$ which gives the "scale" of all the nonlinear effects, arises from the "time advance" of

the "nearby" geodesics, represented by the initial velocity of the Jacobi field in the $u$-direction (cf. initial conditions (108) ).
Note, however, that this "time advance" is not the only effect generically generating deviations from the flat space behavior, but rather arises as an artifact of our initial conditions. One can easily show that different initial conditions on the deviation vector field, even without "time advance", produce kinks and jumps as well. For example, let

$$N^a_\varepsilon(-1) \;=\; (0,0,0,0) \;\text{ and }\; \dot{N}^a_\varepsilon(-1) \;=\; (0,a,b,0)\,,$$

then a similar (but now even simpler) calculation leads to the following distributional shadows of the generalized Jacobi field:

$$\begin{aligned}
&N^u_\varepsilon = 0\ ,\\
&N^v_\varepsilon \approx a(1+u) + \frac{1}{4}\,b\,f'(x_0)\,f''(x_0)\,u_+ + b\,f'(x_0)\,H(u)\ ,\\
&N^x_\varepsilon \approx b(1+u) + \frac{1}{2}\,b\,f''(x_0)\,u_+\ ,\\
&N^y_\varepsilon = 0\ .
\end{aligned} \tag{117}$$

The kink of the $x$-component of the deviation field now arises from the fact that a "nearby" geodesic passes the shock at an $x$-value of $x_0{+}b$, hence according to equation (100) suffers a kink of "strength" $(1/2)\,f'(x_0{+}b)$. Taylor expansion yields $f'(x_0+b) \approx f'(x_0) + f''(x_0)\,b$, such that the kink difference of "nearby" geodesics is given by $(1/2)\,f''(x_0)\,b$, which is exactly the factor given in the third equation of (117). The kink and jump in $v$-direction can be explained by similar heuristic arguments.

### 5.D. The Penrose Transformation

This section is devoted to a detailed study of the interrelations between the distributional form of the pp-wave metric (90) as used above (arising from the Brinkmann form (82)) and its continuous form (associated with the Rosen form (87)) to be introduced in a moment.

**5.32** While the metric (90) very clearly demonstrates the nature of the impulsive wave, i.e., that the spacetime is flat everywhere except for the null hyperplane $u = 0$ where the $\delta$-like shock is located, it has the obvious disadvantage of involving distributional coefficients. However, using the technical apparatus of Colombeau's algebras of generalized functions we were able to describe the spacetime geometry entirely in the distributional picture in sections 5.B and 5.C above.
On the other hand, impulsive pp-waves are frequently described by a different spacetime metric which is actually continuous (see [152], [161] and, for the general case, [6]). It is derived from the Rosen form (87) and in the special case of an impulsive plane wave of constant linear polarization (cf. 5.4) with $h = 1/2$, i.e.,

$$ds^2 \;=\; \frac{1}{2}\,(x^2 - y^2)\,\delta(u)du^2 - dudv + dx^2 + dy^2\,, \tag{118}$$

takes the form [152],

$$ds^2 \;=\; (1+u_+)^2 dX^2 + (1-u_+)^2 dY^2 - dudV\,. \tag{119}$$

where again $u_+$ denotes the kink function. This form of the metric has the advantage that only the curvature tensor involves distributions while the metric can be treated "classically," i.e., it is gt-regular (cf. 2.19). Moreover, if one constructs impulsive pp-waves according to Penrose's "scissors and paste" approach (see 5.7) a theorem of Clarke and Dray [33] ensures the existence of a $\mathcal{C}^1$-atlas in which the metric components are continuous.

**5.33** Clearly a transformation relating the metrics (118) and (119) cannot even be continuous, hence—strictly speaking—it changes the topological structure of the manifold. In the special case envisaged above this discontinuous change of variables was given by R. Penrose in [150] (for more general cases again see [6] and [161])

$$
\begin{aligned}
x &= (1+u_+)X\\
y &= +(1-u_+)Y\\
v &= V + \frac{1}{2}X^2(u_+ + H(u)) + \frac{1}{2}Y^2(u_+ - H(u))\\
u &= u\,.
\end{aligned}
\tag{120}
$$

However, the two mathematically distinct spacetimes are equivalent from a physical point of view, i.e., the geodesics and the particle motion agree on a heuristic level (see also [176]). To demonstrate this in some more detail observe that the metric (119) has the advantage that simple particle motion can bee seen directly. Indeed free particles at fixed values of $X, Y$ and $Z = V - T$ after the shock start to move such that their relative $X$- and $Y$-distance is given by the functions $1 + u_+$ and $1 - u_+$, respectively. Applying formally the transformation (120) one finds total agreement with the distributional Jacobi field derived in theorem 5.30. (Note that it is the coordinate transformation which introduces the motion in $v$-direction.)
Moreover, we can solve the geodesic equations for the metric (119) either by using the method of sec. 5.B or (since these equations only involve Heaviside and kink functions) by solving them separately for $u < 0$ and $u > 0$ and joining them in a $\mathcal{C}^1$-manner. Either way leads to the distributional solutions ($u < 1$, and using analogous initial values as in 5.18)

$$
\begin{aligned}
X(u) &= x_0 + \dot{x}_0(2+u_-) - \frac{\dot{x}_0}{1+u_+}\\
Y(u) &= y_0 + \dot{y}_0 u_- + \frac{\dot{y}_0}{1-u_+}\\
V(u) &= v_0 + \dot{v}_0(1+u) + \frac{\dot{y}_0^2 u_+^2}{1-u} - \frac{\dot{x}_0^2 u_+^2}{1+u}\,,
\end{aligned}
\tag{121}
$$

where $u_- := H(-u)u$. If we now *formally* transform equations (121) according to (120) we again obtain the distributional geodesics of theorem 5.22.

**5.34** By the above we may conclude that *physically* the two approaches to impulsive plane waves, hence the two spacetime structures are *equivalent.* However, the transformation besides changing the manifold structure once more involves products of distributions ill-defined in the linear theory. In the following we are going give a precise meaning to the term "physically equivalent" by interpreting the discontinuous transformation as the distributional shadow of a generalized coordinate transformation, i.e., we are first going to turn the distributional spacetime metric (90) into a generalized one (in the sense of sec. 4.C above), again replacing the $\delta$-distribution by a generalized delta function. This amounts to replacing the distributional metric by a whole sequence of sandwich waves forming (a representative of) the generalized metric (cf. 5.17). Then we are going to apply a generalized change of coordinates modeling the distributional one. To achieve a physically sensible result it is essential not to just arbitrarily regularize the distributional transformation but to take into account the following geometrical consideration: The coordinate lines in the new variables defined in (120) are exactly given by the distributional geodesics of the metric derived in theorem 5.22 with vanishing initial speed in the $x, y$ and $v$-directions. Finally we calculate the distributional shadow of the transformed generalized metric to arrive precisely at the continuous form (which in the special case considered above is given by (118)).
Again all our results are independent of the regularization within the class of generalized delta functions, hence "natural." Physically, in the sandwich wave-picture both forms of the impulsive wave, i.e., the metrics (118) and (119), arise as (distributional) limits in different coordinate systems (see also the remarks in 5.41).

Before actually treating the problem at hand explicitly we first have to clarify what we mean by a generalized coordinate transformation $T$. Besides allowing for a representative $(t_\varepsilon)_\varepsilon$ which for fixed (at least small) $\varepsilon$ is a diffeomorphism we have to ensure moderateness of the componentwise inverse and well-definedness of the resp. compositions in $\mathcal{G}$, formally

**5.35 Definition.** *Let $\Omega$ be an open subset of $\mathbb{R}^n$. We call $T \in \mathcal{G}(\Omega, \mathbb{R}^n)$ a* generalized diffeomorphism *if there exists $\eta > 0$ such that*

*(i) There exists a representative $(t_\varepsilon)_\varepsilon$ such that $t_\varepsilon : \Omega \to t_\varepsilon(\Omega)$ is a diffeomorphism for all $\varepsilon \le \eta$ and there exists $\tilde{\Omega} \subseteq \mathbb{R}^n$ open, $\tilde{\Omega} \subseteq \bigcap_{\varepsilon\le\eta} t_\varepsilon(\Omega)$.*

*(ii) $(t_\varepsilon^{-1})_\varepsilon \in \mathcal{E}_M(\tilde{\Omega}, \mathbb{R}^n)$ and there exists $\Omega_1 \subseteq \mathbb{R}^n$ open, $\Omega_1 \subseteq \bigcap_{\varepsilon\le\eta} t_\varepsilon^{-1}(\tilde{\Omega})$.*

*(iii) Writing* $T^{-1} := \mathrm{cl}[(t_\varepsilon^{-1}|_{\tilde\Omega})_\varepsilon]$, $T \circ T^{-1}$ *as well as* $T^{-1} \circ T|_{\Omega_1}$ *are well-defined elements of* $\mathcal{G}(\tilde\Omega, \mathbb{R}^n)$ *resp.* $\mathcal{G}(\Omega_1, \mathbb{R}^n)$.

It is then clear that $T \circ T^{-1} = \mathrm{id}_{\tilde\Omega}$ resp. $T^{-1} \circ T|_{\Omega_1} = \mathrm{id}_{\Omega_1}$.

**5.36** We now begin by transforming our problem into the generalized functions setting. As already indicated above we shall deal with the following generalized spacetime metric (denoted also by $\hat G_{ab}$)

$$\hat{ds}^2 \;=\; f(x^j) D(u) du^2 - du dv + dx^2 + dy^2\,, \tag{122}$$

where $D$ again is a generalized delta function (definition 5.16). Hence a representative of the metric (in a somewhat sloppy notation) is given by

$$\hat{ds}_\varepsilon^2 \;=\; f(x^j)\rho_\varepsilon(u) du^2 - du dv + dx^2 + dy^2\,,$$

with $\rho_\varepsilon$ a strict delta net (cf. 5.15). We shall also write $\hat G_{ab} = \mathrm{cl}[(\hat g_{ab\,\varepsilon})_\varepsilon]$ (cf. the remarks on notation in sec. 4.C).
An existence and uniqueness result for the corresponding geodesic equations has been given in theorem 5.21 above. Since here we are only interested in geodesics with the special initial conditions

$$\begin{aligned} V(-1) &= v_0, & X^i(-1) &= x_0^i,\\ \dot V(-1) &= 0, & \dot X^i(-1) &= 0, \end{aligned}$$

(with $x_0^i$, $v_0$ again real numbers) we adopt the notation $X^i(x_0^j, u)$ respectively $V(v_0, x_0^j, u)$ for these unique generalized geodesics with vanishing initial speed. Recall that on the level of representatives they obey the following (implicit) set of equations (cf. (101))

$$\begin{aligned} x_\varepsilon^i(x_0^k, u) &= x_0^i + \frac{1}{2}\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} \partial_i f(x_\varepsilon^j(x_0^k, r))\rho_\varepsilon(r)\, dr\, ds \\ v_\varepsilon(v_0, x_0^k, u) &= v_0 + \int_{-\varepsilon}^{u} f(x_\varepsilon^j(x_0^k, s))\rho_\varepsilon(s)\, ds + \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s} \partial_i f(x_\varepsilon^j(x_0^k, r))\dot x_\varepsilon^i(x_0^k, r))\rho_\varepsilon(r)\, dr\, ds\,. \end{aligned} \tag{123}$$

Let us now consider the transformation $T = \mathrm{cl}[(t_\varepsilon)_\varepsilon] : (u, v, x^i) \mapsto (u, V, X^i)$ (with $X^i = (X, Y)$) given implicitly by (123), i.e., depending on the regularization parameter $\varepsilon$ according to

$$\begin{aligned} t_\varepsilon:\quad x^i &= x_\varepsilon^i(X^j, u)\\ v &= v_\varepsilon(V, X^j, u) \end{aligned} \tag{124}$$

which is precisely analogous (and by theorem 5.22 in the limit $\varepsilon \to 0$ and the special case of a plane wave actually reduces) to (120). We now have the following

**5.37 Theorem.** *(Generalized Penrose transformation) The generalized function* $T$ *defined above is a generalized coordinate transformation on a suitable open subset* $\Omega$ *of* $\mathbb{R}^4$ *containing the shock hyperplane at* $u = 0$. *Moreover, the new coordinates are constant along the geodesics given by (123).*

Note that this result in particular implies that for small $\varepsilon$ the geodesics (123) do not cross in the specified region (although they may do so for some finite value of $u$, due to the focusing property already mentioned in 5.4).

**Proof.** Since the latter property follows directly by construction we only have to verify (i)-(iii) of definition 5.35 above.
(i) To show that for fixed small $\varepsilon$ the map $t_\varepsilon$ is a classical coordinate transformation we employ a global univalence theorem by Gale and Nikaido ( [67], thm. 4), stating that any differentiable $F : \Omega \to \mathbb{R}^n$, where $\Omega$ is a closed rectangular region in $\mathbb{R}^n$ is univalent (injective) if all principal minors of its Jacobian $J(x)$ are positive. Since

$$\frac{\partial(\,u,\,x^1,\,x^2,\,v\,)}{\partial(u,X^1,X^2,V)} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ \frac{\partial x^1_\varepsilon}{\partial u} & \frac{\partial x^1_\varepsilon}{\partial X^1} & \frac{\partial x^1_\varepsilon}{\partial X^2} & 0 \\ \frac{\partial x^2_\varepsilon}{\partial u} & \frac{\partial x^2_\varepsilon}{\partial X^1} & \frac{\partial x^2_\varepsilon}{\partial X^2} & 0 \\ \frac{\partial v_\varepsilon}{\partial u} & \frac{\partial v_\varepsilon}{\partial X^1} & \frac{\partial v_\varepsilon}{\partial X^2} & 1 \end{vmatrix} \tag{125}$$

we have to find estimates for

$$\frac{\partial x^i_\varepsilon}{\partial X^j} = \delta^i_j + \frac{1}{2}\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}(\partial_m\partial_i f)(x^k_\varepsilon(X^l,r))\frac{\partial x^m_\varepsilon}{\partial X^j}(X^l,r)\rho_\varepsilon(r)\,dr\,ds\,. \tag{126}$$

If $X^l$ varies in a compact region $K$ of $\mathbb{R}^2$ and $-1 \le p \le u < \infty$ it follows from a straightforward modification of lemma 5.20 that for small $\varepsilon$ the terms $x^k_\varepsilon(X^l,p)$ remain bounded, independently of $X^l$, $p$ and $\varepsilon$. Let $g(u) := \sup\{\sum_{i=1}^{2} |\frac{\partial x^i_\varepsilon}{\partial X^j}(X^k,p)| \,:\, X^k \in K,\, -1 \le p \le u\}$. Then (126) gives ($C$, $C'$, again constants)

$$|g(u)| \le C + C'\int\limits_{-\varepsilon}^{u}|g(s)|\,ds\,,$$

so Gronwall's lemma implies the same boundedness property for $\frac{\partial x^i_\varepsilon}{\partial X^j}$. Using these estimates it follows from (126) that for small $\varepsilon$ and for small $u > 0$ (depending on $\|\partial_i\partial_j f\|_\infty$ in a compact region and on $C$ from 5.15 (c)) $\frac{\partial x^i_\varepsilon}{\partial X^j}$ will remain arbitrarily small (for $i \ne j$) or arbitrarily close to 1 (for $i = j$), respectively. Hence all principal minors of (125) are indeed positive in a suitable rectangular region (independent of $\varepsilon$) containing the shock hypersurface $u = 0$, which establishes our claim.
To see that there exists some open $\tilde\Omega \subseteq \bigcap_{\varepsilon\le\eta} t_\varepsilon(\Omega)$ first note that $t_\varepsilon$ for $u \le -\varepsilon$ is just the identity map. Also $t_\varepsilon$ is independent of $\varepsilon$ for $u \ge \varepsilon$. On the other hand for $-\varepsilon < u < \varepsilon$ (for $\varepsilon$ small) we have from the uniform boundedness properties of $x^i_\varepsilon$ and $\dot x^i_\varepsilon$ established in lemma 5.20

$$\begin{aligned} |x^i_\varepsilon(X^j,u) - X^i| &\le 2\varepsilon C_1 \\ |v_\varepsilon(X^j,V,u) - v| &\le 2\varepsilon C_2 + C_3||\rho||_1 \end{aligned}$$

where $C_i$ $(i = 1,2,3)$ are constants independent of $\varepsilon$. So if we take arbitrary two points in $\Omega$ separated by a finite distance their respective images under $t_\varepsilon$ again are separated by a finite distance independent of $\varepsilon$.
(ii) Denote by $\tilde\Omega$ a subset of $\bigcap_{\varepsilon\le\eta} t_\varepsilon(\Omega)$. Using the above argument we also see that there exists $\Omega_1 \subseteq \bigcap_{\varepsilon\le\eta} t_\varepsilon^{-1}(\tilde\Omega)$. It is a straightforward consequence of theorem 5.21 and the remarks following (126) that $(t_\varepsilon|_{\Omega_1})_\varepsilon$ and $(t_\varepsilon^{-1})_\varepsilon$ are elements of $\mathcal{E}_M(\Omega_1;\mathbb{R}^4)$ and $\mathcal{E}_M(\tilde\Omega;\mathbb{R}^4)$, respectively. Let us denote their classes in $\mathcal{G}(\Omega_1;\mathbb{R}^4)$ (resp. $\mathcal{G}(\Omega_1;\mathbb{R}^4)$) by $T$ and $T^{-1}$.
(iii) To show that the respective compositions are well-defined we use proposition 3.8. It follows immediately from the remarks following (126) that $T^{-1}$ satisfies the assertions of 60. Concerning $T$, suppose that $(u,v,x,y)$ varies in some $K \subset\subset \Omega$ and set $(u,V_\varepsilon,X_\varepsilon,Y_\varepsilon) = t_\varepsilon(u,v,x,y)$. Then

$$x^i = x^i_\varepsilon(X^j_\varepsilon,u) = X^i_\varepsilon + \frac{1}{2}\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k_\varepsilon,r))\rho_\varepsilon(r)\,dr\,ds\,.$$

Since the left hand side and the argument of $\partial_i f$ in this equation are bounded by assumption, $X^i_\varepsilon$ remains in some compact region $L$ in $\mathbb{R}^2$. We already know that on any such set $L$ (times some compact $u$-interval) $\dot x^i_\varepsilon$ is uniformly bounded. Inserting this into

$$V_\varepsilon = v_\varepsilon(V_\varepsilon,X^i_\varepsilon,u) - \int\limits_{-\varepsilon}^{u} f(x^j_\varepsilon(X^k_\varepsilon,s))\rho_\varepsilon(s)\,ds - \int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k_\varepsilon,r))\dot x^i_\varepsilon(X^k_\varepsilon,r))\rho_\varepsilon(r)\,dr\,ds$$

also establishes the desired boundedness property for $V_\varepsilon$. □

**5.38** If we now apply the generalized coordinate transformation $T$ to the metric (122) we find by a straightforward computation that in the new coordinates it is given by the class of

$$\hat{ds}^2_\varepsilon = -dudV + (2\sum_{i=1}^{2}(\dot{x}^i_\varepsilon \partial_j x^i_\varepsilon) - \partial_j v_\varepsilon)dudX^j + \sum_{i=1}^{2}(\partial_j x^i_\varepsilon dX^j)^2\,, \tag{127}$$

where $\dot{}$ and $\partial_i$ denote derivatives with respect to $u$ and $X^i$, respectively. Note that this is the *general form of a sandwich wave in Rosen coordinates.*
Next we are going to derive the associated distribution of the generalized metric (127), thereby recovering the continuous form of the impulsive pp-wave metric.

**5.39 Theorem.** *For the generalized metric (127) above we have the following association relations:*

$$\begin{aligned} \mathrm{cl}[(\hat{ds}^2_\varepsilon)_\varepsilon] \approx -dudV &+ (1+\frac{1}{2}\partial_{11}f(X^j)u_+)^2dX^2 + (1+\frac{1}{2}\partial_{22}f(X^j)u_+)^2dY^2 \\ &+ \frac{1}{2}\partial_{12}f(X^j)\triangle f(X^j)u_+^2 dXdY + 2u_+\partial_{12}f(X^j)dXdY \\ &+ \frac{1}{4}(\partial_{12}f(X^j))^2u_+^2(dX^2+dY^2)\,. \end{aligned} \tag{128}$$

**Proof.** To reduce the notational complexity we are going to suppress the explicit dependence of test functions on the variables $X^i$ and $V$, i.e., we will write $\varphi(u)$ for $\varphi \in \mathcal{D}(\mathbb{R}^4)$ and we will simply drop the integrations with respect to $X^i$ and $V$. This abuse of notation is admissible due to the uniform boundedness properties (in $X^i$ and $V$) established above and serves to reduce the number of integrals in the sequel by three. We start out with the coefficient

$$\begin{aligned} \hat{g}_{uX\,\varepsilon} = &\sum_{i=1}^{2}(\int_{-\varepsilon}^{u}\partial_i f(x^j_\varepsilon(X^k,s))\rho_\varepsilon(s)\,ds\;\partial_1(X^i + \frac{1}{2}\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k,r)\rho_\varepsilon(r)\,drds)) \\ &-\partial_1(V + \int_{-\varepsilon}^{u} f(x^j_\varepsilon(X^k,r))\rho_\varepsilon\,dr + \int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_l f(x^j_\varepsilon(X^k,r)\dot{x}^l_\varepsilon(X^j,r)\rho_\varepsilon(r)\,drds)\,. \end{aligned} \tag{129}$$

The first term of this expression can be written as

$$\sum_{i=1}^{2}(\delta^1_i \underbrace{\int_{-\varepsilon}^{u}\partial_i f(x^j_\varepsilon(X^k,s))\rho_\varepsilon(r)\,dr}_{\to \partial_i f(X^j)H} + \underbrace{\frac{1}{2}\int_{-\varepsilon}^{u}\partial_i f(x^j_\varepsilon(X^k,r))\rho_\varepsilon(r)dr\;\partial_1\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k,r)\rho_\varepsilon(r)\,drds)}_{A_\varepsilon}\,.$$

For any test function $\varphi$ we have

$$\langle 2A_\varepsilon,\varphi\rangle = \int_{-\varepsilon}^{\infty}\varphi(u)\int_{-\varepsilon}^{u}\partial_i f(x^j_\varepsilon(X^k,r))\rho_\varepsilon(r)dr\;\partial_1\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k,r)\rho_\varepsilon(r)\,drdsdu\,.$$

Splitting this integral into a sum of the form

$$\int_{-\varepsilon}^{\varepsilon}\dots\int_{-\varepsilon}^{u}\dots\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\dots + \int_{\varepsilon}^{\infty}\dots\int_{-\varepsilon}^{u}\dots\int_{-\varepsilon}^{\varepsilon}\int_{-\varepsilon}^{s}\dots + \int_{\varepsilon}^{\infty}\dots\int_{-\varepsilon}^{u}\dots\int_{\varepsilon}^{u}\int_{-\varepsilon}^{s}\dots \tag{130}$$

the boundedness arguments following (126) imply that the first two summands converge to 0. Also, by (102) and by (123) we obtain

$$\lim_{\varepsilon\to 0}\sup_{|r|<\varepsilon}|\partial_i(f(x^j_\varepsilon(X^k,r))) - \partial_i f(X^k)| = 0 \tag{131}$$

uniformly for $X^k$ in compact sets. For later use we note that by the same reasoning also

$$\lim_{\varepsilon\to 0}\sup_{|r|<\varepsilon}|\partial_i(\partial_j f(x^j_\varepsilon(X^k,r)))-\partial_{ij}f(X^k)|=0 \tag{132}$$

uniformly for $X^k$ in compact sets. Hence by a direct estimation the limit of the remaining term is

$$\langle \partial_i f(X^j)\partial_1\partial_i f(X^j)u_+,\varphi\rangle\,,$$

so the distributional limit of the first term in (129) is $\partial_1 f(X^j)H+\frac{1}{2}\sum_{i=1}^{2}\partial_i f(X^j)\partial_i\partial_1 f(X^j)u_+$. Turning now to the second term in (129) we obtain from (131) and condition (b) of 5.15

$$\partial_1\int_{-\varepsilon}^{u} f(x^j_\varepsilon(X^k,r))\rho_\varepsilon(r)\,dr \to \partial_1 f(X^j)H\,.$$

Inserting (123) it remains to calculate the distributional limit of

$$\begin{aligned}&\frac{1}{2}\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_1(\partial_i f(x^j_\varepsilon(X^k,r))\rho_\varepsilon(r)\int_{-\varepsilon}^{r}\partial_i f(x^j_\varepsilon(X^k,q)\rho_\varepsilon(q)\,dq\,dr\,ds\\ &+\frac{1}{2}\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k,r)\rho_\varepsilon(r)\int_{-\varepsilon}^{r}\partial_1(\partial_i f(x^j_\varepsilon(X^k,q))\rho_\varepsilon(q)\,dq\,dr\,ds\,.\end{aligned} \tag{133}$$

A splitting scheme as in (130) for both these terms, (131) and the fact that

$$\int_{\varepsilon}^{u}\int_{-\varepsilon}^{s}\rho_\varepsilon(r)\int_{-\varepsilon}^{r}\rho_\varepsilon(q)dq\,dr\,ds\to\frac{1}{2}u_+$$

imply that each summand in (133) converges distributionally to $\frac{1}{4}\partial_i f(X^j)\partial_{1i}f(X^j)u_+$.
An analogous argument holds for $\hat g_{uY\,\varepsilon}$. Summing up, we obtain

$$\hat G_{uX}\approx 0 \qquad\text{and}\qquad \hat G_{uY}\approx 0\,. \tag{134}$$

If we write

$$\hat g_{XX\,\varepsilon}=\sum_{i=1}^{2}(\partial_1(X^i+\frac{1}{2}\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_i f(x^j_\varepsilon(X^k,r))\rho_\varepsilon(r)\,dr\,ds))^2=:\sum_{i=1}^{2}R_{i\,\varepsilon}$$

then

$$R_{i\,\varepsilon}=\delta_i^1+\underbrace{\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_1(\partial_1 f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds}_{\to\partial_{11}f(X^j)u_+}+\underbrace{\frac{1}{4}(\int_{-\varepsilon}^{u}\int_{-\varepsilon}^{s}\partial_1(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds)^2}_{B_\varepsilon}$$

and by a similar argument as above it follows that the distributional limit of the derivative $\partial_u B_\varepsilon$ is $\frac{1}{2}(\partial_1\partial_i f(X^j))^2u_+$. Since taking primitives (i.e., convoluting with $H$) is separately continuous on the convolution algebra of distributions supported in an acute cone ( [87], 4 §9, prop. 7) and since taking tensor products of distributions is separately continuous as well ( [87], 4§8, prop. 7) we get

$$R_{i\,\varepsilon}\to\delta_i^1+\partial_{11}f(X^j)u_++\frac{1}{2}(\partial_1\partial_i f(X^j))^2\frac{u_+^2}{2}\,.$$

Thus

$$\hat{G}_{XX} \approx (1 + \frac{1}{2}\partial_{11} f(X^j)u_+)^2 + \frac{1}{4}(\partial_{12} f(X^j))^2 u_+^2 \tag{135}$$

$$\hat{G}_{YY} \approx (1 + \frac{1}{2}\partial_{22} f(X^j)u_+)^2 + \frac{1}{4}(\partial_{12} f(X^j))^2 u_+^2 \,. \tag{136}$$

Finally, we turn to $\hat{g}_{XY\,\varepsilon} = 2\sum\limits_{i=1}^{2}(\partial_1 x^i_\varepsilon(X^j))(\partial_2 x^i_\varepsilon(X^j)) =: 2\sum\limits_{i=1}^{2} S_{i\,\varepsilon}$. Inserting from (123) we have

$$S_{i\,\varepsilon} = \underbrace{\frac{1}{2}\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_2(\partial_1 f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds}_{\to \partial_{12} f(X^j)u_+} + \underbrace{\frac{1}{2}\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_1(\partial_2 f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds}_{\to \partial_{12} f(X^j)u_+}$$
$$+\frac{1}{4}\underbrace{\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_2(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_1(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds}_{D_\varepsilon}\,.$$

We claim that $D_\varepsilon \to \partial_{1i} f(X^j)\partial_{2i} f(X^j)u_+^2$. To establish this, let $\varphi \in \mathcal{D}$ and consider

$$\langle D,\varphi\rangle - \partial_{1i} f(X^j)\partial_{2i} f(X^j)\int\limits_0^\infty u^2\varphi(u)\,du = \int\limits_{-\varepsilon}^{\infty}\left(\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_1(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds\right)$$
$$\times\left(\int\limits_{-\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_2(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds\right)\varphi(u)\,du - \int\limits_0^\infty \partial_{1i} f(X^j)\partial_{2i} f(X^j)u^2\varphi(u)\,du\,.$$

In order to show that this goes to zero, by a splitting similar to (130) and by the boundedness properties already established it suffices to prove that

$$\int\limits_{\varepsilon}^{\infty}\left(\int\limits_{\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_1(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds\right)\left(\int\limits_{\varepsilon}^{u}\int\limits_{-\varepsilon}^{s}\partial_2(\partial_i f(x^j_\varepsilon(X^k,r)))\rho_\varepsilon(r)\,dr\,ds\right)\varphi(u)\,du$$
$$-\int\limits_{\varepsilon}^{\infty}\partial_{1i} f(X^j)\partial_{2i} f(X^j)u^2\varphi(u)\,du \to 0\,,$$

which in turn is a consequence of (132) and (5.15). Therefore,

$$\hat{G}_{XY} \approx \frac{1}{2}u_+^2\left(\partial_{11} f(X^j)\partial_{21} f(X^j) + \partial_{12} f(X^j)\partial_{22} f(X^j)\right) + 2u_+\partial_{12} f(X^j) \tag{137}$$

□

**5.40** Let us sum up the main result of this section in the following diagram:

$$\begin{tikzcd}
ds^2 = f(x^i)\delta(u)du^2 - dudv + \sum(dx^i)^2 \arrow[r, "\text{reg.}"] \arrow[d, "??"] & \hat{ds}_\varepsilon^2 = f(x^i)\rho_\varepsilon(u)du^2 - dudv + \sum(dx^i)^2 \arrow[d, "T_\varepsilon"] \\
\text{r.h.s. of (128)} = \lim\limits_{\varepsilon\to 0}\hat{ds}_\varepsilon^2 & \begin{array}{l}\hat{ds}_\varepsilon^2 = -dudV + (2\sum\limits_{i=1}^{2}(\dot{x}^i_\varepsilon\partial_j x^i_\varepsilon) - \\ \partial_j v_\varepsilon)dudX^j + \sum\limits_{i=1}^{2}(\partial_j x^i_\varepsilon dX^j)^2\end{array} \arrow[l, "\mathcal{D}'\text{-limit}"']
\end{tikzcd}$$

The left vertical arrow in the above diagram represents the forming of the formal pullback of the distributional metric (90) under the discontinuous coordinate change (cf. (100))

$$\begin{aligned} x^i &= X^i + \frac{1}{2}\partial_i f(X^j) u_+ \\ v &= V + f(X^j) H(u) + \frac{1}{4}\partial_i f(X^j) \partial^i f(X^j) u_+ \,. \end{aligned} \tag{138}$$

Note that (138) is the generalization of the Penrose transformation (120) to the case of an arbitrary wave profile $f$ (cf. (16) in [6]). Although undefined within Schwartz distribution theory, this transformation can be interpreted consistently as the composition of a regularization procedure, a smooth transformation of the regularized metric and a distributional limit. It is precisely in this sense that the pullback is given by the right hand side of (128) or for short

$$ds^2 = (\delta_{ij} + \frac{1}{2} u_+ \partial_{ij} f)\,(\delta^j_k + \frac{1}{2} u_+ \partial^j_k f)\, dX^i dX^k - du dV \,. \tag{139}$$

For the special case of an impulsive plane wave with constant linear polarization and $f(x,y) = (1/2)(x^2 - y^2)$ this exactly reduces to (119) and in the general vacuum case it is equivalent to (2) in [161]. Also, we note that (139) appeared first as (17) in [6]. The regularization procedure employed above justifies *a posteriori* the derivation of (139) by using formal multiplication rules for distributions.

**5.41** From the viewpoint of the generalized curvature framework developed in section 4.C we may also phrase the contents of the above diagram with the following words: The generalized metric (122) is subjected to the generalized coordinate transformation $T$. In either coordinates the distributional shadow is computed giving the distributional resp. the continuous form of the pp-wave metric. Note that although the action of a smooth diffeomorphism is compatible with the notion of association, generalized coordinate transformations clearly are not.

Physically speaking the two forms of the impulsive metric arise as the (distributional) limits of the sandwich wave in different coordinate systems. It is a well known fact [70] that given a one-parameter family of spacetimes one may obtain different limits by introducing a suitable (parameter dependent) change of coordinates. This effect is mainly caused by a different identification of points in the manifold. More precisely, the transformation in the limit sends finite points to infinity, hence is unbounded. However, in our case the limit of the transformation, although discontinuous, remains bounded and the mentioned ambiguities do not arise. This is in fact not surprising since our change of coordinates is adjusted to the geometry of the spacetime which (in the sense of geodesic completeness) is nonsingular, allowing only for a *finite* jump of the geodesics.

*Finally we may say that an impulsive pp-wave is very sensibly modeled by the generalized spacetime metric (122); in different coordinate systems related by a generalized coordinate transformations it reveals different distributional pictures.*

### 5.E. Conclusions and Outlook

Summing up in the previous sections we have established a mathematically rigorous and physically sensible distributional description of impulsive pp-waves using Colombeau's generalized functions. We have explored the geometry of the spacetime via the geodesic and geodesic deviation equation using the more intuitive distributional form of the metric. This choice is also motivated by the fact that impulsive pp-wave metrics physically (i.e., in the impulsive as well as the ultrarelativistic limit) arise that way. Viewed distributionally, the geodesics are given by refracted, broken straight lines which is also reflected in the behavior of the respective Jacobi field. Moreover, the existence and uniqueness results for the geodesic equation in the generalized functions setting have provided us with the key to a mathematical description of the discontinuous change of coordinates from the distributional to the continuous form of the metric as the distributional shadow of a generalized coordinate transformation.
Note that we did not have to impose Einstein's equations throughout our analysis, hence our results apply to all pp-waves, i.e., even to the non-vacuum case.

Future work will be concerned with a similar analysis of *impulsive spherical waves.* In [152] Penrose used his "scissors and paste" approach (see 5.7) also to construct spherical impulsive waves in a Minkowski background but only gave the respective junction conditions explicitly. Later Nutku and Penrose [143] and Hogan [85,86] derived explicitly the continuous form of the metric (analogously to the Rosen form for impulsive pp-waves (139)). These solutions may be interpreted as the "snapping" of a cosmic string or as an expanding cosmic string behind a gravitational wave [143]. Already Penrose in [152] remarked that these spacetimes formally arise as the impulsive limit of type-N Robinson-Trautman solutions ( [111], chap. 24). However, in the standard coordinate system (cf. [68]) the latter are quadratic in the profile function, i.e.,

$$ds^2 \;=\; 2v^2 d\zeta d\bar\zeta - dudv - 2v^2\bar F(\bar\zeta,u)d\zeta du - 2v^2 F(\zeta,u)d\bar\zeta du \;+\; [2v^2F(\zeta,u)\bar F(\bar\zeta,u) + v(F_\zeta + \bar F_{\bar\zeta})]\, du^2\,, \quad (140)$$

with $F$ a holomorphic function of $\zeta$ and smoothly depending on $u$. It is now evident that replacing $F(\zeta,u)$ by $f(\zeta)\delta(u)$ in the metric—which exactly corresponds to the impulsive limit in the case of pp-waves—produces terms formally proportional to the square of the Dirac-$\delta$. However, recently Podolský and Griffiths [162] have given a formal discontinuous transformation relating the continuous form of the metric mentioned above to exactly the form (140) with $F(\zeta,u) = f(\zeta)\delta(u)$. This transformation has to be viewed as the precise analog to the respective transformation (138) in the pp-wave case relating the Brinkmann to the Rosen form and it is desirable to set up a rigorous description analogous to the one in the preceding section. However, the actual calculations in this case are considerably more complicated as might already be guessed from the form of the metric (140).
Also a generalization of the description of the preceding three sections 5.B-5.D to include a non-vanishing cosmological constant, i.e., a treatment of all the non-expanding impulsive waves as classified in [159], seems to be a very tempting program.

Another possible line of research leads into a somewhat different direction which is more mathematical in nature. Since the generalized change of coordinates is the first example of a "Colombeau diffeomorphism" a detailed study may lead to a general inverse function theorem in this framework. Note that of the (equivalent) main theorems of classical calculus, i.e., Frobenius theorem, inverse function theorem, implicit function theorem and existence and uniqueness of solutions to ODEs up to now only the ODE-theorem has been generalized to the Colombeau framework ( [83]).

The generalized curvature framework presented in section 4.C is a very promising tool for applications in general relativity. First of all the quest for more detailed consistency results with respect to the classical theory, in particular the framework of Geroch and Traschen, arises. Moreover, since it provides such an easy-to-use setting it seems to be very well suited to the study of classical singular spacetimes.
On the other hand there is still no (diffeomorphism invariant) full version of the theory of generalized tensor fields on manifolds, hence no full framework for generalized general relativity. However, research in that area based upon the preprint [184] of Vickers and Wilson has just been started.

## BIBLIOGRAPHY


[1] R. Abraham, J. E. Marsden, "Foundations of Mechanics," (Benjamin/Cummings, Reading, 2nd ed., 1978).

[2] R. A. Adams, "Sobolev Spaces," (Academic Press, New York, 1975).

[3] P. C. Aichelburg, R. U. Sexl, "On the gravitational field of a massless particle," J. Gen. Rel. Grav. **2**, 303-312 (1971).

[4] P. C. Aichelburg, H. Balasin, "A note on the symmetries of the gravitational field of a massless particle," Class. Quant. Grav. **11**, L121-L123 (1994).

[5] P. C. Aichelburg, H. Balasin, "Symmetries of pp-waves with distributional profile," Class. Quant. Grav. **13**, 723-729 (1996).

[6] P. C. Aichelburg, H. Balasin, "Generalized symmetries of impulsive gravitational waves," Class. Quant. Grav. **14**, A31-A41 (1997).

[7] P. C. Aichelburg, H. Balasin, "ADM and Bondi four-momenta for the ultrarelativistic Schwarzschild black hole," Class. Quant. Grav. **15**, 3841-3844 (1998).

[8] P. Antosik, J. Mikusinski, R. Sikorski, "Theory of Distributions - The sequential Approach," (Elsevier, Amsterdam, 1973).

[9] J. Aragona, H. A. Biagioni, "Intrinsic formulation of the Colombeau algebra of generalized functions," Analysis Mathematica, **17**, 75-132 (1991).

[10] H. Balasin, H. Nachbagauer, "On the distributional nature of the energy-momentum tensor of a black hole or What curves the Schwarzschild geometry ?" Class. Quant. Grav. **10**, 2271-2278 (1993).

[11] H. Balasin, H. Nachbagauer, "Distributional energy-momentum tensor of the Kerr-Newman space-time family," Class. Quant. Grav. **11**, 1453-1461 (1994).

[12] H. Balasin, H. Nachbagauer, "The ultrarelativistic Kerr-geometry and its energy-momentum tensor," Class. Quant. Grav. **12**, 707-713 (1995).

[13] H. Balasin, H. Nachbagauer, "Boosting the Kerr-geometry into an arbitrary direction," Class. Quant. Grav. **13**, 731-737 (1995).

[14] H. Balasin, "Colombeau's generalized functions on arbitrary manifolds," preprint gr-qc/9610017 [1] (1996).

[15] H. Balasin, "Geodesics for impulsive gravitational waves and the multiplication of distributions," Class. Quant. Grav. **14**, 455-462 (1997).

[16] H. Balasin, "Distributional energy-momentum tensor of the extended Kerr geometry," Class. Quant. Grav. **14**, 3353-3362 (1997).

[17] H. Balasin, "Distributional aspects of general relativity: the expample of the energy-momentum tensor of the extended Kerr-geometry," in *Nonlinear Theory of Generalized Functions,* Chapman & Hall/CRC Research Notes in Mathematics **401**, 275-290, eds. M. Grosser, G. Hörmann, M. Kunzinger, M. Oberguggenberger (Chapman & Hall/CRC, Boca Raton 1999).

[18] O. R. Baldwin, G. B. Jefferey, "The relativity theory of plane waves," Proc. Roy. Soc. London **A 111**, 95 (1926).

[19] J. Barros-Neto, "An Introduction to the Theory of Distributions," (Marcel Dekker, New Jork, 1973).

[20] F. Berger, J. F. Colombeau, "Numerical solutions of one-pressure models in multifluid flows," SIAM J. Numer. Anal. **32**, 1139–1154 (1995).

[21] F. Berger, J. F. Colombeau, M. Moussaoui, "Solutions mesures de Dirac de systemes de lois de conservation et applications numeriques," C. R. Acad. Sci. Paris Sér. I Math. **316**, 989–994 (1993).

[22] H. A. Biagioni, "A Nonlinear Theory of Generalized Functions," Lecture Notes in Mathematics **1421**, (Springer, Berlin 1990).

[23] H. A. Biagioni, M. Oberguggenberger, "Generalized solutions to the Korteweg-de Vries and the regularized long-wave equations," SIAM J. Math. Anal. **23**, 923–940 (1992).

[24] H. A. Biagioni, M. Oberguggenberger, "Generalized solutions to Burgers' equation," J. Differential Equations **97**, 263–287 (1992).

[25] J. Bičák, J. Podolský, "Gravitational waves in vacuum spacetimes with cosmological constant. I. Classification and geometrical properties of non-twisting type-N solutions," J. Math. Phys. to appear, gr-qc/9907048 [1] (1999).

[26] S. Bochner, "Vorlesungen über Fouriersche Integrale," (Leipzig, 1932).

[27] H. Bondi, "Plane gravitational waves in general relativity," Nature **179**, 1072-1073 (1957).

[28] H. Bondi, F. A. E. Pirani, I. Robinson, "Gravitational waves in general relativity III: Exact plane waves," Proc. R. Soc. Lond., Ser. A **251**, 519-533 (1959).


---

[1] gr-qc preprints are available electronically at `http://xxx.lanl.gov/archive/gr-qc`.

[29] H. Bondi, F. A. E. Pirani, "Gravitational waves in general relativity. XIII: Caustic property of plane waves," Proc. Roy. Soc. A, **421**, 519-533 (1989).

[30] M.W. Brinkmann, "On Riemann spaces conformal to Euclidean space," Proc. Natl. Acad. Sci. U.S.A. **9**, 1-3 (1923).

[31] Y. Choquet-Bruhat, C. De Witt-Morette, M. Dillard-Bleick, "Analysis, Manifolds and Physics," (North Holland, Amsterdam, 1977).

[32] Y. Choquet-Bruhat, "Applications of generalized functions to shocks and discrete models," in *Generalized Functions and Their Applications*, 37-49, ed. R. S. Pathak (Plenum Press, New York, 1993).

[33] C. J. S. Clarke, T. Dray, "Junction conditions for null hypersurfaces," Class. Quant. Grav. **4**, 265-275 (1987).

[34] C. J. S. Clarke, "The Analysis of Space-Time Singularities," Cambridge Lecture Notes in Physics, (Cambridge, 1993).

[35] C. J. S. Clarke, "Singularities: boundaries or internal points," in *Singularities, Black Holes and Cosmic Censorship* edited by P. S. Joshi, 24-32 (IUCCA, Bombay, 1996).

[36] C. J. S. Clarke, J. A. Vickers, J. P. Wislon, "Generalised functions and distributional curvature of cosmic strings," Class. Quant. Grav. **13**, 2485-2498 (1996).

[37] J. F. Colombeau, "Differential Calculus and Holomorphy, Real and Complex Analysis in Locally Convex Spaces," (North Holland, Amsterdam 1982.)

[38] J. F. Colombeau, "New Generalized Functions and Multiplication of Distributions," (North Holland, Amsterdam, 1984).

[39] J. F.Colombeau, "Elementary Introduction to New Generalized Functions," (North Holland, Amsterdam 1985).

[40] J. F. Colombeau, M. Oberguggenberger, "On a hyperbolic system with a compatible quadratic term: Generalized solutions, delta waves, and multiplication of distributions," Comm. Partial Differential Equations **15**, 905–938 (1990).

[41] J. F. Colombeau, "Multiplication of Distributions," (Springer, Berlin, 1992).

[42] J. F. Colombeau, A. Heibig, M. Oberguggenberger, "Le probleme de Cauchy dans un espace de fonctions generalisees I," C. R. Acad. Sci. Paris Sér. I Math. **317**, 851–855 (1993).

[43] J. F. Colombeau, A. Heibig, M. Oberguggenberger, "Le probleme de Cauchy dans un espace de fonctions generalisees II," C. R. Acad. Sci. Paris Sér. I Math. **319**, 1179–1183 (1994).

[44] J. F. Colombeau, A. Meril, "Generalized functions and multiplication of distributions on $\mathcal{C}^\infty$-manifolds," J. Math. Anal. Appl. **186**, 357-354 (1994).

[45] G. Darmois, "Lés equations de la gravitation Einsteinienne," Mémorial des sciences mathematiques XXV (Gauthier-Villars, Paris, 1927).

[46] N. Dapić, S. Pilipović, "Microlocal analysis of Colombeau's generalized functions on a manifold," Indag. Math, New Ser. **7** (3), 293-309 (1996).

[47] N. Dapić, S. Pilipović D. Scarpalézos, "Microlocal analysis of Colombeau's generalized functions—Propagation of singularities," J. Anal. Math. **75**, 51-66 (1998).

[48] S. de Haro, "Planckian scattering and black holes," J. High Energy Physics **10**, 023 (1998).

[49] G. de Rham, "Differentiable Manifolds," (Springer, Berlin, 1984).

[50] J. W. de Roever, M. Damsma, "Colombeau Algebras on a $\mathcal{C}^\infty$-manifold," Indag. Mathem., N.S., **2** (3), 341-358 (1991).

[51] H. J. De Vega, N. Sánchez, "Quantum scattering in the Aichelburg-Sexl geometry," Nucl. Phys. **B317**, 706-730 (1989).

[52] H. J. De Vega, N. Sánchez, "Particle scattering at the Planck scale an the Aichelburg-Sexl geometry," Nucl. Phys. **B317**, 731-756 (1989).

[53] J. Dieudonné, "Treatise on Analysis," Vol. 3, (Academic Press, New York, 1972).

[54] P. A. M. Dirac, "The physical interpretation of the quantum dynamics," Proc. Roy. Soc. A **113**, 612-641 (1926).

[55] P. A. M. Dirac, "The Principles of Quantum Mechanics," (Oxford, 1947, 1st ed. 1930).

[56] T. Dray, G. 't Hooft, "The gravitational shockwave of a massless particle," Nucl. Phys. B **232**, 173-188 (1985).

[57] J. Ehlers, W. Kundt, "Exact solutions of the gravitational field equations," in *Gravitation: An Itroduction into current research,* 49-101, ed. L. Witten (Wiley, New York, 1962).

[58] L. Ehrenpreis, "Solutions of some problems of division I," Amer. J. Math. **76**, 883-903 (1954).

[59] Yu. V. Egorov, "On a new theory of generalized functions," (russian) Vestnik Moskov. Univ. Ser. **I** (4), 69-99 (1989).

[60] Yu. V. Egorov, "A contribution to the theory of generalized functions," Russ. Math. Surveys **45** (5), 3-40 (1990).

[61] E. Farkas, M. Grosser, M. Kunzinger, R. Steinbauer, "On the foundations of nonlinear generalized functions I," ESI preprint **811** [2], Memoirs of the AMS, to appear (2000).


[2] ESI preprints are available electronically at `http://www.esi.ac.at`.

[62] V. Ferrari, P. Pendenza, G. Veneziano, "Beam like gravitational waves and their geodesics," J. Gen. Rel. Grav. **20** (11), 1185-1191 (1988).

[63] V. Ferrari, P. Pendenza, "Boosting the Kerr metric," Gen. Rel. Grav. **22** (10), 1105-1117 (1990).

[64] J. B. J. Fourier, "Theorie Analytique de la Chaleur," (Oeuvers I.) ed. Darboux, (Paris, 1822).

[65] F. G. Friedlander, "The Wave Equation on a Curved Space-Time," (Cambridge University Press, 1975).

[66] F. G. Friedlander, "Introduction to the Theory of Distributions," (Cambridge University Press, 1982).

[67] D. Gale, H. Nikaido, "The Jacobian matrix and global univalence of mappings," Math. Ann., **159**, 81-93 (1965).

[68] A. García Díaz, J. F. Plebański, "All nontwisting N's with cosmological constant," Jour. Math. Phys. **22**, 2655-2658 (1981).

[69] D. Garfinkle, "Metrics with distributional curvature," Class.Quant.Grav. **16**, 4101-4109 (1999).

[70] R. Geroch, "Limits of spacetimes," Comm. Math. Phys. **13**, 180-193, (1969).

[71] R. Geroch, J. Traschen, "Strings and other distributional sources in general relativity," Phys. Rev. D **36**, 1017-1031 (1987).

[72] J. R. III Gott, "Gravitational lensing effects of vacuum strings: exact solutions.," Astrophys. J. **288** (2), 422-427 (1985).

[73] J. B. Griffiths, "Colliding plane waves in general relativity," Oxford Mathematical Monographs (Oxford Science Publications, 1991).

[74] H. Grosse, M. Oberguggenberger, I. T. Todorov, "Generalized functions for quantum fields obeying quadratic exchange relations," ESI preprint [2] **653**, (1999).

[75] M. Grosser, "Neue verallgemeinerte Funktionen," Seminar Notes, (University of Vienna, 1995.)

[76] M. Grosser, "On the foundations of nonlinear generalized functions II," ESI preprint **812** [2], Memoirs of the AMS, to appear (2000).

[77] M. Grosser, M. Kunzinger, R. Steinbauer, J. Vickers, "A global theory of algebras of generalized functions," ESI-preprint **813** [2] (1999).

[78] M. Grosser, M. Kunzinger, R. Steinbauer, H. Urbantke, J. Vickers, ,,Diffeomorphism-invariant construction of nonlinear generalized functions," to appear in *Proceedings of Journées Relativistes 99,* Annalen der Physik (2000).

[79] O. Hájek, Bull. AMS **12**, 272-279 (1985).

[80] S. W. Hawking, G. F. R. Ellis, "The Large Scale Structure of Space-Time," (Cambridge University Press, Cambridge, 1973).

[81] K. Hayashi, T. Samura, "Gravitational shock waves for Schwarzschild and Kerr black holes," Phys. Rev. D **50**, 3666-3675 (1994).

[82] J. Hély, "Modèles d'univers en état de radiation pure," Compt. Rend. **249**, 1867-1874 (1959).

[83] R. Hermann, M. Oberguggenberger, "Ordinary differential equations and generalized functions," in *Nonlinear Theory of Generalized Functions,* Chapman & Hall/CRC Research Notes in Mathematics **401**, 85-98, eds. M. Grosser, G. Hörmann, M. Kunzinger, M. Oberguggenberger (Chapman & Hall/CRC, Boca Raton 1999).

[84] W. Hiscock, "Exact gravitational field of a string," Phys. Rev. D **31**, 3288-3290 (1985).

[85] P. A. Hogan, "A spherical impulsive gravity wave," Phys. Rev. Lett. **70** (2), 117-118 (1993).

[86] P. A. Hogan, "Lorentz group and spherical impulsive gravity waves," Phys. Rev. D **49** (12), 6521-6525 (1994).

[87] J. Horváth, "Topological Vector Spaces and Distributions," Vol. 1, (Addison-Wesly, Reading, 1966).

[88] L. Hörmander, "The Analysis of Linear Partial Differential Opertors I," (Springer, Berlin, 2nd ed., 1990).

[89] L. Hörmander, "The Analysis of Linear Partial Differential Opertors II," (Springer, Berlin, 2nd ed., 1990).

[90] L. Hörmander, "Local and global properties of fundamental solutions," Math. Scand. **5**, 27-39 (1957).

[91] L. Hörmander, "On the division of distributions by polynomials," Ark. Mat. **3**, 555-568 (1958).

[92] L. Hörmander, "On the existence and the regularity of solutions of linear pseudo-differenatial equations," Ens. Math. **17**, 99-163 (1971).

[93] G. Hörmann, M. Kunzinger, "Nonlinearity and Self-Interaction in Physical Field Theories with Singularities," Int. Transf. Spec. Funct. **6**, 190-199 (1997).

[94] G. Hörmann, M. Kunzinger, "Regularized derivatives in a 2-dimensional model of self-interacting fields with singular data," Zeitschr. Anal. Anw. to appear (1999).

[95] M. Hotta, M. Tanaka, "Shock wave geometry with nonvanishing cosmological constant," Class. Quant. Grav. **10**, 307-314 (1993).

[96] C. J. Isham, "Some Quantum field theory aspects of the superspace quantization of general relativity," Proc. R. Soc. **A 351**, 209-232 (1976).

[97] W. Israel, "Singular hypersurfaces and thin shells in general relativity," Nouv. Cim. **44B**, (1) 1-14 (1966).

[98] W. Israel, "Line sources in general relativity," Phys. Rev. D **15**, (4), 935-941, (1977).

[2] ESI preprints are available electronically at `http://www.esi.ac.at`.

[99] J. Jelínek, "An intrinsic definition of the Colombeau generalized functions," Comment. Math. Univ. Carolinae. **40**, 71–95 (1999).

[100] P. Jordan, J. Ehlers, W. Kundt, "Strenge Lösungen der Feldgleichungen der allgemeinen Relativitätstheorie," Akad. Wiss. Lit. Mainz, Abh. Math.-Naturw. Kl. 1960, 21-85 (1960).

[101] P. Jordan, J. Ehlers, W. Kundt, "Beiträge zur Theorie der Gravitations-Strahlungsfelder. Strenge Lösungen der Feldgleichungen der allgemeinen Relativitätstheorie," Akad. Wiss. Lit. Mainz. Abh. Math.-Naturw. Kl. 1962, 969-1000 (1963).

[102] A. Kager et. al., Europ. Jour. Psy. to appear (2000).

[103] K. Kahn, R. Penrose, "Scattering of two impulsive gravitational plane waves," Nature **229**, 185-186 (1971).

[104] W. Kamleh, "Signature changing space-times and the new generalised functions," preprint, gr-qc/0004057 [1] (2000).

[105] T. Kawai, E. Sakane, "Distributional energy-momentum densities of Schwarzschild space-time," Prog. Theor. Phys. **98**, 69-86 (1997).

[106] B. L. Keyfitz, H. C. Kranzer, "Spaces of weighted measures for conservation laws with singular shock solutions," J. Differential Equations **118**, 420–451 (1995).

[107] G. Kirchhoff, "Zur Theorie der Lichstrahlen," Sitz. K. Pruess. Akad. Wiss. 641-669 (Berlin, 1882).

[108] H. König, "Neue Begründung der Theorie der Distributionen von L. Schwartz," Math. Nachr. **9**, 129-148 (1953).

[109] H. König, "Multiplikation von Distributionen. I.," Math. Ann. **128**, 420-452 (1955).

[110] G. Köthe, "Topological vector spaces, II.," (Springer, Berlin Heidelberg, New York 1979).

[111] D. Kramer, H. Stephani, E. Herlt, M. MacCallum, "Exact solutions of Einstein's field equations," Cambridge Monographs on Mathematical Physics **6**, (Cambridge University Press, 1980).

[112] A. Kriegl, "Differentialgeometrie," Seminar Notes, (University of Vienna, 1996).

[113] A. Kriegl, P. W. Michor, "The Convenient Setting of Global Analysis," Mathematical Surveys and Monographs **53**, (Amer. Math. Soc., Providence, 1997.)

[114] W. Kundt, "Methoden zur Charakterisierung von Lösungen der Einsteinschen Gravitationsfeldgleichungen," thesis (Hamburg, 1958).

[115] W. Kundt, A. Thompson, "Le tenseur de Weyl et une congruence associée de géodésique isotropes sans distortion," C. R. Acad. Sci. Paris **254**, 4257-4259 (1962).

[116] M. Kunzinger, "Barrelledness, Baire-like and (LF)-Spaces," Pitman Research Notes in Mathematics **298**, (Longman Scientific and Technical, Harlow, 1993).

[117] M. Kunzinger, "Lie Transformation Groups in Colombeau Algebras," Ph.D. thesis (University of Vienna, 1996).

[118] M. Kunzinger, R. Steinbauer, "A rigorous solution concept for geodesic and geodesic deviation equations in impusive gravitational waves," J. Math. Phys. **40**, 1479-1489 (1999).

[119] M. Kunzinger, R. Steinbauer, "A note on the Penrose junction conditions," Class. Quant. Grav. **16**, 1255-1264 (1999).

[120] M. Kunzinger, "Lie symmetries of differential equations in a generalized functions setting," in *Nonlinear Theory of Generalized Functions,* Chapman & Hall/CRC Research Notes in Mathematics **401**, 241-250, eds. M. Grosser, G. Hörmann, M. Kunzinger, M. Oberguggenberger (Chapman & Hall/CRC, Boca Raton 1999).

[121] K. Lanczos, "Bemerkungen zur de Sitterschen Welt," Phys. Z. **32**, 539-543 (1922).

[122] K. Lanczos, "Flächenhafte Verteilung der Materie in der Einsteinschen Gravitationstheorie," Ann. d. Phsy. **74**, 518-540 (1924).

[123] H. Lewy, "An Example of a smooth linear partial differential equation without solution," Ann. Math. **66** (2), 155-158 (1957).

[124] A. Lichnerowicz, "Théories Relativistes de la Gravitation et de l'Electromangnétisme," (Masson, Paris 1955).

[125] A. Lichnerowicz, "Relativity and mathematical physics," in *Relativity, Quanta and Cosmology in the Development of the Scientific thought of Albert Einstein* Vol.2, 403-472, eds. M. Pentalo, I. de Finis (Johnson, New York, 1979).

[126] A. Lichnerowicz, "Sur les ondes de choc gravitationnelles," C. R. Acad. Sc. Paris **273**, 528-532 (1971).

[127] S. Łojasiewicz, "Sur le probléme de division," Studia Math. **18**, 87-136 (1959).

[128] S. Łojasiewicz, "Sur la valeur et la limite d'une distribution en un point," Studia Math. **16**, 1-36 (1957).

[129] C. O. Loustó, N. Sánchez, "Gravitational shock waves of ultra-high energetic particles on curved space-time," Pys. Lett. B **220** (1,2), 55-60 (1989).

[130] C. O. Loustó, N. Sánchez, "The ultrarelativistic limit of the Kerr-Newman geometry and particle scattering at the Planck scale," Pys. Lett. B **232** (4), 462-466 (1989).

[131] C. O. Loustó, N. Sánchez, "The curved shock wave space-time of ultrarelativistic charged particles and their scattering," Int. J. Mod. Phys. **A5**, 915-938 (1990).


[1] gr-qc preprints are available electronically at `http://xxx.lanl.gov/archive/gr-qc`.

[132] C. O. Loustó, N. Sánchez, "Gravitational shock waves generated by extended sources: Ultrarelativistic cosmic strings, monopoles and domain walls," Nucl. Phys. B **355**, 231-249 (1991).

[133] C. O. Loustó, N. Sánchez, "The ultrarelativistic limit of the boosted Kerr-Newman geometry and the scattering of spin-$\frac{1}{2}$ particles," Nucl. Phys. B **383**, 377-394 (1992).

[134] J. Lützen, "The Prehistory of the Theory of Distributions," (Springer, New York, 1982).

[135] B. Malgrange, "Existence et approximation des solutions des éuqations aux dérivées partielles et des équations de convolution," Ann. Inst. Fourier (Grenoble) **6**, 271-335 (1955-56).

[136] R. Mansouri, M. Khorrami, "Equivalence of Darmois-Israel and distributional-methods for thin shells in general relativity," J. Math. Phys. **37**, 5672-5683 (1996).

[137] R. Mansouri, K. Nozari, "Signature change, inflation, and the cosmological constant," preprint, gr-qc/9806109 [1] (1998).

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

[139] P. W. Michor, "Foundations of Differential Geometry," Seminar Notes (University of Vienna, 1991).

[140] J. Mikusinski, "Sur la méthode de généralisation de M. Laurent Schwartz sur la convergence faible," Fund. Math. **35**, 235-239 (1948).

[141] C. W. Misner, K. S. Thorne, J. A. Wheeler, "Gravitation," (Freeman, San Francisco, 1973).

[142] M. Nedelkov, S. Pilipović, D. Scarpalézos, "The Linear Theory of Colombeau Generalized Functions," Pitman Research Notes in Mathematics **385**, (Longman Scientific and Technical, Harlow, 1998.)

[143] Y. Nutku, R. Penrose, "On impulsive gravitational waves," Twistor Newsletter **34**, 9-12 (1992).

[144] M. Oberguggenberger, "Multiplication of Distributions and Applications to Partial Differenatial Equations," Pitman Research Notes in Mathematics **259**, (Longman Scientific and Technical, Harlow, 1992).

[145] M. Oberguggenberger, M. Kunzinger, "Characterization of Colombeau functions by their pointvalues," Math. Nachr. to appear (1999).

[146] S. O'Brien, J. L. Synge, Commun. Dublin Inst. Adv. Stud. **A9** (1952).

[147] B. O'Neill, "Semi-Riemannian Geometry (With Applications to Relativity)," (Academic Press, New York, 1983).

[148] N. Ortner, P. Wagner, "A short proof of the Malgrange-Ehrenpreis theorem," in *Functional analysis. Proceedings of the first international workshop held at Trier University, Germany, September 26–October 1, 1994*, 343-352, eds. S. Dierolf et al. (de Gruyter, Berlin) (1996).

[149] P. Parker, "Distributional geometry," J. Math. Phys. **20** (7), 1423-1426, (1979).

[150] R. Penrose, "Twistor Quantization and curved space-time," Int. Jour. Theor. Phys. **1** (1), 61-99, (1968).

[151] R. Penrose, "Structure of space-time," in *Battelle Recontres*, 121-235, ed. C. M. DeWitt and J. A. Wheeler (Benjamin, New York, 1968).

[152] R. Penrose, "The geometry of impulsive gravitational waves," in *General Relativity, Papers in Honour of J. L. Synge,* 101-115, ed. L. O'Raifeartaigh (Clarendon Press, Oxford, 1972).

[153] R. Penrose, "A remarkable property of plane waves in general relativity," Rev. Mod. Phys, **3**, 215-220 (1965).

[154] A. Peres, "Some gravitational waves," Phys. Rev. Lett. **3**, 571-572 (1959).

[155] J. Persson, "Fundamental theorems for linear measure differenatial equations," Math. Scand. **62**, 19-43 (1988).

[156] J. Persson, "Distribution differenatial equations. Regularization and bending of a rod," Math. Scand. **66**, 291-300 (1990).

[157] J. Podolský, J. B. Griffiths, "Impulsive waves in de Stitter and anti-de Sitter spacetimes generated by null particles with an arbitrary multipole structure," Class. Quant. Grav. **15**, 453-463 (1998).

[158] J. Podolský, J. B. Griffiths, "Boosted static multipole particles as sources of impulsive gravitational waves," Phys. Rev. D **58**, 124024 (1998).

[159] J. Podolský, "Non-expanding impulsive gravitational waves," Class. Quant. Grav. **15**, 3229-3239 (1998).

[160] J. Podolský, K. Vesely, "New examples of sandwich gravitational waves and their impulsive limit," Czech. J. Phys. **48**, 871 (1998).

[161] J. Podolský, K. Vesely, "Continuous coordinates for all impulsive pp-waves," Phys. Lett. A **241**, 145-147 (1998).

[162] J. Podolský, J. B. Griffiths, "Expanding impulsive gravitational waves," Class. Quant. Grav. **16**, 2937-2946 (1999).

[163] N. Rosen, "Plane polarized waves in the general theory of relativity," Phys. Z. Sowjet. **12**, 366-372 (1937).

[164] E. E. Rosinger, "Nonlinear Partial Differential Equations. Sequential and Weak Solutions," (North Holland, Amsterdam, 1980).

[165] E. E. Rosinger, "Generalized Solutions of Nonlinear Partial Differential Equations," (North Holland, Amsterdam, 1987).

[166] E. E. Rosinger, "Non-Linear Partial Differential Equations. An Algebraic View of Generalized Solutions," (North Holland, Amsterdam, 1990).

[1] gr-qc preprints are available electronically at `http://xxx.lanl.gov/archive/gr-qc`.

[167] L. Schwartz, “Théorie des Distributions,” (Herman, Paris, 1966).

[168] L. Schwartz, “Sur l’impossibilité de la multiplication des distributions,” C. R. Acad. Sci. Paris, **239**, 847-848 (1954).

[169] S. R. Simanca, “Pseudo-differential Operators,” Pitman Research Notes in Mathematics **236**, (Longman Scientific and Technical, Harlow, 1990).

[170] S. L. Sobolev, “Méthode nouvelle á résoudre le problème de Cauchy pour les équations linéaires hyperboliques normale,” Mat. Sb. **1**, (43), 39-71 (1936).

[171] R. Steinbauer, “Colombeau-Theorie und ultrarelativistischer Limes,” master thesis, (University of Vienna, 1995).

[172] R. Steinbauer, “Distributionelle Krümmung von Strings,” Seminar Notes (1997).

[173] R. Steinbauer, “The ultrarelativistic Reissner-Nordstrøm field in the Colombeau algebra,” J. Math. Phys. **38**, 1614-1622 (1997).

[174] R. Steinbauer, “Geodesics and geodesic deviation for impulsive gravitational waves,” J. Math. Phys. **39**, 2201-2212 (1998).

[175] R. Steinbauer, “Distributional description of impulsive gravitational waves,” in *Nonlinear Theory of Generalized Functions,* Chapman & Hall/CRC Research Notes in Mathematics **401**, 267-274, eds. M. Grosser, G. Hörmann, M. Kunzinger, M. Oberguggenberger (Chapman & Hall/CRC, Boca Raton 1999).

[176] R. Steinbauer, “On the geometry of impulsive gravitational waves,” *Proceedings of the VIII Romanian Conference on General Relativity,* ed. I. Cotaescu, D. Vulcanov (Mirton Publishing House, Timisoara, 1999).

[177] H. A. Taub, “Space-times with distribution valued curvature tensors,” J. Math. Phys. **21** (6), 1423-1431 (1980).

[178] H. Takeno, “On plane wave solutions of field equations in general relativity,” Tensor, n. Ser. **7**, 97-102 (1957).

[179] G. Temple, “Theories and applications of generalized functions,” J. London Math. Soc. **28**, 134-148 (1953).

[180] G. t’Hooft, “The scattering matrix approach for the quantum black hole,” Int. J. Mod. Phys. A **11**, 4623-4688 (1996).

[181] F. Treves, “Fundamental solutions of linear partial differential equations with constant coefficients depending on parameters,” Amer. J. Math. **84**, 561-577 (1962).

[182] H. Verlinde, E. Verlinde, “Scattering at Planckian energies,” Nucl. Phys. B **371**, 246-268 (1992).

[183] H. Verlinde, E. Verlinde, “High-energy scattering in Quantum gravity,” Class. Quant. Grav. **10**, 175-184 (1993).

[184] J. A. Vickers, J. P. Wilson, “A nonlinear theory of tensor distributions,” ESI preprint[1] **566**, gr-qc/9807068 [2] (1998).

[185] J. A. Vickers, J. P. Wilson, “Invariance of the distributional curvature of the cone under smooth diffeomorphisms,” Class. Quant. Grav. **16**, 579-588 (1999).

[186] J. A. Vickers, “Nonlinear generalised functions in general relativity,” in *Nonlinear Theory of Generalized Functions,* Chapman & Hall/CRC Research Notes in Mathematics **401**, 275-290, eds. M. Grosser, G. Hörmann, M. Kunzinger, M. Oberguggenberger (Chapman & Hall/CRC, Boca Raton 1999).

[187] R. Wald, “General Relativity,” (The University of Chicago Press, Chicago, 1984).

[188] J. P. Wilson, “Regularity of Axisymmetric Space-times in General Relativity,” Ph.D. thesis (University of Southampton, 1997).

[189] J. P. Wilson, “Distributional curvature of time dependent cosmic strings,” Class. Quant. Grav. **14**, 3337-3351 (1997).

[190] S. Yamamuro, “Differenatial Calculus in Linear Topological Spaces,” (Hermann, Paris 1966).

[191] U. Yurtsever, “Colliding almost-plane gravitational waves: Colliding exact plane waves and general properties of almost-plane-wave space-times, ” Pys. Rev. D, **37**, 2803-2817 (1988).

[192] V. D. Zakharov, “Gravitational waves in Einstein’s theory,” (Halsted Press, New York, 1973).


[1] ESI preprints are available electronically at `http://www.esi.ac.at`.

[2] gr-qc preprints are available electronically at `http://xxx.lanl.gov/archive/gr-qc`.

# Index

# Curriculum Vitae

### Persönliche Daten

| | |
|---|---|
| Name: | Roland Steinbauer |
| Geburtdatum: | 23. Februar 1970, Wien |
| Staatsbürgerschaft: | Österreich |
| Familienstand: | ledig |
| Eltern: | Christl und DI. Wolfgang Steinbauer |
| Bruder: | Mag. Christoph Steinbauer |

### Ausbildung

| | |
|---|---|
| 1976 - 1980: | Volksschule, Marktg. 31, A-1090 Wien |
| 1980 - 1988: | BRG IX, Glaserg. 25, A-1090 Wien, Matura mit ausgezeichnetem Erfolg |
| Okt. 88 - Mai 89: | Präsenzdienst in Wien und Stockerau |
| Okt. 89 - Jun. 91: | 1. Studienabschnitt der Diplomstudien Mathematik und Physik an der Universität Wien |
| Juni 1991: | 1. Diplomprüfung beider Studien mit ausgezeichnetem Erfolg |
| Okt. 91 - Jan. 96: | 2. Studienabschnitt Mathematik und Physik |
| 29-01-1996: | 2. Diplomprüfung (Theoretische Physik) mit Auszeichung |
| 18-03-1996: | Sponsion zum Magister der Natuwissenschaften |
| seit März 96: | Doktoratsstudium Mathematik |

### Wissenschaftliche Laufbahn

| | |
|---|---|
| WS 1992/93: | Tutor am Institut für Mathematik der Universität Wien |
| Jan. 93 - Jun. 93: | Studienassistent (halbbeschäftigt) am Institut für Mathematik der Universität Wien |
| 1993-1998: | Regelmäßig Tutor am Institut für Mathematik und am Institut für Theoretische Physik der Universität Wien |
| Nov. 96 - April 98: | Doktoratsstipendium der Österreichischen Akademie der Wissenschaften |
| Mai 98 - Okt 99: | Wissenschaftlicher Mitarbeiter des FWF-Projekts P-12023MAT „Distributionelle Methoden in Einsteins Gravitationstheorie" (tw. halbbeschäftigt) |
| seit Juli 98: | Vertragsassistent am Institut für Mathematik der Univeristät Wien (tw. halbbeschäftigt) |

### Publikationen

1. R. Steinbauer, *The ultrarelativistic Reissner-Nordstrom field in the Colombeau algebra,* J. Math. Phys. **38** (3), 1614-1622 (1997).

2. S. Husa, A. Nairz, R. Steinbauer und B. Pocza, *ELFI, A gravitational wave detector for cosmology,* in *Fundamental Physics in Space, Proceedings of the Alpbach Summer School 1997*, Hrsg. A. Wilson, 45-47 (ESA Publications Division, Noordwijk, 1997).

3. R. Steinbauer, *Geodesics and geodesic deviation for impulsive gravitational waves,* J. Math. Phys. **39** (4), 2201-2212 (1998).

4. R. Steinbauer, *On the geometry of impulsive gravitational waves,* in *Proccedings of the VIII Romanian Conference on General Relativity and Gravitation,* Hrsg. I. Cotaescu, D. Vulcanov (Mirton Publishing House, Timisoara, Rumänien, 1998).

5. M. Kunzinger, R. Steinbauer, *A rigorous solution concept for geodesic equations in impulsive gravitational waves,* J. Math. Phys, **40** (3), 1479-1489 (1999).

6. M. Kunzinger, R. Steinbauer, *A note on the Penrose junction conditions,* Class. Quant. Grav. **16**, 1255-1264 (1999).

7. R. Steinbauer, *Distributional description of impulsive gravitational waves,* in *Nonlinear Theory of Generalized Functions,* Hrsg. M. Grosser, G. Hörmann, M. Kunzinger und M. Oberguggenberger (Chapman & Hall/CRC Research Notes in Mathematics, Boca Raton, 1999).

8. R. Steinbauer, *On the impulsive limit of gravitational waves,* in *Proceedings of the Spanish Relativity Meeting* (Bilbao, 1999).

9. E. Farkas, M. Grosser, M. Kunzinger, R. Steinbauer, *On the foundations of nonlinear generalized functions I,* Memoirs of the AMS, akzeptiert (2000).

10. M. Grosser, M. Kunzinger, R. Steinbauer, J. Vickers, *A global theory of algebras of generalized functions,* preprint (1999).

11. S. Husa, B. Pocza, R. Steinbauer, A. Nairz, K. Danzmann, *LARS, Large Space Interferometer for Cosmology,* Proposal to the European Space Agency for a Flexi Mission in Fundamental Physics (2000).

12. M. Grosser, M. Kunzinger, R. Steinbauer, H. Urbantke, J. Vickers, *Diffeomorphism-invariant construction of nonlinear generalized functions,* erscheint in *Proceedings of Journées Relativistes 99,* Annalen der Physik (2000).

**Forschungsaufenthalte, Konferenzen, Sommerschulen**

seit Feb. 1996: Regelmäßige Forschungsaufenthalte am Institut für Mathematik und Geometrie, Universität Innsbruck.

Juni 1997: II Workshop on Low-dimensional Topology and Quantum Gravity, Stefan Banach Center, Warschau, Polen.

Juli 1997: ESA Summer School 1997 “Fundamental Physics in Space,” Alpbach.

Aug. 1997: 179. WE-Heraeus Semiar “Black Holes: Theory and Observation,” Bad Honnef, Deutschland.

Nov. 1997: ESI Workshop “Nonlinear Theory of Generalized Functions,” Erwin Schrödinger Institut, Wien.

Dez. 1997: GR15 (15th International Conference on General Relativity and Gravitation), IUCAA, Pune, Indien.

Juni 1998: VIII Rumanian Conference on General Relativity and Gravitation, Bistrita, Rumänien.

Juli 1998: Workshop “Nonlinear Theory of Generalized Functions,” Universität Novi Sad, Jugoslawien.

Aug. 1998: 2nd Samos Meeting on Cosmology, Geometry and Relativity “Mathematical & Quantum Aspects of Relativity and Cosmology,” Pythagoreon, Griechenland.

Sep. 1998: 201. WE-Heraeus Seminar “Mathematical Problems in General Relativity,” Bad Honnef, Deutschland.

Sep. 1999: Spanish Relativity Meeting, University of the Basque Country, Bilbao, Spanien.

Nov. 1999: Institut für Theoretische Physik, Karlsuniversität Prag.

April 2000: International Conference on Generalized Functions (ICGF 2000), Université des Antilles et de la Guyane, Guadeloupe, Frankreich.