DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

Hyperbolic systems of partial differential equations: Generalised solutions, regularity, and regularisations

verfasst von / submitted by

Mag. Christian Spreitzer

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 091 411
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Physik
Betreut von / Supervisor:	ao. Univ.-Prof. Mag. Dr. Günther Hörmann

Abstract

We study the Cauchy problem for symmetric hyperbolic systems of partial differential equations in Colombeau algebras of generalised functions. This framework allows to consider equations with highly singular coefficients, for which distributional solution concepts fail. In particular, it is possible to deal with coefficients incorporating the delta distribution or its derivatives on a mathematically sound basis. Differential equations with coefficients of this type occur in various contexts in physics. The essential methods for our existence and uniqueness results are based on energy estimates on lens-shaped domains in space-time. We investigate the regularity of generalised solutions and their distributional limits when the coefficients are obtained as regularisations of non-smooth Sobolev functions. Specifically, we show that distributional limits of generalised solutions exist even for coefficient regularities below those required within classical weak solution concepts. Furthermore, we establish the transformation from scalar second-order wave equations to related symmetric hyperbolic first-order systems in a generalised functions setting, thereby obtaining existence and uniqueness results for second-order wave equations. The operations involved in the transformation algorithm are not defined classically for coefficients of low regularity. As applications in physics, we consider the Dirac equation and the Klein-Gordon equation with delta distribution potentials, as well as the wave equation derived from the Laplace-Beltrami operator of a Lorentzian metric of low regularity. Motivated by the significant role of convolution regularisations in the embedding of distributions in Colombeau algebras, we develop an explicit construction of Schwartz functions with prescribed moments and support contained in an arbitrarily chosen cone. Our construction provides tailor-made mollifiers, satisfying adaptable support and convergence conditions. In addition, it yields explicit solutions to several variants of the Stieltjes moment problem for rapidly decreasing smooth functions.

Acknowledgements

First and foremost, I am deeply indebted to my thesis supervisor Günther Hörmann. I am grateful for his continuous support, encouragement, patience, and mentorship. Our meetings were not only most helpful for the work on my thesis project, but also gave me a lot of inspiration and confidence. I feel very privileged as his doctoral student.

I want to express my gratitude to all members and friends of the DiANA group at the Faculty of Mathematics, who have contributed immensely to my academic development. Special thanks go to my co-authors Clemens Hanel, Günther Hörmann, Roland Steinbauer, and to Michael Kunzinger. Our joint work and discussions have been a great pleasure and I benefited considerably from reading mathematical texts together. Michael Grosser and his style of teaching added a lot to my mathematical education and my personal objectives as a teacher. Working together with Franz Embacher for the project 'Math-Bridge' was a very nice experience I would not want to have missed. I am grateful to my physics professors Peter C. Aichelburg and Markus Arndt, who had a significant influence on my scientific interests and my fascination for physics, and to my student colleagues for numerous stimulating discussions and all the joyful time and experiences we shared.

My sincere thanks go to Alfred S. Posamentier for continuously encouraging and reminding me to work on the completion of my thesis, despite my professional duties and teaching obligations. I am indebted to my colleagues at the University College of Teacher Education, in particular I want to thank my co-authors Evelyn Süss-Stepancik and Elisabeth Mürwald-Scheifinger. During my teaching experiences over the last years, I also learned a lot from my students and I appreciate the time we spent together.

I gratefully acknowledge the funding sources that made the writing of this thesis possible. My work has been supported by the Initiativkolleg of the University of Vienna (Time-frequency analysis and microlocal analysis) and by the research projects Y237 (Nonlinear distributional geometry) and P20525 (Global analysis in algebras of generalized functions) of the Austrian Science Fund (FWF). I want to thank the project leaders Michael Kunzinger and Hans Georg Feichtinger for their support and their efforts to provide the best possible working environment.

This thesis would not have been possible without the love and support of my whole family. Thank you! My unbounded gratitude goes to Katharina for her deep support, love, care, and understanding. I also want to thank her for proofreading the manuscript and for all her comments and suggestions that substantially improved this thesis.

Contents

Chapter 1

Introduction and preliminaries

Partial differential equations (PDEs) are at the heart of many physical models. In particular, our fundamental theories of nature are most beautifully represented in the form of PDEs. Famous and often celebrated examples include the Schrödinger equation in quantum mechanics, the Dirac equation and the Klein-Gordon equation in relativistic quantum mechanics, the Maxwell equations in electrodynamics, and the Einstein equations in general relativity.

Solutions of these and many other PDEs in physics are quantities that assign numbers to points in space and time. In physics, such quantities are called (classical) fields, a term that was originally introduced by M. Faraday (cf. [McM02]). However, not all fields in physics are directly measurable in experiments and some are, in addition, even subject to a gauge freedom. It has been debated among physicists whether certain fields constitute any kind of physical reality or whether they are merely theoretical objects simplifying the mathematical description of a physical problem (cf. [WF98, p. 163] and [EI38, p. 152]). A famous contribution to this discussion is the Ehrenberg-Siday-Ahoronov-Bohm effect [ES49, AB59], in which an electrically charged particle is influenced by the presence of an electromagnetic field enclosed in a solenoid, although the particle is completely shielded from the electromagnetic field, in the sense that the overlap between the particle's wave function and the electromagnetic field is negligible. The particle seems to be affected by the electromagnetic potential field which extends beyond the solenoid, but was believed to be a purely mathematical construct without physical reality. Moreover, the meaning or interpretation of fields used in physical theories may change as the theories evolve or are replaced by more accurate or more elegant descriptions of nature. For example, the fields in general relativity represent the geometry of space-time itself, whereas in Newton's theory of gravity, space and time were separated and considered to be the everlasting and unchangeable arena in which the gravitational field acts. In any case, the notion of a field has been and still is of utmost importance in modern physics.

Since nothing can travel faster than the speed of light, disturbing a field in some region in space and time cannot be felt instantaneously by distant observers. Physical particles or disturbances of physical fields always propagate at finite speed. Disturbing an initially static field gives rise to a 'wave-like' phenomenon travelling through space at a certain speed. In fact, waves or 'wave-like' phenomena are nothing but travelling disturbances of physical fields. Finite propagation speed of disturbances is an inherent property of hyperbolic PDEs, the wave equation being an archetypical example. Electromagnetic waves travel essentially at the speed of light, that is why telecommunication is such a successful technology. In relativistic quantum mechanics, moving particles are described by 'wave packets' whose speed is limited by the speed of light. Wave-like phenomena are ubiquitous in modern physics. As a matter of fact, describing any physical quantity that can be represented by a field in a way consistent with special relativity, will most of the time lead to a hyperbolic PDE or a system of such equations.

Considering only first-order systems does not pose a restriction inasmuch as any system of PDEs can be written as a first-order system. Establishing existence and uniqueness of solutions of the initial value problem (or Cauchy problem) for hyperbolic systems used in physical models is indispensable from a mathematical point of view, but it is also important for the physical interpretation of solutions.

1.1 Partial differential equations with distributions as coefficients

PDEs with discontinuous or distributional coefficients are widely used in physics. The prototype of a discontinuous function is the Heaviside function and the most famous example of a distribution is Dirac's delta distribution (also called delta function, δ-distribution, $\delta(x)$). It is an efficient and practical tool to model sharp impulses or 'point-like' concentrations of energy. By employing the delta distribution to represent a narrow spike in a coefficient of a differential equation, one often obtains a good approximation of the real physical situation without having to deal with the detailed structure of the spike. If, for instance, the coefficient represents an energy density, then only the total energy contained in the spike has to be known, while its inner structure is submersed in the delta distribution.

However, using a delta distribution in the coefficient of a differential equation is not always tantamount to ignoring the details of an interaction on purpose, in order to simplify a physical model. Elementary particles like electrons or quarks are assumed to be physical point particles. As far as we know today, they do not possess any spatial extension in a classical sense. Consequently, the charge density of an electron within the framework of classical electrodynamics has to be a delta distribution. This is not a deliberate simplification, but the best possible description of the charge density of an electron that is available in the classical theory of electrodynamics. The delta distribution may also arise in certain limiting cases of physical models, as for instance in the Aichelburg-Sexl boost, describing the physical experience of an observer moving past a spherically symmetric gravitating object as it approaches the speed of light [AS71]. Discontinuous functions as coefficients occur even more very frequently in PDEs in physics, for instance in continuum mechanics, where the Heaviside function provides an accurate description of the behaviour of certain material parameters across a boundary surface between two regions made up of different compounds. Of course, Dirac's delta distribution is just a theoretical construct. But in a less obvious way, the same is true for the Heaviside function or any other function, smooth or not. Ultimately, the mathematical formulation of any physical theory cannot be more than a theoretical model designed to describe nature. Distributions appear as coefficients in PDEs in many branches of physics, sometimes for the purpose of simplifying an otherwise too complicated problem, sometimes just because it is the most authentic description of a physical reality, at least within our current theories of nature. Therefore, the study of discontinuous functions or distributions as coefficients of (hyperbolic) PDEs is not only an issue in mathematics, but also of substantial significance in physics, as numerous applications show.

The δ-distribution was given its name by P.A.M. Dirac in 1926 in analogy to the 'Kronecker-δ' (cf. [Dir27, p. 625] and [Dir30, p. 58]), but concepts which are at least closely related go back to the 1820ies. Different representations of the delta distribution had already been used by Fourier in 1822 (cf. [Fou09, §235]) and, around the same time, by Poisson and Cauchy [Cau27]. An in-depth discussion of the origins of the delta distribution is given in [KT13], [Lau89], [Lau92], and [Lüt82]. Dirac's delta distribution soon proved to be a valuable tool in manifold applications and gained great popularity in physics. It is now widely used in many circumstances and physics students often encounter it already in introductory courses.

A popular example that can be found in many undergraduate textbooks on quantum mechanics is the one-dimensional Schrödinger equation with a potential barrier proportional to the delta distribution:

$$i\hbar\frac{\partial}{\partial t}\psi(t,x) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \alpha\,\delta(x)\right)\psi(t,x). \tag{1.1}$$

For $\alpha > 0$, this simple model demonstrates the 'quantum tunnelling effect', as travelling plane waves hitting the barrier from one side are partially transmitted through the barrier. Interpreting the transmission coefficient (that is, the ratio of the intensity of the transmitted wave to that of the incident wave) associated to this model as the probability for a single particle to penetrate the barrier, a quantum mechanical particle can 'tunnel' through a potential barrier of arbitrary height, provided the barrier is sufficiently narrow. In this standard example, the wave function is usually constructed by joining 'plane wave solutions' on both sides of the barrier by means of a continuity condition derived from the equation. The so-obtained solution candidate $(t,x)\mapsto\psi(t,x)$ is continuous across the barrier, but its x-derivative exhibits a jump at $x=0$. Although a mathematically educated reader would be concerned about regularity issues and well-posedness of the problem, to begin with, physics textbooks typically ignore such things or bypass them in a more or less elegant way. Yet, from a mathematical point of view we can still make sense of the equation by employing a weak solution concept. The product of the delta distribution and the continuous function ψ is defined and the solution candidate indeed satisfies the equation in a weak sense, the equation being interpreted as an equality in the space of distributions. However, the theory of distributions does not provide a general framework for dealing with PDEs whose coefficients are allowed to be as singular as Dirac's delta distribution. Suppose we replace the potential barrier $\delta(x)$ in (1.1) by a 'shock in time' $\delta(t)$. Then it is easy to see that any nonzero wave function that is continuous at $t = 0$ cannot be a (weak) solution in any classical sense. Exchanging the roles of space and time in the potential term drastically affects the regularity here, since space derivatives and time derivatives are not represented symmetrically in the Schrödinger equation. Nevertheless, if a solution still exists, then it must satisfy the free Schrödinger equation for all $t \neq 0$. Therefore, a reasonable strategy to find a solution candidate, would be, once again, to first 'solve' the equation for $t < 0$ (before the shock) as well as for $t > 0$ (after the shock) and then use the equation and, if necessary, additional input from the underlying physical theory, to determine how these two wave functions should be glued together at $t = 0$. Whatever the details of such a procedure, the obtained solution candidate will have to be discontinuous at $t = 0$. As a consequence, testing this candidate by inserting it into the equation, we inevitably have to deal with an ill-defined product, that is, the product of the delta distribution $\delta(t)$ and a function which is discontinuous at $t = 0$.

As this elementary textbook example for the Schrödinger already indicates, products of distributions occur very frequently in physics. Physicists have long ago developed and probably always been using heuristic multiplications of distributions, often adapted to the physical problem under consideration. Some of these methods and the occasionally involved ad-hoc-assumptions can be justified and explained mathematically by invoking distribution theory and suitable concepts of weak solutions of differential equations, but some go beyond the theory of distributions. For instance, the latter is an inadequate and insufficient setting for problems involving singular products of the delta distribution with a discontinuous function or with another distribution.

Algebras of generalised functions as introduced by Colombeau in [Col84, Col85] constitute a framework in which singular problems arising from nonlinear operations on distributions can be dealt with on a mathematically rigorous basis. Although Colombeau theory has still not been widely adopted by the mathematical physics community, it is a flourishing field of research in mathematics. It has proven to be a valuable tool for studying PDEs with coefficients of low regularity, including examples such as (1.1), thereby not only securing paths already travelled by physicists on a heuristic level, but also contributing new mathematical aspects in various contexts. In particular, the theory provides solution concepts for PDEs with coefficients whose regularity is substantially below the regularity required in classical existence and uniqueness theories. Distributions can be embedded in Colombeau algebras by convolution with a mollifier function, satisfying certain moment conditions.

Singular equations have been used as model equations in physics for a long time, not only in quantum mechanics. Examples range from classical mechanics over electrodynamics to general relativity. Moreover, singular products of distributions are omnipresent in quantum field theory, rendering algebras of generalised functions in the sense of Colombeau a promising tool that might eventually contribute to the development of a mathematical foundation for physically relevant but still mathematically consistent quantum field theories (cf. [GOT00]). Dirac was a strong advocate for mathematical rigour in physical theories, as can be inferred from this quote:

> I want to emphasize the necessity for a sound mathematical basis for any fundamental physical theory. Any philosophical ideas that one may have play only a subordinate role. Unless such ideas have a mathematical basis they will be ineffective. [Dir78, p. 1]

Over the last three decades, numerous interesting applications of Colombeau theory have emerged, not the least in mathematical physics, including PDE models for acoustic and elastic or seismic waves in low regularity settings, for example in [CL88, Obe89, Col92, Obe92, HdH01, dHHO08, Hör11], as well as manifold contributions to general relativity in a non-smooth setting [CVW96, Bal97, Vic99, VW00, GKOS01, KSV05, SV06, GMS09, Han11, HKS12, Vic12]. Clearly, the theory has proven its strength and capability in various fields of modern analysis. However, in spite of the important results mentioned, its full scope with regard to applications in physics has not yet been exploited, as there are many more possible applications to consider and pursue.

1.2 Main results and outline

We briefly summarise the main results of the thesis:

- We provide three variants of existence and uniqueness results for generalised solutions of the Cauchy problem for first-order symmetric or Hermitian hyperbolic systems of PDEs (Theorems 2.7, 2.8, 2.13). These results were published in [HS12]. Furthermore, distributional limits of generalised solutions are obtained, if the coefficients are defined as regularisations of functions which are locally integrable in the time variable and of a certain Sobolev regularity in the spatial variables (Proposition 2.28). However, the required regularity assumptions can be substantially relaxed, even including the case of a delta distribution in the time variable, if additional conditions on the regularisation methods or on the structure of the coefficients are imposed (Propositions 2.32, 2.34, 2.36).

- We establish the transformation from a scalar second-order wave equation to a related symmetric hyperbolic first-order system in a generalised functions setting (Theorem 3.5), thereby obtaining existence and uniqueness results for second-order wave equations (Theorem 3.7). The required regularity assumptions are sufficiently general to cover wave equations derived from the Laplace-Beltrami operator of a metric of Geroch-Traschen class. These results were presented in [HHSS13]. We also discuss further examples and give a refined analysis of the necessary regularity assumptions.

- We present an explicit construction of Schwartz functions with prescribed moments and support contained in an arbitrarily chosen cone (Theorems 4.4, 4.12), providing Colombeau mollifiers with special convergence and support properties. In addition, our construction yields explicit solutions to several variants of the Stieltjes moment problem for rapidly decreasing smooth functions (Corollaries 4.7, 4.14). This immediately implies a density result for Schwartz functions with vanishing moments in spaces of Lebesgue integrable functions.

The outline of the remaining parts of the thesis is as follows:

Section 1.3 introduces and briefly reviews the required notions from Colombeau's theory of generalised functions and numbers, including basic concepts from linear algebra in a generalised setting. For convenience of the reader we also provide a proof of specific versions of Gronwall inequalities in Section 1.4, with emphasis on explicit expressions for all occurring constant. These inequalities are an essential tool in several proofs and will be used repeatedly to generate different variants of energy estimates.

In Chapter 2 we study existence, uniqueness, and distributional limits of generalised solutions of the Cauchy problem for first-order symmetric or Hermitian hyperbolic systems of partial differential equations with Colombeau generalised functions as coefficients and data. The proofs of solvability are based on energy estimates on lens-shaped domains with spacelike boundaries. We obtain several variants of existence and uniqueness results, including partial extensions of previous results in [Obe89, LO91, Hör04]. In addition, we investigate the regularity of generalised solutions in terms of Sobolev norms of their distributional limits, thereby providing aspects accompanying related work in [Obe09, GO14, GO15, Gar15]. We recover distributional limits of the generalised solution in situations where the regularity of the coefficients of the equation is below the usual regularity assumptions in corresponding classical theorems on distributional solutions. Under certain structural assumptions on the coefficient matrices, we obtain distributional limits of generalised solutions also for coefficient regularities which do not allow for a weak interpretation of the differential equation outside the framework of Colombeau generalised functions. As an application in physics, we discuss the Dirac equation with coefficients proportional to the delta distribution, modelling an impulsive electromagnetic field. Finally, we address the question of propagation of coefficient singularities by carrying out a case study in which we consider an explicitly solvable scalar first-order hyperbolic PDE. We calculate the distributional limit of the generalised solution and its singular support.

Chapter 3 is devoted to the transformation of a second-order hyperbolic PDE with generalised coefficients into a hyperbolic first-order system, using an explicit method that guarantees the symmetry of the system. The operations involved in this process are not defined classically for coefficients of low regularity. In particular, the low regularities covered by the existence and uniqueness results obtained in Chapter 2 do not even allow to extend these operations to a distribution theoretic setting. Within the framework of generalised functions, well-definedness of the transformation algorithm can be restored. In particular, we arrive at clear statements about the transfer of questions concerning solvability and uniqueness from wave equations to symmetric hyperbolic systems and vice versa. This enables us to reformulate our results on first-order systems for wave equations, thereby establishing unique solvability of the Cauchy problem for second-order wave equations with generalised coefficients. The admissible coefficient regularities are below those required in similar existence and uniqueness results for wave equations. As applications, we consider the Klein-Gordon equation with an impulsive coupling term and the wave equation derived from a Lorentzian metric of low regularity. In Section 3.3, we briefly discuss aspects of the Hadamard parametrix construction for second-order elliptic PDEs, based on [Hör07, Chapter XVII, Section 17.4].

In the final Chapter 4 we provide an explicit construction of Schwartz functions with prescribed moments and support bounded from below. The presented algorithm automatically yields an elementary and constructive proof of the Stieltjes moment problem for rapidly decreasing smooth functions, which has been discussed and solved in [Dur89, DE94, LS08]. The main result of this chapter is a similar construction for the multidimensional case, extending and refining the techniques used in the proof of the one-dimensional case. We show that any test function φ_0 can be approximated in Sobolev norms by a function $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with prescribed moments. Moreover, given an arbitrary open cone Γ in $\mathbb{R}^n$, it is possible to have $\operatorname{supp}(\varphi) \subseteq \operatorname{supp}(\varphi_0) \cup \Gamma$. Schwartz functions with vanishing moments are required as mollifiers for the embedding of distributions in the special Colombeau algebra. Our construction provides tailor-made mollifier functions with additional features such as satisfying certain support conditions or approximating a given test function in a particular norm. In addition, the method can be employed to create generalised mollifiers with special properties, as considered and used in [NPS98, OT98, OV08, SV09, Hör15].

1.3 Colombeau algebras of generalised functions

This section contains a short introduction to Colombeau's theory of generalised functions, focussing on aspects that will be of relevance in the following chapters. It also serves to fix some notational conventions.

In 1954, L. Schwartz proved the 'impossibility of the multiplication of distributions' [Sch54], an important result we want to briefly review. If we want to have an (associative and commutative) algebra $\mathcal{A}$ containing objects such as the delta distribution and equipped with a product extending the pointwise product of continuous functions, then $\mathcal{A}$ should have the following properties:

(i) There exists a linear embedding of the space of distributions $\mathcal{D}'$ into the algebra $\mathcal{A}$ such that the constant function with value 1 is the multiplicative unit in $\mathcal{A}$.

(ii) There exists a linear map $D : \mathcal{A} \to \mathcal{A}$, generalising the differential operator for continuously differentiable functions and satisfying the Leibniz rule.

(iii) If the generalised derivative operator D on $\mathcal{A}$ is restricted to the space $\mathcal{D}'$, then it coincides with the distributional derivative.

(iv) If we restrict the product in $\mathcal{A}$ to continuous functions, then it coincides with the usual pointwise product of functions. In other words, $\mathcal{A}$ contains the algebra $\mathcal{C}^0$ of all continuous functions as a subalgebra.

All of these requirements appear to be very natural and even necessary, if $\mathcal{A}$ ought to be of any practical value. Yet they are impossible to fulfil. More precisely, L. Schwartz showed that an algebra $\mathcal{A}$ satisfying (ii)-(iv) cannot contain the delta distribution. There are actually no distributions involved in the proof, it only uses the differential algebra property of $\mathcal{A}$ and the pointwise product of continuous functions. In fact, the usual differentiation rules are already not compatible with the pointwise multiplication of functions of finite differentiability. For instance, if H denotes the Heaviside function, then we clearly have $H^n = H$ in L^1_{loc} for all $n \in \mathbb{N}$ (since this is true pointwise almost everywhere). Differentiating this equation yields $H^{n-1}H' = \frac{1}{n}H'$ for arbitrary $n \in \mathbb{N}$, in particular $HH' = \frac{1}{2}H'$ and $H^2H' = \frac{1}{3}H'$. Upon the substitution $H^2 = H$ in the last equation we obtain $HH' = \frac{1}{3}H'$ and since $HH' = \frac{1}{2}H'$ as well, we end up with $\frac{1}{2}H' = \frac{1}{3}H'$, which is a contradiction for $H' \neq 0$. If associativity and commutativity of $\mathcal{A}$ are not to be given up, then the only way to avoid such contradictions is to replace (iv) by the weaker condition

(iv') $\mathcal{A}$ contains the algebra $\mathcal{C}^\infty$ of smooth functions as a subalgebra.

Differential algebras satisfying (i)-(iii) and (iv') exist and were introduced by J.-F. Colombeau [Col83, Col84, Col85, Col92], who called their elements 'new generalised functions'. For a comprehensive introduction to Colombeau algebras we also refer to [GKOS01].

Construction of the special Colombeau algebra

The underlying principle for constructing a differential algebra containing the distributions is regularisation by sequences (nets) of smooth functions. The sequential approach to distribution theory has already been developed around 1950 by Mikusinski [Mik48] and Temple [Tem53], but it was lacking a concept of multiplication compatible with differentiation in all circumstances.

Colombeau generalised functions are basically (equivalence classes of) nets of smooth functions indexed by a parameter $\varepsilon \in\,]0,1]$ and obeying certain asymptotic growth estimates as ε tends to zero. For such objects we shall write $(u_\varepsilon)_\varepsilon$ in place of $(u_\varepsilon)_{\varepsilon\in\,]0,1]}$ to simplify notation. Throughout this thesis we will use the 'special' (or 'simplified') version of the theory, since it is much more convenient in the context of differential equations compared to the 'full' Colombeau algebra as it is presented in [Col85]. Moreover, the 'special' Colombeau algebra possesses all essential features of the full theory, except allowing a canonical embedding of $\mathcal{D}'$.

Definition 1.1 (Moderate and negligible nets)**.** Let Ω be an open subset of $\mathbb{R}^n$. Set

$$\mathcal{E}(\Omega) := \mathcal{C}^\infty(\Omega)^{]0,1]} = \{(u_\varepsilon)_{\varepsilon\in\,]0,1]} \,|\, u_\varepsilon \in \mathcal{C}^\infty(\Omega)\}.$$

Then the elements of

$$\mathcal{E}_M(\Omega) := \{(u_\varepsilon)_\varepsilon \in \mathcal{E}(\Omega) \,|\, \forall K \subset\subset \Omega\; \forall \alpha \in \mathbb{N}_0^n\; \exists N \in \mathbb{N}_0 : \\ \|\partial^\alpha u_\varepsilon\|_{L^\infty(K)} = O(\varepsilon^{-N}) \text{ as } \varepsilon \to 0\}$$

are called *moderate* and those of

$$\mathcal{N}(\Omega) := \{(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(\Omega) \,|\, \forall K \subset\subset \Omega \,\forall \alpha \in \mathbb{N}_0^n \,\forall m \in \mathbb{N}_0 : \\ \|\partial^\alpha u_\varepsilon\|_{L^\infty(K)} = O(\varepsilon^m) \text{ as } \varepsilon \to 0\}$$

are called *negligible*.

Defining vector space operations, partial derivatives, and multiplication componentwise, the space $\mathcal{E}_M(\Omega)$ of moderate nets on Ω is an associative and commutative differential algebra with $\mathcal{N}(\Omega)$ a differential ideal in it.

Definition 1.2 (Colombeau generalised functions)**.** The Colombeau algebra of generalised functions on Ω is defined as the quotient space

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

A generalised function $u \in \mathcal{G}(\Omega)$ is written as

$$u = [(u_\varepsilon)_\varepsilon] = (u_\varepsilon)_\varepsilon + \mathcal{N}(\Omega),$$

where the functions u_ε are assumed to be complex-valued, if not stated otherwise.

A function $f \in \mathcal{C}^\infty(\Omega)$ is canonically embedded into $\mathcal{G}(\Omega)$ via the trivial assignment $f \mapsto [(f)_\varepsilon]$. A compactly supported distribution $w \in \mathcal{E}'(\Omega)$ can be embedded by regularisation via convolution with a net $(\rho_\varepsilon)_\varepsilon$ where

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon^n}\rho\left(\frac{x}{\varepsilon}\right)$$

and the mollifier function $\rho \in \mathcal{S}(\mathbb{R}^n)$ satisfies

$$\int\limits_{\mathbb{R}^n} \rho(x)\mathrm{d}x = 1, \qquad \int\limits_{\mathbb{R}^n} x^\alpha \rho(x)\mathrm{d}x = 0 \quad \text{for all } \alpha \in \mathbb{N}_0^n,\ |\alpha| > 0.$$

Extending w to $\mathbb{R}^n$ by setting $w = 0$ on $\mathbb{R}^n \setminus \Omega$, the map $w \mapsto [((w * \rho_\varepsilon)|_\Omega)_\varepsilon]$ defines a linear embedding of $\mathcal{E}'(\Omega)$ into $\mathcal{G}(\Omega)$ that coincides with the canonical embedding of smooth functions, if restricted to $\mathcal{D}(\Omega)$. It is exactly this compatibility with the trivial embedding of smooth functions that makes it necessary to take a mollifier with all moments of order $|\alpha| > 0$ vanishing. Distributions in $\mathcal{D}'(\Omega)$ can be embedded with the help of an open covering $(\Omega_\lambda)_{\lambda\in\Lambda}$ of Ω (all Ω_λ relatively compact) and a subordinated partition of unity by smooth cut-off functions. The so-obtained embedding $\iota : \mathcal{D}'(\Omega) \hookrightarrow \mathcal{G}(\Omega)$ does not depend on the particular choice of $(\Omega_\lambda)_{\lambda\in\Lambda}$ and is consistent with the previously defined embeddings of $\mathcal{E}'(\Omega)$ and $\mathcal{C}^\infty(\Omega)$. However, this construction is mainly of theoretical value, but not very practicable to actually compute $\iota(w)$ for a given $w \in \mathcal{D}'(\Omega)$. At least, a direct calculation of $\iota(w)$ is possible for all distributions on Ω that admit an extension as a tempered distribution to all of $\mathbb{R}^n$. Obviously, the embedding considered here depends on the chosen mollifier. This is remedied by the construction of the full

Colombeau algebra where one uses the set of all admissible mollifiers rather than picking a particular one. Yet, this comes at the cost of more technicalities and a less transparent notation.

Colombeau algebras contain the distributions as a linear subspace and enable us to give a meaning to ill-defined or ambiguous products as, for instance, $H\delta$ or δ^2. The ill-definedness on the distributional level is rectified in $\mathcal{G}(\Omega)$, since in contrast to distributions, Colombeau generalised functions representing H or δ contain precise information on the microscopic structure of their singularities. This information is lost in the distributional limit, but it is essential when we want to multiply distributions. Representing the delta distribution by a generalised function always means to choose a certain regularisation, thereby endowing this model of the delta distribution with a 'history' of how its singularity formed or where it evolved from.

The concept of association

We say that a Colombeau function u is *associated with a distribution* $u_0 \in \mathcal{D}'(\Omega)$ if some (and hence every) representative $(u_\varepsilon)_\varepsilon$ converges to u_0 in $\mathcal{D}'(\Omega)$, that is, if

$$\forall \varphi \in \mathcal{D}(\Omega): \quad \lim_{\varepsilon \to 0} \int_\Omega u_\varepsilon(x)\varphi(x)\mathrm{d}x = \langle u_0, \varphi \rangle.$$

The distribution u_0 represents the macroscopic behaviour of u and is called the *distributional shadow* of u. Apparently, different Colombeau functions can have the same distributional shadow. For example, taking an arbitrary function $\rho \in \mathcal{D}(\mathbb{R})$ with $\int_{-\infty}^{\infty} \rho(x)\mathrm{d}x = 1$ and a net $(\sigma_\varepsilon)_\varepsilon$ with $\sigma_\varepsilon \to 0$ and $\sigma_\varepsilon^{-1} = O(\varepsilon^{-m})$ as $\varepsilon \to 0$ for some $m \in \mathbb{N}$, the generalised function $[(\sigma_\varepsilon^{-1}\rho(\sigma_\varepsilon^{-1}\,\cdot\,))_\varepsilon]$ has the delta distribution as its shadow.

A *delta net* is any net $(\rho_\varepsilon)_\varepsilon$ that converges to the delta distribution. Since there are many ways to create a delta net, there are obviously also many generalised functions in $\mathcal{G}(\Omega)$ representing the delta distribution. They differ in their 'microstructure' and behave differently under nonlinear operations. To illustrate this, the distributional limit (if it exists) of a product in $\mathcal{G}(\mathbb{R})$ corresponding to $H\delta$, H^2, or δ^2 will strongly depend on the very regularisations of H and δ that we use.

Any delta net $(\varphi_\varepsilon)_\varepsilon$ for which $\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi(\frac{x}{\varepsilon})$ is called a *model delta net*, whereas a *strict delta net* $(\psi_\varepsilon)_\varepsilon$, $\psi_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$, satisfies the following conditions (cf. [GKOS01, Section 1.1]):

(i) $\mathrm{supp}(\psi_\varepsilon) \to 0$ as $\varepsilon \to 0$,

(ii) $\lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \psi_\varepsilon(x)\mathrm{d}x = 1$,

(iii) $\int_{\mathbb{R}^n} |\psi_\varepsilon(x)|\mathrm{d}x$ is bounded uniformly in ε.

Association defines an equivalence relation on $\mathcal{G}(\Omega)$, where all representatives of an equivalence class share the same distributional shadow. But not every element of a Colombeau algebra is associated with a distribution. For example, the generalised function $[(\frac{1}{\varepsilon})_\varepsilon]$ has no distributional shadow. Association is stable under linear operations, in particular differentiation. It is thus an important concept when studying differential equations in $\mathcal{G}(\Omega)$. For instance, when modeling a differential equation with distributional coefficients in $\mathcal{G}(\Omega)$, the generalised coefficients must be associated with the original distributional coefficients, but there is some freedom in choosing a special regularisation satisfying certain additional properties. Most significantly, if a generalised solution is associated with a distribution of a certain regularity, the latter can directly be compared with a weak solution obtained from other solution concepts for differential equations with non-smooth coefficients. Different solution concepts for first-order hyperbolic differential equations with coefficients of regularity below Lipschitz continuity are discussed in [HH08]. Among these, the framework of generalised functions can deal with the lowest coefficient regularities and provides unique generalised solutions even for coefficient regularities substantially below those required within other solution concepts. If, in this situation, the generalised solution admits a distributional limit that is independent of the particular regularisations of the coefficients, the transfer to and from $\mathcal{G}(\Omega)$ can be viewed as a technique to extend classical results on weak solutions. However, even if the generalised solution is not associated with a distribution, it may still be of physical relevance, since the solution of a differential equation is not directly observable anyway. Only certain quantities derived or computed from the solution can be measured, and predictions for the outcomes of measurements may be also be derived from generalised solutions.

Variants of Colombeau algebras

The special Colombeau algebra introduced in Definition 1.1 is based on the space $\mathcal{C}^\infty(\Omega)$. However, analogous constructions are possible for any locally convex topological vector space [Gar05b].

Definition 1.3 (Generalised functions based on a locally convex vector space)**.** Let E be a locally convex topological vector space whose topology is given by the family of semi-norms $\{p_j\}_{j\in J}$. The elements of

$$\mathcal{M}_E := \{(u_\varepsilon)_\varepsilon \in E^{\,]0,1]} \mid \forall j \in J \ \exists N \in \mathbb{N}_0 : \ p_j(u_\varepsilon) = O(\varepsilon^{-N}) \ \text{ as } \varepsilon \to 0\}$$

and

$$\mathcal{N}_E := \{(u_\varepsilon)_\varepsilon \in E^{\,]0,1]} \mid \forall j \in J \ \forall m \in \mathbb{N}_0 : \ p_j(u_\varepsilon) = O(\varepsilon^{m}) \ \text{ as } \varepsilon \to 0\}$$

are called *E-moderate* and *E-negligible*, respectively. The *generalised functions based on E* are then defined as the quotient $\mathcal{G}_E := \mathcal{M}_E/\mathcal{N}_E$. If E is a differential algebra, then $\mathcal{N}_E$ is an ideal in $\mathcal{M}_E$ and $\mathcal{G}_E$ is a differential algebra as well, called the Colombeau algebra based on E.

In this thesis we will also employ variants of Colombeau algebras with other basic spaces than $\mathcal{C}^\infty(\Omega)$. Specifically, we will frequently use the Sobolev spaces $E =$

$H^\infty(\Omega) = \{h \in \mathcal{C}^\infty(\overline{\Omega}) \,|\, \forall\alpha \in \mathbb{N}_0^n : \partial^\alpha h \in L^2(\Omega)\}$ with the family of semi-norms

$$\|h\|_{H^k} = \Big(\sum_{|\alpha|\le k} \|\partial^\alpha h\|_{L^2}^2\Big)^{1/2} \quad (k \in \mathbb{N}_0),$$

as well as $E = W^{\infty,\infty}(\Omega) = \{h \in \mathcal{C}^\infty(\overline{\Omega}) \,|\, \forall\alpha \in \mathbb{N}_0^n : \partial^\alpha h \in L^\infty(\Omega)\}$ with the family of semi-norms

$$\|h\|_{W^{k,\infty}} = \max_{|\alpha|\le k} \|\partial^\alpha h\|_{L^\infty} \quad (k \in \mathbb{N}_0),$$

and $E = \mathcal{C}^\infty(\bar{I} \times \mathbb{R}^n)$, where I is an open interval, equipped with semi-norms

$$\|h\|_{m,K} = \max_{|\alpha|\le m} \|\partial^\alpha h\|_{L^\infty(\bar{I}\times K)} \quad (m \in \mathbb{N}_0),$$

where $K \subset\subset \mathbb{R}^n$. To simplify notation and avoid overloaded subscripts, we use notations as in [Hör11] and denote

$$\mathcal{G}_{L^2}(\Omega) := \mathcal{G}_{H^\infty(\Omega)}, \quad \mathcal{G}_{L^\infty}(\Omega) := \mathcal{G}_{W^{\infty,\infty}(\Omega)} \quad \text{and} \quad \mathcal{G}(\bar{I} \times \mathbb{R}^n) := \mathcal{G}_{\mathcal{C}^\infty(\bar{I}\times\mathbb{R}^n)}.$$

Regularity theory in Colombeau algebras

In [Obe92], the subalgebra $\mathcal{G}^\infty(\Omega)$ of *regular generalised functions* in $\mathcal{G}(\Omega)$ was introduced to develop an intrinsic regularity theory in $\mathcal{G}(\Omega)$. The subalgebra $\mathcal{G}_E^\infty$ of a Colombeau algebra $\mathcal{G}_E$ is obtained by demanding that the inverse ε-power N in the moderateness estimates (see Definition 1.1) can be chosen uniformly over all derivatives (cf. [Obe92, Definition 25.1, Chapter VII]). For instance, an element $u = [(u_\varepsilon)_\varepsilon] \in \mathcal{G}(\Omega)$ belongs to $\mathcal{G}^\infty(\Omega)$ if and only if

$$\forall K \subset\subset \Omega \;\exists N \in \mathbb{N}_0 \;\forall \alpha \in \mathbb{N}_0^n : \; \|\partial^\alpha u_\varepsilon\|_{L^\infty(K)} = O(\varepsilon^{-N}) \quad \text{as} \quad \varepsilon \to 0.$$

Conversely, the derivatives of a generalised function which is not $\mathcal{G}^\infty$-regular must become more and more singular in terms of inverse ε-powers as the order of the derivatives increases. For instance, taking a function $\rho \in \mathcal{D}(\mathbb{R})$ with $\int_{-\infty}^\infty \rho(x)\mathrm{d}x = 1$ and considering the model delta net $(\rho_\varepsilon)_\varepsilon = \left(\frac{1}{\varepsilon}\rho\left(\frac{\cdot}{\varepsilon}\right)\right)_\varepsilon$, its n-th derivative will satisfy $\|\rho_\varepsilon^{(n)}\|_{L^\infty(\mathbb{R})} = O(\varepsilon^{-(n+1)})$ as $\varepsilon \to 0$, but will not behave better.

The subalgebra $\mathcal{G}^\infty(\Omega)$ plays the same role within $\mathcal{G}(\Omega)$ as $\mathcal{C}^\infty(\Omega)$ does within $\mathcal{D}'(\Omega)$ and one obtains the important compatibility relation (cf. [Obe92, Theorem 25.2, Chapter VII])

$$\mathcal{G}^\infty(\Omega) \cap \mathcal{D}'(\Omega) = \mathcal{C}^\infty(\Omega).$$

Here, $\mathcal{D}'(\Omega)$ and $\mathcal{C}^\infty(\Omega)$ are identified with the corresponding subspace $\iota(\mathcal{D}'(\Omega))$ and $\iota(\mathcal{C}^\infty(\Omega))$ of $\mathcal{G}(\Omega)$. Note that not every generalised function belonging to $\mathcal{G}^\infty(\Omega)$ is associated with a smooth function. For instance, with $\rho \in \mathcal{D}(\mathbb{R})$ as before, the class $\left[\left(\log\frac{1}{\varepsilon}\rho\left(\log\frac{1}{\varepsilon}\,\cdot\right)\right)_\varepsilon\right]$ is both an element of $\mathcal{G}^\infty(\Omega)$ as well as associated with the delta distribution.

A generalised function $u \in \mathcal{G}(\Omega)$ is called *of L^∞-type* if it has a representative $(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(\Omega)$ such that

$$\|u_\varepsilon\|_{L^\infty(\Omega)} = O(\varepsilon^{-m})$$

as $\varepsilon \to 0$. It is called *locally of logarithmic growth* or *locally of log-type*, if it has a representative $(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(\Omega)$ such that for all $K \subset\subset \Omega$,

$$\|u_\varepsilon\|_{L^\infty(K)} = O(\log(1/\varepsilon))$$

as $\varepsilon \to 0$ (cf. [Obe88, Definition 1.1]). It is called *of L^∞-log-type* if

$$\|u_\varepsilon\|_{L^\infty(\Omega)} = O(\log(1/\varepsilon))$$

as $\varepsilon \to 0$ (cf. [GKOS01, Definition 1.5.1]). Moreover, we call $u \in \mathcal{G}(\Omega)$ *locally uniformly bounded* if it has a representative $(u_\varepsilon)_\varepsilon \in \mathcal{E}_M(\Omega)$ such that

$$\forall K \subset\subset \Omega \ \exists C > 0 \ \exists \varepsilon_0 : \|u_\varepsilon\|_{L^\infty(K)} \leq C \quad \forall \varepsilon < \varepsilon_0.$$

Generalised numbers and point values of generalised functions

Although a concept of point values for Schwartz distributions had been developed prior to Colombeau algebras [Loj57], an arbitrary distribution need not have a point value in this sense at every point. For instance, the Dirac delta distribution cannot be assigned a point value at 0. However, point values of a Colombeau generalised function arise naturally by evaluating its representatives, yielding nets of complex numbers. Thus, point values of Colombeau generalised functions are not elements in $\mathbb{C}$ (or $\mathbb{R}$), but represent *generalised* numbers. For $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, the ring of generalised numbers, $\widetilde{\mathbb{K}}$, is defined analogous to the algebra of Colombeau generalised functions as the quotient ring of moderate nets of complex numbers modulo negligible nets, that is,

$$\begin{aligned}
\mathcal{E}_M &:= \{(r_\varepsilon)_\varepsilon \in \mathbb{K}^{]0,1]} \mid \exists N \in \mathbb{N}_0 : |r_\varepsilon| = O(\varepsilon^{-N}) \text{ as } \varepsilon \to 0\} \\
\mathcal{N} &:= \{(r_\varepsilon)_\varepsilon \in \mathbb{K}^{]0,1]} \mid \forall m \in \mathbb{N}_0 : |r_\varepsilon| = O(\varepsilon^{m}) \text{ as } \varepsilon \to 0\} \\
\widetilde{\mathbb{K}} &:= \mathcal{E}_M / \mathcal{N},
\end{aligned}$$

and all operations defined componentwise on representatives. Denoting the class of a net $(r_\varepsilon)_\varepsilon$ in $\widetilde{\mathbb{K}}$ by $[(r_\varepsilon)_\varepsilon]$, the generalised number $[(\sin(1/\varepsilon))_\varepsilon]$ is an example for a zero divisor in $\widetilde{\mathbb{K}}$. With

$$c_\varepsilon := \begin{cases} 1 & \text{if} \quad \varepsilon = \frac{1}{k\pi} \text{ for some } k \in \mathbb{N} \\ 0 & \text{else,} \end{cases}$$

the product $[(\sin(1/\varepsilon))_\varepsilon][(c_\varepsilon)_\varepsilon]$ vanishes in $\widetilde{\mathbb{K}}$. Thus $\mathbb{K}$ is not a field.

The basis for studying algebraic properties of $\widetilde{\mathbb{K}}$ was laid in [AJ01], further analysis of the ring of generalised numbers in different contexts was carried out in [GKOS01, May08, AJOS08, AFJ09, Ver10, BK12, HKK14]. For a thorough compilation of properties of $\widetilde{\mathbb{K}}$ we refer to [HKK14]. Here, we only gather some basic notions required for our investigations.

For any $r \in \widetilde{\mathbb{K}}$, the following statements are equivalent (cf. [AJ01, Theorem 2.18] or [GKOS01, Theorem 1.2.39]):

(i) r is non-invertible.

(ii) there exists a representative $(r_\varepsilon)_\varepsilon$ and a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ with $\varepsilon_k \to 0$ as $k \to \infty$ and $r_{\varepsilon_k} = 0$ for all $k \in \mathbb{N}$.

(iii) r is a zero divisor.

On the converse, r is invertible if and only if there exists some $m \in \mathbb{N}$ such that $|r_\varepsilon| \geq \varepsilon^m$ for ε sufficiently small. Then r is called *strictly nonzero*. Furthermore, an element of $r \in \widetilde{\mathbb{R}}$ is called *strictly positive* if there exists some $m \in \mathbb{N}$ such that $r_\varepsilon \geq \varepsilon^m$ for small ε.

A generalised function is not uniquely determined by prescribing its point values. However, a function $u \in \mathcal{G}(\Omega)$ is characterised by its generalised point values, a concept introduced in [OK99]. For any open subset $\Omega \subseteq \mathbb{R}^n$, the corresponding set of generalised points is defined as $\widetilde{\Omega} := \Omega_M / \sim$, with

$$\Omega_M := \{(x_\varepsilon)_\varepsilon \in \Omega^{]0,1]} \,|\, \exists p > 0 \; \exists e_0 > 0 : |x_\varepsilon| \leq \varepsilon^{-p} \; (0 < \varepsilon < \varepsilon_0)\},$$

where $|x|$ is the Euclidean norm of a vector $x \in \mathbb{R}^n$, and

$$(x_\varepsilon)_\varepsilon \sim (y_\varepsilon)_\varepsilon \quad :\Longleftrightarrow \quad \forall q > 0 \; \exists \varepsilon_0 > 0 : \; |x_\varepsilon - y_\varepsilon| \leq \varepsilon^q \; (0 < \varepsilon < \varepsilon_0).$$

Then $u = 0$ in $\mathcal{G}(\Omega)$ if and only if $u(\tilde{x}) = 0$ in $\widetilde{\mathbb{C}}$ for each $\tilde{x} \in \widetilde{\Omega}_c$, where $\widetilde{\Omega}_c$ denotes the set of compactly supported generalised points,

$$\widetilde{\Omega}_c = \{\tilde{x} \in \widetilde{\Omega} \,|\, \exists \text{ representative } (x_\varepsilon)_\varepsilon \; \exists K \subset\subset \Omega \; \exists \eta > 0 : \; x_\varepsilon \in K \; (0 < \varepsilon < \eta)\}.$$

Subsequently, characterisations of generalised functions by their generalised point values were given for Colombeau algebras on manifolds in [KS02a] and for diffeomorphism invariant full Colombeau algebras in [Nig13].

Matrices and linear algebra in Colombeau algebras

Since $\widetilde{\mathbb{K}}$ is not a field, the translation of standard concepts from linear algebra to a generalised numbers setting deserves special attention. In this section we collect some basic notions and results for matrices over $\widetilde{\mathbb{K}}$ that will be important for our considerations.

We denote the set of square matrices of size n with entries in $\widetilde{\mathbb{K}}$ by $M_n(\widetilde{\mathbb{K}})$. Then, for any $A \in M_n(\widetilde{\mathbb{K}})$, the following assertions are equivalent (cf. [GKOS01, Lemma 1.2.41]):

(i) A is non-degenerate, that is, $\forall x \in \widetilde{\mathbb{K}}^n : \; x^{\mathsf{T}} A y = 0 \; \forall y \in \widetilde{\mathbb{K}}^n \Rightarrow x = 0$.

(ii) A is injective as a linear operator on $\widetilde{\mathbb{K}}^n$.

(iii) A is bijective as a linear operator on $\widetilde{\mathbb{K}}^n$.

(iv) $\det(A)$ is invertible.

Symmetric matrices over $\widetilde{\mathbb{R}}$ were studied in [May08], mainly with regard to applications in low-regularity Riemannian and Lorentzian geometry. Linear algebra over the ring of generalised numbers has been re-investigated and further developed in [HKK14], including revised notions of eigenvalues and eigenvectors.

Adopting the definitions given in [HKK14], we call a generalised number λ an eigenvalue of $A \in M_n(\widetilde{\mathbb{C}})$, if there exists a vector $x \in \widetilde{\mathbb{C}}^n$ such that $Ax = \lambda x$ and $\langle x, x\rangle$ is strictly positive. Here, $\langle \cdot\, , \cdot \rangle$ is the standard unitary inner product. Note that this condition is stronger than the non-invertibility of $\det(A - \lambda I_n)$, where I_n denotes the n-dimensional identity matrix. The vector x is then called an eigenvector for A associated with λ. Furthermore, we call $A \in M_n(\widetilde{\mathbb{C}})$ Hermitian, if it is equal to its Hermitian transpose, that is, if $A^* = A$. Under these premises, Hermitian matrices in $M_n(\widetilde{\mathbb{C}})$ possess the usual properties. More precisely, if $A \in M_n(\widetilde{\mathbb{C}})$ is Hermitian and $\lambda \in \widetilde{\mathbb{C}}$ is an eigenvalue of A, then $\bar{\lambda} = \lambda$, that is, $\lambda \in \widetilde{\mathbb{R}}$ (cf. [HKK14, Proposition 3.18]). In addition, by [HKK14, Proposition 3.18], there exists a unitary matrix $U \in M_n(\widetilde{\mathbb{C}})$ such that $U^*AU = \operatorname{diag}(\lambda_1, ..., \lambda_n)$ with $\lambda_j \in \widetilde{\mathbb{R}}$. The matrix A is called positive definite, if all its eigenvalues are strictly positive (cf. [May08, Definition 4.8]).

1.4 Gronwall inequalities

Gronwall's lemma and similar estimates play a central role in results on existence and uniqueness of solutions of PDEs. For convenience of the reader we provide a proof of specific versions of Gronwall inequalities, based on [DL92, Chapter XVIII, §5, Section 2.2]. Here we focus on explicit expressions for the involved constants. For a comprehensive collection of comparable estimates for differential and integral equations of various types we refer to [Pac98].

Lemma 1.4 (Gronwall estimates)**.** *Let u , α, and β be real-valued and continuous functions on the interval* $[0, T]$*. We assume*

$$u(t) \leq \alpha(t) + \int_0^t \beta(s)u(s)\mathrm{d}s \qquad \text{for all } t \in [0, T].$$

Then we have the following inequalities:

(i) *If β is non-negative, then*

$$u(t) \leq \alpha(t) + \int_0^t \alpha(s)\beta(s)\, e^{\int_s^t \beta(r)\mathrm{d}r}\, \mathrm{d}s \qquad \text{for all } t \in [0, T].$$

(ii) *If β is non-negative and α is constant, then*

$$u(t) \leq \alpha\, e^{\int_0^t \beta(s)\mathrm{d}s} \qquad \text{for all } t \in [0, T].$$

(iii) *If β is a positive constant and α is absolutely continuous, then*

$$u(t) \le e^{\beta t}\left(\alpha(0) + \int\limits_0^t e^{-\beta s}\alpha'(s)\mathrm{d}s\right) \qquad \text{for all } t \in [0,T].$$

Proof. We set

$$v(s) := e^{-\int\limits_0^s \beta(r)\mathrm{d}r} \int\limits_0^s \beta(r)u(r)\mathrm{d}r$$

and obtain

$$\begin{aligned} v'(s) = -\beta(s)e^{-\int\limits_0^s \beta(r)\mathrm{d}r} \int\limits_0^s \beta(r)u(r)\mathrm{d}r + e^{-\int\limits_0^s \beta(r)\mathrm{d}r}\beta(s)u(s) \\ = \Big(u(s) - \int\limits_0^s \beta(r)u(r)\mathrm{d}r\Big)\beta(s)e^{-\int\limits_0^s \beta(r)\mathrm{d}r} \le \alpha(s)\beta(s)e^{-\int\limits_0^s \beta(r)\mathrm{d}r} \end{aligned}$$

since β is non-negative. Integration leads to

$$v(t) \le \int\limits_0^t \alpha(s)\beta(s)e^{-\int\limits_0^s \beta(r)\mathrm{d}r}\mathrm{d}s. \tag{1.2}$$

Rewriting the assumed inequality for $u(t)$ as follows,

$$u(t) \le \alpha(t) + \int\limits_0^t \beta(s)u(s)\mathrm{d}s = \alpha(t) + e^{\int\limits_0^t \beta(r)\mathrm{d}r}v(t), \tag{1.3}$$

and using (1.2), we get

$$u(t) \le \alpha(t) + e^{\int\limits_0^t \beta(r)\mathrm{d}r}\int\limits_0^t \alpha(s)\beta(s)e^{-\int\limits_0^s \beta(r)\mathrm{d}r}\mathrm{d}s = \int\limits_a^t \alpha(s)\beta(s)e^{\int\limits_s^t \beta(r)\mathrm{d}r}\mathrm{d}s,$$

which proves assertion (i). If α is constant, we find from (1.2) that

$$v(t) \le \alpha\int\limits_0^t \beta(s)e^{-\int\limits_0^s \beta(r)\mathrm{d}r}\mathrm{d}s = \alpha\left(-e^{-\int\limits_0^t \beta(r)dr} + 1\right).$$

In combination with (1.3), this yields the desired inequality (ii). Turning to (iii), integration by parts in (i) yields

$$u(t) \le \alpha(t) + e^{\beta t}\left(-\alpha(s)\,e^{-\beta s}\Big|_{s=0}^{s=t} + \int\limits_0^t \alpha'(s)\,e^{-\beta s}\mathrm{d}s\right) \le e^{\beta t}\left(\alpha(0) + \int\limits_0^t e^{-\beta s}\alpha'(s)\mathrm{d}s\right),$$

since α' exists and is integrable in the Lebesgue sense. □

Chapter 2

Symmetric hyperbolic systems in algebras of generalised functions

The motivation for the work presented in this chapter originated from the study of a physical model describing the scattering of neutrons by an ultra-short laser pulse, that is, the collision of a neutron and an impulsive electromagnetic wave whose time duration is of the order of femtoseconds or less. A neutron is not affected by electric fields, but it is susceptible to the magnetic field of the laser pulse due to its magnetic moment. The Dirac equation, a hyperbolic first-order system of PDEs, is used to model the interaction of a neutron with the laser pulse, the latter being represented by a coupling term proportional to the delta distribution, corresponding to a high concentration of energy within a very short time or space interval. The solution of the corresponding initial value problem then represents the quantum mechanical wave function of a neutron colliding with the pulse. Because of the delta distribution appearing in the coefficients of the model equation and, as a consequence, the occurrence of ill-defined products of distributions when physically reasonable solution candidates are inserted in the equation, classical concepts of weak solutions are not applicable in this setting. The nonlinear theory of generalised functions in the sense of Colombeau provides solution concepts even for PDEs with coefficients of very low regularity, including coefficients proportional to the delta distribution.

In this chapter we establish existence and uniqueness of a generalised solution of the hyperbolic Cauchy problem

$$\begin{aligned} \partial_t u + \sum_{j=1}^{n} A^j \partial_{x_j} u + Bu &= f \qquad \text{on }]0,T[\times\mathbb{R}^n, && (2.1)\\ u|_{t=0} &= g, && (2.2) \end{aligned}$$

where u, f, and g are vectors of length m and A^j and B are $m \times m$-matrices. All coefficients and data are Colombeau generalised functions and may therefore represent functions or distributions of low regularity.

Apart from the Dirac equation, which is a special case of a hyperbolic first-order system, problem (2.1–2.2) includes PDE models for the propagation of electromagnetic, acoustic, and elastic waves. In these PDEs, non-smooth coefficients typically arise when highly heterogeneous media are considered. For example, at the boundary layer between two different material compounds, physical properties of the medium determining the velocity, direction, or attenuation of propagating waves may be discontinuous. In general, this will also lead to discontinuous coefficients in the model equations, but may even result in regularities below L^1_{loc}, if derivatives of the position-dependent material parameters appear the coefficients.

The content of this chapter is mainly based on [HS12], supplemented by additional results on distributional limits in Section 2.3 and a case study on the propagation of singularities in Section 2.4. The main focus of our investigations as well as the essential methods are following up along the lines of the seminal papers [Obe88, Obe89, LO91].

Initial value problems of the form (2.1–2.2) with generalised functions as coefficients and data have also been studied in [CO90, KH01, HdH01, Hör04, Kmi06, Kmi07, Obe09, GO14, GO15, Gar15], to which we refer the reader for closely related research and further mathematical aspects in the context of generalised functions. Second-order wave equations in a similar mathematical setting have been discussed in [Gra94, VW00, GMS09, Han11, HKS12, DHO13, GR15].

In addition to existence and uniqueness of generalised solutions to the Cauchy problem, we are interested in existence and regularity of distributional limits of generalised solutions. If the coefficients represent regularisations of non-smooth functions or distributions whose regularity properties allow to interpret (2.1–2.2) as a Cauchy problem within a classical weak or distributional solution concept, then the classical solution can be compared with the distributional shadow of the unique Colombeau solution. The analysis of the relation between generalised and classical weak solutions, as well as several convergence results in this context, are going back to investigations in [Obe88, Obe89, LO91, HdH01, NOP05]. More recently, convergence properties of generalised solutions to the initial value problem have been studied in [NO12] for ordinary differential equations, and in [DHO13, GR15, Gar15] for hyperbolic PDEs. Different solutions concepts for linear first-order hyperbolic differential equations with non-smooth coefficients have been discussed and compared in [HH08].

The outline of this chapter is as follows. First we devote a section to the details of the construction of lens-shaped domains and to energy estimates on these domains with explicit expressions for the involved constants. The energy estimates derived in Section 2.1 are essential ingredients in proving three variants of existence and uniqueness results in Section 2.2. We obtain generalised solutions for coefficients and data with substantially weaker regularities than those typically required in classical existence and uniqueness results. Section 2.3 investigates regularity properties of generalised solutions, both in terms of intrinsic regularity in the Colombeau algebra and in terms of the regularity of their associated distributional shadows. The distributional limit of the unique generalised solution of the Cauchy problem is shown to be consistent with classical results on weak solutions for non-smooth coefficients and data. More precisely, if coefficients and data of the required regularity are regularised to obtain corresponding generalised objects, then the distributional limit of the generalised solution coincides with the classical weak solution. However, the unique generalised solution may possess a distributional limit even in circumstances in which classical results on weak solutions are not applicable, for instance when the involved regularities do not prevent the occurrence of products of the delta distribution with a discontinuous function. We present different variants of results on distributional shadows of generalised solutions, aiming at an exploration of this range of regularity constellations. If the distributional limit of the unique generalised solution only depends on the macroscopic aspects of the coefficients and data, it may be interpreted as a (classical) weak solution, obtained by transferring the problem to a generalised functions setting, thereby eliminating the ill-definedness of the original PDE. In Section 2.4, we consider scalar first-order hyperbolic PDEs and discuss an explicit example as a case study for the propagation of coefficient singularities, that is, singularities propagating along characteristic curves, but originating from singularities in the coefficients and not from initial data singularities.

2.1 Standard lenses and basic energy estimates

Central elements of our proofs of unique solvability of the Cauchy problem will be L^2-estimates, carried out on lens-shaped subsets of the domain $[0,T] \times \mathbb{R}^n$. Similar constructions or domains have been used, for example, in [BGS07, Part I, Section 2.2], in [HE73, Section 4.3], in [Tay11, Chapter 6, Section 5], and in [Fri75, Section 4.4]. Since we are working in a generalised functions setting, it is essential to have precise information on all dependencies of various constants involved in these estimates. For this reason we devote this section to the construction of a special variant of lens-shaped domains and to some basic estimates for these types of lenses.

Definition 2.1 (Standard lenses)**.** A standard lens $\mathcal{L} \subseteq \mathbb{R}^{1+n}$ of thickness $T > 0$, inner radius $R_1 > 0$ and outer radius $R_2 > R_1$ is the image set of the map

$$\psi : [0,1] \times B_{R_2} \quad \to \quad [0,T] \times \mathbb{R}^n, \quad \psi(\Theta, y) = \begin{cases} \big(\Theta T, y\big)^{\mathsf{T}} & \text{for} \quad |y| \leq R_1 \\ \big(\Theta T \, \frac{R_2 - |y|}{R_2 - R_1}, y\big)^{\mathsf{T}} & \text{for} \quad |y| > R_1 \end{cases}$$

where B_R denotes the closed ball of radius R_2, centred at the origin. We introduce sections or slices of a lens,

$$\mathcal{H}_\Theta := \psi(\Theta, B_{R_2}), \quad \text{where} \quad \Theta \in [0,1],$$

as well as partial lenses of $\mathcal{L}$, defined as

$$\mathcal{L}_\Theta = \bigcup_{0 \leq \tau \leq \Theta} \mathcal{H}_\tau \quad \text{for} \quad \Theta \in]0,1].$$

Partial lenses $\mathcal{L}_\Theta$ are compact convex subsets of $[0,T] \times \mathbb{R}^n$ with Lipschitz continuous boundary $\partial \mathcal{L}_\Theta = \mathcal{H}_0 \cup \mathcal{H}_\Theta$ (see Figure 2.1). Note that $\mathcal{L}(1) = \mathcal{L}$.

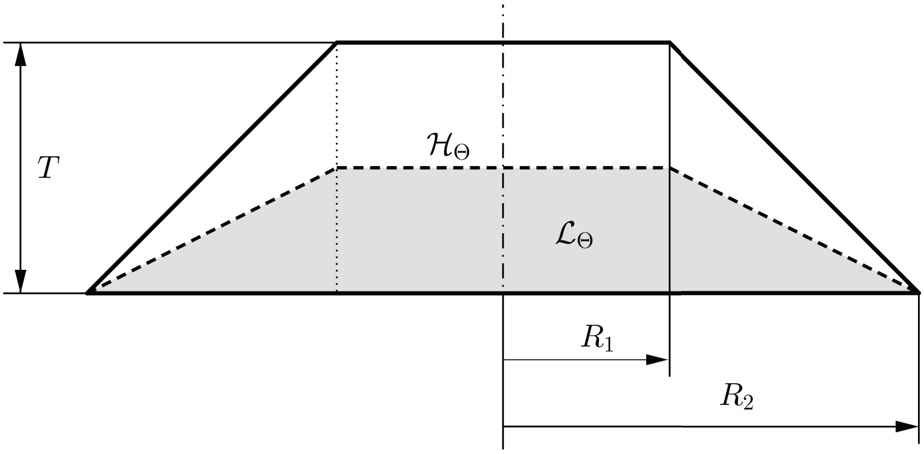

Figure 2.1: A standard lens in $[0,T] \times \mathbb{R}$ with a section $\mathcal{H}_\Theta$ and a partial lens $\mathcal{L}_\Theta$.

A standard lens consists of a cylinder (the inner part) and a ring (the outer part), enclosing the cylinder. Each part is responsible for one of two properties that are important for our considerations. Firstly, for any $K \subset\subset [0,T] \times \mathbb{R}^n$, there exists a cylinder $[0,T] \times B_{R_1}$ (and thus a standard lens) containing K. Secondly, by increasing the outer radius of the lens we can control the slope of its outer part.

The standard lens map ψ introduced in Definition 2.1 as well as its restrictions $\psi_\Theta : B_{R_2} \to \mathcal{H}_\Theta$, $y \mapsto \psi(\Theta, y)$ are Lipschitz continuous, but not differentiable at points (Θ, y) with $|y| = R_1$. More precisely, we have

$$D\psi(\Theta,y) = \big(D_\Theta\psi(\Theta,y), D_y\psi(\Theta,y)\big) = \begin{cases} \begin{pmatrix} T & 0 \\ 0 & I_n \end{pmatrix} & \text{for } |y| \leq R_1, \\[2em] \begin{pmatrix} T\frac{R_2-|y|}{R_2-R_1} & \frac{\Theta T}{(R_2-R_1)}\frac{y}{|y|} \\ 0 & I_n \end{pmatrix} & \text{for } |y| > R_1. \end{cases}$$

However, the collection of points where ψ or ψ_Θ is not differentiable is of Lebesgue measure zero with respect to $\psi([0,\Theta] \times B_{R_2}) = \mathcal{L}_\Theta$ or $\psi_\Theta(B_{R_2}) = \mathcal{H}_\Theta$ respectively. Therefore these points can be ignored when using the lens map to transform integrals over the lens $\mathcal{L}$ or the section $\mathcal{H}_\Theta$ into integrals over the cylinder $[0,T] \times B_{R_2}$ or the ball B_{R_2} (cf. [Rud87, Lemma 7.25 and 7.26]).

Remark 2.2 (Smooth lenses)**.** By a slight modification of the lens map introduced in Definition 2.1, lenses with differentiable or even smooth sections $\mathcal{H}_\Theta$ can be constructed. However, since smooth sections are not necessary for our considerations, we use the most elementary variant in order to keep the expressions for the derivatives of the lens map as simple as possible.

Basic estimates for standard lenses

Standard lenses and their sections will later serve as domains for energy estimates. Therefore we first gather two basic estimates relating integrals over a lens $\mathcal{L} \subseteq \mathbb{R}^{1+n}$

to integrals over its sections $\mathcal{H}_\Theta$.

Lemma 2.3 (Integrals over lenses and their sections)**.** *For $u \in \mathcal{C}^0([0,T]\times\mathbb{R}^n)$ and $\mathcal{L}$ a standard lens of thickness T and radii $R_2 > R_1 > 0$, we have*

(i) $$\int_{\mathcal{L}} |u|\,\mathrm{d}V_{1+n} \le T \sup_{\Theta\in[0,1]} \int_{\mathcal{H}_\Theta} |u(\Theta,\,\cdot\,)|\mathrm{d}V_n.$$

(ii) $$0 \le \frac{d}{d\Theta}\int_{\mathcal{L}_\Theta} |u|\,\mathrm{d}V_{1+n} \le T \int_{\mathcal{H}_\Theta} |u(\Theta,\,\cdot\,)|\,\mathrm{d}V_n.$$

Proof. First note that the map $\psi_\Theta : B_{R_2} \to \mathcal{H}_\Theta$ is a (global) parametrisation of the section $\mathcal{H}_\Theta$ and that $D\psi_\Theta(y) = D_y\psi(\Theta,y)$. The volume density on $\mathcal{H}_\Theta$ is then given by

$$\rho_\Theta(y) = \sqrt{\det(D\psi_\Theta(y)^\mathsf{T} D\psi_\Theta(y))} = \sqrt{\det(D_y\psi(\Theta,y)^\mathsf{T} D_y\psi(\Theta,y))},$$

where $D_y\psi$ is obtained from the Jacobian matrix $D\psi$ by removing the first column and $(D_y\psi)^\mathsf{T} D_y\psi$ is the Gramian matrix of the vectors $D_{y_j}\psi$ $(1 \le j \le n)$. Hence

$$\begin{aligned} |\det D\psi(\Theta,y)| &= \langle D_{y_1}\psi(\Theta,y)\times\ldots\times D_{y_n}\psi(\Theta,y), D_\Theta\psi(\Theta,y)\rangle \\ &\le |(D_{y_1}\psi\times\ldots\times D_{y_n}\psi)(\Theta,y)|\,|D_\Theta\psi(\Theta,y)| \\ &\le \rho_\Theta(y)T, \end{aligned}$$

where we have used that $|D_\Theta\psi| \le \max\left(T, T\frac{R_2-|y|}{R_2-R_1}\right) \le T$. Employing this estimate in the transformation formula for integrals we may write

$$\begin{aligned} \int_{\mathcal{L}} |u|\,\mathrm{d}V_{1+n} &= \int_0^1 \int_{B_{R_2}} |u\circ\psi(\Theta,y))|\,|\det D\psi(\Theta,y)|\,\mathrm{d}y\,\mathrm{d}\Theta \\ &\le T\int_0^1 \int_{B_{R_2}} |u\circ\psi(\Theta,y))|\,\rho_\Theta(y)\,\mathrm{d}y\,\mathrm{d}\Theta \le T \sup_{\Theta\in[0,1]} \int_{\mathcal{H}_\Theta} |u(\Theta,\cdot)|\,\mathrm{d}V_n, \end{aligned}$$

which proves inequality (i). To show the second inequality, we again transform the integral by means of the standard lens map ψ and obtain

$$\frac{d}{d\Theta}\int_{\mathcal{L}_\Theta} |u|\,\mathrm{d}V_{1+n} = \frac{d}{d\Theta}\int_0^\Theta \int_{B_{R_2}} |u\circ\psi(\varepsilon,y)|\,|\det D\psi(\varepsilon,y)|\,\mathrm{d}y\,\mathrm{d}\varepsilon$$

$$= \int_{B_{R_2}} |u\circ\psi(\varepsilon,y)|\,|\det D\psi(\varepsilon,y)|\,\mathrm{d}y \le T\int_{B_{R_2}} |u\circ\psi(\Theta,y)|\,\rho_\Theta(y)\,\mathrm{d}y = T\int_{\mathcal{H}_\Theta} |u(\Theta,\cdot)|\,\mathrm{d}V_n.$$

□

Preparatory for the generalised setting we pursue, we give a full proof of a Gronwall-type estimate in [BGS07, Appendix A, Lemma A.3], based on the elementary Gronwall inequality derived in Lemma 1.4 (iii). Here we focus on explicit expressions for all constants appearing in the calculation.

Lemma 2.4 (Basic estimate relating different sections of a lens)**.** *Let $\mathcal{L}$ be a standard lens of thickness T and suppose that u and f are functions of class $\mathcal{C}^0([0,T]\times\mathbb{R}^n)$ such that*

$$\frac{1}{2}\int\limits_{\mathcal{H}_\Theta} |u(\Theta,\cdot)|\,\mathrm{d}V_n \le \int\limits_{\mathcal{H}_0} |u(0,\cdot)|\,\mathrm{d}V_n + \alpha\int\limits_{\mathcal{L}_\Theta} |u|\,\mathrm{d}V_{1+n} + \int\limits_{\mathcal{L}} |f|\,\mathrm{d}V_{1+n} \tag{2.3}$$

for all $\Theta\in\,]0,1]$ with a constant $\alpha>0$, independent of Θ. Then we have

$$\frac{1}{2}\int\limits_{\mathcal{H}_\Theta} |u(\Theta,\cdot)|\,\mathrm{d}V_n \le e^{2T\alpha\Theta}\Big(\int\limits_{\mathcal{H}_0} |u(0,\cdot)|\,\mathrm{d}V_n + \int\limits_{\mathcal{L}} |f|\,\mathrm{d}V_{1+n}\Big) \tag{2.4}$$

for all $\Theta\in[0,1]$.

Proof. We introduce an auxiliary function $v\in\mathcal{C}^0([0,1])$,

$$v(\Theta) := \frac{1}{2T}\int\limits_{\mathcal{L}_\Theta} |u|\,\mathrm{d}V_{1+n} \quad \text{for } \Theta\in\,]0,1] \quad \text{and} \quad v(0):=0.$$

In addition we abbreviate the terms independent of Θ on the right-hand side of (2.3) by

$$a := \int\limits_{\mathcal{H}_0} |u(0,\cdot)|\,\mathrm{d}V_n + \int\limits_{\mathcal{L}} |f|\,\mathrm{d}V_{1+n}.$$

Applying Lemma 2.3 (ii) yields

$$0 \le v'(\Theta) \le \frac{1}{2}\int\limits_{\mathcal{H}_\Theta} |u|\,\mathrm{d}V_n \qquad \text{for all } \Theta\in\,]0,1].$$

By assumption (2.3) this implies $v'(\Theta)\le a + Cv(\Theta)$, where $C := 2T\alpha$. Integrating the last inequality, we find

$$v(\Theta) \le a\Theta + C\int_0^\Theta v(\tau)\mathrm{d}\tau,$$

which, upon using Gronwall's inequality (see Lemma 1.4 (iii)), leads to

$$Cv(\Theta) \le Ce^{C\Theta}\int_0^\Theta ae^{-C\tau}\mathrm{d}\tau \le \big(e^{C\Theta}-1\big)a \qquad \text{for all } \Theta\in[0,1].$$

Expressing the result in terms of u and f, we obtain

$$\alpha\int\limits_{\mathcal{L}_\Theta} |u|\,\mathrm{d}V_{1+n} \le \big(e^{C\Theta}-1\big)\Big(\int\limits_{\mathcal{H}_0} |u(0,\cdot)|\,\mathrm{d}V_n + \int\limits_{\mathcal{L}} |f|\,\mathrm{d}V_{1+n}\Big) \tag{2.5}$$

for all $\Theta\in\,]0,1]$. Finally, using this estimate in assumption (2.3) and recalling that

$C = 2T\alpha$, we arrive at

$$\frac{1}{2}\int\limits_{\mathcal{H}_\Theta} |u(\Theta,\cdot)|\,\mathrm{d}V_n \le e^{C\Theta}\Big(\int\limits_{\mathcal{H}_0} |u(0,\cdot)|\,\mathrm{d}V_n + \int\limits_{\mathcal{L}} |f|\,\mathrm{d}V_{1+n}\Big)$$

for all $\Theta \in [0,1]$. $\square$

Spacelike standard lenses for hyperbolic partial differential operators

The aim of this section is to construct standard lenses which are tailored to the principal symbol of a first-order partial differential operator (PDO). For convenience of the reader, we first review some of the relevant notions in this context. To a first-order PDO

$$P(t,x;\partial_t,\partial_x) = \partial_t + \sum_{j=1}^{n} A^j(t,x)\partial_{x_j} + B(t,x),$$

with A_j and B matrices of size m and $(t,x) \in \mathbb{R}^{1+n}$ (we will generally omit the identity matrix in front of ∂_t), we assign its principal symbol

$$\sigma(t,x;\tau,\xi) = \tau\, I_m + \sum_{j=1}^{n} \xi_j A^j(t,x).$$

If the matrices A^j are Hermitian, then the principal symbol of P is a linear combination of Hermitian matrices and therefore Hermitian for all directions $(\tau,\xi) \in \mathbb{R}^{1+n}$. At a point $(t,x) \in \mathbb{R}^{1+n}$, one may then define the *forward cone* $\Gamma(t,x)$ as the set of all directions (τ,ξ) where $\sigma(t,x;\tau,\xi)$ is a positive definite matrix. A hypersurface is called *spacelike* (with respect to the principal symbol of P) if its normal vector is almost everywhere contained in the forward cone. In the following lemma we will construct a lens whose individual sections are spacelike hypersurfaces with boundary $\partial\mathcal{H}_\Theta = \{(0,x) : |x| = R_2\}$.

Lemma 2.5 (Constructing a spacelike lens)**.** *Let $(A^j)_{1\le i\le n}$ be Hermitian matrices of size m such that $\|A^j(t,x)\|_{\mathrm{op}} \le C$ for $|x| \ge R_A$. Then a standard lens $\mathcal{L}$ with radii $R_1 \ge R_A$ and $R_2 \ge R_1 + T(1 + 2\,C\sqrt{n})$ has unit normal vector fields ν_Θ on $\mathcal{H}_\Theta$ (pointing outwards with respect to $\mathcal{L}_\Theta$) satisfying the inequality*

$$\langle \eta, \sigma(t,x;\nu_\Theta(t,x))\eta\rangle \ge \frac{1}{2}|\eta|^2 \quad \forall (t,x) \in \mathcal{H}_\Theta\ \forall \Theta \in]0,1]\ \forall \eta \in \mathbb{C}^m. \tag{2.6}$$

Proof. For a lens of thickness T and radii R_1 and R_2, the section $\mathcal{H}_\Theta$ can be represented as the graph of the map $\psi_\Theta : \mathbb{R}^n \to \mathbb{R}$, $\psi_\Theta(x) = \psi(\Theta,x)$ (see Definition 2.1). We have

$$\mathbb{R}^n \ni \operatorname{grad}\psi_\Theta(x) = \begin{cases} (0,\dots,0)^{\mathsf{T}} & \text{for} \quad |x| \le R_1 \\ -\frac{\Theta T}{R_2 - R_1}\frac{x}{|x|} & \text{for} \quad |x| > R_1 \end{cases}$$

and the unit normal vector field on $\mathcal{H}_\Theta$, $(t,x) \mapsto \nu_\Theta(t,x) \in \mathbb{R}^{1+n}$ $(\Theta \in]0,1])$, pointing outwards with respect to $\mathcal{L}_\Theta$, is given by

$$\nu_\Theta(t,x) = \begin{cases} (1,0,...,0)^\mathsf{T} & \text{for} \quad |x| \le R_1 \\ \frac{1}{\sqrt{(\Theta T)^2 + (R_2 - R_1)^2}} \begin{pmatrix} R_2 - R_1 \\ \Theta T \frac{x}{|x|} \end{pmatrix} & \text{for} \quad |x| > R_1. \end{cases}$$

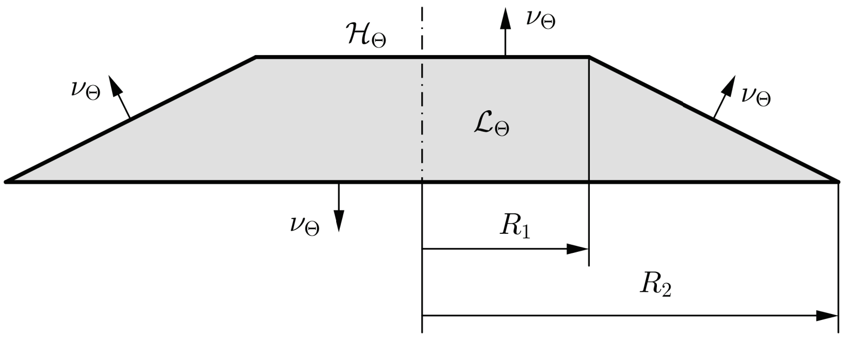

Figure 2.2: The unit normal vector field ν_Θ on $\partial\mathcal{L}_\Theta$.

The condition $R_1 \ge R_A$ ensures that the plateau of the lens extends beyond the region where $\|A^j(t,x)\|_{\mathrm{op}}$ may be unbounded. The larger the outer radius R_2 of the lens, the smaller the spatial components of the normal vector of the lens for $R_1 < |x| < R_2$, where we have $\|A^j(t,x)\|_{\mathrm{op}} \le C$. The inequality in (2.6) is obviously satisfied on the plateau of $\mathcal{L}_\Theta$, since $\langle \eta, \sigma(t,x;\nu_\Theta(t,x))\eta\rangle = \langle \eta,\eta\rangle = |\eta|^2$ for $|x| \le R_1$. If $|x| > R_1 \ge R_A$, we find by virtue of (2.7) that

$$\begin{aligned} \Big\langle \eta, \Big(\nu^0_\Theta(t,x) I_m + \sum_{j=1}^n \nu^j_\Theta(t,x) A^j\Big)\eta\Big\rangle &\ge \nu^0_\Theta(t,x)|\eta|^2 - C\sum_{j=1}^n |\nu^j_\Theta(t,x)|\,|\eta|^2 \\ &= \frac{R_2 - R_1 - C\|\Theta T\frac{x}{|x|}\|_1}{\sqrt{(\Theta T)^2 + (R_2-R_1)^2}}\,|\eta|^2 \ge \frac{(R_2 - R_1 - TC\sqrt{n})}{\sqrt{(\Theta T)^2 + (R_2-R_1)^2}}\,|\eta|^2 \end{aligned}$$

since $\|\Theta T\frac{x}{|x|}\|_1 \le T\,\|\frac{x}{|x|}\|_1 \le T\sqrt{n}\,\big|\frac{x}{|x|}\big| = T\sqrt{n}$ (where $\|x\|_1 = \sum_{j=1}^n |x_j|$). The condition $R_2 \ge R_1 + T(1 + 2\,C\sqrt{n})$ ensures that the numerator is positive and we may proceed to estimate

$$\frac{R_2 - R_1 - TC\sqrt{n}}{\sqrt{(\Theta T)^2 + (R_2-R_1)^2}}\,|\eta|^2 \ge \frac{R_2 - R_1 - \frac{1}{2}(R_2 - R_1 - T)}{T + R_2 - R_1}\,|\eta|^2 = \frac{1}{2}\,|\eta|^2.$$

□

Energy estimates for symmetric hyperbolic operators

A first-order partial differential operator $P = \partial_t + \sum_{j=1}^n A^j\partial_{x_j} + B$ with smooth coefficient matrices is called *symmetric hyperbolic* if

(i) the principal coefficients A^j are Hermitian and

(ii) the matrices A^j and B are uniformly bounded together with all their derivatives.

With applications in Colombeau theory in mind, we perform energy estimates for symmetric hyperbolic operators on standard lenses. It is important to keep explicit expressions for all involved constants, in order to have precise information on their ε-dependence in a generalised setting later on. We provide L^2-estimates in two versions, the second of which can be interpreted as the limiting case for lenses with infinite radius.

Lemma 2.6 (Energy estimates for spacelike lenses). *Let $P = \partial_t + \sum_{j=1}^n A^j \partial_{x_j} + B$ be a symmetric hyperbolic PDO, where A^j and B are matrices of dimension m, depending on variables $(t,x) \in [0,T] \times \mathbb{R}^n$. Then, for any $u \in \mathcal{C}^1([0,T], L^2(\mathbb{R}^n))^m \cap \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m$, the following statements are true:*

(i) *Let $\mathcal{L} \subseteq [0,T] \times \mathbb{R}^n$ be a standard lens of thickness T that satisfies inequality (2.6) and set*
$$\alpha(\mathcal{L}_\Theta) := 1 + \|\mathrm{div}A - B - B^*\|_{L^\infty(\mathcal{L}_\Theta)},$$
where $\mathrm{div}A = \sum_{j=1}^n \partial_{x_j} A^j$ and the L^∞-norm is understood as taking the operator norm first and then the supremum over all $(t,x) \in \mathcal{L}_\Theta$. Then we have:
$$\|u\|^2_{L^2(\mathcal{L})} \le 2Te^{2T\alpha(\mathcal{L})}\big(\|u\|^2_{L^2(\mathcal{H}_0)} + \|Pu\|^2_{L^2(\mathcal{L})}\big). \tag{2.7}$$

(ii) *Denote $\Omega_t :=]0,t[\times \mathbb{R}^n$ and*
$$\beta(t) := 1 + \|\mathrm{div}A(t,\cdot) - B(t,\cdot) - B^*(t,\cdot)\|_{L^\infty(\mathbb{R}^n)}.$$
Then u satisfies the inequality
$$\|u(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \le e^{\int_0^t \beta(s)\mathrm{d}s}\big(\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \|Pu\|^2_{L^2(\Omega_t)}\big) \tag{2.8}$$
for all $t \in [0,T]$, where $\|v\|^2_{L^2(\Omega_t)} = \int_0^t \|v(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\mathrm{d}s$ for any $v \in L^2(\Omega_T)^m$.

Both assertions are also true for non-smooth coefficients $A \in M_m(W^{1,\infty}([0,T]\times\mathbb{R}^n))$ and $B \in M_m(L^\infty([0,T] \times \mathbb{R}^n))$.

Proof. Let $\langle\,\cdot\,,\,\cdot\,\rangle$ be the standard unitary inner product on $\mathbb{C}^m$ and let $\|\cdot\|$ denote the associated norm. Applying the operator P to an arbitrary $u \in \mathcal{C}^\infty([0,T]\times\mathbb{R}^n)^m$, we may write
$$\langle Pu, u\rangle = \langle \partial_t u, u\rangle + \sum_{j=1}^n \langle A^j \partial_{x_j} u, u\rangle + \langle Bu, u\rangle. \tag{2.9}$$

Aiming at an expression in divergence form, we calculate
$$\partial_t\|u\|^2 + \sum_{j=1}^n \partial_{x_j}\langle A^j u, u\rangle = 2\mathrm{Re}\langle \partial_t u, u\rangle + \langle(\mathrm{div}A)u, u\rangle + \sum_{j=1}^n 2\mathrm{Re}\langle A^j \partial_{x_j} u, \partial_t u\rangle,$$

implying that

$$2\mathrm{Re}\Big(\langle\partial_t u,u\rangle+\sum_{j=1}^{n}\langle A^j\partial_{x_j}u,u\rangle\Big)=\partial_t\|u\|^2+\sum_{j=1}^{n}\partial_{x_j}\langle A^j u,u\rangle-\langle(\mathrm{div}A)u,u\rangle, \quad (2.10)$$

where $\mathrm{div}A=\sum_{j=1}^{n}\partial_{x_j}A^j$ is again a Hermitian matrix. Taking two times the real part of (2.9) and using (2.10), we obtain

$$2\mathrm{Re}\langle Pu,u\rangle=\partial_t\|u\|^2+\sum_{i=1}^{n}\partial_{x_j}\langle A^j u,u\rangle-\langle(\mathrm{div}A)u,u\rangle+\,2\mathrm{Re}\langle Bu,u\rangle\,,$$

which can be written as

$$\mathrm{div}(\|u\|^2,\langle A_1u,u\rangle,\ldots,\langle A_nu,u\rangle)^{\mathsf{T}}=\langle(\mathrm{div}A-B-B^*)u,u\rangle+2\mathrm{Re}\langle Pu,u\rangle\,. \quad (2.11)$$

Integrating equation (2.11) over a partial lens $\mathcal{L}_\Theta\subseteq\mathbb{R}^{1+n}$ and using the divergence theorem yields

$$\int\limits_{\partial\mathcal{L}_\Theta}\Big(\nu_\Theta^0\|u\|^2+\sum_{j=1}^{n}\nu_\Theta^j\langle A^j u,u\rangle\Big)\,\mathrm{d}S=\int\limits_{\mathcal{L}_\Theta}\langle(\mathrm{div}A-B-B^*)u,u\rangle\,\mathrm{d}V+2\mathrm{Re}\int\limits_{\mathcal{L}_\Theta}\langle Pu,u\rangle\,\mathrm{d}V,$$

where ν_Θ is the unit normal vector field on $\partial\mathcal{L}_\Theta$, pointing outwards with respect to $\mathcal{L}_\Theta$. Since $\nu_0=(-1,0,...,0)^{\mathsf{T}}$, we conclude from inequality (2.6) that

$$\int\limits_{\partial\mathcal{L}_\Theta}\Big(\nu_\Theta^0\|u\|^2+\sum_{j=1}^{n}\nu_\Theta^j\langle A^j u,u\rangle\Big)\,\mathrm{d}S\geq\frac{1}{2}\|u(\Theta,\cdot)\|^2_{L^2(\mathcal{H}_\Theta)}-\|u(0,\cdot)\|^2_{L^2(\mathcal{H}_0)}.$$

Hence we have for all $\Theta\in\,]0,1]$:

$$\frac{1}{2}\|u(\Theta,\cdot)\|^2_{L^2(\mathcal{H}_\Theta)}\leq\|u(0,\cdot)\|^2_{L^2(\mathcal{H}_0)}+\int\limits_{\mathcal{L}_\Theta}\langle(\mathrm{div}A-B-B^*)u,u\rangle\,\mathrm{d}V+2\mathrm{Re}\int\limits_{\mathcal{L}_\Theta}\langle Pu,u\rangle\,\mathrm{d}V.$$

The terms on the right-hand side can be estimated by means of the Cauchy-Schwarz inequality, leading to

$$\begin{aligned}\Big|\int\limits_{\mathcal{L}_\Theta}\langle(\mathrm{div}A-B-B^*)u,u\rangle\,\mathrm{d}V\Big| &\leq \int\limits_{\mathcal{L}_\Theta}\|(\mathrm{div}A-B-B^*)u\|\,\|u\|\,\mathrm{d}V\\ &\leq \|\mathrm{div}A-B-B^*\|_{L^\infty(\mathcal{L})}\|u\|^2_{L^2(\mathcal{L}_\Theta)}\end{aligned}$$

and

$$\begin{aligned}\Big|2\mathrm{Re}\int\limits_{\mathcal{L}_\Theta}\langle Pu,u\rangle\,\mathrm{d}V\Big|\leq 2\int\limits_{\mathcal{L}_\Theta}\|Pu\|\,\|u\|\,\mathrm{d}V &\leq \int\limits_{\mathcal{L}_\Theta}\big(\,\|Pu\|^2+\|u\|^2\,\big)\,\mathrm{d}V\\ &= \|Pu\|^2_{L^2(\mathcal{L}_\Theta)}+\|u\|^2_{L^2(\mathcal{L}_\Theta)}.\end{aligned}$$

Putting everything together yields the estimate

$$\frac{1}{2}\|u(\Theta,\cdot)\|^2_{L^2(\mathcal{H}_\Theta)} \leq \|u(0,\cdot)\|^2_{L^2(\mathcal{H}_0)} + \alpha(\mathcal{L})\|u\|^2_{L^2(\mathcal{L}_\Theta)} + \|Pu\|^2_{L^2(\mathcal{L}_\Theta)}$$

where $\alpha(\mathcal{L}) = 1 + \|\mathrm{div}A - B - B^*\|_{L^\infty(\mathcal{L})}$. By Lemma 2.4 this implies

$$\frac{1}{2}\|u(\Theta,\cdot)\|^2_{L^2(\mathcal{H}_\Theta)} \leq e^{2T\Theta\alpha(\mathcal{L})}\big(\|u(0,\cdot)\|^2_{L^2(\mathcal{H}_0)} + \|Pu\|^2_{L^2(\mathcal{L})}\big).$$

Finally, applying Lemma 2.3, we obtain

$$\begin{aligned}
\|u\|^2_{L^2(\mathcal{L})} &\leq T \sup_{\Theta\in[0,1]} \|u(\Theta,\cdot)\|^2_{L^2(\mathcal{H}_\Theta)} \\
&\leq T \sup_{\Theta\in[0,1]} \Big(2\, e^{2T\Theta\alpha(\mathcal{L})}\big(\|u(0,\cdot)\|^2_{L^2(\mathcal{H}_0)} + \|Pu\|^2_{L^2(\mathcal{L})}\big)\Big) \\
&\leq 2T\, e^{2T\alpha(\mathcal{L})}\big(\|u(0,\cdot)\|^2_{L^2(\mathcal{H}_0)} + \|Pu\|^2_{L^2(\mathcal{L})}\big).
\end{aligned}$$

To prove (ii), we consider time t as a parameter and integrate equation (2.11) over the spatial domain $\mathbb{R}^n$, leading to

$$\begin{aligned}
\frac{d}{dt}\|u(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \sum_{j=1}^{n} \int_{\mathbb{R}^n} \partial_{x_j}\langle A^j u, u\rangle \,\mathrm{d}V = \int_{\mathbb{R}^n} \langle(\mathrm{div}A - B - B^*)u, u\rangle \,\mathrm{d}V \\
+ 2\mathrm{Re}\int_{\mathbb{R}^n} \langle Pu, u\rangle \,\mathrm{d}V.
\end{aligned}$$

After integration with respect to the the time variable between 0 and t, we get with $\beta(t) := 1 + \|\mathrm{div}A(t,\cdot) - B(t,\cdot) - B^*(t,\cdot)\|_{L^\infty(\mathbb{R}^n)}$:

$$\|u(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \leq \|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \int_0^t \beta(s)\|u(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\mathrm{d}s + \int_0^t \|Pu(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\mathrm{d}s.$$

Here we have used that $\int_{\mathbb{R}^n} \partial_{x_j}\langle A^j(s,\cdot)u(s,\cdot), u(s,\cdot)\rangle \mathrm{d}V = 0$ for all $j = 1, ..., n$ and for all $s \in [0,T]$, since $x \mapsto u(s,x)$ belongs to $H^1(\mathbb{R}^n)$ and the latter possesses $\mathcal{C}^\infty_c(\mathbb{R}^n)$ as a dense subspace. More precisely, regularising $s \mapsto u(s,\cdot)$ and applying proper cut-off functions yields a $\mathcal{C}^\infty_c$-sequence $u_n(s,\cdot) \to u(s,\cdot) \in H^1(\mathbb{R}^n)$. Since $\int_{\mathbb{R}^n} \partial_{x_j}\langle A^j(s,\cdot)u_n(s,\cdot), u_n(s,\cdot)\rangle \mathrm{d}V = 0$ and $\int_{\mathbb{R}^n} \partial_{x_j}\langle A^j(s,\cdot)u_n(s,\cdot), u_n(s,\cdot)\rangle \mathrm{d}V \to \int_{\mathbb{R}^n} \partial_{x_j}\langle A^j(s,\cdot)u(s,\cdot), u(s,\cdot)\rangle \mathrm{d}V$, the limit must vanish as well. Employing the Gronwall estimate in Lemma 1.4 (ii), we turn the last inequality into the desired estimate

$$\|u(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \leq e^{\int_0^t \beta(s)\mathrm{d}s}\big(\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \|Pu\|^2_{L^2(\Omega_t)}\big).$$

□

2.2 Generalised solutions of the Cauchy problem

Having all necessary prerequisites at hand we now draw our attention to the initial value problem (2.1–2.2) on the space-time domain $\Omega_T =]0, T[\times\mathbb{R}^n$. We will give three versions of an existence and uniqueness result, each using different spaces of initial data and right-hand side. Working with a smaller space in this respect allows to relax the asymptotic conditions on the coefficient matrices A^j and B.

We call a partial differential operator $P = \partial_t + \sum_{j=1}^n A^j \partial_{x_j} + B$ with Colombeau generalised coefficients *symmetric hyperbolic*, if

(i) all matrices A^j are Hermitian and

(ii) the entries of A^j and B together with all their derivatives are of L^∞-type, that is, $A^j, B \in M_m(\mathcal{G}_{L^\infty}(\Omega))$.

This ensures that there exists $\varepsilon_0 > 0$ and representatives $(A^j_\varepsilon)_\varepsilon$ and $(B_\varepsilon)_\varepsilon$ such that $P_\varepsilon = \partial_t + \sum_{j=1}^n A^j_\varepsilon \partial_{x_j} + B_\varepsilon$ is a classical symmetric hyperbolic operator for all $\varepsilon < \varepsilon_0$ (the existence of Hermitian representatives of A^j is clear by [HKK14, Lemma 3.23]). The corresponding family of smooth solutions of the classical Cauchy problem for fixed ε then represents the natural candidate for the generalised solution. Yet some additional asymptotic growth conditions in ε have to be imposed on the coefficients to obtain a moderate family of solutions. In particular, certain log-type conditions on the coefficient matrices are essential in order to use a Gronwall-type argument in the proof (cf. [Obe88, Obe89, LO91, Hör04, Hör11]).

The first theorem admits the most general initial data and right-hand side, but requires the principal coefficients $A^j_\varepsilon(t, x)$ to be bounded uniformly in ε and (t, x) for large $|x|$. Given a compact set $K \subseteq \Omega_T$, we use a suitable standard lens containing K and apply the energy estimates derived in Section 2.1 to show moderateness of the ε-wise constructed solution candidate. More precisely, the uniform boundedness of A^j outside an arbitrarily large cylinder of radius R_A (independent of ε) guarantees that the generalised forward cones $[(\Gamma_\varepsilon(t, x))_\varepsilon]$ do not collapse outside this cylinder as $\varepsilon \to 0$. This in turn enables us to construct a fixed standard lens enclosing $K \subseteq \Omega_T$, whose sections $\mathcal{H}_\Theta$ satisfy the inequality (2.6) uniformly in ε. However, the forward cones are allowed to collapse inside the cylindrical region of the lens, hence the boundedness condition actually only restricts the behaviour of $A^j_\varepsilon(t, x)$ at infinity, that is, in the limit $|x| \to \infty$.

Theorem 2.7 (Existence and uniqueness in $\mathcal{G}$)**.** *The Cauchy problem (2.1–2.2) for a symmetric hyperbolic operator with generalised coefficents $A^j, B \in M_m(\mathcal{G}_{L^\infty}(\mathbb{R}^{1+n}))$,*

$$\begin{aligned} \partial_t u + \sum_{j=1}^n A^j \partial_{x_j} u + Bu &= f \qquad \text{on } \Omega_T =]0, T[\times \mathbb{R}^n \\ u|_{t=0} &= g, \end{aligned}$$

has a unique solution $u \in \mathcal{G}(\Omega_T)^m$, if the following conditions are satisfied:

(1) *The initial data g and the right-hand side f are Colombeau generalised functions, that is, $g \in \mathcal{G}(\mathbb{R}^n)^m$ and $f \in \mathcal{G}(\overline{\Omega}_T)^m$.*

(2) *All spatial derivatives $\partial_{x_i} A^j$ as well as the Hermitian part of B are locally of log-type.*

(3) *There exists $R_A > 0$ such that $\|A^j_\varepsilon(t,x)\|_{\mathrm{op}} = O(1)$ as $\varepsilon \to 0$ uniformly on $]0,T[\, \times \{x \in \mathbb{R}^n : |x| > R_A\}$.*

Proof. We pick Hermitian representatives $(A^j_\varepsilon)_\varepsilon$ and representatives of B, f, and g. There exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, the initial value problem

$$\partial_t u_\varepsilon + \sum_{j=1}^{n} A^j_\varepsilon \partial_{x_j} u_\varepsilon + B_\varepsilon u_\varepsilon \;=\; f_\varepsilon \tag{2.12}$$

$$u_\varepsilon|_{t=0} \;=\; g_\varepsilon \tag{2.13}$$

has a unique solution $u_\varepsilon \in \mathcal{C}^\infty(\overline{\Omega}_T)^m$ (cf. [BGS07, Theorem 2.12]). We claim that the equivalence class $u = [(u_\varepsilon)_\varepsilon]$ is the unique Colombeau solution. Hence we must show that the net $(u_\varepsilon)_\varepsilon$ is moderate and that negligible variations of the coefficients or the data yield the same generalised solution. Note that if $u \in \mathcal{G}(\Omega_T)^m$, we also have $u|_{t=0} = [(u_\varepsilon|_{t=0})_\varepsilon] \in \mathcal{G}(\mathbb{R}^n)^m$.

Let $K \subset\subset \Omega_T$ and assume that $K \subseteq]0,T[\, \times \{x \in \mathbb{R}^n : |x| < R_K\}$. We choose $R_1 \geq \max(R_A, R_K)$ and $R_2 \geq R_1 + T(1 + 2C\sqrt{n})$. Then a standard lens $\mathcal{L}$ of thickness T, inner radius R_1 and outer radius R_2 will contain K and satisfy the positivity condition (2.6) by Lemma 2.5. Subsequently we apply the energy estimate (2.7) in Lemma 2.6 (i) ε-wise and obtain

$$\|u_\varepsilon\|^2_{L^2(\mathcal{L})} \leq 2T\, e^{2T\alpha_\varepsilon(\mathcal{L})} \big(\|g_\varepsilon\|^2_{L^2(\mathcal{H}_0)} + \|f_\varepsilon\|^2_{L^2(\mathcal{L})} \big) = O(\varepsilon^{-m}) \tag{2.14}$$

as $\varepsilon \to 0$ for some $m \in \mathbb{N}_0$, since $\|g_\varepsilon\|_{L^2(\mathcal{H}_0)}$ and $\|f_\varepsilon\|_{L^2(\mathcal{L})}$ grow only with some inverse power of ε as $\varepsilon \to 0$ and assumption (2) implies that

$$\alpha_\varepsilon(\mathcal{L}) = 1 + \|\mathrm{div} A_\varepsilon - B_\varepsilon - B^*_\varepsilon\|_{L^\infty(\mathcal{L})} = O(\log(1/\varepsilon)) \quad \text{as} \quad \varepsilon \to 0.$$

We attempt to show that all derivatives of u_ε satisfy an estimate similar to (2.14). For this purpose we introduce some convenient notations. For $A \in M_m(\mathcal{C}^\infty(\overline{\Omega}_T))$ and $u \in \mathcal{C}^\infty(\Omega_T)^m$ we define

$$\begin{aligned} \widetilde{\nabla}^1 A &:= \mathrm{diag}(\partial_{x_1} A, ..., \partial_{x_n} A) & \nabla^1 u &:= (\partial_{x_1} u, ..., \partial_{x_n} u)^{\mathsf{T}} \\ \widetilde{\nabla}^{r+1} A &:= \widetilde{\nabla}^1 \widetilde{\nabla}^r A & \nabla^{r+1} u &:= \nabla^1 \nabla^r u \\ \widetilde{\Sigma}^r A &:= \mathrm{diag}(\mathrm{A}, ..., \mathrm{A})_{n^r} & \Sigma^r u &:= (u, ..., u)^{\mathsf{T}}_{n^r}\,, \end{aligned}$$

For example, $\widetilde{\nabla}^r A$ is a blockdiagonal matrix of dimension $n^r m$, built from all spatial derivatives $\partial^\alpha_x A$ of length $|\alpha| = r$. Correspondingly, $\widetilde{\Sigma}^r A$ is a blockdiagonal matrix whose blocks are just n^r copies of A itself. The notation for vectors is analogous, that is, if u is a vector of length m, then the vector $\nabla^r u$ has $n^r m$ components, representing all partial derivatives $\partial^\alpha_x u$ of length $|\alpha| = r$, and $\Sigma^r u$ is built by concatenating n^r copies of u to a vector of size $n^r m$.

Claim 1. The vector $\nabla^r u_\varepsilon$ satisfies an equation of the form

$$\partial_t \nabla^r u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\Sigma}^r A_\varepsilon^j)\, \partial_{x_j} \nabla^r u_\varepsilon + \widetilde{B}_\varepsilon^r \nabla^r u_\varepsilon = Q_\varepsilon^{r-1} \Sigma^r u_\varepsilon + \nabla^r f_\varepsilon \tag{2.15}$$

where $\widetilde{B}_\varepsilon^1 = \widetilde{\Sigma}^1 B_\varepsilon + (\partial_{x_i} A_\varepsilon^j)_{1\le i,j\le n}$, $\widetilde{B}_\varepsilon^{r+1} = \widetilde{\Sigma}^1 \widetilde{B}_\varepsilon^r + (\partial_{x_i} \widetilde{\Sigma}^r A_\varepsilon^j)_{1\le i,j\le n}$ for $r \ge 1$, and Q_ε^r is a purely spatial partial differential operator of order r with coefficients depending linearly on spatial derivates of A_ε^j and B_ε up to order $r+1$.

Proof of Claim 1. We present the base case $r = 1$ in detail and proceed by induction. Differentiating equation (2.12) with respect to x_k yields

$$\partial_t \partial_{x_k} u_\varepsilon + \sum_{j=1}^{n} A_\varepsilon^j \partial_{x_j} \partial_{x_k} u_\varepsilon + \sum_{j=1}^{n} (\partial_{x_k} A_\varepsilon^j) \partial_{x_j} u_\varepsilon + B_\varepsilon \partial_{x_k} u_\varepsilon + (\partial_{x_k} B_\varepsilon) u_\varepsilon = \partial_{x_k} f_\varepsilon.$$

One would like to read this as an equation for $\partial_{x_k} u_\varepsilon$, but the third term on the left-hand side spoils the idea by introducing a coupling between the k systems. However, combining all partial derivatives $\partial_{x_k} u_\varepsilon$ to the vector $\nabla^1 u_\varepsilon$, we may consider this as one big system of PDEs and write, using our special notation,

$$\begin{aligned}\partial_t \nabla^1 u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\Sigma}^1 A_\varepsilon^j)\, \partial_{x_j} \nabla^1 u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\nabla}^1 A_\varepsilon^j)\, \partial_{x_j} \Sigma^1 u_\varepsilon + (\widetilde{\Sigma}^1 B_\varepsilon)\, \nabla^1 u_\varepsilon \\ = -(\widetilde{\nabla}^1 B_\varepsilon)\, \Sigma^1 u_\varepsilon + \nabla^1 f_\varepsilon.\end{aligned}$$

There are no derivatives of u_ε present on the right-hand side and the only term on the left-hand side that does not fit into our concept can be rewritten in a suitable form, since

$$\sum_{j=1}^{n} (\widetilde{\nabla}^1 A_\varepsilon^j)\, \partial_{x_j} \Sigma^1 u_\varepsilon = \sum_{j=1}^{n} (\widetilde{\nabla}^1 A_\varepsilon^j)\, \Sigma^1 \partial_{x_j} u_\varepsilon = (\partial_{x_i} A_\varepsilon^j)_{1\le i,j\le n} \nabla^1 u_\varepsilon.$$

Consequently, $\nabla^1 u_\varepsilon$ satisfies the system of PDEs

$$\partial_t \nabla^1 u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\Sigma}^1 A_\varepsilon^j)\, \partial_{x_j} \nabla^1 u_\varepsilon + \widetilde{B}_\varepsilon^1 \nabla^1 u_\varepsilon = Q_\varepsilon^0 \Sigma^1 u_\varepsilon + \nabla^1 f_\varepsilon,$$

where $\widetilde{B}_\varepsilon^1 = \widetilde{\Sigma}^1 B_\varepsilon + (\partial_{x_i} A_\varepsilon^j)_{1\le i,j\le n}$ and $Q_\varepsilon^0 = -\widetilde{\nabla}^1 B_\varepsilon$, which can be interpreted as a PDO of order 0. We proceed by induction with respect to the differentiation index r. Applying ∂_{x_k} to the induction hypothesis (2.15), we obtain

$$\begin{aligned}\partial_t \partial_{x_k} \nabla^r u_\varepsilon + \sum_{j=1}^{n} (\partial_{x_k} \widetilde{\Sigma}^r A_\varepsilon^j) \partial_{x_j} \nabla^r u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\Sigma}^r A_\varepsilon^j)\, \partial_{x_j} \partial_{x_k} \nabla^r u_\varepsilon + \widetilde{B}_\varepsilon^r \partial_{x_k} \nabla^r u_\varepsilon \\ = -(\partial_{x_k} \widetilde{B}_\varepsilon^r) \nabla^r u_\varepsilon + \partial_{x_k} (Q_\varepsilon^{r-1} \Sigma^r u_\varepsilon) + \partial_{x_k} \nabla^r f_\varepsilon.\end{aligned} \tag{2.16}$$

Just as in the case $r = 1$ we try to write these k coupled systems of PDEs as one big system. It is easy to see that the right-hand sides of the k equations above can altogether be written as $Q^r_\varepsilon \Sigma^{r+1} u_\varepsilon + \nabla^{r+1} f_\varepsilon$ with Q^r_ε representing a purely spatial partial differential operator of order r with coefficients depending linearly on spatial derivates of A^j_ε and B_ε up to order $r+1$. More precisely, writing Q^{r-1}_ε in a general form as $Q^{r-1}_\varepsilon = \sum_{|\alpha|\le r-1} C^\alpha_\varepsilon \partial^\alpha_x$ with $n^r m$-dimensional matrices C^α_ε, we have

$$Q^r_\varepsilon \Sigma^{r+1} u_\varepsilon = -(\widetilde{\nabla}^1 \widetilde{B}^r_\varepsilon)\,\Sigma^1(\nabla^r u_\varepsilon) + \sum_{|\alpha|\le r-1} (\widetilde{\nabla}^1 C^\alpha_\varepsilon) \Sigma^{r+1} u_\varepsilon + \sum_{|\alpha|\le r-1} (\widetilde{\Sigma}^1 C^\alpha_\varepsilon) \nabla^1(\Sigma^r u_\varepsilon),$$

hence the components of the coefficient matrices of the PDO Q^r_ε emerge from those of Q^{r-1}_ε essentially by differentiating them and adding derivatives of components of A^j and B. In particular,

$$Q^1_\varepsilon \Sigma^2 u_\varepsilon = -(\widetilde{\nabla}^1 \widetilde{B}^1_\varepsilon)\,\Sigma^1(\nabla^1 u_\varepsilon) - (\widetilde{\nabla}^1(\widetilde{\nabla}^1 B_\varepsilon))\Sigma^2 u_\varepsilon - (\widetilde{\Sigma}^1(\widetilde{\nabla}^1 B_\varepsilon))\nabla^1(\Sigma^1 u_\varepsilon),$$

where $\widetilde{B}^1_\varepsilon = \widetilde{\Sigma}^1 B_\varepsilon + (\partial_{x_i} A^j_\varepsilon)_{1\le i,j\le n}$. Thus the coefficients of Q^r_ε are built from spatial derivatives of the matrices A^j_ε and B_ε of highest order $r+1$ and we get

$$\|Q^{r-1}_\varepsilon \Sigma^r u_\varepsilon\|_{L^2(\mathcal{L})} \le C_r \Big(\sum_{1\le j\le n} \|A^j_\varepsilon\|_{W^{r,\infty}(\mathcal{L})} + \|B_\varepsilon\|_{W^{r,\infty}(\mathcal{L})} \Big) \|u_\varepsilon\|_{H^{r-1}(\mathcal{L})}.$$

Furthermore we can rewrite the lower-order terms on the left-hand sides of the k systems in (2.16) as

$$\begin{aligned} \widetilde{B}^{r+1}_\varepsilon \nabla^{r+1} u_\varepsilon &:= \sum_{j=1}^n (\widetilde{\nabla}^1 \widetilde{\Sigma}^r A^j_\varepsilon) \Sigma^1(\partial_{x_j} \nabla^r u_\varepsilon) + (\widetilde{\Sigma}^1 \widetilde{B}^r_\varepsilon) \nabla^{r+1} u_\varepsilon \\ &= \sum_{j=1}^n (\partial_{x_i} \widetilde{\Sigma}^r A^j)_{1\le i,j\le n} \nabla^{r+1} u_\varepsilon + (\widetilde{\Sigma}^1 \widetilde{B}^r_\varepsilon) \nabla^{r+1} u_\varepsilon, \end{aligned}$$

finally leading to

$$\partial_t \nabla^{r+1} u_\varepsilon + \sum_{j=1}^n (\widetilde{\Sigma}^{r+1} A^j_\varepsilon)\, \partial_{x_j} \nabla^{r+1} u_\varepsilon + \widetilde{B}^{r+1}_\varepsilon \nabla^{r+1} u_\varepsilon = Q^r_\varepsilon \Sigma^{r+1} u_\varepsilon + \nabla^{r+1} f_\varepsilon\,,$$

where indeed $\widetilde{B}^{r+1}_\varepsilon = \widetilde{\Sigma}^1 \widetilde{B}^r_\varepsilon + (\partial_{x_i} \widetilde{\Sigma}^r A^j)_{1\le i,j\le n}$. This completes the inductive step and thus the proof of Claim 1.

Since $\|\widetilde{\Sigma}^r A^j_\varepsilon\|_{\mathrm{op}} = \|A^j_\varepsilon\|_{\mathrm{op}}$, the positivity condition (2.6) is also valid for the symmetric hyperbolic operator $P^r_\varepsilon = I_{n^r m}\partial_t + \sum_{j=1}^n (\widetilde{\Sigma}^r A^j_\varepsilon)\partial_{x_j} + \widetilde{B}^r_\varepsilon$ and the same standard lens as used in (2.14). By (2.7) in Lemma 2.6 (i) we therefore have

$$\|\nabla^r u_\varepsilon\|^2_{L^2(\mathcal{L})} \le C_{\mathcal{L},r,\varepsilon} \left(\|\nabla^r g_\varepsilon\|^2_{L^2(\mathcal{H}_0)} + \|Q^{r-1}_\varepsilon \Sigma^r u_\varepsilon + \nabla^r f_\varepsilon\|^2_{L^2(\mathcal{L})} \right) \tag{2.17}$$

where $C_{\mathcal{L},r,\varepsilon} := 2T\, e^{2T\alpha^r_\varepsilon(\mathcal{L})}$ and

$$\alpha^r_\varepsilon(\mathcal{L}) := 1 + \|\mathrm{div}\, \widetilde{\Sigma}^r A_\varepsilon - \widetilde{B}^r_\varepsilon - (\widetilde{B}^r_\varepsilon)^*\|_{L^\infty(\mathcal{L})}.$$

Note that the Hermitian part of $\widetilde{B}^r_\varepsilon$ only involves the Hermitian part of B_ε as well as first-order derivatives $\partial_{x_i} A^j_\varepsilon$. By assumption (2), all constituents $\partial_{x_i} A^j_\varepsilon$ and $B_\varepsilon + B^*_\varepsilon$ are locally of log-type, hence we have $C_{\mathcal{L},r,\varepsilon} = O(\varepsilon^{-N})$ as $\varepsilon \to 0$ for some $N \in \mathbb{N}_0$. By (2.14) we know that $\|u_\varepsilon\|^2_{L^2(\mathcal{L})} = O(\varepsilon^{-m})$ as $\varepsilon \to 0$ and since $\|Q^{r-1}_\varepsilon \Sigma^r u_\varepsilon\|_{L^2(\mathcal{L})} = O(\varepsilon^{-M})\|u_\varepsilon\|_{W^{r-1,2}(\mathcal{L})}$ for some $M \in \mathbb{N}_0$ (depending on r), we may apply (2.17) iteratively for $r = 1, 2, 3, ...$ to conclude that

$$\forall r \in \mathbb{N}_0 \, \exists m \in \mathbb{N}_0 : \, \|\nabla^r u_\varepsilon\|^2_{L^2(\mathcal{L})} = O(\varepsilon^{-m}) \quad \text{as} \quad \varepsilon \to 0. \tag{2.18}$$

This automatically implies moderateness of the net $(\partial_t u_\varepsilon)_\varepsilon$ by (2.12). It remains to show a similar asymptotic estimate for mixed derivatives.

Claim 2. All mixed derivatives $\partial^l_t \nabla^r u_\varepsilon$ satisfy an equation of the form

$$\partial^l_t \nabla^r u_\varepsilon = R^{r,l}_\varepsilon u_\varepsilon + \partial^{l-1}_t \nabla^r f_\varepsilon \tag{2.19}$$

where $R^{r,l}_\varepsilon : \mathcal{C}^\infty(\Omega_T)^m \to \mathcal{C}^\infty(\Omega_T)^{n^r m}$ is a linear partial differential operator of order $r + l$ involving time derivatives only up to order $l - 1$. Moreover, the coefficients of $R^{r,l}_\varepsilon$ are linear combinations of mixed derivatives of A^j_ε and B_ε, including spatial derivatives up to order r and time derivatives up to order $l - 1$.

Proof of Claim 2. Again we prove this assertion by induction. The base case $l = 1$ follows immediately from (2.15) by setting

$$R^{r,1}_\varepsilon u_\varepsilon := Q^{r-1}_\varepsilon \Sigma^r u_\varepsilon - \sum_{j=1}^n (\widetilde{\Sigma}^r A^j_\varepsilon)\partial_{x_j} \nabla^r u_\varepsilon - \widetilde{B}^r_\varepsilon \nabla^r u_\varepsilon.$$

By this definition, $R^{r,1}_\varepsilon$ is a PDO of order $r+1$ with respect to the spatial coordinates and of order 0 with respect to the time variable. Applying the operator ∂_t to equation (2.19) gives

$$\partial^{l+1}_t \nabla^r u_\varepsilon = \partial_t(R^{r,l}_\varepsilon u_\varepsilon) + \partial^l_t \nabla^r f_\varepsilon$$

and $R^{r,l+1}_\varepsilon u_\varepsilon := \partial_t(R^{r,l}_\varepsilon u_\varepsilon) = \partial^l_t(R^{r,1}_\varepsilon u_\varepsilon)$ amounts to the action of a PDO of order $r + l + 1$ containing time derivatives only up to order l and spatial derivatives up to order $r + 1$. As an obvious implication of the Leibniz rule, the coefficients of $R^{r,l}_\varepsilon$ are linear combinations of derivatives of A^j_ε and B_ε including spatial derivatives of order r at most, since the coefficients of the PDOs Q^{r-1}_ε depend (only) on spatial derivatives of A^j_ε and B_ε up to order r. Moreover, by construction, the coefficients of $R^{r,l}_\varepsilon$ carry time derivatives of (spatial derivatives of) A^j_ε and B_ε of order $l - 1$. This completes the proof of Claim 2.

Successively making use of (2.19) for $l = 1, 2, 3, ...$ implies in combination with (2.17) that

$$\forall l, r \in \mathbb{N}_0 \, \exists m \in \mathbb{N}_0 : \, \|\partial^l_t \nabla^r u_\varepsilon\|_{L^2(\mathcal{L})} = O(\varepsilon^{-m}) \quad \text{as} \quad \varepsilon \to 0. \tag{2.20}$$

Employing the Sobolev embedding theorem on domains with a locally Lipschitz continuous boundary (cf. [AF03, Theorem 4.12, Part II]) and the fact that $K \subseteq \mathcal{L}$, this yields

$$\forall \alpha \in \mathbb{N}^{1+n}_0 \, \exists m \in \mathbb{N}_0 : \, \|\partial^\alpha u_\varepsilon\|_{L^\infty(K)} = O(\varepsilon^{-m}) \quad \text{as} \quad \varepsilon \to 0, \tag{2.21}$$

that is, the class $[(u_\varepsilon)_\varepsilon]$ is moderate, since $K \subset\subset \Omega_T$ was arbitrary.

Uniqueness of the generalised solution amounts to stability under negligible perturbations of the data. To show uniqueness of the ε-wise constructed generalised solution, we choose negligible nets $(\widetilde{f}_\varepsilon)_\varepsilon$ and $(\widetilde{g}_\varepsilon)_\varepsilon$ to represent right-hand side and initial data. From the energy estimates (2.14), (2.17), and equation (2.19) it is then easy to see that the corresponding solution $[(\widetilde{u}_\varepsilon)_\varepsilon]$ of the Cauchy problem will also be negligible. By [GKOS01, Theorem 1.2.3], it actually suffices to show the negligibility estimate for the zeroth derivative only, that is, in terms of L^2-estimates, for all derivatives of $\widetilde{u}_\varepsilon$ with order less than or equal to $\lceil (n+1)/2 \rceil$. □

If condition (3) in Theorem 2.7 is dropped, then the coefficient matrices $A^j_\varepsilon(t,x)$ will not bounded independently of ε for large $|x|$ anymore. This means that the generalised forward cones $[(\Gamma_\varepsilon(t,x))_\varepsilon]$ might not only collapse in a bounded region, but also at spatial infinity as $\varepsilon \to 0$. Thus it is impossible to construct a fixed standard lens whose normal vector field remains everywhere in the local forward cone for all $\varepsilon < \varepsilon_0$. Therefore it seems that the energy estimates obtained in Section 2.1 can no longer be applied. However, giving up the idea of using a fixed standard lens (for each $K \subset\subset \Omega_T$) which is independent of ε and working with an ε-indexed family of standard lenses $(\mathcal{L}_\varepsilon)_\varepsilon$ instead, we can increase the diameter of the lens $\mathcal{L}_\varepsilon$ as $\varepsilon \to 0$, thereby steadily tilting its normal vector field as the forward cones get narrower. The slices of this 'generalised lens' remain spacelike for all $\varepsilon < \varepsilon_0$ and since the growth of its volume as $\varepsilon \to 0$ is under control in terms of inverse powers of ε, we can still establish a generalised solution by restricting the initial data g to $\mathcal{G}_{L^\infty}(\mathbb{R}^n)^m$, interpreted as a subset of $\mathcal{G}(\mathbb{R}^n)^m$, and the right-hand side f to $\mathcal{G}_{L^\infty}(\Omega_T)^m$, interpreted as a subset of $\mathcal{G}(\Omega_T)^m$.

Theorem 2.8 (Existence and uniqueness in $\mathcal{G}$ without uniform bounds on the coefficients)**.** *Consider the initial value problem (2.1–2.2) with the following alternative conditions:*

(1') *The initial data $g \in \mathcal{G}(\mathbb{R}^n)^m$ and the right-hand side $f \in \mathcal{G}(\Omega_T)^m$ satisfy the asymptotic estimates of $\mathcal{G}_{L^\infty}(\mathbb{R}^n)^m$ and $\mathcal{G}_{L^\infty}(\Omega_T)^m$ respectively.*[1]

(2') *All spatial derivatives $\partial_{x_i} A^j$ as well as the Hermitian part of B are of L^∞-log-type.*

Then there exists a unique solution $u \in \mathcal{G}(\Omega_T)^m$.

Proof. The loss of condition (3) and the alternative version (1') of (1) have no effect on the applicability of Theorem 2.12 in [BGS07] for fixed $\varepsilon < \varepsilon_0$. Hence we still get a solution candidate $[(u_\varepsilon)_\varepsilon]$ from the ε-wise construction of a family $(u_\varepsilon)_\varepsilon$.

After choosing a compact set $K \subset\subset \Omega_T$, we again aim at building a standard lens around K such that the positivity condition (2.6) in Lemma 2.5 is satisfied. Assuming that $K \subseteq]0,T[\times \{x \in \mathbb{R}^n : |x| < R_K\}$, we fix the thickness T of the lens and its inner radius $R_1 \geq R_K$. Since the coefficient matrices are no longer

[1]Every representative $(f_\varepsilon)_\varepsilon$ of a generalised function $f \in \mathcal{G}_{L^\infty}(\Omega)^m$ is also moderate in the sense of $\mathcal{G}(\Omega)^m$, thus $\mathcal{G}_{L^\infty}(\Omega)^m$ can be interpreted as a subset of $\mathcal{G}(\Omega)^m$ if we allow the difference of two representatives of f to be in the ideal $\mathcal{N}(\Omega)^m$ instead of $\mathcal{N}_{L^\infty}(\Omega)^m$.

bounded independently of ε at spatial infinity, but satisfy $\|A_\varepsilon^j\|_{L^\infty(\Omega_T)} \leq C\varepsilon^{-m}$ for some $C > 0$ and for all $\varepsilon < \varepsilon_0$, we use a family of radii $R_2^\varepsilon = R_1 + T(1 + 2C\varepsilon^{-m}\sqrt{n})$ to meet the positivity condition (2.6) for all $\varepsilon < \varepsilon_0$. We thus obtain a family of standard lenses $(\mathcal{L}_\varepsilon)_\varepsilon$ with outer radii $R_2^\varepsilon = O(\varepsilon^{-m})$ as $\varepsilon \to 0$, all containing the compact set K and all staying spacelike as $\varepsilon \to 0$.

We put $\alpha'_\varepsilon := 1 + \|\mathrm{div} A_\varepsilon - B_\varepsilon - B_\varepsilon^*\|_{L^\infty(\Omega_T)}$, that is, the supremum is taken over the whole domain Ω_T. Employing the energy estimate (2.7) ε-wise with the integration domains now changing with ε as well, leads to

$$\begin{aligned}
\|u_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} &\leq 2T\, e^{2T\alpha'_\varepsilon}\Big(\|g_\varepsilon\|^2_{L^2(\mathcal{H}_0^\varepsilon)} + \|f_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)}\Big) \\
&\leq 2T\, e^{2T\alpha'_\varepsilon}\Big(\mathrm{Vol}_n(\mathcal{H}_0^\varepsilon)\|g_\varepsilon\|^2_{L^\infty(\mathbb{R}^n)} + \mathrm{Vol}_{1+n}(\mathcal{L}_\varepsilon)\|f_\varepsilon\|^2_{L^\infty(\Omega_T)}\Big) \\
&= O(\varepsilon^{-N}) \quad (\varepsilon \to 0)
\end{aligned}$$

for some $N \in \mathbb{N}_0$, since $\mathrm{Vol}_n(\mathcal{H}_0^\varepsilon) = \mathrm{Vol}_n(B_{R_2^\varepsilon}) = O((R_2^\varepsilon)^n) = O((\varepsilon^{-m})^n)$ and $\mathrm{Vol}_{1+n}(\mathcal{L}_\varepsilon) < T\,\mathrm{Vol}_n(B_{R_2^\varepsilon}) = O((\varepsilon^{-m})^n)$ as $\varepsilon \to 0$. The analogue of the energy estimate (2.17) for spatial derivatives of u_ε now yields

$$\|\nabla^r u_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} \leq 2T\, e^{2T\alpha'^r_\varepsilon}\Big(\|\nabla^r g_\varepsilon\|^2_{L^2(\mathcal{H}_0^\varepsilon)} + \|Q_\varepsilon^{r-1}\Sigma^r u_\varepsilon + \nabla^{r+1} f_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)}\Big)$$

where $\alpha'^r_\varepsilon := 1 + \|\mathrm{div}\,\widetilde{\Sigma}^r A_\varepsilon - \widetilde{B}^r_\varepsilon - (\widetilde{B}^r_\varepsilon)^*\|_{L^\infty(\Omega_T)} = O(\log(1/\varepsilon))$ as $\varepsilon \to 0$ by assumption (2'). Using the triangle inequality, we may write

$$\begin{aligned}
\|Q_\varepsilon^{r-1}\Sigma^r u_\varepsilon + \nabla^{r+1} f_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} &\leq \Big(\|Q_\varepsilon^{r-1}\Sigma^r u_\varepsilon\|_{L^2(\mathcal{L}_\varepsilon)} + \|\nabla^{r+1} f_\varepsilon\|_{L^2(\mathcal{L}_\varepsilon)}\Big)^2 \\
&\leq 2\Big(\|Q_\varepsilon^{r-1}\Sigma^r u_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} + \|\nabla^{r+1} f_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)}\Big).
\end{aligned}$$

Hence there exist $C > 0$ and $l \in \mathbb{N}_0$ such that

$$\|\nabla^r u_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} \leq C\varepsilon^{-l}\Big(\|\nabla^r g_\varepsilon\|^2_{L^2(\mathcal{H}_0^\varepsilon)} + \|\nabla^{r+1} f_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} + \|Q_\varepsilon^{r-1}\Sigma^r u_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)}\Big)$$

and all norms on the right-hand side are of moderate growth: We have

- $\|\nabla^r g_\varepsilon\|^2_{L^2(\mathcal{H}_0^\varepsilon)} \leq \mathrm{Vol}_n(\mathcal{H}_0^\varepsilon)\|\nabla^r g_\varepsilon\|^2_{L^\infty(\mathbb{R}^n)} = O(\varepsilon^{-N_1})$ as $\varepsilon \to 0$,
- $\|\nabla^{r+1} f_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} \leq \mathrm{Vol}_{1+n}(\mathcal{L}_\varepsilon)\|\nabla^{r+1} f_\varepsilon\|^2_{L^\infty(\mathbb{R}^n)} = O(\varepsilon^{-N_2})$ as $\varepsilon \to 0$,

for some $N_1, N_2 \in \mathbb{N}_0$. Moreover, since the coefficient matrices C_ε^α of Q_ε^{r-1} are built from spatial derivates of A_ε^j and B_ε of order less than or equal to r and $\|C_\varepsilon^\alpha\|_{W^{r,\infty}(\mathcal{L}_\varepsilon)} \leq \|C_\varepsilon^\alpha\|_{W^{r,\infty}(\Omega_T)}$, we have

- $\|Q_\varepsilon^{r-1}\Sigma^r u_\varepsilon\|^2_{L^2(\mathcal{L}_\varepsilon)} = O(\varepsilon^{-N_3})\|u_\varepsilon\|_{W^{r-1,2}(\mathcal{L}_\varepsilon)}$ as $\varepsilon \to 0$.

Combining these observations, we consecutively obtain moderate growth of all spatial derivatives $\nabla^r u_\varepsilon$ with respect to $\|\cdot\|_{L^2(\mathcal{L}_\varepsilon)} \geq \|\cdot\|_{L^2(K)}$.

With the help of (2.19) it is then easy to see that

$$\forall K \subset\subset \Omega_T\ \forall l, r \in \mathbb{N}_0\ \exists m \in \mathbb{N}_0 :\ \|\partial_t^l \nabla^r u_\varepsilon\|_{L^2(K)} = O(\varepsilon^{-m}) \quad (\varepsilon \to 0).$$

By virtue of the Sobolev embedding theorem we conclude that for all $\alpha \in \mathbb{N}_0^{1+n}$, there exists $m \in \mathbb{N}_0$ such that $\|\partial^\alpha u_\varepsilon\|_{L^\infty(K)} = O(\varepsilon^{-m})$ as $\varepsilon \to 0$. The uniqueness part of the proof is similar to the corresponding part in the proof of Theorem 2.7. □

Remark 2.9 (On the solution concept in Theorem 2.8)**.** In assumption (2') of Theorem 2.8, we consider g as an element of $\mathcal{G}(\mathbb{R}^n)^m$, additionally satisfying the moderateness estimates of $\mathcal{G}_{L^\infty}(\mathbb{R}^n)^m$. In particular, the initial condition $u|_{t=0} = g$ has to be understood as an equality in $\mathcal{G}(\mathbb{R}^n)^m$, that is, we have $(u_\varepsilon|_{t=0} - g_\varepsilon)_\varepsilon \in \mathcal{N}(\mathbb{R}^n)$ for any representative $(u_\varepsilon)_\varepsilon$ of the solution $u \in \mathcal{G}(\Omega_T)^m$ and any representative $(g_\varepsilon)_\varepsilon$ of g, but not necessarily $(u_\varepsilon|_{t=0} - g_\varepsilon)_\varepsilon \in \mathcal{N}_{L^\infty}(\mathbb{R}^n)$.

Remark 2.10 (Dirac measures as coefficients)**.** In contrast to Theorem 2.7, uniform boundedness of the principal coefficients $A^j_\varepsilon(t,x)$ for $|x| \to \infty$ is no longer required in the assumptions of Theorem 2.8. Existence and uniqueness of the solution is also guaranteed for principal coefficients associated with the delta distribution, concentrated on a hypersurface. For instance, we may consider matrices $A^j(t,x)$ including entries of the form $\delta(x_1) \otimes h(t, x_2, ..., x_n)$, in the sense of association (with h a bounded function). Then condition (3) in Theorem 2.7 is violated, since $\|A^j_\varepsilon(t,x)\|_{\mathrm{op}}$ is not uniformly bounded in ε as $\varepsilon \to 0$ for $|x| \to \infty$. Notwithstanding this broader range of applicability compared to Theorem 2.7, condition (2') imposes a log-type asymptotic growth of the generalised function modelling $\delta(x_1)$, meaning that we have to work with a logarithmically rescaled delta net. Only if the entries of $A^j(t,x)$ are of the form $c_{ij}\delta(t)$ with constants $c_{ij} = (c_{ji})^*$, the log-type scaling is not necessary, since $\partial_{x_i} A^j = 0$ in this case.

Apart from the possibility to consider coefficients proportional to the delta distribution concentrated on an arbitrary hypersurface, Theorem 2.8 also incorporates the case of coefficients with periodic singularities, as the following remark illustrates.

Remark 2.11 (Lattices of Dirac measures)**.** Theorem 2.8 provides a unique generalised solution even for coefficients and data which are both associated to highly singular and periodic distributions. Denoting the Dirac measure at a point $y \in \mathbb{R}^n$ by $\delta_y = \delta(\cdot - y)$, it is possible to consider, for example, generalised coefficients $B(t,x) \approx \sum_{\kappa \in \mathbb{Z}^n} \delta_{l\cdot\kappa}(x)\widetilde{B}(t,x)$ and initial data $g(x) \approx \sum_{\kappa\in\mathbb{Z}^n} \delta_{q\cdot\kappa}(x)\widetilde{g}(x)$ with $\widetilde{B}$ and $\widetilde{g}$ smooth, and real numbers $l, q > 0$, representing n-dimensional lattices of Dirac measures with lattice constants $1/l$ and $1/q$, respectively. However, the scaling of the corresponding delta nets still has to be of logarithmic type. Interpreting B as an electromagnetic potential, such periodic structures could be used to model diffraction gratings or crystal lattices in certain limiting cases of the involved physical parameters.

Remark 2.12 (Infinite propagation speed)**.** The conditions in Theorem 2.8 allow for infinite propagation speed near spatial infinity as $\varepsilon \to 0$. In general this may cause non-uniqueness of solutions, see [Obe92, Example 17.1]. Furthermore, non-vanishing initial data may yield the solution $u = 0$ in $\mathcal{G}(\Omega_T)$. Consider the scalar equation $\partial_t u + \frac{1}{\varepsilon}\partial_x u = 0$ with $u|_{t=0} = \varphi$, where $\varphi \in \mathcal{D}(\mathbb{R})$ and $\varphi(0) = 1$. Then the unique Colombeau solution is the class $[((t,x) \mapsto \varphi(x - \frac{1}{\varepsilon}t))_\varepsilon] = 0$ in $\mathcal{G}(]0,T[\times\mathbb{R})$, since for all $K \subset\subset]0,T[\times\mathbb{R}$ we have $\operatorname{supp}((t,x) \mapsto \varphi(x - \frac{1}{\varepsilon}t)) \cap K = \emptyset$ for ε small enough. Imposing the asymptotic estimates of $\mathcal{G}_{L^\infty}$ for initial data and right-hand side saves uniqueness of the solution in $\mathcal{G}$, but non-vanishing initial data yielding the solution $u = 0$ are still possible. In Theorem 2.7, infinite propagation speed may also occur, but only inside a cylindrical tube $S_A = \{(t,x) : 0 \leq t \leq T \wedge |x| \leq R_A\}$ with R_A independent of ε. As time starts to evolve, initial data $(g_\varepsilon)_\varepsilon$, concentrated in the ball B_{R_A}, may propagate away infinitely fast, leaving the tube S_A (almost) instantaneously as $\varepsilon \to 0$. However, since propagation speed is limited outside S_A by condition (3) in Theorem 2.7, the generalised solution $[(u_\varepsilon)_\varepsilon]$ cannot be zero.

Inspecting the structure of L^2-based energy estimates more closely, one often finds some freedom to play with the coefficienties regularity of the PDE, balancing them against the regularity of initial data and right-hand side. In fact, initial data and right-hand side decaying at spatial infinity ($|x| \to \infty$) make it possible to relax the conditions on A^j and B even a bit further. In particular, the required asymptotic behaviour of $(A^j_\varepsilon)_\varepsilon$ and $((B + B^*)_\varepsilon)_\varepsilon$ can be made less restrictive with respect to the time variable. To set up a convenient notation first, we introduce the mixed norm

$$\|A\|_{L^{1,\infty}(\Omega_T)} := \int_0^T \|A(s,\cdot)\|_{L^\infty(\mathbb{R}^n)} \mathrm{d}s$$

for any $A \in M_m(C_b^\infty(\overline{\Omega}_T))$. We say that an element $A \in M_m(\mathcal{G}_{L^\infty}(\Omega_T))$ is of $L^{1,\infty}$*-log-type*, if it has a representative $(A_\varepsilon)_\varepsilon$ such that

$$\|A_\varepsilon\|_{L^{1,\infty}}(\Omega_T) = O(\log(1/\varepsilon)) \quad \text{as} \quad \varepsilon \to 0.$$

A similar norm was introduced in [CO90, Definition 2.1]. Using the alternative energy estimate (2.8), we obtain a refined existence and uniqueness result for generalised hyperbolic systems of PDEs with regard to the time dependence of the coefficients. Here, the logarithmic growth condition with respect to the time variable can be dispensed with.

Theorem 2.13 (Existence and uniqueness in $\mathcal{G}_{L^2}$)**.** *The initial value problem (2.1–2.2) together with the regularity conditions*

(1") *initial data $g \in \mathcal{G}_{L^2}(\mathbb{R}^n)^m$ and right-hand side $f \in \mathcal{G}_{L^2}(\Omega_T)^m$,*

(2") *all spatial derivatives $\partial_{x_i} A^j$ as well as the Hermitian part of B are of $L^{1,\infty}$-log-type,*

has a unique generalised solution $u \in \mathcal{G}_{L^2}(\Omega_T)^m$.

Proof. Having fixed Hermitian representatives of A^j and representatives of B, f and g, we may use Theorem 2.6 in [BGS07] to provide smooth solutions $u_\varepsilon \in \mathcal{C}^\infty([0,T], H^\infty(\mathbb{R}^n))^m$ to the corresponding classical initial value problems for each $\varepsilon < \varepsilon_0$. To show moderateness, we apply inequality (2.8) and estimate

$$\|u_\varepsilon\|^2_{L^2(\Omega_T)} \leq T \sup_{0 \leq t \leq T} \|u_\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \leq T e^{\beta_\varepsilon} \Big(\|g_\varepsilon\|^2_{L^2(\mathbb{R}^n)} + \|f_\varepsilon\|^2_{L^2(\Omega_T)} \Big) = O(\varepsilon^{-m}),$$

as $\varepsilon \to 0$, where $\beta_\varepsilon := \|\mathrm{div} A_\varepsilon - B_\varepsilon - B_\varepsilon^*\|_{L^{1,\infty}(\Omega_T)} = O(\log(1/\varepsilon))$ as $\varepsilon \to 0$. In Claim 1 in the proof of Theorem 2.7 it has been shown that the $n^r m$-dimensional vector $\nabla^r u_\varepsilon$ satisfies the equation

$$\partial_t \nabla^r u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\Sigma}^r A_\varepsilon^j)\, \partial_{x_j} \nabla^r u_\varepsilon + \widetilde{B}_\varepsilon^r \nabla^r u_\varepsilon = Q_\varepsilon^{r-1} \Sigma^r u_\varepsilon + \nabla^r f_\varepsilon$$

where Q_ε^r is a purely spatial partial differential operator of order r, the matrices $\widetilde{\Sigma}^r A_\varepsilon^j$ are blockdiagonal matrices consisting of A_ε^j-blocks and $\widetilde{B}_\varepsilon^r$ depends solely on B_ε and first-order spatial derivatives $\partial_{x_i} A_\varepsilon^j$. Thus, using inequality (2.8) once again, we get

$$\sup_{0 \le t \le T} \|\nabla^r u_\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \le e^{\beta_\varepsilon^r} \Big(\|\nabla^r g_\varepsilon\|^2_{L^2(\mathbb{R}^n)} + \|Q_\varepsilon^{r-1} \Sigma^r u_\varepsilon + \nabla^r f_\varepsilon\|^2_{L^2(\Omega_T)} \Big),$$

where $\beta_\varepsilon^r = 1 + \|\mathrm{div}\, \widetilde{\Sigma}^r A_\varepsilon - \widetilde{B}_\varepsilon^r - (\widetilde{B}_\varepsilon^r)^*\|_{L^{1,\infty}(\Omega_T)} = O(\log(1/\varepsilon))$ as $\varepsilon \to 0$. This implies that for all $r \in \mathbb{N}_0$ there exists $m \in \mathbb{N}_0$ such that $\|\nabla^r u_\varepsilon\|^2_{L^2(\Omega_T)} = O(\varepsilon^{-m})$ as $\varepsilon \to 0$. Finally, iteratively employing equation (2.19) for $l = 1, 2, 3, \ldots$ one finds that for all $r \in \mathbb{N}_0$ and for all $l \in \mathbb{N}_0$, $\|\partial_t^l \nabla^r u_\varepsilon\|^2_{L^2(\Omega_T)}$ is of moderate growth as well. Thus we again obtain moderateness of the representative $(u_\varepsilon)_\varepsilon$ of the solution candidate, that is,

$$\forall \alpha \in \mathbb{N}_0^{1+n}\ \exists m \in \mathbb{N}_0 : \ \|\partial^\alpha u_\varepsilon\|_{L^2(\Omega_T)} = O(\varepsilon^{-m}) \quad \text{as} \quad \varepsilon \to 0$$

and hence $[(u_\varepsilon)_\varepsilon] \in \mathcal{G}_{L^2}(\Omega_T)$. To show uniqueness of the generalised solution in $\mathcal{G}_{L^2}(\Omega_T)$, we assume negligible initial data $(\widetilde{g}_\varepsilon)_\varepsilon \in \mathcal{N}_{L^2}(\mathbb{R}^n)$ and right-hand side $(\widetilde{f}_\varepsilon)_\varepsilon \in \mathcal{N}_{L^2}(\Omega_T)$. Then the same energy estimates we used to prove moderateness immediately imply negligibility of the solution $[(\widetilde{u}_\varepsilon)_\varepsilon]$. □

Remark 2.14 (Model delta nets are admissible)**.** In Theorem 2.13, thanks to the $L^{1,\infty}$-norms in condition (2”), a logarithmic rescaling of the mollifier is not required anymore to model singular principal coefficient matrices $A^j(t,x)$ containing entries associated with $\delta(t) \otimes h(x)$, where h is Lipschitz continuous, and lower-order coefficient matrices $B(t,x)$ incorporating entries associated with $\delta(t) \otimes k(x)$, where k is a bounded function. Within the conditions of Theorem 2.8, avoiding a logarithmic scaling of the delta net in these circumstances would only be possible if h is a (generalised) constant and $k = 0$.

Remark 2.15 (Extending solutions to infinite time)**.** For the existence and uniqueness results presented in this section, there exist versions which are also global in time. In correspondence with Theorem 2.7, given coefficients A^j and B in $M_m(\mathcal{G}_{L^\infty}(\mathbb{R}^{1+n}))$, initial data $g \in \mathcal{G}(\mathbb{R}^n)^m$, and right-hand side $f \in \mathcal{G}(\mathbb{R}^{1+n})^m$, one obtains a global solution $u \in \mathcal{G}(\mathbb{R}^{1+n})^m$, if all $\partial_{x_i} A^j$ as well as the Hermitian part of B are locally of log-type and $\|A_\varepsilon^j(t,x)\|_{\mathrm{op}} = O(1)$ as $\varepsilon \to 0$ outside $\mathbb{R} \times B_{R_A}$ for some $R_A > 0$. Considering the generalised initial value problem (2.1–2.2) on $\mathbb{R}^{1+n}$, we briefly sketch how to obtain a representative $(u_\varepsilon)_\varepsilon$ of the global solution:

First we pick Hermitian representatives of the matrices A^j and representatives of B, F, and G. Then we construct $(u_\varepsilon)_\varepsilon$ separately for $t > 0$ and $t < 0$ as follows:

- For any $t > 0$, we pick some $T > t$ and construct $(u_\varepsilon^T)_\varepsilon \in \mathcal{C}^\infty(\Omega_T)^{]0,\varepsilon_0[}$ as in the proof of Theorem 2.7. The energy estimates therein together with the Sobolev embedding theorem imply that for any $T_2 > T_1$, we must have $(u_\varepsilon^{T_2}|_{\Omega_{T_1}})_\varepsilon = (u_\varepsilon^{T_1})_\varepsilon$, since for all $\varepsilon < \varepsilon_0$, $v_\varepsilon := u_\varepsilon^{T_2}|_{\Omega_{T_1}} - u_\varepsilon^{T_1}$ solves the initial value problem in $\mathcal{C}^\infty(\Omega_T)$ with initial data $g_\varepsilon = 0$ and right-hand side $f_\varepsilon = 0$. Thus, by setting $u_\varepsilon(t,x) := u_\varepsilon^T(t,x)$ for some $T > t$, we are able to consistently define $u_\varepsilon(t,x)$ for all $t > 0$.

- For any $t < 0$, we pick some $T > -t > 0$. Putting $\widetilde{A}_\varepsilon^j(t,x) := -A_\varepsilon^j(-t,x)$, $\widetilde{B}_\varepsilon(t,x) := -B_\varepsilon(-t,x)$, and $\widetilde{f}_\varepsilon(t,x) = -f_\varepsilon(-t,x)$, we again construct a solution representative $(\widetilde{u}_\varepsilon^T)_\varepsilon \in \mathcal{C}^\infty(\Omega_T)^{]0,\varepsilon_0[}$ for the corresponding initial value problem on Ω_T, just as in the proof of Theorem 2.7. This in turn yields a 'negative-time' solution in $\mathcal{C}^\infty(\Omega_{-T})$ for the original initial value problem, considered on Ω_{-T}, by setting $u_\varepsilon^{-T}(t,x) := \widetilde{u}_\varepsilon^T(-t,x)$. We then have

$$\Big(\partial_t + \sum_{j=1}^n A_\varepsilon^j(t,x)\partial_{x_j} + B_\varepsilon(t,x)\Big)u_\varepsilon^{-T}(t,x)$$
$$= \Big(-\partial_{-t} - \sum_{j=1}^n \widetilde{A}_\varepsilon^j(-t,x)\partial_{x_j} - \widetilde{B}_\varepsilon(-t,x)\Big)\widetilde{u}_\varepsilon^T(-t,x) = -\widetilde{f}_\varepsilon(-t,x) = f_\varepsilon(t,x).$$

for $-T < t < 0$. Finally, we put $u_\varepsilon(t,x) := u_\varepsilon^{-T}(t,x)$.

It is evident that the constructed family $(u_\varepsilon)_\varepsilon$ is smooth for $t \neq 0$. Smoothness at the initial surface $t = 0$ is also clear since $P_\varepsilon u_\varepsilon = f_\varepsilon$ everywhere and $(\partial_x^\alpha u_\varepsilon)\big|_{t=0} = \partial_x^\alpha g_\varepsilon$ for all $\alpha \in \mathbb{N}_0^n$ and $\varepsilon < \varepsilon_0$.

Similarly extending the respective asymptotic growth conditions from Ω_T to $\mathbb{R}^{1+n}$, we also obtain 'global in time' variants of Theorems 2.8 and 2.13. The construction of global solutions as outlined in Remark 2.15 reveals that the conditions on the coefficients with regard to their dependence on the time variable can even be slightly relaxed. It suffices to demand the asymptotic estimates required in Theorems 2.7, 2.8, and 2.13 only locally in time, that is for all compact intervals $I \subset\subset \mathbb{R}$. A similar observation was made in [GKOS01, Remark 1.5.3].

Having established existence and uniqueness of generalised solutions, we revisit the physical model for the interaction of a relativistic neutron with an ultra-short laser pulse, briefly discussed at the beginning of this chapter. The Dirac equation, a first-order system 'by nature', is used to account for the neutron's spin. The laser pulse propagates at the speed of light and is described mathematically by the delta distribution, concentrated on a null hyperplane in Minkowski space, with an additional matrix-valued 'profile function', characterising the transversal shape of the pulse. We may now consider this problem in a generalised framework.

Example 2.16 (Dirac equation with impulsive wave potential)**.** We use the Dirac equation (describing spin-$\frac{1}{2}$ massive particles) with a non-minimal coupling term to model a neutron in an electromagnetic field (cf. [BD64, Chapter 3]). The Dirac equation for the spinor wave function ψ of a relativistic neutron in an impulsive electromagnetic wave can formally be written as

$$\partial_t \psi(t,x) + \sum_{j=1}^{3} \gamma^0 \gamma^j \partial_{x_j} \psi(t,x) + im\gamma^0 \psi(t,x) + \delta(t-x_1) F(x_2,x_3)\psi(t,x) = 0, \quad (2.22)$$

where γ^k are the Dirac gamma matrices, m is the mass of the neutron, i is the imaginary unit, and we are using units of measure such that $\hbar = c = 1$ (where $\hbar$ is Planck's constant and c is the speed of light). The matrix F is derived from the electromagnetic field tensor and satisfies additional conditions determined by the Maxwell equations. In order to obtain a symmetric hyperbolic system of the form (2.1), Dirac's gamma matrices are represented in the Weyl basis, that is

$$\gamma^0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

yielding Hermitian principal coefficients $\gamma^0\gamma^j$ in system (2.22). There is no rigorous solution concept for this equation within a distribution-theoretic setting, as ψ cannot be continuous across the hyperplane $t - x_1 = 0$, if it ought to be a 'solution' in any (weak) sense and therefore the product $\delta(t-x_1)F(x_2,x_3)\psi(t,x)$ will in general be ill-defined. However, interpreting equation (2.28) as an equality in the differential algebra of Colombeau generalised functions and imposing the initial condition $\psi|_{t=0} = \varphi \in \mathcal{G}(\mathbb{R}^3)^4$, the corresponding initial value problem in $\mathcal{G}(]0,T[\times\mathbb{R}^3)^4$ is in fact well-defined, if we use a logarithmically rescaled mollifier to model the delta distribution.[2] This guarantees that condition (2) in Theorem 2.7 is satisfied and we thus obtain a unique solution $\psi \in \mathcal{G}(]0,T[\times\mathbb{R}^3)^4$ of the initial value problem for the generalised symmetric hyperbolic system (2.22).

2.3 Regularity theory and distributional limits

In this section we explore regularity properties of the unique generalised solution of the initial value problem (2.1–2.2) upon comparison with solutions obtained from classical existence and uniqueness theorems, in cases where the coefficients are of compatible regularity. Moreover, we investigate the regularity of the generalised solution in terms of an intrinsic notion of regularity in the algebra of generalised

[2] For the generalised coefficients to belong to $\mathcal{G}_{L^\infty}(\mathbb{R}^{1+n})$, an additional cut-off is necessary, if the entries of F are not bounded functions. This issue can be resolved by using a net $(\chi_\varepsilon)_\varepsilon$ of ε-dependent cut-off functions approaching 1 as $\varepsilon \to 0$. Moreover, a suitable cut-off procedure can also be justified physically, as long as it occurs sufficiently 'far away' from the particle's wave function.

functions, developed in [Obe92]. We also present several variants of results on distributional limits of the generalised solution, if the generalised coefficients A^j and B are constructed by regularising functions of low regularity, or even distributions.

To begin with, when the coefficients are smooth, the unique Colombeau solution should in a certain sense be equal to the respective classical solution. As in [Hör04] and [LO91], the following two propositions establish coherence with classical results for smooth and distributional data respectively.

Proposition 2.17 (Smooth coefficients and data)**.** *In Theorem 2.7, additionally assume that A^j and B have components in $\mathcal{C}_b^\infty(\overline{\Omega}_T)$. If $f \in \mathcal{C}^\infty(\overline{\Omega}_T)^m$ and $g \in \mathcal{C}^\infty(\mathbb{R}^n)^m$, then the generalised solution $u \in \mathcal{G}(\Omega_T)^m$ is equal to the classical smooth solution, that is, the classical solution is one of its representatives.*

Proof. Since we may choose the constant nets $(f)_\varepsilon$ and $(g)_\varepsilon$ as representatives of the classes of f and g in $\mathcal{G}(\Omega_T)^m$ and $\mathcal{G}(\mathbb{R}^n)^m$, we obtain the classical smooth solution to problem (2.1–2.2) as a representative of the unique Colombeau solution. □

In the case of non-smooth initial data g_0 and right-hand side f_0, we obtain corresponding generalised objects g and f via regularisation. Then, any representative $(u_\varepsilon)_\varepsilon$ of the generalised solution u converges to the distributional solution u_0.

Proposition 2.18 (Smooth coefficients and data in Sobolev spaces)**.** *Suppose that A^j and B in Theorem 2.13 are smooth. For some $s \in \mathbb{R}$, let $f_0 \in L^2([0,T], H^s(\mathbb{R}^n))^m$ and $g_0 \in H^s(\mathbb{R}^n)^m$. Denote by u_0 the unique distributional solution to the initial value problem (2.1–2.2) in the space $\mathcal{C}^0([0,T], H^s(\mathbb{R}^n))^m$. Define generalised initial data $g := [(g_\varepsilon)_\varepsilon] \in \mathcal{G}_{L^2}(\mathbb{R}^n)^m$ and right-hand side $f := [(f_\varepsilon)_\varepsilon] \in \mathcal{G}_{L^2}(\Omega_T)^m$, where g_ε and f_ε are moderate regularisations such that $g_\varepsilon \to g_0$ in $H^s(\mathbb{R}^n)^m$ and $f_\varepsilon \to f_0$ in $L^2([0,T], H^s(\mathbb{R}^n))^m$ as $\varepsilon \to 0$. If $u = [(u_\varepsilon)_\varepsilon]$ is the corresponding generalised solution in $\mathcal{G}_{L^2}(\Omega_T)^m$ provided by Theorem 2.13, then $u_\varepsilon \to u_0$ in $\mathcal{C}^0([0,T], H^s(\mathbb{R}^n))^m$.*

Proof. By Theorem 2.6 in [BGS07], we have the inequality

$$\|u_\varepsilon(t,\cdot) - u_0(t,\cdot)\|^2_{H^s(\mathbb{R}^n)} \leq C\Big(\|g_\varepsilon - g_0\|^2_{H^s(\mathbb{R}^n)} + \int_0^t \|f_\varepsilon(s,\cdot) - g_0(s,\cdot)\|^2_{H^s(\mathbb{R}^n)} \mathrm{d}s\Big)$$

for all $t \in [0,T]$, where the constant C does not depend on ε, since all coefficient matrices A^j and B are assumed have components in $\mathcal{C}^\infty(\overline{\Omega}_T)$. Taking the supremum over all $t \in [0,T]$ and considering ε tending to 0 yields the convergence in $\mathcal{C}^0([0,T], H^s(\mathbb{R}^n))^m$. □

We may interpret Proposition 2.18 as a statement about the regularity of the generalised solution, measured in terms of H^s-norms of the associated distributional shadow. Yet this concept of regularity is restricted to situations in which distributional limits exist and therefore not applicable if initial data or right-hand side are not associated with any distribution.

Intrinsic regularity theory in Colombeau algebras is based on the subalgebra $\mathcal{G}^\infty(\Omega)$ of regular generalised functions in $\mathcal{G}(\Omega)$, introduced in [Obe92], and has been investigated in the context of hyperbolic PDEs in [HK01, HOP06, Obe09, GO14, GO15]. In the study of intrinsic regularity of generalised solutions of PDEs, the notion of *slow-scale nets* was introduced in [HO04] and has proven to be essential in many circumstances (cf. [HOP06, GO14, GO15]).

A net $(r_\varepsilon)_\varepsilon$ of complex numbers is said to be of *slow scale* if

$$|r_\varepsilon|^p = O(1/\varepsilon) \quad \text{as} \quad \varepsilon \to 0 \quad \text{for all} \quad p \geq 0.$$

As in [Hör04] we call a net $(s_\varepsilon)_\varepsilon$ of complex numbers a *slow-scale log-type net* if there is a slow-scale net $(r_\varepsilon)_\varepsilon$ of real numbers, $r_\varepsilon \geq 1$, such that

$$|s_\varepsilon| = O(\log(r_\varepsilon)) \quad \text{as} \quad \varepsilon \to 0.$$

Generalised functions satisfying the moderateness estimate in Definition 1.1 with slow scale nets in place of inverse powers ε^{-m} are called *slow-scale regular*. They represent a different notion of regularity than regular generalised functions.

Proposition 2.19 (Intrinsic regularity). *In Theorem 2.7, assume all coefficients A^j and B to be slow-scale regular generalised functions. In addition, replace the log-type conditions in assumption* (2) *by the corresponding slow-scale log-type estimates. Then, given regular initial data $g \in \mathcal{G}^\infty(\mathbb{R}^n)^m$ and regular right-hand side $f \in \mathcal{G}^\infty([0,T] \times \mathbb{R}^n)^m$, the solution will also be regular, that is, $u \in \mathcal{G}^\infty(\Omega_T)^m$.*

Proof. Fix a standard lens $\mathcal{L} \subseteq \overline{\Omega}_T$ of thickness T with initial surface $\mathcal{H}_0$. Since both g and f are regular generalised functions, there exists $M \in \mathbb{N}_0$ such that for all $r, l \in \mathbb{N}_0$, we have for $\varepsilon \to 0$:

$$\|\partial_t^l \nabla^r f_\varepsilon\|_{L^2(\mathcal{L})} = O(\varepsilon^{-M}) \quad \text{as well as} \quad \|\nabla^r g_\varepsilon\|_{L^2(\mathcal{H}_0)} = O(\varepsilon^{-M}).$$

With A^j and B satisfying slow scale log-type estimates, the generalised bounds in the L^2-estimates (2.14) and (2.17) will now be of slow scale. Hence, for each $\alpha \in \mathbb{N}_0^n$ we obtain a certain slow-scale net $(r_\varepsilon)_\varepsilon$ of positive real numbers such that $\|\partial_x^\alpha u_\varepsilon\|_{L^2(\mathcal{L})} = O(r_\varepsilon \varepsilon^{-M})$ as $\varepsilon \to 0$. Finally, by (2.19) we will only pick up slow-scale factors with each additional time derivative we apply to $\partial_x^\alpha u_\varepsilon$, and so we have

$$\|\partial_t^l \partial_x^\alpha u_\varepsilon\|_{L^2(\mathcal{L})} = O(\varepsilon^{-M-1})$$

as $\varepsilon \to 0$, for all $l \in \mathbb{N}_0$ and all $\alpha \in \mathbb{N}_0^n$. This proves the assertion. □

The intrinsic regularity property holds also for the generalised solutions obtained from Theorems 2.8 and 2.13 respectively, if all log-type conditions are replaced by slow-scale log-type estimates and all coefficients are slow-scale regular generalised functions. Note that all these conditions are automatically satisfied when the coefficients are smooth (as in Proposition 2.17).

Despite its $\mathcal{G}^\infty$-regularity, the solution in Proposition 2.19 may not be associated to any distribution. In the remaining part of this section we want to investigate distributional limits of the generalised solution when the coefficients are non-smooth. To this end we consider the generalised Cauchy problem (2.1–2.2) with initial data, right-hand side, and coefficients representing regularisations of corresponding non-smooth objects. For convenience of the reader we first gather some basic properties of regularisations of $W^{k,p}$-functions obtained by convolution with a mollifier.

Lemma 2.20 (Convolution regularisations). *Let $\rho \in \mathcal{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \rho(x)\mathrm{d}x = 1$, $\mathrm{supp}(\rho) \subseteq B_1(0)$, and put $\rho_\varepsilon(x) = \sigma_\varepsilon^{-n}\rho(x/\sigma_\varepsilon)$ where $\sigma_\varepsilon \to 0$ as $\varepsilon \to 0$ and $\sigma_\varepsilon^{-1} = O(\varepsilon^{-m})$ as $\varepsilon \to 0$ for some $m \in \mathbb{N}$. Given $u \in W^{k,p}(\mathbb{R}^n)$ with $k \in \mathbb{N}_0$ and $1 \leq p \leq \infty$, we put $u_\varepsilon := u * \rho_\varepsilon$. Then the family of regularisations $(u_\varepsilon)_\varepsilon$ has the following properties:*

(i) *If $u \in L^p(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$, then $[(u_\varepsilon)_\varepsilon] \in \mathcal{G}_{L^\infty}(\mathbb{R}^n)$ and $\|\partial^\alpha u_\varepsilon\|_{L^\infty(\mathbb{R}^n)} = O(\sigma_\varepsilon^{-|\alpha|-\frac{n}{p}})$ as $\varepsilon \to 0$.*

(ii) *If $u \in W^{k,p}(\mathbb{R}^n)$, then $\|\partial^\alpha u_\varepsilon\|_{W^{k,p}(\mathbb{R}^n)} = O(\sigma_\varepsilon^{-|\alpha|})$ as $\varepsilon \to 0$.*

(iii) *If $u \in W^{k,q}(\mathbb{R}^n)$ with $1 \leq q < \infty$, then $\|u_\varepsilon - u\|_{W^{k,q}(\mathbb{R}^n)} \to 0$ as $\varepsilon \to 0$.*

(iv) *If $u \in L^\infty(\mathbb{R}^n)$ is uniformly continuous, then we have $\|u_\varepsilon - u\|_{L^\infty(\mathbb{R}^n)} \to 0$ as $\varepsilon \to 0$. In particular, if $u \in W^{k,\infty}(\mathbb{R}^n)$, then its regularisations $(u_\varepsilon)_\varepsilon$ satisfy $\|u_\varepsilon - u\|_{W^{k-1,\infty}(\mathbb{R}^n)} \to 0$ as $\varepsilon \to 0$.*

Proof. (i): If $u \in L^p(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$, we may use Young's inequality to estimate

$$\|\partial^\alpha u_\varepsilon\|_{L^\infty(\mathbb{R}^n)} = \|(\partial^\alpha \rho_\varepsilon) * u\|_{L^\infty(\mathbb{R}^n)} \leq \|\partial^\alpha \rho_\varepsilon\|_{L^q(\mathbb{R}^n)} \|u\|_{L^p(\mathbb{R}^n)}$$

with $1/p + 1/q = 1$. Moreover we have

$$\|\partial^\alpha \rho_\varepsilon\|_{L^q(\mathbb{R}^n)} = \sigma_\varepsilon^{-|\alpha|-n+\frac{n}{q}} \|\partial^\alpha \rho\|_{L^q(\mathbb{R}^n)}$$

for $1 \leq q \leq \infty$. Hence $\|\partial^\alpha u_\varepsilon\|_{L^\infty(\mathbb{R}^n)} = O(\sigma_\varepsilon^{-|\alpha|-\frac{n}{p}})$ as $\varepsilon \to 0$ for all $\alpha \in \mathbb{N}_0^n$ and since $\sigma_\varepsilon^{-1} = O(\varepsilon^{-m})$ as $\varepsilon \to 0$, this implies $[(u_\varepsilon)_\varepsilon] \in \mathcal{G}_{L^\infty}(\mathbb{R}^n)$.
(ii): For $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$, we obtain

$$\begin{aligned} \|\partial^\alpha u_\varepsilon\|_{W^{k,p}(\mathbb{R}^n)} = \max_{|\beta| \leq k} \|\partial^{\alpha+\beta} u_\varepsilon\|_{L^p(\mathbb{R}^n)} &= \max_{|\beta| \leq k} \|(\partial^\alpha \rho_\varepsilon) * \partial^\beta u\|_{L^p(\mathbb{R}^n)} \\ &\leq \sigma_\varepsilon^{-|\alpha|} \|\partial^\alpha \rho\|_{L^1(\mathbb{R}^n)} \|u\|_{W^{k,p}(\mathbb{R}^n)}. \end{aligned}$$

For (iii) and (iv) we refer to [Fol99, Theorem 8.14]. □

The propositions to follow are concerned with the generalised Cauchy problem (2.1–2.2) when A^j, B, f, and g are not generic generalised functions, but represent regularisations as in Lemma 2.20. Thus we consider the (formal) Cauchy problem

$$\partial_t u_0 + \sum_{j=1}^{n} A_0^j \partial_{x_j} u_0 + B_0 u_0 = f_0 \qquad \text{on } \Omega_T =]0,T[\times \mathbb{R}^n \tag{2.23}$$

$$u_0|_{t=0} = g_0, \tag{2.24}$$

where A_0^j (Hermitian), B_0, f_0, and g_0 are classical, non-smooth objects. We translate (2.23–2.24) to a generalised initial value problem (2.1–2.2) by means of regularising

A_0^j, B_0, f_0, and g_0 to obtain corresponding generalised functions A^j, B, f, and g. Then, our aim is to identify sufficient regularity conditions on A_0^j, B_0, f_0, and g_0, such that the unique generalised solution u has a distributional limit u_0, the regularity of which can be measured in terms of Sobolev norms. We refer to the regularity of u_0 as 'extrinsic regularity' of the generalised solution. Note that u_0 need not necessarily represent a solution of the original problem (2.23–2.24) within any classical concept of 'solution'.

The following Propositions 2.21, 2.25, 2.28 constitute a series of successive refinements of extrinsic regularity results, where the required regularity of the coefficients with respect to the time variable is more and more weakened, without producing a substantial loss of regularity of the distributional limit of the solution. Proposition 2.30 establishes a stronger convergence result for the generalised solution obtained in Proposition 2.25 by employing a refined energy estimate.

Subsequently, Propositions 2.32 and 2.34 are concerned with distributional shadows of the generalised solution in case of coefficients whose regularity with respect to the spatial variables is (slightly) below Lipschitz. Finally, Proposition 2.36 establishes a distributional shadow of the generalised solution even for lower-order coefficients proportional to a Dirac measure in the time variable, at the cost of additional structural assumptions on these coefficients.

Proposition 2.21 (Extrinsic regularity for non-smooth coefficients and data)**.** *In the initial value problem (2.23–2.24), assume that the components of A_0^j and B_0 belong to the space $\mathcal{C}^0([0,T], W^{1,\infty}(\mathbb{R}^n))$. Let initial data $g_0 \in H^1(\mathbb{R}^n)^m$ and right-hand side $f_0 \in L^2([0,T], H^1(\mathbb{R}^n))^m$ be given. Define corresponding generalised coefficients and data for the initial value problem (2.1–2.2), satisfying the assumptions in Theorem 2.13 (for instance by regularisation with a mollifier as in Lemma 2.20). Then there exists $u_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$ such that $u_\varepsilon \to u_0$ as $\varepsilon \to 0$ in the norm $L^\infty([0,T], L^2(\mathbb{R}^n))$ for every representative $(u_\varepsilon)_\varepsilon$ of the unique Colombeau solution $u \in \mathcal{G}_{L^2}(\Omega_T)^m$.*

Proof. First we verify that the ε-wise constructed family $(u_\varepsilon)_\varepsilon$ of solutions of the Cauchy problem (2.12–2.13) satisfies $\|u_\varepsilon\|_{H^1(\Omega_T)} = O(1)$ as $\varepsilon \to 0$. Using inequality (2.8), it was shown in the proof of Theorem 2.13 that

$$\sup_{0 \le t \le T} \|u_\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \le e^{\beta_\varepsilon}\left(\|g_\varepsilon\|^2_{L^2(\mathbb{R}^n)} + \|f_\varepsilon\|^2_{L^2(\Omega_T)}\right),$$

where $\beta_\varepsilon = \|\mathrm{div} A_\varepsilon - B_\varepsilon - B_\varepsilon^*\|_{L^{1,\infty}(\Omega_T)}$, implying that $\|u_\varepsilon\|_{L^2(\Omega_T)} = O(1)$ as $\varepsilon \to 0$ by virtue of the regularity assumptions on A_0^j and B_0 (see Lemma 2.20). This is true for all representatives of the generalised solution. Moreover, since

$$\sup_{0 \le t \le T} \|\nabla^1 u_\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \le e^{\beta^1_\varepsilon}\Big(\|\nabla^1 g_\varepsilon\|^2_{L^2(\mathbb{R}^n)} + \|(\widetilde{\nabla}^1 B_\varepsilon)\, \Sigma^1 u_\varepsilon + \nabla^1 f_\varepsilon\|^2_{L^2(\Omega_T)}\Big),$$

where $\beta^1_\varepsilon = 1 + \|\mathrm{div}\, \widetilde{\Sigma}^1 A_\varepsilon - \widetilde{B}^1_\varepsilon - (\widetilde{B}^1_\varepsilon)^*\|_{L^{1,\infty}(\Omega_T)}$, the uniform boundedness of $(u_\varepsilon)_\varepsilon$, $(\nabla^r g_\varepsilon)_\varepsilon$, $(\nabla^r f_\varepsilon)_\varepsilon$, and $(\widetilde{\nabla}^1 B_\varepsilon)_\varepsilon$ in their respective norms also yields $\|\partial_{x_j} u_\varepsilon\|_{L^2(\Omega_T)} = O(1)$ as $\varepsilon \to 0$. As an immediate consequence, $\|\partial_t u_\varepsilon\|_{L^2(\Omega_T)}$ must obey the same asymptotic estimate, since we have $\partial_t u_\varepsilon + \sum_{j=1}^n A^j_\varepsilon \partial_{x_j} u_\varepsilon + B_\varepsilon u_\varepsilon = f_\varepsilon$. We aim at showing that the family $(u_\varepsilon)_\varepsilon$ represents a Cauchy net in $\mathcal{C}^0([0,T],L^2(\mathbb{R}^n))^m$. With this in mind, we consider $P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon})$ for $0<\tilde\varepsilon<\varepsilon$ small enough and apply the energy estimate (2.8) to obtain

$$\sup_{0\le t\le T}\|u_\varepsilon(t,\cdot)-u_{\tilde\varepsilon}(t,\cdot)\|_{L^2(\mathbb{R}^n)} \le C_T\big(\|g_\varepsilon - g_{\tilde\varepsilon}\|_{L^2(\mathbb{R}^n)} + \|f_\varepsilon - P_\varepsilon u_{\tilde\varepsilon}\|_{L^2(\Omega_T)}\big),$$

where we have used that $|a|^2\le|b|^2+|c|^2$ implies $|a|\le|b|+|c|$, to get rid of the squares. The constant C_T does not depend on ε. We may write

$$\begin{aligned}P_\varepsilon u_{\tilde\varepsilon} = (P_\varepsilon - P_{\tilde\varepsilon} + P_{\tilde\varepsilon})u_{\tilde\varepsilon} &= \sum_{j=1}^n \big(A_\varepsilon^j - A_{\tilde\varepsilon}^j\big)\partial_{x_j}u_{\tilde\varepsilon} + \big(B_\varepsilon - B_{\tilde\varepsilon}\big)u_{\tilde\varepsilon} + P_{\tilde\varepsilon}u_{\tilde\varepsilon}\\ &= \sum_{j=1}^n \big(A_\varepsilon^j - A_{\tilde\varepsilon}^j\big)\partial_{x_j}u_{\tilde\varepsilon} + \big(B_\varepsilon - B_{\tilde\varepsilon}\big)u_{\tilde\varepsilon} + f_{\tilde\varepsilon}.\end{aligned}$$

Substituting this for $P_\varepsilon u_{\tilde\varepsilon}$ in the energy estimate, we find

$$\begin{aligned}\|u_\varepsilon - u_{\tilde\varepsilon}\|_{L^\infty([0,T],L^2(\mathbb{R}^n))} \le C_T\Big(&\|g_\varepsilon - g_{\tilde\varepsilon}\|_{L^2(\mathbb{R}^n)} + \|f_\varepsilon - f_{\tilde\varepsilon}\|_{L^2(\Omega_T)}\\ &+\sum_{j=1}^n\|A_\varepsilon^j - A_{\tilde\varepsilon}^j\|_{L^\infty(\Omega_T)}\|\partial_{x_j}u_{\tilde\varepsilon}\|_{L^2(\Omega_T)} + \|B_\varepsilon - B_{\tilde\varepsilon}\|_{L^\infty(\Omega_T)}\|u_{\tilde\varepsilon}\|_{L^2(\Omega_T)}\Big).\end{aligned}$$

The convergence properties of the nets $(g_\varepsilon)_\varepsilon$, $(f_\varepsilon)_\varepsilon$, $(A_\varepsilon^j)_\varepsilon$, $(B_\varepsilon)_\varepsilon$, and the uniform boundedness of $\|\partial_{x_j}u_{\tilde\varepsilon}\|_{L^2(\Omega_T)}$ imply that $(u_\varepsilon)_\varepsilon$ is a Cauchy net in the norm $L^\infty([0,T],L^2(\mathbb{R}^n))$, thus there exists a function $u_0\in\mathcal{C}([0,T],L^2(\mathbb{R}^n))^m$ such that $\|u_\varepsilon - u_0\|_{L^\infty([0,T],L^2(\mathbb{R}^n))}\to 0$ as $\varepsilon\to 0$. For any $t\in[0,T]$ and any $\varphi\in\mathcal{D}(\mathbb{R}^n)^m$, we have

$$|\langle\partial_{x_j}u_0(t,\cdot),\varphi\rangle| = |\lim_{\varepsilon\to 0}\langle\partial_{x_j}u_\varepsilon(t,\cdot),\varphi\rangle| \le \lim_{\varepsilon\to 0}|\langle\partial_{x_j}u_\varepsilon(t,\cdot),\varphi\rangle| \le C\|\varphi\|_{L^2(\mathbb{R}^n)},$$

hence $\partial_{x_j}u_0(t,\cdot)\in L^2(\mathbb{R}^n)$ and therefore $u_0\in\mathcal{C}^0([0,T],H^1(\mathbb{R}^n))$. In addition, for any $\varphi\in\mathcal{D}(\Omega_T)^m$ and ∂_k representing ∂_t or spatial derivatives ∂_{x_j}, we obtain

$$|\langle\partial_k u_0,\varphi\rangle| = |\lim_{\varepsilon\to 0}\langle\partial_k u_\varepsilon,\varphi\rangle| \le \lim_{\varepsilon\to 0}|\langle\partial_k u_\varepsilon,\varphi\rangle| \le C\|\varphi\|_{L^2(\mathbb{R}^n)}$$

since $\|\partial_k u_\varepsilon\|_{L^2(\Omega_T)}\le C$. Hence all first-order derivatives $\partial_k u_0$ belong to $L^2(\mathbb{R}^n)^m$ and therefore $u_0\in\mathcal{C}^0([0,T],H^1(\mathbb{R}^n))^m\cap H^1(\Omega_T)^m$. $\square$

The conditions on the coefficient matrices in Proposition 2.21 are typical assumptions in results on weak solutions. For a comparison of solution concepts for linear first-order hyperbolic differential equations with non-smooth coefficients we refer to [HH08].

Proposition 2.21 provides a distributional shadow of the generalised solution which is continuous in the time variable and enjoys a certain Sobolev space regularity with respect to the spatial variables. We may interpret this distributional shadow as a weak solution of the original first-order system with coefficients A_0^j and B_0. However, weak solutions for PDEs with non-smooth coefficients can also be obtained without Colombeau theory. Therefore it should be checked whether both solution concepts are compatible, that is, wether they yield 'the same' solution. The following Proposition establishes the desired consistency.

Proposition 2.22 (Coherence with weak solution concepts). *The initial value problem (2.23–2.24) with A_0^j, B_0, g_0, and f_0 as in Proposition 2.21 has a unique weak solution $u_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$, which is equal to the distributional limit of the generalised solution.*

Proof. We show that the distributional limit of the generalised solution $u = [(u_\varepsilon)_\varepsilon]$ is the unique weak solution.

First, observe that the limit u_0 of u_ε is continuous in time, hence it satisfies the initial condition $u_0\big|_{t=0} = g_0$. Moreover, since u_ε converges to u_0 in $L^2(\Omega_T)^m$, we also get the convergence $B_\varepsilon u_\varepsilon \to B_0 u_0$ in $L^2(\Omega_T)^m$ as $\varepsilon \to 0$. It remains to calculate the limits for the terms including derivatives of the generalised solution.

We claim that $\partial_t u_\varepsilon \to \partial_t u_0$ and $A_\varepsilon^j \partial_{x_j} u_\varepsilon \to A_0^j \partial_{x_j} u_0$ weakly$*$ in $L^2(\Omega_T)^m$ as $\varepsilon \to 0$. Since there exists $C > 0$ such that $\|u_\varepsilon\|_{H^1(\Omega_T)} \leq C$ for all $\varepsilon < \varepsilon_0$ (see the Proof of Proposition 2.21), we know that the set $\{u_{1/n} | n \in \mathbb{N}\}$ is bounded in $H^1(\Omega_T)^m$. By the weak compactness theorem, this implies that there exists a weakly$*$ convergent subsequence $(u_{1/n_k})_{k\in\mathbb{N}}$ with limit $\widetilde{u}_0 \in H^1(\Omega_T)^m$ (cf. [RR04, Theorem 6.64]). Therefore we have $\partial_t u_{1/n_k} \to \partial_t \widetilde{u}_0$ $(k \to \infty)$ weakly$*$ in $L^2(\Omega_T)^m$ and indeed $\widetilde{u}_0 = u_0$ since $u_{1/n_k} \to u_0$ in $L^2(\Omega_T)^m$ $(k \to \infty)$. Moreover, for any $\psi \in L^2(\Omega_T)^m$, we find

$$\begin{aligned}
\big|\langle A_{1/n_k}^j \partial_{x_j} u_{1/n_k}, \psi\rangle - \langle A_0^j\, \partial_{x_j} u_0, \psi\rangle\big| &= \big|\langle \partial_{x_j} u_{1/n_k}, (A_{1/n_k}^j - A_0^j)\psi\rangle\big| \\
&\leq \|\partial_{x_j} u_{1/n_k}\|_{L^2\Omega_T} \|(A_{1/n_k}^j - A_0^j)\psi\|_{L^2(\Omega_T)} \\
&\to 0 \quad (k \to \infty).
\end{aligned}$$

Hence, applying the operator $P_0 = \partial_t + \sum_{j=1}^n A_0^j \partial_{x_j} + B_0$ to the distributional shadow of the generalised solution, $u_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$, we obtain for any $\psi \in L^2(\Omega_T)^m$,

$$\langle P_0 u_0, \psi\rangle = \lim_{\varepsilon \to 0} \Big(\langle \partial_t u_\varepsilon, \psi\rangle + \sum_{j=1}^n \langle A_\varepsilon^j \partial_{x_j} u_\varepsilon, \psi\rangle + \langle B_\varepsilon u_\varepsilon, \psi\rangle \Big) = \lim_{\varepsilon \to 0} \langle f_\varepsilon, \psi\rangle = \langle f_0, \psi\rangle.$$

This in turn shows that u_0 is a weak solution of the initial value problem (2.23–2.24) with non-smooth coefficients. To prove uniqueness, suppose that there exists another solution $v_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$ such that $P_0 v_0 = f_0$ and $v_0|_{t=0} = g_0$. We may regularise this solution so that $v_\varepsilon \to v_0$ as $\varepsilon \to 0$ in $\mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m$ as well as in $H^1(\Omega_T)^m$. Then we must have $v_\varepsilon\big|_{t=0} \to g_0$ in $L^2(\mathbb{R}^n)^m$ as $\varepsilon \to 0$ and, moreover,

$$\lim_{\varepsilon \to 0} \Big(\partial_t v_\varepsilon + \sum_{j=1}^n A_\varepsilon^j \partial_{x_j} v_\varepsilon + B_\varepsilon v_\varepsilon \Big) = f_0$$

in $L^2(\Omega_T)^m$, irrespective of the explicit regularisations we choose for A_0, B_0, and v_0. To see this, consider otherwise arbitrary regularisations $A_\varepsilon^j \stackrel{\varepsilon \to 0}{\longrightarrow} A_0^j$ in $L^\infty(\Omega_T)$,

$B_\varepsilon \stackrel{\varepsilon\to 0}{\longrightarrow} B_0$ in $L^\infty(\Omega_T)$, and $v_\varepsilon \stackrel{\varepsilon\to 0}{\longrightarrow} v_0$ in $H^1(\Omega_T)$. Then we have

$$\begin{aligned}
& \|P_\varepsilon v_\varepsilon - P_0 v_0\|_{L^2(\Omega_T)} = \|P_\varepsilon v_\varepsilon - P_\varepsilon v_0 + P_\varepsilon v_0 - P_0 v_0\|_{L^2(\Omega_T)} \\
\leq\ & \|\partial_t(v_\varepsilon - v_0)\|_{L^2(\Omega_T)} + \|A^j_\varepsilon(\partial_{x_j} v_\varepsilon - \partial_{x_j} v_0)\|_{L^2(\Omega_T)} + \|(A^j_\varepsilon - A^j_0)\partial_{x_j} v_0\|_{L^2(\Omega_T)} \\
& + \|B_\varepsilon(v_\varepsilon - v_0)\|_{L^2(\Omega_T)} + \|(B_\varepsilon - B_0)v_0\|_{L^2(\Omega_T)} \\
\leq\ & \|v_\varepsilon - v_0\|_{H^1(\Omega_T)} + \|A^j_\varepsilon\|_{L^\infty(\Omega_T)}\|v_\varepsilon - v_0\|_{H^1(\Omega_T)} + \|A^j_\varepsilon - A^j_0\|_{L^\infty(\Omega_T)}\|v_0\|_{L^2(\Omega_T)} \\
& + \|B_\varepsilon\|_{L^\infty(\Omega_T)}\|v_\varepsilon - v_0\|_{L^2(\Omega_T)} + \|B_\varepsilon - B_0\|_{L^\infty(\Omega_T)}\|v_0\|_{L^2(\Omega_T)} \\
\to\ & 0 \quad (\varepsilon \to 0)
\end{aligned}$$

and therefore $P_\varepsilon v_\varepsilon \stackrel{\varepsilon\to 0}{\longrightarrow} f_0$ in $L^2(\Omega_T)$. Denoting by $(u_\varepsilon)_\varepsilon$ the representative of the generalised solution constructed in the proof of Theorem 2.13 and applying the basic energy estimate (2.8) to the difference $u_\varepsilon - v_\varepsilon$ yields

$$\sup_{0\leq t\leq T} \|u_\varepsilon(t,\cdot) - v_\varepsilon(t,\cdot)\|_{L^2(\mathbb{R}^n)} \leq C_T\Big(\|g_\varepsilon - v_\varepsilon|_{t=0}\|_{L^2(\mathbb{R}^n)} + \|f_\varepsilon - P_\varepsilon v_\varepsilon\|_{L^2(\Omega_T)}\Big),$$

which converges to zero as $\varepsilon \to 0$. Thus we have $u_0 = v_0$ in $\mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m$. □

Remark 2.23 (Weak solutions via generalised solutions)**.** Note that the statement in Proposition 2.22 is not just a compatibility result, in fact we directly obtain a unique weak solution of the initial value problem (2.23–2.24) with non-smooth coefficients, merely by studying properties of the generalised solution when the coefficients and data arise as regularisations of Sobolev functions. In case of smooth coefficients, the methods applied in the proofs of Proposition 2.21 and Proposition 2.22 would lead to the corresponding classical existence and uniqueness results as well (see Propositions 2.17 and 2.18).

The asymptotic estimates for the generalised coefficients demanded in Theorems 2.7, 2.8, and 2.13 are substantially weaker than those obtained for the regularised versions of the Sobolev-type coefficients used in Proposition 2.21. This is a typical situation for existence and uniqueness results for PDEs in Colombeau algebras. While generalised solutions frequently exist for much rougher coefficients and initial data than those required in corresponding classical results on weak solutions, distributional limits of the generalised solution can often only be shown to exist for coefficients whose regularity is similar or equal to the regularity required in results within other weak solution concepts. In this sense, Colombeau theory, albeit very successful in many circumstances, does not always provide the means to considerably extend results on distributional solutions beyond what is already known. However, not being able to show that the generalised solution has a distributional shadow does not mean that a distributional limit does not exist. In any case, there is almost always a regime of coefficient regularities, for which generalised solutions exist, but the existence of a distributional limit of the solution is unclear.

The generalised solution $[(u_\varepsilon)_\varepsilon]$ obtained in Theorem 2.13 satisfies $\|u_\varepsilon\|_{L^2(\Omega_T)} = O(1)$ as $\varepsilon \to 0$ even for coefficients with components of the form $\delta(t-t_0)\otimes h(x_1, ..., x_n)$ with $h \in W^{1,\infty}(\mathbb{R}^n)$. This suggests that there might exist a distributional shadow of the solution even for regularity assumptions which are significantly weaker than those in Proposition 2.21. However, to show that $(u_\varepsilon)_\varepsilon$ is a Cauchy net and therefore possesses a weak limit, asymptotic boundedness of the coefficients was needed. Yet this is mainly because the corresponding energy estimates relied on L^2-norms of u_ε and Pu_ε. By inspecting the proof of Theorem 2.13, it becomes clear that as soon as asymptotic boundedness of the solution is established, there is actually no need to keep working with the same energy estimates that led to the existence of a solution satisfying certain asymptotic estimates. In particular, there is no need to stick to using L^2-norms of Pu_ε. Lemma 2.24 gathers some alternative estimates involving mixed norms, which will be the basis for a variant of Proposition 2.21, providing a distributional limit of the generalised solution for matrices A_0^j and B_0 belonging to $M_m(L^2([0,T],W^{1,\infty}(\mathbb{R}^n)))$. This will enable us to consider coefficient matrices with entries proportional to $H(t-t_0)\otimes h(x_1,...,x_n)$, where H denotes the Heaviside function and $h \in W^{1,\infty}(\mathbb{R}^n)$.

Lemma 2.24 (Energy estimates with mixed norms). *Let $P = \partial_t + \sum_{j=1}^n A^j\partial_{x_j} + B$ be a symmetric hyperbolic PDO and $u, w \in \mathcal{C}^1([0,T], L^2(\mathbb{R}^n))^m \cap \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m$. For $1 \le p, q \le \infty$, let*

$$\|A\|_{L^{p,q}(\Omega_T)} := \big\|s \mapsto \|A(s,\cdot)\|_{L^q(\mathbb{R}^n)}\big\|_{L^p([0,T])}.$$

Then we have

(i) $$\|u\|^2_{L^{\infty,2}(\Omega_T)} \le e^{\int_0^T \beta(s)\mathrm{d}s}\Big(\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + 2\,\mathrm{Re}\int_0^T \langle Pu, u\rangle_{L^2(\mathbb{R}^n)}(s)\mathrm{d}s\Big).$$

(ii) $$\|\langle Bu, w\rangle\|_{L^1(\Omega_T)} \le \|B\|_{L^{1,\infty}(\Omega_T)}\Big(\|u\|^2_{L^{\infty,2}(\Omega_T)} + \|w\|^2_{L^{\infty,2}(\Omega_T)}\Big).$$

(iii) $$\|\langle A^i\partial_{x_j}u, w\rangle\|_{L^1(\Omega_T)} \le \|A^i\|_{L^{1,\infty}(\Omega_T)}\Big(\|\partial_{x_j}u\|^2_{L^{\infty,2}(\Omega_T)} + \|w\|^2_{L^{\infty,2}(\Omega_T)}\Big).$$

Proof. The proof of (i) follows along the lines of the proof of inequaltiy (2.8) in Lemma 2.6 until we integrate equation (2.11) over the domain Ω_T. From here on, we follow a slightly different route regarding the integral of $2\mathrm{Re}\langle Pu, u\rangle$. Instead of estimating this term by

$$\Big|2\int_0^T \mathrm{Re}\int_{\mathbb{R}^n}\langle Pu(s,\cdot), u(s,\cdot)\rangle\,\mathrm{d}V\mathrm{d}s\Big| \le \|Pu\|^2_{L^2(\Omega_T)} + \|u\|^2_{L^2(\Omega_T)},$$

we directly apply Gronwall's lemma to obtain

$$\|u\|^2_{L^{\infty,2}(\Omega_T)} \le e^{\int_0^T \beta(s)\mathrm{d}s}\Big(\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + 2\,\mathrm{Re}\int_0^T \langle Pu, u\rangle_{L^2(\mathbb{R}^n)}(s)\mathrm{d}s\Big).$$

To prove (ii), we consider

$$
\begin{aligned}
2\int_0^T\int_{\mathbb{R}^n} \|Bu(s,\cdot)\|\,\|w(s,\cdot)\|\,\mathrm{d}V\mathrm{d}s &\leq \int_0^T \|B(s,\cdot)\|_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n}\big(\|u(s,\cdot)\|^2 + \|w(s,\cdot)\|^2\big)\mathrm{d}V\mathrm{d}s \\
&\leq \|B\|_{L^{1,\infty}(\Omega_T)}\Big(\|u\|^2_{L^{\infty,2}(\Omega_T)} + \|w\|^2_{L^{\infty,2}(\Omega_T)}\Big),
\end{aligned}
$$

where we have used that $|ab| \leq \frac{1}{2}(|a|^2+|b|^2)$. Finally, assertion (iii) is proved in a similar fashion, that is,

$$
\begin{aligned}
&2\int_0^T\int_{\mathbb{R}^n} \|A^i\,\partial_{x_j}u(s,\cdot)\|\,\|w(s,\cdot)\|\,\mathrm{d}V\mathrm{d}s \\
&\leq \int_0^T \|A^i(s,\cdot)\|_{L^\infty(\mathbb{R}^n)}\Big(\|\partial_{x_j}u(s,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \|w(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\Big)\mathrm{d}s \\
&\leq \|A^i\|_{L^{1,\infty}(\Omega_T)}\Big(\sup_{0\leq s\leq T}\|\partial_{x_j}u(s,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \sup_{0\leq s\leq T}\|w(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\Big).
\end{aligned}
$$

□

At first sight, the estimate in Lemma 2.24 (i) seems to be futile, if we are interested in bounds of $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)}$ (or any other norm of u_ε on Ω_T) in terms of initial data and right-hand side of the PDE system (for instance to show moderateness). We keep the term $\langle Pu, u\rangle_{L^2(\mathbb{R}^n)}$ unaltered throughout the estimate and thus do not get rid of u on the right-hand side of the inequality. However, if boundedness of the generalised solution has already been established by some other method, this inequality turns out to be very valuable. The crucial point is that the term $\int_0^T \langle Pu, u\rangle_{L^2(\mathbb{R}^n)}(s)\,\mathrm{d}s$ in the Cauchy net estimate (see the proof of Proposition 2.21) can now be estimated using L^1-norms of the coefficient matrices with respect to the time variable, instead of L^∞-norms, thereby leading to expressions as on the left-hand sides of inequalities (ii) and (iii) in Lemma 2.24. In fact, since we already have a representative $(u_\varepsilon)_\varepsilon$ of the generalised solution with certain uniform boundedness properties at hand, we can employ Lemma 2.24 to prove convergence of the generalised solution in $L^{\infty,2}(\Omega_T)$.

Proposition 2.25 (Extrinsic regularity, coefficients L^2 in the time variable)**.** *In the initial value problem (2.23–2.24), assume that the components of A_0^j and B_0 belong to the space $L^2([0,T], W^{1,\infty}(\mathbb{R}^n))$. Let initial data $g_0 \in H^1(\mathbb{R}^n)^m$ and right-hand side $f_0 \in L^2([0,T], H^1(\mathbb{R}^n))^m$ be given. Define corresponding generalised coefficients and data for the initial value problem (2.1–2.2), satisfying the assumptions in Theorem 2.13. Then there exists $u_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$ such that $u_\varepsilon \to u_0$ as $\varepsilon \to 0$ in the norm $L^\infty([0,T], L^2(\mathbb{R}^n))$ for every representative $(u_\varepsilon)_\varepsilon$ of the unique Colombeau solution $u \in \mathcal{G}_{L^2}(\Omega_T)^m$.*

Proof. Let $(u_\varepsilon)_\varepsilon$ denote the ε-wise constructed family $(u_\varepsilon)_\varepsilon$ of solutions to the Cauchy problem (2.12–2.13). First, we have $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$ by means of inequality (2.8),

$$\sup_{0\le t\le T}\|u_\varepsilon(t,\cdot)\|_{L^2(\mathbb{R}^n)} \le e^{\beta_\varepsilon/2}\Big(\|g_\varepsilon\|_{L^2(\mathbb{R}^n)} + \|f_\varepsilon\|_{L^2(\Omega_T)}\Big),$$

where $\beta_\varepsilon = \|\mathrm{div}A_\varepsilon - B_\varepsilon - B_\varepsilon^*\|_{L^{1,\infty}(\Omega_T)}$. Now we turn our attention to the spatial derivatives of u_ε, for which

$$\begin{aligned}\|\nabla^1 u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} &\le e^{\beta^1_\varepsilon/2}\Big(\|\nabla^1 g_\varepsilon\|_{L^2(\mathbb{R}^n)} + \|(\widetilde{\nabla}^1 B_\varepsilon)\,\Sigma^1 u_\varepsilon + \nabla^1 f_\varepsilon\|_{L^2(\Omega_T)}\Big)\\ &\le e^{\beta^1_\varepsilon/2}\Big(\|\nabla^1 g_\varepsilon\|_{L^2(\mathbb{R}^n)} + \|(\widetilde{\nabla}^1 B_\varepsilon)\,\Sigma^1 u_\varepsilon\|_{L^2(\Omega_T)} + \|\nabla^1 f_\varepsilon\|_{L^2(\Omega_T)}\Big),\end{aligned}$$

where $\|\nabla^1 g_\varepsilon\|_{L^2(\mathbb{R}^n)}, \|\nabla^1 f_\varepsilon\|_{L^2(\Omega_T)} = O(1)$ as $\varepsilon\to 0$ by Lemma 2.20. Recalling that $\beta^1_\varepsilon = 1 + \|\mathrm{div}\,\widetilde{\Sigma}^1 A_\varepsilon - \widetilde{B}^1_\varepsilon - (\widetilde{B}^1_\varepsilon)^*\|_{L^{1,\infty}(\Omega_T)}$ and $\widetilde{B}^1_\varepsilon = \widetilde{\Sigma}^1 B_\varepsilon + (\partial_{x_i} A^j_\varepsilon)_{1\le i,j\le n}$, we conclude that $e^{\beta^1_\varepsilon/2} = O(1)$ as $\varepsilon\to 0$ as well. Finally, we inspect the only remaining term $\|(\widetilde{\nabla}^1 B_\varepsilon)\,\Sigma^1 u_\varepsilon\|_{L^2(\Omega_T)}$ more closely and estimate

$$\begin{aligned}\|(\widetilde{\nabla}^1 B_\varepsilon)\,\Sigma^1 u_\varepsilon\|^2_{L^2(\Omega_T)} &= \int_0^T \|(\widetilde{\nabla}^1 B_\varepsilon(s,\cdot))\,\Sigma^1 u_\varepsilon(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\mathrm{d}s\\ &\le \int_0^T \|\widetilde{\nabla}^1 B_\varepsilon(s,\cdot)\|^2_{L^\infty(\mathbb{R}^n)}\|\Sigma^1 u_\varepsilon(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\mathrm{d}s\\ &\le \sup_{0\le s\le T}\|\Sigma^1 u_\varepsilon(s,\cdot)\|^2_{L^2(\mathbb{R}^n)}\int_0^T \|\widetilde{\nabla}^1 B_\varepsilon(s,\cdot)\|^2_{L^\infty(\mathbb{R}^n)}\mathrm{d}s\\ &\le C\,\|u_\varepsilon\|^2_{L^{\infty,2}(\Omega_T)}\|\widetilde{\nabla}^1 B_\varepsilon\|^2_{L^{2,\infty}(\Omega_T)},\end{aligned}$$

implying that $\|\nabla^1 u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon\to 0$. Next we want to show that the family $(u_\varepsilon)_\varepsilon$ represents a Cauchy net in $L^\infty([0,T],L^2(\mathbb{R}^n))^m$. Considering $P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon})$ for $0<\tilde\varepsilon<\varepsilon$ small enough and applying Lemma 2.24 (i), we obtain

$$\|u_\varepsilon - u_{\tilde\varepsilon}\|^2_{L^{\infty,2}(\Omega_T)} \le C\Big(\|g_\varepsilon - g_{\tilde\varepsilon}\|^2_{L^2(\mathbb{R}^n)} + 2\,\|\langle P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon}), u_\varepsilon - u_{\tilde\varepsilon}\rangle\|_{L^1(\Omega_T)}\Big),$$

where C is independent of ε. Since $P_\varepsilon u_\varepsilon = f_\varepsilon$ and recalling that we can write $P_\varepsilon u_{\tilde\varepsilon} = \sum_{j=1}^n \big(A^j_\varepsilon - A^j_{\tilde\varepsilon}\big)\partial_{x_j} u_{\tilde\varepsilon} + \big(B_\varepsilon - B_{\tilde\varepsilon}\big)u_{\tilde\varepsilon} + f_{\tilde\varepsilon}$, we estimate the integrand in the last term by

$$|\langle P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon}), u_\varepsilon - u_{\tilde\varepsilon}\rangle| = \Big|\langle f_\varepsilon - \sum_{j=1}^n \big(A^j_\varepsilon - A^j_{\tilde\varepsilon}\big)\partial_{x_j}u_{\tilde\varepsilon} - \big(B_\varepsilon - B_{\tilde\varepsilon}\big)u_{\tilde\varepsilon} - f_{\tilde\varepsilon}, u_\varepsilon - u_{\tilde\varepsilon}\rangle\Big|$$

$$\le |\langle f_\varepsilon - f_{\tilde\varepsilon}, u_\varepsilon - u_{\tilde\varepsilon}\rangle| + \sum_{j=1}^n |\langle A^j_\varepsilon - A^j_{\tilde\varepsilon}\big)\partial_{x_j}u_{\tilde\varepsilon}, u_\varepsilon - u_{\tilde\varepsilon}\rangle| + |\langle (B_\varepsilon - B_{\tilde\varepsilon})u_{\tilde\varepsilon}, u_\varepsilon - u_{\tilde\varepsilon}\rangle|.$$

Using this representation of the integrand and applying Lemma 2.24, (ii) and (iii), we get

$$
\begin{aligned}
\|u_\varepsilon - u_{\tilde\varepsilon}\|^2_{L^{\infty,2}(\Omega_T)} \le C\Bigg(& \|g_\varepsilon - g_{\tilde\varepsilon}\|^2_{L^2(\mathbb{R}^n)} + \|f_\varepsilon - f_{\tilde\varepsilon}\|_{L^2(\Omega_T)}\|u_\varepsilon - u_{\tilde\varepsilon}\|_{L^2(\Omega_T)} \\
& + \sum_{j=1}^n \|A^j_\varepsilon - A^j_{\tilde\varepsilon}\|_{L^{1,\infty}(\Omega_T)}\Big(\|\partial_{x_j} u_{\tilde\varepsilon}\|^2_{L^{\infty,2}(\Omega_T)} + \|u_\varepsilon - u_{\tilde\varepsilon}\|^2_{L^2(\Omega_T)}\Big) \\
& \qquad + \|B_\varepsilon - B_{\tilde\varepsilon}\|_{L^{1,\infty}(\Omega_T)}\Big(\|u_{\tilde\varepsilon}\|^2_{L^{\infty,2}(\Omega_T)} + \|u_\varepsilon - u_{\tilde\varepsilon}\|^2_{L^{\infty,2}(\Omega_T)}\Big)\Bigg).
\end{aligned}
$$

The convergence properties of the nets $(g_\varepsilon)_\varepsilon$, $(f_\varepsilon)_\varepsilon$, $(A^j_\varepsilon)_\varepsilon$, $(B_\varepsilon)_\varepsilon$ and the uniform boundedness of $\|\partial_{x_j} u_{\tilde\varepsilon}\|_{L^2(\Omega_T)}$ as $\tilde\varepsilon \to 0$ imply that $(u_\varepsilon)_\varepsilon$ is a Cauchy net in the norm $L^\infty([0,T], L^2(\mathbb{R}^n))$. Thus there exists a function $u_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m$ such that $\|u_\varepsilon - u_0\|_{L^\infty([0,T],L^2(\mathbb{R}^n))} \to 0$ as $\varepsilon \to 0$. For any $t \in [0,T]$ and any $\varphi \in \mathcal{D}(\mathbb{R}^n)^m$, we have

$$
|\langle \partial_{x_j} u_0(t,\cdot), \varphi\rangle| = |\lim_{\varepsilon\to 0}\langle \partial_{x_j} u_\varepsilon(t,\cdot), \varphi\rangle| \le \lim_{\varepsilon\to 0} |\langle \partial_{x_j} u_\varepsilon(t,\cdot), \varphi\rangle| \le C\|\varphi\|_{L^2(\mathbb{R}^n)},
$$

hence $\partial_{x_j} u_0(t,\cdot) \in L^2(\mathbb{R}^n)$ and therefore $u_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))$. In addition, we have

$$
\begin{aligned}
\|\partial_t u_\varepsilon\|_{L^2(\Omega_T)} &\le \sum_{j=1}^n \|A^j_\varepsilon \partial_{x_j} u_\varepsilon\|_{L^2(\Omega_T)} + \|B_\varepsilon u_\varepsilon\|_{L^2(\Omega_T)} + \|f_\varepsilon\|_{L^2(\Omega_T)} \\
&\le \sum_{j=1}^n \|A^j_\varepsilon\|_{L^{2,\infty}(\Omega_T)}\|\partial_{x_j} u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} + \|B_\varepsilon\|_{L^{2,\infty}(\Omega_T)}\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} + \|f_\varepsilon\|_{L^2(\Omega_T)}.
\end{aligned}
$$

Therefore, since $\|A^j_\varepsilon\|_{L^{2,\infty}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$, we have $\partial_t u_0 \in L^2(\Omega_T)$ as well. Considering the distributional action of $\partial_j u_0$ on a test function $\varphi \in \mathcal{D}(\Omega_T)$, we obtain $u_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$. □

Remark 2.26 (Weaker regularity assumptions are possible)**.** In Proposition 2.25, using the slightly weaker assumption $A^j_0 \in M_m(L^1([0,T], W^{1,\infty}(\mathbb{R}^n)))$ instead of $A^j_0 \in M_m(L^2([0,T], W^{1,\infty}(\mathbb{R}^n)))$, we would still obtain a distributional limit $u_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m$ with $\partial_{x_j} u_0 \in L^2(\Omega_T)^m$, but the time derivative $\partial_t u_0$ would not necessarily belong to $L^2(\Omega_T)^m$. However, the proof of Proposition 2.25 requires B_0 to belong to $M_m(L^2([0,T], W^{1,\infty}(\mathbb{R}^n)))$, in order to get asymptotic boundedness of $(\nabla^1 u_\varepsilon)_\varepsilon$, so we cannot work with $B_0 \in M_m(L^1([0,T], W^{1,\infty}(\mathbb{R}^n)))$.

With the regularity conditions we imposed on the coefficient matrices A^j_0 and B_0 in Proposition 2.25, we could still interpret the PDE system $P_0 u_0 := \big(\partial_t + \sum_{j=1}^n A^j_0 \partial_{x_j} + B_0\big) u_0 = f_0$ in a weak sense without having to transfer the problem to a differential algebra of generalised functions. For example, since u_0 is continuous in the time variable, we can make sense of products of $u_0(t,x)$ and the Heaviside function $H(t - t_0)$ in the space of distributions. Naturally, it would be nicer to obtain a distributional shadow of the generalised solution for regularity assumptions which do not easily allow for a weak interpretation of the formal expression $P_0 u_0$. However, the energy estimate we used so far (see inequality (ii) in Lemma 2.6) is essentially an estimate of the $L^{\infty,2}$-norm of the solution. But since the limit of a net of continuous functions converging in the L^∞-norm is continuous, it is clear that if we obtain a distributional limit of the generalised solution by using this estimate, the limit must automatically be continuous. Yet if we want to allow for coefficients whose regularity with respect to the time variable is below L^1_{loc}, we cannot expect the limit of the solution (if it exists) to be continuous. This implies that for coefficient regularities below L^1_{loc}, it is impossible to show that the solution has a distributional limit, if our method relies on the the basic estimate (ii) in Lemma 2.6, even if a limit exists. The content of the following Lemma 2.27 is a 'refined' energy estimate, which is coarser with respect to the time variable, preparing the stage for another variant of Proposition 2.21.

Lemma 2.27 (Adapted energy estimates)**.** *Let $P = \partial_t + \sum_{j=1}^n A^j \partial_{x_j} + B$ be a symmetric hyperbolic PDO and $u, w \in \mathcal{C}^1([0,T], L^2(\mathbb{R}^n))^m \cap \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m$. Then, for $\gamma(\tau) := \sup\limits_{0 \le s \le \tau} \|\mathrm{div}A(s,\cdot) - B(s,\cdot) - B^*(s,\cdot)\|_{L^\infty(\mathbb{R}^n)}$, we have*

$$\text{(i)}\ \ \|u\|^2_{L^2(\Omega_T)} \le e^{\int_0^T \gamma(\tau)\mathrm{d}\tau} \Big(T\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + 2\mathrm{Re} \int_0^T \int_0^\tau \langle Pu(s,\cdot), u(s,\cdot)\rangle_{L^2(\mathbb{R}^n)} \mathrm{d}s\, \mathrm{d}\tau \Big),$$

$$\text{(ii)}\ \ \|u\|^2_{L^2(\Omega_T)} \le T e^{T + \int_0^T \gamma(\tau)\mathrm{d}\tau} \Big(\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \|Pu\|^2_{L^2(\Omega_T)} \Big),$$

$$\text{(iii)}\ \ 2\mathrm{Re} \int_0^T \int_0^\tau \Big(\sum_{j=1}^n \langle A^j \partial_{x_j} u, w\rangle_{L^2(\mathbb{R}^n)}(s) + \langle Bu, w\rangle_{L^2(\mathbb{R}^n)}(s) \Big) \mathrm{d}s\, \mathrm{d}\tau$$

$$\le \Big(\|\partial_{x_j} u\|^2_{L^{\infty,2}(\Omega_T)} + \|w\|^2_{L^{\infty,2}(\Omega_T)} \Big) \int_0^T \int_0^\tau \|A^j(s,\cdot)\|_{L^\infty(\mathbb{R}^n)} \mathrm{d}s\, \mathrm{d}\tau$$
$$+ \Big(\|u\|^2_{L^{\infty,2}(\Omega_T)} + \|w\|^2_{L^{\infty,2}(\Omega_T)} \Big) \int_0^T \int_0^\tau \|B(s,\cdot)\|_{L^\infty(\mathbb{R}^n)} \mathrm{d}s\, \mathrm{d}\tau.$$

Proof. Starting from equation (2.11), integration over the spatial domain $\mathbb{R}^n$ yields

$$\frac{d}{d\tau} \|u(\tau,\cdot)\|^2_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \langle (\mathrm{div}A - B - B^*)u, u\rangle \, \mathrm{d}V_n \, + \, 2\mathrm{Re} \int_{\mathbb{R}^n} \langle Pu, u\rangle \, \mathrm{d}V_n,$$

where $\tau \in [0,T]$. With $\beta(s) := 1 + \|\mathrm{div}A(s,\cdot) - B(s,\cdot) - B^*(s,\cdot)\|_{L^\infty(\mathbb{R}^n)} \ge 1$ as in Lemma 2.6 and setting $v(\tau) := \|u(\tau,\cdot)\|^2_{L^2(\mathbb{R}^n)}$, we integrate the last equation over the interval $[0,\tau]$ and get

$$v(\tau) - v(0) \le \int_0^\tau (\beta(s) - 1) v(s) \mathrm{d}s \, + 2\mathrm{Re} \int_0^\tau \langle Pu(s,\cdot), u(s,\cdot)\rangle_{L^2(\mathbb{R}^n)} \mathrm{d}s.$$

Denoting $V(t) := \int_0^t v(\tau)d\tau$, we deliberately integrate once more and find

$$V(t) \leq v(0)t + \int_0^t \int_0^\tau (\beta(s)-1)v(s)\mathrm{d}s\,\mathrm{d}\tau + 2\mathrm{Re}\int_0^t \int_0^\tau \langle Pu(s,\cdot), u(s,\cdot)\rangle_{L^2(\mathbb{R}^n)}\mathrm{d}s\,\mathrm{d}\tau.$$

Before using Gronwall's lemma to get an upper bound for $V(t)$, we estimate the terms in the last inequality as follows:

$$V(t) \leq v(0)T + \int_0^t \gamma(\tau)V(\tau)\mathrm{d}\tau \; + 2\mathrm{Re}\int_0^T \int_0^\tau \langle Pu(s,\cdot), u(s,\cdot)\rangle_{L^2(\mathbb{R}^n)}\mathrm{d}s\,\mathrm{d}\tau$$

where $\tau \mapsto \gamma(\tau) := \sup_{0\leq s\leq \tau}(\beta(s)-1) \geq 0$ is continuous, if β is. Now applying Gronwall's inequality, Lemma 1.4 (ii), we arrive at

$$\|u\|^2_{L^2(\Omega_T)} \leq e^{\int_0^T \gamma(\tau)\mathrm{d}\tau}\left(v(0)T + 2\mathrm{Re}\int_0^T \int_0^\tau \langle Pu(s,\cdot), u(s,\cdot)\rangle_{L^2(\mathbb{R}^n)}\mathrm{d}s\,\mathrm{d}\tau \right).$$

To prove (ii), we estimate

$$2\mathrm{Re}\int_0^\tau \langle Pu(s,\cdot), u(s,\cdot)\rangle_{L^2(\mathbb{R}^n)}\mathrm{d}s \leq \|u\|^2_{L^2(\Omega_\tau)} + \|Pu\|^2_{L^2(\Omega_\tau)}$$

just before the second integration with respect to the time variable, then apply Gronwall's inequality, and end up with

$$\|u\|^2_{L^2(\Omega_T)} \leq Te^{T+\int_0^T \gamma(\tau)d\tau}\Big(\|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)} + \|Pu\|^2_{L^2(\Omega_T)}\Big).$$

Finally, the proof of (iii) is similar to the proof of inequalities (ii) and (iii) in Lemma 2.24. ☐

Using the refined estimates of Lemma 2.27, the required regularity assumptions for the matrices A_0 and B_0 can be slightly weakened without losing the distributional limit of the solution.

Proposition 2.28 (Extrinsic regularity, coefficients L^1 in the time variable)**.** *In the initial value problem (2.23–2.24), assume that the components of A_0^j and B_0 belong to the space $L^1([0,T], W^{1,\infty}(\mathbb{R}^n))$, and that $g_0 \in H^1(\mathbb{R}^n)^m$ and $f_0 \in L^1([0,T], H^1(\mathbb{R}^n))^m$. Define corresponding generalised coefficients and data for the initial value problem (2.1–2.2), satisfying the assumptions in Theorem 2.13. Then there exists a function $u_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m \cap W^{1,1}(\Omega_T)^m$ such that $u_\varepsilon \to u_0$ as $\varepsilon \to 0$ in the norm $L^\infty([0,T], L^2(\mathbb{R}^n))$ for every representative $(u_\varepsilon)_\varepsilon$ of the unique generalised solution $u \in G_{L^2}(\Omega_T)^m$.*

Proof. The coefficients fulfil assumptions (1") and (2") in Theorem 2.13 and we immediately obtain a unique generalised solution $u \in \mathcal{G}_{L^2}(\Omega_T)^m$ with $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} =$

$O(1)$ as $\varepsilon \to 0$. As was shown in Theorem 2.7, the spatial derivative vector $\nabla^1 u_\varepsilon$ satisfies the PDE system

$$\partial_t \nabla^1 u_\varepsilon + \sum_{j=1}^{n} (\widetilde{\Sigma}^1 A^j_\varepsilon)\, \partial_{x_j} \nabla^1 u_\varepsilon + \widetilde{B}^1_\varepsilon \nabla^1 u_\varepsilon = -(\widetilde{\nabla}^1 B_\varepsilon)\Sigma^1 u_\varepsilon + \nabla^1 f_\varepsilon,$$

where $\widetilde{B}^1_\varepsilon = \widetilde{\Sigma}^1 B_\varepsilon + (\partial_{x_i} A^j_\varepsilon)_{1\le i,j\le n}$. We now use the alternative energy estimate in Lemma 2.24 (i) and find

$$\begin{aligned}\|\nabla^1 u_\varepsilon\|^2_{L^{\infty,2}(\Omega_T)} \le{}& e^{\|\mathrm{div}(\widetilde{\Sigma}^1 A_\varepsilon) - \widetilde{B}^1_\varepsilon - (\widetilde{B}^1_\varepsilon)^*\|_{L^{1,\infty}(\Omega_T)}} \Bigg(\|\nabla^1 g_\varepsilon\|^2_{L^2(\mathbb{R}^n)} \\ &+ 2\mathrm{Re} \int_0^T \langle -(\widetilde{\nabla}^1 B_\varepsilon(s,\cdot))\Sigma^1 u_\varepsilon(s,\cdot) + \nabla^1 f_\varepsilon(s,\cdot), \nabla^1 u_\varepsilon(s,\cdot)\rangle_{L^2(\mathbb{R}^n)} \mathrm{d}s \Bigg).\end{aligned}$$

This implies that there exists a constant $C > 0$ such that

$$\begin{aligned}\|\nabla^1 u_\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \le C \Bigg(&\|g_\varepsilon\|^2_{H^1(\mathbb{R}^n)} \\ &+ \int_0^T \Big(\|(\widetilde{\nabla}^1 B_\varepsilon(s,\cdot))\Sigma^1 u_\varepsilon(s,\cdot)\|_{L^2(\mathbb{R}^n)} + \|f_\varepsilon(s,\cdot)\|_{H^1(\mathbb{R}^n)} \Big) \|\nabla^1 u_\varepsilon(s,\cdot)\|_{L^2(\mathbb{R}^n)} \mathrm{d}s \Bigg).\end{aligned}$$

We may now apply Gronwall's inequality as in Lemma 1.4 (ii) to obtain

$$\begin{aligned}\|\nabla^1 u_\varepsilon\|^2_{L^{\infty,2}(\Omega_T)} &\le C' \|g_\varepsilon\|^2_{H^1(\mathbb{R}^n)} \\ &\quad \cdot \exp\Bigg(\int_0^T \Big(\|(\widetilde{\nabla}^1 B_\varepsilon(s,\cdot))\Sigma^1 u_\varepsilon(s,\cdot)\|_{L^2(\mathbb{R}^n)} + \|f_\varepsilon(s,\cdot)\|_{H^1(\mathbb{R}^n)} \Big) \mathrm{d}s \Bigg) \\ &\le C'' \|g_\varepsilon\|^2_{H^1(\mathbb{R}^n)} \cdot \exp\Big(\|\widetilde{\nabla}^1 B_\varepsilon\|_{L^{1,\infty}(\Omega_T)} \|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} + \|f_\varepsilon(s,\cdot)\|_{L^1([0,T],H^1(\mathbb{R}^n))} \Big)\end{aligned}$$

which is uniformly bounded as $\varepsilon \to 0$, since we already have $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$ and, by the regularity assumptions for f_0, g_0, and B_0, the same is true for the other norms on the right-hand side. Thus we have shown that $\|\nabla^1 u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$.

Aiming at a Cauchy net argument to show the existence of a distributional shadow of the generalised solution of $[(u_\varepsilon)_\varepsilon]$, we will now make use of the refined energy estimate in Lemma 2.27 (i) to get

$$\begin{aligned}\|u_\varepsilon - u_{\tilde\varepsilon}\|^2_{L^2(\Omega_T)} \le{}& \exp\Bigg(\int_0^T \sup_{0\le s\le\tau} \|(\mathrm{div} A_\varepsilon - B_\varepsilon - B^*_\varepsilon)(s)\|_{L^\infty(\mathbb{R}^n)} \mathrm{d}\tau \Bigg) \\ &\cdot \Bigg(T\|g_\varepsilon - g_{\tilde\varepsilon}\|^2_{L^2(\mathbb{R}^n)} + 2\mathrm{Re} \int_0^T \int_0^\tau \langle P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon})(s,\cdot), (u_\varepsilon - u_{\tilde\varepsilon})(s,\cdot)\rangle_{L^2(\mathbb{R}^n)} \mathrm{d}s\, \mathrm{d}\tau \Bigg).\end{aligned}$$

We proceed as in the proof of Proposition 2.25, but use the refined estimate in 2.27 (iii) for the right-hand side. Thus we can essentially work with $L^{1,\infty}$-norms of the coefficients, instead of $L^{2,\infty}$-norms, and with $L^{1,2}$-norms of the right-hand side, showing that $(u_\varepsilon)_\varepsilon$ is still a Cauchy net in $L^2(\Omega_T)$. Moreover, the $O(1)$-estimate for $\partial_{x_j} u_\varepsilon$ remains true, so we still get $\partial_{x_j} u_0 \in L^2(\Omega_T)^m$. For $\partial_t u_\varepsilon$ we only obtain boundedness in the $L^1(\Omega_T)$-norm, since the coefficients and the right-hand side are no longer in L^2 with respect to the time variable and we have to estimate them against the $L^{\infty,2}$-norm of the solution. However, since $L^2([0,T]) \subseteq L^1([0,T])$ by Hölder's inequality, we obtain at least $u_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m \cap W^{1,1}(\Omega_T)^m$. $\square$

As Proposition 2.28 shows, weakening the regularity assumptions for A_0^j and B_0 from $L^2([0,T], W^{1,\infty}(\mathbb{R}^n))$ to $L^1([0,T], W^{1,\infty}(\mathbb{R}^n))$ does not significantly alter the regularity of the distributional limit. However, changes in the coefficient regularities should somehow be reflected in the regularity of the solution. Lemma 2.29 provides yet another version of an energy estimate, which will be used to prove that the solution actually belongs to $L^4([0,T], W^{1,\infty}(\mathbb{R}^n))^m$, if the coefficient matrices have components in $L^2([0,T], W^{1,\infty}(\mathbb{R}^n))$, as in Proposition 2.25.

Lemma 2.29 (Refined energy estimate)**.** *Let $P = \partial_t + \sum_{j=1}^n A^j \partial_{x_j} + B$ be a symmetric hyperbolic PDO and $u, w \in \mathcal{C}^1([0,T], L^2(\mathbb{R}^n))^m \cap \mathcal{C}^0([0,T], H^1(\mathbb{R}^n))^m$. Then, for $\mu(\tau) = 2\|\mathrm{div}A - B - B^*\|^2_{L^{2,\infty}(\Omega_\tau)}$, we have*

$$\|u\|^4_{L^{4,2}(\Omega_T)} \leq 4\, e^{\int_0^T \mu(\tau)\mathrm{d}\tau} \int_0^T \Big(\Big(2\mathrm{Re} \int_{\Omega_\tau} \langle Pu, u\rangle \,\mathrm{dV} \Big)^2 + \|u(0,\cdot)\|^4_{L^2(\mathbb{R}^n)} \Big) \mathrm{d}\tau.$$

Proof. We start as in the proof of Lemma 2.27, integrating

$$\int_0^\tau \frac{d}{ds} \|u(s,\cdot)\|^2_{L^2(\mathbb{R}^n)} \mathrm{d}s = \int_{\Omega_\tau} \langle (\mathrm{div}A - B - B^*)u, u\rangle \,\mathrm{dV} \; + 2\mathrm{Re} \int_{\Omega_\tau} \langle Pu, u\rangle \,\mathrm{dV}$$

and estimate this as follows,

$$\begin{aligned} \|u(\tau,\cdot)\|^2_{L^2(\mathbb{R}^n)} \leq \|\mathrm{div}A - B - B^*\|_{L^{2,\infty}(\Omega_\tau)} & \left(\int_0^\tau \|u(s,\cdot)\|^4_{L^2(\mathbb{R}^n)} \right)^{\frac{1}{2}} \\ & + 2\mathrm{Re} \int_{\Omega_\tau} \langle Pu, u\rangle \,\mathrm{dV} + \|u(0,\cdot)\|^2_{L^2(\mathbb{R}^n)}, \end{aligned}$$

where we have used that

$$\begin{aligned} \|au^2\|_{L^1(\Omega_\tau)} &\leq \int_0^\tau \big(\|au\|_{L^2(\mathbb{R}^n)} \|u\|_{L^2(\mathbb{R}^n)} \big)(s) \mathrm{d}s \\ &\leq \int_0^\tau \big(\|a\|_{L^\infty(\mathbb{R}^n)} \|u\|^2_{L^2(\mathbb{R}^n)} \big)(s) \mathrm{d}s \leq \|a\|_{L^{2,\infty}(\Omega_\tau)} \left(\int_0^\tau \|u(s,\cdot)\|^4_{L^2(\mathbb{R}^n)} \mathrm{d}s \right)^{\frac{1}{2}} \end{aligned}$$

with $a(t,x) = \|(\mathrm{div}A - B - B^*)(t,x)\|_{\mathrm{op}}$. Taking the square on both sides of the inequality for $\|u(\tau,\cdot)\|^2_{L^2(\mathbb{R}^n)}$ and using twice that $(c+d)^2 \leq 2(c^2+d^2)$, we find

$$\|u(\tau,\cdot)\|^4_{L^2(\mathbb{R}^n)} \leq 2\|\mathrm{div}A - B - B^*\|^2_{L^{2,\infty}(\Omega_\tau)} \int_0^\tau \|u(s,\cdot)\|^4_{L^2(\mathbb{R}^n)} \mathrm{d}s + 4\Big(2\mathrm{Re}\int_{\Omega_\tau} \langle Pu, u\rangle \,\mathrm{d}V\Big)^2 + 4\|u(0,\cdot)\|^4_{L^2(\mathbb{R}^n)}.$$

A second integration with respect to the time variable yields

$$\|u\|^4_{L^{4,2}(\Omega_t)} \leq 2\int_0^t \|\mathrm{div}A - B - B^*\|^2_{L^{2,\infty}(\Omega_\tau)} \|u\|^4_{L^{4,2}(\Omega_\tau)} \mathrm{d}\tau + 4\int_0^t \Big(\Big(2\mathrm{Re}\int_{\Omega_\tau} \langle Pu, u\rangle \,\mathrm{d}V\Big)^2 + \|u(0,\cdot)\|^4_{L^2(\mathbb{R}^n)}\Big)\mathrm{d}\tau.$$

Finally, applying Gronwall's inequality, Lemma 1.4 (ii), we get

$$\|u\|^4_{L^{4,2}(\Omega_T)} \leq 4\, e^{\int_0^T \mu(\tau)\mathrm{d}\tau} \int_0^T \Big(\Big(2\mathrm{Re}\int_{\Omega_\tau} \langle Pu, u\rangle \,\mathrm{d}V\Big)^2 + \|u(0,\cdot)\|^4_{L^2(\mathbb{R}^n)}\Big)\mathrm{d}\tau.$$

where $\mu(\tau) = 2\|\mathrm{div}A - B - B^*\|^2_{L^{2,\infty}(\Omega_\tau)} \leq \mu(T)$. □

Proposition 2.30 (Improved regularity result)**.** *With the assumptions of Proposition 2.25, the limit u_0 of the generalised solution u satisfies $u_0 \in L^4([0,T], H^1(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$, where $u_\varepsilon \to u_0$ as $\varepsilon \to 0$ in the norm $L^4([0,T], L^2(\mathbb{R}^n))$ for any representative $(u_\varepsilon)_\varepsilon$ of u.*

Proof. First note that conditions (i) and (ii) in Theorem 2.13 are satisfied and we get a unique generalised solution $u \in \mathcal{G}_{L^2}(\Omega_T)^m$ with $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$. Considering the spatial derivatives of u_ε, we proceed just as in the proof of Proposition 2.28, leading again to $\|\nabla^1 u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$ (by the very same estimate, since $\|v\|_{L^{1,\infty}(\Omega_T)} \leq C\|v\|_{L^{2,\infty}(\Omega_T)}$ and thus the upper bound for $\|\nabla^1 u_\varepsilon\|_{L^{\infty,2}(\Omega_T)}$ in terms of $L^2([0,T])$-norms automatically produces an upper bound in terms of $L^1([0,T])$-norms). For the Cauchy net argument, we employ the energy estimate in Lemma 2.29 to obtain

$$\|u_\varepsilon - u_{\tilde\varepsilon}\|^4_{L^{4,2}(\Omega_T)} \leq 4\exp\Big(2\int_0^T \|(\mathrm{div}A_\varepsilon - B_\varepsilon - B^*_\varepsilon)\|^2_{L^{2,\infty}(\Omega_\tau)}\mathrm{d}\tau\Big) \cdot \int_0^T \Big(\Big(2\mathrm{Re}\int_{\Omega_\tau} \langle P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon})(s,\cdot), (u_\varepsilon - u_{\tilde\varepsilon})(s,\cdot)\rangle \,\mathrm{d}V\Big)^2 + \|g_\varepsilon - g_{\tilde\varepsilon}\|^4_{L^2(\mathbb{R}^n)}\Big)\mathrm{d}\tau.$$

The integral of $\langle P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon})(s,\cdot), (u_\varepsilon - u_{\tilde\varepsilon})(s,\cdot)\rangle$ over the domain Ω_τ can be estimated similarly to Lemma 2.27 (iii), taking the square after the integration does not cause any problems. It is then straightforward to see that the convergence properties of $(A^j_\varepsilon)_\varepsilon$ and $(B_\varepsilon)_\varepsilon$ imply that $(u_\varepsilon)_\varepsilon$ is a Cauchy net in $L^4([0,T], L^2(\mathbb{R}^n))^m$. $\square$

In order to get more information on the regularity of the distributional shadow u_0 and how it is influenced by the regularity of the coefficients A^j_0, B_0, initial data g_0, and right-hand side f_0, it is essential to have precise estimates on the speed of convergence of the regularised objects (cf. the notion of 'strong association' introduced in [Sca98]). To show that the conditions in the proposition to follow are realistic, we briefly discuss estimates on the convergence speed of convolution regularisations.

Lemma 2.31 (Convergence rate of convolution regularisations)**.** *Let $\rho \in \mathcal{S}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \rho(x)\mathrm{d}x = 1$ and $\int_{\mathbb{R}^n} x^\alpha \rho(x)\mathrm{d}x = 0$ for all $\alpha \in \mathbb{N}_0^n$, $0 < |\alpha| < m$, $s \in \mathbb{R}$, and put $\rho_\varepsilon := \sigma_\varepsilon^{-n}\rho(\frac{\cdot}{\sigma_\varepsilon})$ where $\sigma_\varepsilon \to 0$ as $\varepsilon \to 0$.*

(i) *If $u \in W^{m,\infty}(\mathbb{R}^n)$, then $\|\rho_\varepsilon * u - u\|_{L^\infty(\mathbb{R}^n)} = O(\sigma_\varepsilon^m)$ as $\varepsilon \to 0$.*

(ii) *If $u \in L^1(\mathbb{R}^n)$, then $\|\rho_\varepsilon * u - u\|_{H^s(\mathbb{R}^n)} = O(\sigma_\varepsilon^m)$ as $\varepsilon \to 0$ for all $s < -\frac{n}{2} - m$.*

(iii) *If $u \in H^s(\mathbb{R}^n)$, $s \in \mathbb{R}$, then $\|\rho_\varepsilon * u - u\|_{H^s(\mathbb{R}^n)} \to 0$ and $\|\rho_\varepsilon * u - u\|_{H^{s-m}(\mathbb{R}^n)} = O(\sigma_\varepsilon^m)$ as $\varepsilon \to 0$.*

Proof. The claimed asymptotic estimates can be obtained as follows:

(i): Writing $\|\rho_\varepsilon * u - u\|_{L^\infty(\mathbb{R}^n)} = \int (u(x - \sigma_\varepsilon y) - u(x))\rho(y)\mathrm{d}y$ and using Taylor's formula to represent $u(x - \sigma_\varepsilon y)$ as a polynomial of order $m-1$ plus a remainder term, only the latter will survive thanks to the fact that ρ has vanishing moments up to order $m-1$.

(ii): For $\varphi \in \mathcal{D}(\mathbb{R}^n)$ we find $\langle \rho_\varepsilon * u - u, \varphi\rangle = \int\int u(y)\rho(z)\big(\varphi(y+\varepsilon z) - \varphi(y)\big)\mathrm{d}z\mathrm{d}y$ by a transformation of variables and using Fubini's theorem. Taylor expansion then yields

$$
\begin{aligned}
|\langle \rho_\varepsilon * u - u, \varphi\rangle| &\le m \sum_{|\alpha|=m} \Big| \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(y)\rho(z)\frac{(\sigma_\varepsilon z)^\alpha}{\alpha!}\int_0^1 (1-s)^{m-1}\partial^\alpha\varphi(y + s\sigma_\varepsilon z)\mathrm{d}s\mathrm{d}z\mathrm{d}y\Big| \\
&\le C_\rho \sigma_\varepsilon^m \|u\|_{L^1(\mathbb{R}^n)}\|\varphi\|_{W^{m,\infty}(\mathbb{R}^n)} \le C_{\rho,u}\sigma_\varepsilon^m \|\varphi\|_{H^{m+n/2+\alpha}(\mathbb{R}^n)}
\end{aligned}
$$

for all $\alpha > 0$ and therefore $\|\rho_\varepsilon * u - u\|_{H^s(\mathbb{R}^n)} \le C_{\rho,u}\sigma_\varepsilon^m$ as $\varepsilon \to 0$ for all $s < -\frac{n}{2} - m$.

(iii): Denoting the Fourier transform of $u \in \mathcal{S}'(\mathbb{R}^n)$ by $\widehat{u}$ we estimate

$$
\begin{aligned}
(2\pi)^n \|\rho_\varepsilon * u - u\|^2_{H^{s-m}} &\le \int_{\mathbb{R}^n} (1+|\xi|^2)^{s-m}|\widehat{u}(\xi)|^2|\widehat{\rho}(\sigma_\varepsilon\xi) - 1|^2\mathrm{d}\xi = \int_{\mathbb{R}^n}(1+|\xi|^2)^{s-m}|\widehat{u}(\xi)|^2 \\
&\cdot \Big| \sum_{1\le|\alpha|<m} \frac{(\partial^\alpha\widehat{\rho})(0)}{\alpha!}(\sigma_\varepsilon\xi)^\alpha + m\sigma_\varepsilon^m \int_0^1 (1-t)^{m-1}\sum_{|\alpha|=m}(\partial^\alpha\rho)(t\sigma_\varepsilon\xi)\frac{\xi^\alpha}{\alpha!}\mathrm{d}t\Big|^2 \mathrm{d}\xi \\
&\le C\sigma_\varepsilon^{2m}\|\rho\|^2_{W^{m,\infty}(\mathbb{R}^n)} \int_{\mathbb{R}^n} (1+|\xi|^2)^{s-m}|\widehat{u}(\xi)|^2|\xi|^{2m}\mathrm{d}\xi \le C_\rho \sigma_\varepsilon^{2m}\|u\|^2_{H^s(\mathbb{R}^n)}
\end{aligned}
$$

since $\partial^\alpha \widehat{\rho}(0) = 0$ for $0 < |\alpha| < m$. The assertion $\|\rho_\varepsilon * u - u\|_{H^s(\mathbb{R}^n)} \to 0$ as $\varepsilon \to 0$ follows directly from the estimate

$$(2\pi)^n \|\rho_\varepsilon * u - u\|^2_{H^s(\mathbb{R}^n)} \leq \int_{\mathbb{R}^n} (1 + |\xi|^2)^s |\widehat{u}(\xi)|^2 |\widehat{\rho}(\sigma_\varepsilon \xi) - 1|^2 \mathrm{d}\xi$$

by dominated convergence. □

Fast convergence of the regularised coefficients and data results in stronger association of the generalised solution to its distributional limit, measured in terms of ε-powers. The following proposition shows that rapidly converging principal coefficients yield the existence of a distributional shadow under weaker regularity assumptions on the initial data g_0 and right-hand side f_0 than in Proposition 2.21.

Proposition 2.32 (Lower-order coefficients below Lipschitz regularity in the spatial variables)**.** *In the initial value problem (2.23–2.24), let the entries of A_0^j belong to $L^1([0,T], W^{1,\infty}(\mathbb{R}^n))$ and those of B_0 belong to $L^2([0,T], L^\infty(\mathbb{R}^n))$ and be uniformly continuous with respect to the spatial variables. For given initial data $g_0 \in L^2(\mathbb{R}^n)^m$ and right-hand side $f_0 \in L^2(\Omega_T)^m$, define corresponding generalised coefficients and data for the initial value problem (2.1–2.2), satisfying the assumptions in Theorem 2.13 and denote*

$$\gamma^\varepsilon := \sum_{1 \leq j \leq n} \Big(\|\partial_{x_j} B_\varepsilon\|_{L^{2,\infty}(\Omega_T)} + \|\partial_{x_j} g_\varepsilon\|_{L^2(\mathbb{R}^n)} + \|\partial_{x_j} f_\varepsilon\|_{L^1([0,T],L^2(\mathbb{R}^n))} \Big).$$

If the condition

$$\gamma^\varepsilon \|A_\varepsilon^j - A_0^j\|_{L^{1,\infty}(\Omega_T)} \to 0 \quad \text{as} \quad \varepsilon \to 0, \tag{2.25}$$

is satisfied, then the corresponding generalised solution $u = [(u_\varepsilon)_\varepsilon]$ has a distributional shadow $u_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m$ such that for any representatives $(u_\varepsilon)_\varepsilon$ of u, $u_\varepsilon \to u_0$ in the norm $L^\infty([0,T], L^2(\Omega_T))$.

Proof. By Theorem 2.13, we obtain a generalised solution $u = [(u_\varepsilon)_\varepsilon]$ with the property $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$. Moreover, as in the proof of Proposition 2.21, we have

$$\sup_{0 \leq t \leq T} \|\nabla^1 u_\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^n)} \leq e^{\beta_\varepsilon^1} \Big(\|\nabla^1 g_\varepsilon\|^2_{L^2(\mathbb{R}^n)} + \|(\widetilde{\nabla}^1 B_\varepsilon)\, \Sigma^1 u_\varepsilon + \nabla^1 f_\varepsilon\|^2_{L^2(\Omega_T)} \Big),$$

and since $\|(\widetilde{\nabla}^1 B_\varepsilon)\, \Sigma^1 u_\varepsilon\|^2_{L^2(\Omega_T)} \leq C\, \|u_\varepsilon\|^2_{L^{\infty,2}(\Omega_T)} \|\widetilde{\nabla}^1 B_\varepsilon\|^2_{L^{2,\infty}(\Omega_T)}$, as shown in the proof of Proposition 2.25, we obtain $\|\widetilde{\nabla}^1 u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} \leq C\gamma^\varepsilon$. Once again invoking a Cauchy net argument by applying the energy estimate in Lemma 2.24 (i) to $u_\varepsilon - u_{\tilde{\varepsilon}}$, where $\varepsilon < \tilde{\varepsilon} < \varepsilon_0$ and ε_0 small enough, we find

$$\|u_\varepsilon - u_{\tilde{\varepsilon}}\|^2_{L^{\infty,2}(\Omega_T)} \leq C e^{\int_0^T \beta(s) ds} \Big(\|g_\varepsilon - g_{\tilde{\varepsilon}})\|^2_{L^2(\mathbb{R}^n)} + \|f_\varepsilon - f_{\tilde{\varepsilon}}\|_{L^{1,2}(\Omega_T)}$$
$$+ \|B_\varepsilon - B_{\tilde{\varepsilon}}\|_{L^{1,\infty}(\Omega_T)} + \Big(\|A_\varepsilon^j - A_0^j\|_{L^{1,\infty}(\Omega_T)} + \|A_{\tilde{\varepsilon}}^j - A_0^j\|_{L^{1,\infty}(\Omega_T)} \Big) \|\widetilde{\nabla}^1 u_{\tilde{\varepsilon}}\|_{L^{\infty,2}(\Omega_T)} \Big),$$

where we have used that $\|u_\varepsilon - u_{\tilde{\varepsilon}}\|_{L^{\infty,2}(\Omega_T)} \leq C$. Since we already know that $\|\widetilde{\nabla}^1 u_{\tilde{\varepsilon}}\|_{L^{\infty,2}(\Omega_T)} = O(\sigma_{\tilde{\varepsilon}})$, the family $(u_\varepsilon)_\varepsilon$ will be a Cauchy net if the product $\gamma^\varepsilon \|A^j_\varepsilon - A^j_0\|_{L^{1,\infty}(\Omega_T)}$ vanishes in the limit $\varepsilon \to 0$. □

Condition (2.25) can be fulfilled easily by using a suitably rescaled mollifier to regularise the principal coefficients, see Lemma 2.31. One may first regularise B_0, g_0, and f_0, and choose a proper scaling for the convolution regularisation of the coefficient matrices A^j_0 afterwards. If the principal coefficients A^j_0 enjoy a better regularity than $L^1([0,T], W^{1,\infty}(\mathbb{R}^n))$, and therefore a faster rate of convergence after regularisation, the usual scaling of the mollifier can also be sufficient. In particular, if the A^j_0 are smooth, there will always be convergence, as long as we use the standard embedding into the Colombeau algebra for all coefficients and data. On the other hand, lower regularity of the principal coefficients (but not below $L^1([0,T], W^{1,\infty}(\mathbb{R}^n))$) can be compensated by a proper rescaling of the mollifier. However, the adverse effect of this approach is that using different regularisation methods or scalings for the coefficients will generate additional constraints, which are not inherent to the original problem, but only required to endow the generalised solution with certain asymptotic bounds.

While the required regularity of the coefficients with respect to the time variable could be substantially weakened to L^1_{loc} for A^j_0 and B_0 in Proposition 2.28 (compared to the classical result in Proposition 2.22, requiring the coefficients to be continuous), the achievements in regularity reduction for the spatial variables is less encouraging. Here, the principal coefficients have to be at least Lipschitz continuous and the condition on the lower-order coefficients could only be slightly relaxed to uniform continuity in Proposition 2.32. Given the structure of the PDE system (2.23), energy estimates will inevitably involve spatial derivatives of the coefficients. Thus the margins between the regularity assumptions hitherto used for the x-dependence and the minimal regularity needed to obtain a distributional limit of the solution, are probably much smaller than for the t-dependence. However, a little improvement might still be possible by a fine-tuning of the involved regularisations. If we want to avoid spatial L^∞-norms of the coefficient matrices in the Cauchy net estimate, we have to work with higher-order Sobolev norms of the solution.

Lemma 2.33 (Basic estimates with H^s-norms). *Let $s \in \mathbb{R}$, $s > \frac{n}{2}$, and assume that $A^j, B \in M_m(L^1([0,T], H^{-s}(\mathbb{R}^n)))$, $f \in L^1([0,T], L^2(\mathbb{R}^n))^m$, and that $u \in \mathcal{C}^0([0,T], H^{s+1}(\mathbb{R}^n))^m$, $w \in \mathcal{C}^0([0,T], H^s(\mathbb{R}^n))^m$. Then we may estimate*

$$\sum_{1 \leq j \leq n} \langle A^j \partial_{x_j} u, w\rangle_{L^2(\Omega_T)} + \langle Bu + f, w\rangle_{L^2(\Omega_T)}$$
$$\leq \sum_{1 \leq j \leq n} \|A^j\|_{L^1([0,T],H^{-s}(\mathbb{R}^n))} \|u\|_{L^\infty([0,T],H^{s+1}(\mathbb{R}^n))} \|w\|_{L^\infty([0,T],H^s(\mathbb{R}^n))}$$
$$+ \|B\|_{L^1([0,T],H^{-s}(\mathbb{R}^n))} \|u\|_{L^\infty([0,T],H^s(\mathbb{R}^n))} \|w\|_{L^\infty([0,T],H^s(\mathbb{R}^n))} + \|f\|_{L^{1,2}(\Omega_T)} \|w\|_{L^{\infty,2}(\Omega_T)}.$$

Proof. The proof is a straightforward calculation, based on the inequalities

$$\|au'w\|_{L^1(\mathbb{R}^n)} \leq \|a\|_{H^{-s}(\mathbb{R}^n)} \|u'w\|_{H^s(\mathbb{R}^n)} \leq \|a\|_{H^{-s}(\mathbb{R}^n)} \|u\|_{H^{s+1}(\mathbb{R}^n)} \|w\|_{H^s(\mathbb{R}^n)},$$

which are valid for $s > \frac{n}{2}$ by [AF03, Theorem 3.12 and Theorem 4.39]. $\square$

Proposition 2.34 (Lower-order coefficients bounded and square integrable in the spatial variables)**.** *In the initial value problem (2.23–2.24), let the components of A_0^j belong to $L^1([0,T], W^{1,\infty}(\mathbb{R}^n)) \cap L^1([0,T], L^2(\mathbb{R}^n))$ and those of B_0 belong to $L^1([0,T], L^\infty(\mathbb{R}^n)) \cap L^1([0,T], L^2(\mathbb{R}^n))$. Let initial data $g_0 \in L^2(\mathbb{R}^n)^m$ and right-hand side $f_0 \in L^1([0,T], L^2(\mathbb{R}^n))^m$ be given. Define corresponding generalised functions A^j, B, f, g for the initial value problem (2.1–2.2), satisfying the assumptions in Theorem 2.13. For $l \in \mathbb{N}$, denote*

$$\gamma_l^\varepsilon := (1-\delta_{l,1}) \sum_{j=1}^{n} \|A_\varepsilon^j\|_{L^1([0,T],W^{l,\infty}(\mathbb{R}^n))} + \|B_\varepsilon\|_{L^1([0,T],W^{l,\infty}(\mathbb{R}^n))} + \|g_\varepsilon\|_{H^l(\mathbb{R}^n)} + \|f_\varepsilon\|_{L^1([0,T],H^l(\mathbb{R}^n))},$$

where $\delta_{l,1}$ represents the Kronecker delta ($\delta_{l,1} = 1$ for $l = 1$ and $\delta_{l,1} = 0$ otherwise). If there exists $k \in \mathbb{N}$, such that

(i) $(\gamma_k^\varepsilon)^2 \, \|B_\varepsilon - B_0\|_{L^1([0,T],H^{-k}(\mathbb{R}^n))} \to 0 \quad \text{as} \quad \varepsilon \to 0,$

(ii) $\gamma_k^\varepsilon \gamma_{k+1}^\varepsilon \sum_{j=1}^n \|A_\varepsilon^j - A_0^j\|_{L^1([0,T],H^{-k}(\mathbb{R}^n))} \to 0 \quad \text{as} \quad \varepsilon \to 0,$

then the corresponding generalised solution $u = [(u_\varepsilon)_\varepsilon]$ has a distributional limit $u_0 \in \mathcal{C}^0([0,T], L^2(\mathbb{R}^n))^m$ such that $u_\varepsilon \to u_0$ in the norm $L^\infty([0,T], L^2(\mathbb{R}^n))$ for any of its representatives $(u_\varepsilon)_\varepsilon$.

Proof. As in Proposition 2.28, we obtain a generalised solution $u = [(u_\varepsilon)_\varepsilon]$ with $\|u_\varepsilon\|_{L^{\infty,2}(\Omega_T)} = O(1)$ as $\varepsilon \to 0$. Moreover, invoking energy estimates for the spatial derivatives $\|\widetilde{\nabla}^k u_\varepsilon\|_{L^{\infty,2}(\Omega_T)}$, as in the proof of Theorem 2.13, we obtain $\|u_\varepsilon\|_{L^\infty([0,T],H^k(\mathbb{R}^n))} \le C\gamma_k^\varepsilon$ for all $\varepsilon < \varepsilon_0$ (note that the exponents β_ε and β_ε^r as defined in the proof of Theorem 2.13 are still uniformly bounded as $\varepsilon \to 0$). Turning to the usual Cauchy net argument, we rewrite the term $|\langle P_\varepsilon(u_\varepsilon - u_{\tilde\varepsilon}), u_\varepsilon - u_{\tilde\varepsilon}\rangle|$ as in the proof of Proposition 2.25 and use Lemma 2.33 to estimate the resulting term $\langle \sum_{j=1}^n \big(A_{\tilde\varepsilon}^j - A_\varepsilon^j\big)\partial_x u_{\tilde\varepsilon} + \big(B_{\tilde\varepsilon} - B_\varepsilon\big)u_{\tilde\varepsilon} + f_{\tilde\varepsilon}, u_\varepsilon - u_{\tilde\varepsilon}\rangle$ on the right-hand side of the Cauchy net inequality. The convergence conditions (i) and (ii) then guarantee that the family $(u_\varepsilon)_\varepsilon$ is a Cauchy net in the norm $L^\infty([0,T], L^2(\mathbb{R}^n))$. $\square$

Note that in Proposition 2.34, it is not required that A_0^j, B_0, f, or g to belong to some higher-order Sobolev space related to k. For instance, regularisation by convolution with a mollifier as in Lemma 2.20 (ii) guarantees that for arbitrary $l \in \mathbb{N}$, the norms used in the definition of γ_l^ε are finite and therefore $\|u_\varepsilon\|_{L^\infty([0,T],H^l(\mathbb{R}^n))} \le C\gamma_l^\varepsilon$ for all $\varepsilon < \varepsilon_0$, that is, $u_\varepsilon \in L^\infty([0,T], H^l(\mathbb{R}^n))^m$ for each $\varepsilon < \varepsilon_0$. Similarly, the assumption $A_0^j, B_0 \in M_m(L^1([0,T], L^2(\mathbb{R}^n)))$ ensures that the regularisations $(A_\varepsilon^j)_\varepsilon$ and $(B_\varepsilon)_\varepsilon$ are convergent nets in the norm $L^1([0,T], H^{-k}(\mathbb{R}^n))$ for any $k \in \mathbb{N}_0$, where the rate of convergence increases with k, see Lemma 2.31 (iii).

Remark 2.35 (Admissible coefficients in Proposition 2.34)**.** Conditions (i) and (ii) in Proposition 2.34 are automatically satisfied for $k = 1$, if, in addition to the other regularity assumptions, $A_0^j \in M_m(L^1([0,T], H^3(\mathbb{R}^n)))$ and the lower-order coefficient matrix $B_0 \in M_m(L^1([0,T], H^2(\mathbb{R}^n)))$. More precisely, defining regularised objects $(A^j_\varepsilon)_\varepsilon$ and $(B_\varepsilon)_\varepsilon$ by convolution with a model delta net[3] $\rho_\varepsilon(\cdot) = \frac{1}{\varepsilon^n}\rho\left(\frac{\cdot}{\varepsilon}\right)$, we obtain $\gamma^\varepsilon_k = O(\varepsilon^{-k})$ as $\varepsilon \to 0$. By Lemma 2.31, we have $\|B_\varepsilon - B_0\|_{L^1([0,T],H^{-k}(\mathbb{R}^n))} = O(\varepsilon^{k+2})$ and $\|A_\varepsilon - A_0\|_{L^1([0,T],H^{-k}(\mathbb{R}^n))} = O(\varepsilon^{k+3})$ as $\varepsilon \to 0$. Hence, for $k = 1$ we get

- $(\gamma^\varepsilon_1)^2 \|B_\varepsilon - B_0\|_{L^1([0,T],H^{-1}(\mathbb{R}^n))} = O(\varepsilon)$ as $\varepsilon \to 0$,
- $\gamma^\varepsilon_1 \gamma^\varepsilon_2 \sum_{j=1}^n \|A^j_\varepsilon - A^j_0\|_{L^1([0,T],H^{-1}(\mathbb{R}^n))} = O(\varepsilon)$ as $\varepsilon \to 0$.

However, there is no $k > 1$ for which (i) and (ii) are satisfied, since for $k > 1$, $(\gamma^\varepsilon_k)^2$ grows faster in terms of inverse powers of ε than $\|B_\varepsilon - B_0\|_{L^1([0,T],H^{-k}(\mathbb{R}^n))}$ tends to zero. Note that the Sobolev index in the assumed regularities of the coefficients is independent of the spatial dimension n. Thus, with regard to Sobolev embeddings, the required coefficient regularities for the existence of a distributional shadow of the generalised solution get weaker as the dimension n increases.

It is clear that our methods do not allow to consider regularity assumptions for A^j_0 and B_0 below $W^{1,\infty}(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n)$ with respect to the spatial domain, since otherwise their regularisations $(A^j_\varepsilon)_\varepsilon$ and $(B_\varepsilon)_\varepsilon$ would not satisfy condition (2") in Theorem 2.13, which is essential to get a generalised solution $u \in \mathcal{G}(\Omega_T)^m$ with $\|u_\varepsilon\|_{L^2(\Omega_T)} = O(1)$ as $\varepsilon \to 0$, giving at least hope for the existence of a distributional shadow. Thus, concerning lower regularities with respect to the spatial variables, we cannot expect much further progress compared to Proposition 2.34 without having an alternative way to establish the generalised solution in the first place. On the other hand, as was already pointed out in Remark 2.14, condition (2") is satisfied for coefficients A^j_0 and B_0 proportional to $\delta(t - t_0)$, so there is still some room for improvement with respect to the dependence on the time variable, to which we will now turn our attention again.

Whereas the energy estimate in Lemma 2.24 provides a bound for the $L^{\infty,2}(\Omega_T)$-norm of the solution, the refined estimate in Lemma 2.27 only yields an inequality for the $L^2(\Omega_T)$-norm. If the generalised solution converges in $L^2(\Omega_T)$, but not in $L^{\infty,2}(\Omega_T)$, then the estimate in Lemma in 2.24 cannot be used to prove this convergence by means of a Cauchy net argument, but the modified estimate in Lemma 2.27 may still be applied in this situation. However, the required convergence properties of the coefficient matrices are still too strong to allow for a term proportional to $\delta(t - t_0)$ in their components. We will now use a different approach to include this case, thereby keeping the convergence of the generalised solution to a distributional limit. The basic idea is to impose certain structural conditions on the coefficient matrices in order to split the solution into a product of a continuous part and a possibly discontinuous, but explicit part. Since discontinuities of the solution are then represented explicitly and in terms of the coefficients, the energy estimates are only needed to prove convergence of the regular part of the solution.

Proposition 2.36 (Lower-order coefficients including time derivatives of a bounded function)**.** *Let* $A^j_0, B_0 \in M_m(L^\infty([0,T], W^{1,\infty}(\mathbb{R}^n)))$, $f_0 \in L^2([0,T], H^1(\mathbb{R}^n))^m$, $g_0 \in$

[3]Here, $\rho \in \mathcal{S}(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} x^\alpha \rho(x)\mathrm{d}x = \delta_{0,|\alpha|}$.

$H^1(\mathbb{R}^n)^m$, and let $C_0 \in M_m(L^\infty([0,T], W^{2,\infty}(\mathbb{R}^n)))$ be continuous at $t = 0$. Define corresponding generalised coefficients, right-hand side and data A^j, B, C, f, g for the generalised initial value problem

$$\partial_t u + \sum_{j=1}^{n} A^j \partial_{x_j} u + Bu + (\partial_t C)u = f \qquad \text{on } \Omega_T =]0,T[\times \mathbb{R}^n \tag{2.26}$$

$$u|_{t=0} = g \tag{2.27}$$

satisfying the assumptions in Theorem 2.13. For square matrices X, Y, denote the commutator as $[X,Y] := XY - YX$. If

(i) $[C, \partial_t C] = [C, \partial_{x_j} C] = 0$,

(ii) $e^{\pm C} = [(e^{\pm C_\varepsilon})_\varepsilon] \in M_m(\mathcal{G}_{L^\infty}(\Omega_T))$,

and, for any representative $(C_\varepsilon)_\varepsilon$ of C,

(iii) $\|e^{\pm C_\varepsilon} - e^{\pm C_0}\|_{L^2([0,T],L^\infty(\mathbb{R}^n))} \to 0$, *as* $\varepsilon \to 0$,

(iv) $\|\partial_t C_\varepsilon\|_{L^1([0,T],W^{1,\infty}(\mathbb{R}^n))} = O(1)$ *as* $\varepsilon \to 0$,

then the unique generalised solution $u \in \mathcal{G}_{L^2}(\Omega_T)$ has a distributional shadow $u_0 \in L^\infty([0,T], H^1(\mathbb{R}^n))^m$.

Proof. Let $u \in \mathcal{G}_{L^2}(\Omega_T)^m$ be the unique generalised solution of the initial value problem (2.26–2.27), established by Theorem 2.13, and put $v = e^C u$. We compute $Pu = P(e^{-C}v)$ in $\mathcal{G}$, yielding

$$Pu = e^{-C}\partial_t v - \sum_{j=1}^{n} A^j(\partial_{x_j} C)e^{-C} v + \sum_{j=1}^{n} A^j e^{-C}\partial_{x_j} v + Be^{-C} v,$$

since C commutes with its derivatives by (i) and, as a consequence, we have $\partial_t u = -(\partial_t C)u + e^{-C}\partial_t v$. Multiplying by e^C from the left, we obtain

$$\partial_t v - \sum_{j=1}^{n} e^C A^j(\partial_{x_j} C)e^{-C} v + \sum_{j=1}^{n} e^C A^j e^{-C}\partial_{x_j} v + e^C Be^{-C} v = e^C f,$$

where we have used that $[C, \partial_{x_j} C] = 0$. Thus

$$Pu = f \iff \partial_t v + \sum_{j=1}^{n} e^C A^j e^{-C}\partial_{x_j} v + e^C\Big(B - \sum_{j=1}^{n} A^j \partial_{x_j} C\Big)e^{-C} v = e^C f$$

and regarding the last equation together with the initial condition $v|_{t=0} = (e^C|_{t=0})g \in \mathcal{G}_{L^2}(\mathbb{R}^n)^m$ as an initial value problem for v, we may apply Theorem 2.13 and obtain a unique solution $v \in \mathcal{G}_{L^2}(\Omega_T)^m$. The regularities of A_0^j, B_0, C_0 imply that for any representatives of A^j, B, C, we get the asymptotic estimates

- $\|e^{C_\varepsilon} A_\varepsilon^j e^{-C_\varepsilon}\|_{L^\infty([0,T],W^{1,\infty}(\mathbb{R}^n))} = O(1)$ as $\varepsilon \to 0$,
- $\|e^{C_\varepsilon}\Big(B_\varepsilon - \sum_{j=1}^{n} A_\varepsilon^j \partial_{x_j} C_\varepsilon\Big)e^{-C_\varepsilon}\|_{L^\infty([0,T],W^{1,\infty}(\mathbb{R}^n))} = O(1)$ as $\varepsilon \to 0$,

which, together with condition (iv), yield $\|v_\varepsilon\|_{L^2([0,T],H^1(\mathbb{R}^n))} = O(1)$ as $\varepsilon \to 0$ for any representative $(v_\varepsilon)_\varepsilon$ of v. In addition, we obtain convergence in $L^{1,\infty}(\Omega_T)$ of

- $e^{C_\varepsilon} A^j_\varepsilon e^{-C_\varepsilon} \to e^{C_0} A^j_0 e^{-C_0}$ and
- $e^{C_\varepsilon}\Big(B_\varepsilon - \sum\limits_{j=1}^n A^j_\varepsilon \partial_{x_j} C_\varepsilon\Big)e^{-C_\varepsilon} \to e^{C_0}\Big(B_0 - \sum\limits_{j=1}^n A^j_0 \partial_{x_j} C_0\Big)e^{-C_0}$

as $\varepsilon \to 0$, a consequence of (iii) and the fact that $(A^j_\varepsilon)_\varepsilon$, $(B_\varepsilon)_\varepsilon$ and $(\partial_{x_j} C_\varepsilon)_\varepsilon$ converge in $L^p([0,T], L^\infty(\mathbb{R}^n))$ for $1 \leq p < \infty$. Thus, following the steps in the proof of Proposition 2.25, we conclude that v must have a distributional limit $v_0 \in C([0,T], H^1(\mathbb{R}^n))^m \cap H^1(\Omega_T)^m$, where $\|v_\varepsilon - v_0\|_{L^\infty([0,T],L^2(\mathbb{R}^n))} \to 0$ as $\varepsilon \to 0$. Finally, since $u = e^{-C}v$, we may estimate

$$\begin{aligned}
&\|e^{-C_\varepsilon} v_\varepsilon - e^{-C_0} v_0\|_{L^2(\Omega_T)} \leq \\
&\qquad \|e^{-C_\varepsilon}\|_{L^{2,\infty}(\Omega_T)} \|v_\varepsilon - v_0\|_{L^{\infty,2}(\Omega_T)} + \|(e^{-C_\varepsilon} - e^{-C_0})\|_{L^{2,\infty}(\Omega_T)} \|v_0\|_{L^{\infty,2}(\Omega_T)},
\end{aligned}$$

showing that $u_\varepsilon \to e^{-C_0} v_0 \in L^\infty([0,T], L^2(\mathbb{R}^n))^m$ in the norm $L^2(\Omega_T)$ for any representative $(u_\varepsilon)_\varepsilon$ of u. □

Remark 2.37 (Dirac measures as coefficients)**.** Proposition 2.36 establishes a distributional shadow of the generalised solution even in cases where the lower-order coefficients include terms proportional to the delta distribution, concentrated at the hyperplane $t = t_0$. More precisely, representing the Dirac measure in $\mathcal{G}$ as a non-negative model delta net $(\rho_\varepsilon)_\varepsilon$, where $\rho \in \mathcal{D}(]0,T[)$, we have $\|\rho_\varepsilon\|_{L^1([0,T])} = \|\rho\|_{L^1([0,T])} = 1$ and may therefore consider (2.26) with $\partial_t C$ of the form $\delta(t-t_0)S$, where S is a constant matrix. Then conditions (i)-(iv) are automatically satisfied. The same is true for $\partial_t C$ being a diagonal matrix $\Lambda(t,x)$ with entries containing terms such as $\delta(t-t_0) \otimes h(x)$ and h sufficiently regular. Less restrictive conditions for matrices to commute with their derivatives can be found in [Eva85].

In Example 2.16 we have considered the Dirac equation with a coupling term proportional to $\delta(t-x)$, representing an impulsive electromagnetic wave. Modelling the delta distribution therein as an L^∞-log-type generalised function, the corresponding generalised initial value problem has a unique generalised solution. Unfortunately, our results on distributional limits of the generalised solution fail for coefficients of the form $\delta(t-x)$, as they unavoidably require the representatives of the generalised coefficients to be bounded uniformly in ε in the $L^{1,\infty}$-norm. However, this condition can be fulfilled if the delta distribution occurs in the time variable only. In the following example we discuss a physical model that falls into the scope of Proposition 2.36 and preserves at least the impulsive character of the interaction term in the Dirac equation.

Example 2.38 (Dirac equation with delta potential)**.** As a simplified model for a spin-$\frac{1}{2}$ particle in an impulsive electromagnetic field (see Example 2.16), we now consider the Dirac equation with a coupling term of the form $\delta(t-t_0)S$, where $S \neq 0$

is a constant matrix, that is

$$\partial_t\psi(t,x) + \sum_{j=1}^{3}\gamma^0\gamma^j\partial_{x_j}\psi(t,x) + im\psi(t,x) + \delta(t-t_0)S\psi(t,x) = 0. \tag{2.28}$$

Theorem 2.13 establishes a unique solution $\psi \in \mathcal{G}_{L^2}([0,T]\times\mathbb{R}^3)^4$ for a broad spectrum of generalised function models for the delta distribution appearing in equation (2.28). In addition, Proposition 2.36 provides a distributional shadow of the generalised solution under fairly general circumstances. To meet the assumptions in Proposition 2.36, we may define initial data $\psi|_{t=0} = \varphi \in \mathcal{G}_{L^2}(\mathbb{R}^3)^4$ via regularisation of an initial spinor wave function $\varphi_0 \in H^1(\mathbb{R}^3)^4$ and model the delta distribution as a non-negative delta net. For this purpose we may take any non-negative function $\rho \in \mathcal{D}(\mathbb{R})$ with unit integral and represent δ in $\mathcal{G}$ by the class of $(\rho_\varepsilon)_\varepsilon = (\sigma_\varepsilon^{-1}\rho(\sigma_\varepsilon^{-1}\,\cdot\,))_\varepsilon$ where $\sigma_\varepsilon \to 0$ as $\varepsilon \to 0$ and $\sigma_\varepsilon^{-1} = O(\varepsilon^{-m})$ for some $m \in \mathbb{N}$. Then the matrices

$$C_\varepsilon(t) = e^{S\int_0^t \rho_\varepsilon(s-t_0)\mathrm{d}s}$$

satisfy conditions (i)-(iv) in Proposition 2.36 and thus there exists a function $\psi_0 \in L^\infty([0,T], L^2(\mathbb{R}^3))^4$ such that $\|\psi_\varepsilon - \psi_0\|_{L^2([0,T]\times\mathbb{R}^3)} \to 0$ as $\varepsilon \to 0$ for any representative $(\psi_\varepsilon)_\varepsilon$ of the unique generalised solution ψ. Moreover, following the construction in the proof of Proposition 2.36, we even obtain detailed information on the singularity structure of ψ_0, since

$$\psi_0(t,x) = (I_4 + (e^{-S} - I_4)H(t-t_0))\widetilde{\psi}_0(t,x)$$

where H is the Heaviside function and $\widetilde{\psi}_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^3))^4$. In other words, we have $\psi_0(t,x) = \widetilde{\psi}_0(t,x)$ for $t < t_0$ and $\psi_0(t,x) = e^{-S}\widetilde{\psi}_0(t,x)$ for $t > t_0$.

Discussing equation (2.28) physically, one would expect the electromagnetic shock to affect at least the spin state of the neutron, amounting to a linear transformation of the spinor wave function at time $t = t_0$. Yet this transformation is not directly given by the operator e^{-S} acting on $\widetilde{\psi}_0$, since the electromagnetic field matrix S is also incorporated in the 'reduced' initial value problem for $\widetilde{\psi} \in \mathcal{G}_{L^2}([0,T], L^2(\mathbb{R}^3))^4$ itself and will therefore already influence the behaviour of $\widetilde{\psi}_0 \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^3))^4$ at $t = t_0$. For $\widetilde{\psi}_0$ to be completely independent of S, the latter would have to commute with all gamma matrices, which is not the case for a coupling term describing a realistic electromagnetic field.

2.4 Propagation of coefficient singularities

In the results presented in Section 2.3, only global regularity of the generalised solution of the Cauchy problem has been investigated. It has been shown that for sufficiently regular coefficients and data, the solution will be of a certain global regularity, measured intrinsically as $\mathcal{G}^\infty$-regularity or extrinsically in terms of Sobolev norms of the distributional limit of the solution. Of course, the regularity of the solution will depend on the regularity properties of the coefficients as well as the initial data and right-hand side. In some of the results, this relationship can be seen very clearly, for instance in Proposition 2.18, where the coefficients are assumed to be smooth and the solution inherits the regularity of the data. Similarly, for smooth data but non-smooth coefficients, the regularity of the solution would be in some precise correspondence with the regularity of the coefficients. This is obvious, since energy estimates for higher-order derivatives of the solution involve higher-order derivatives of the coefficients (the detailed relationship was discussed in the proof of Theorem 2.7). Yet, none of the results in Section 2.3 provides any information on how discontinuities or singularities of the coefficients and data are propagated or imprinted in the solution, that is, how the subset of the domain Ω_T, where the generalised solution is non-smooth or not $\mathcal{G}^\infty$-regular, is related to the set of singular points of the data and coefficients.

For partial differential operators with smooth coefficients, the propagation of initial data singularities is well understood, based on Hörmander's notion of the wave front set and the techniques of microlocal analysis (cf. [Hör03, Chapter VIII]). In Cauchy problems for PDEs with non-smooth coefficients, the presence and propagation of singularities in the solution is affected by the coefficient singularities as well. However, the influence of coefficient singularities on the singularity structure of the solution and their possible propagation is less understood, even in the case of smooth initial data and right-hand side. Colombeau theory provides a promising framework to study such questions, especially in situations where both coefficients and data are non-smooth.

The concept of the wave front set was transferred to Colombeau algebras in [NPS98] and microlocal analysis in this setting was further developed in [Hör99, HK01, GH05]. Microlocal analysis of propagation of singularities for hyperbolic PDEs with discontinuous coefficients in Colombeau algebras goes back to [HdH01], more recent results can be found in [Deg11, DHO13, DO16].

In general, precise information on the singularity structure of the solution to an initial value problem is very difficult to obtain, when the coefficients are non-smooth. One approach to explore and test possible dependencies and effects is to study equations for which the solution can be represented in explicit form in terms of the initial data and right-hand side. In this section we discuss an example of a scalar PDE with principal coefficients depending on time only, intended as a case study for the propagation of coefficient singularities.

We consider the initial value problem

$$\partial_t u(t,x) + a(t)\partial_x u(t,x) + b(t,x)u(t,x) = f(t,x), \tag{2.29}$$

$$u(t_0,x) = g(x). \tag{2.30}$$

for $(t,x) \in]t_0,T[\, \times \mathbb{R}$ with smooth data f, g, and smooth coefficients a, b. The solution to (2.29–2.30) is given by (cf. [Trè75, Chapter II, Section 16])

$$u(t,x) = e^{-\int\limits_{t_0}^{t} b\left(s,T_s^t(x)\right)\mathrm{d}s} \left(g\left(T_{t_0}^t(x)\right) + \int\limits_{t_0}^{t} e^{\int\limits_{t_0}^{s} b\left(r,T_r^t(x)\right)\mathrm{d}r} f\left(s,T_s^t(x)\right)\mathrm{d}s \right),$$

where

$$T_s^t(x) := x - \int_s^t a(r)\mathrm{d}r.$$

In fact, $s \mapsto \left(s, T_s^t(x)\right)^\mathsf{T}$ is the projected characteristic curve γ running through the point (t,x) at parameter value $s = t$, that is,

$$\gamma(t,x,\cdot) : \mathbb{R} \to \mathbb{R}^2, \qquad s \mapsto \left(\gamma^0(t,x,s), \gamma^1(t,x,s)\right)^\mathsf{T} = \left(s, T_s^t(x)\right)^\mathsf{T}.$$

Using this notation we may also write

$$u(t,x) \quad = \quad e^{-\int\limits_{t_0}^{t} b\circ\gamma(t,x,s)\mathrm{d}s} \left(g \circ \gamma^1(t,x,s) + \int\limits_{t_0}^{t} e^{\int\limits_{t_0}^{s} b\circ\gamma(t,x,r)\mathrm{d}r} f \circ \gamma(t,x,s)\mathrm{d}s \right).$$

The proof is a straight-forward calculation. The advantage of studying the exemplary PDE in (2.29) lies in the fact that a more or less explicit solution formula exists. This in turn allows not only to obtain distributional limits for much rougher coefficients than those admissible in general results on weak solutions, but also to investigate relations between singularities in the coefficients of the PDE and singularities in the solution. A particularly interesting subject in this respect is the possibility of coefficient singularities propagating along the characteristics of the PDE, although initial data and right-hand side are smooth. The solution might then be singular along characteristic curves, while the only non-smoothness in the PDE stems, for instance, from an isolated singularity in a coefficient of the PDE (that is, the solution will be singular at points where all coefficients are smooth).

To illustrate this phenomenon, we examine a special case of (2.29–2.30) in more detail and consider the generalised initial value problem

$$\begin{aligned} \partial_t u(t,x) - \partial_x u(t,x) + b(t,x)u(t,x) &= 0 \quad \text{in} \quad \mathcal{G}(I \times \mathbb{R}) \\ u(-1,x) &= g(x) \quad \text{in} \quad \mathcal{G}(\mathbb{R}), \end{aligned}$$

where $I =]-1, T[$ and $T > 0$. Here b and g are Colombeau generalised functions and we are looking for a solution $u \in \mathcal{G}(I \times \mathbb{R})$. For fixed ε, the classical smooth solution at (t,x) is given by

$$u_\varepsilon(t,x) = \exp\left(- \int\limits_x^{1+t+x} b_\varepsilon(t+x-s,s)\mathrm{d}s \right) g_\varepsilon(1+t+x). \tag{2.31}$$

We thus obtain a natural candidate for the generalised solution. From the structure of $u_\varepsilon(t,x)$ it is clear that any $g \in \mathcal{G}(\mathbb{R})$ is admissible as initial data. Then, for the family of ε-wise solutions $(u_\varepsilon)_\varepsilon$ to be moderate, we basically need the integral

$$\mathcal{F}_\varepsilon(t,x) := \int\limits_x^{1+t+x} b_\varepsilon(t+x-s,s)\mathrm{d}s$$

to be of L^∞-log-type, that is, $\|\mathcal{F}_\varepsilon\|_{L^\infty(I\times\mathbb{R})} = O(\log\frac{1}{\varepsilon})$ as $\varepsilon \to 0$. We aim at deriving asymptotic estimates that guarantee moderateness of the solution. Considering only coefficients with tensor product structure, that is $b(t,x) = \rho(t) \otimes h(x)$, we have

$$\mathcal{F}_\varepsilon(t,x) = \int\limits_x^{1+t+x} \rho_\varepsilon(t+x-s)h_\varepsilon(s)\mathrm{d}s.$$

Hence, the following estimates are possible:

(i) $|\mathcal{F}_\varepsilon(t,x)| \leq \|\rho_\varepsilon\|_{L^\infty(I)}\|h_\varepsilon\|_{L^1(\mathbb{R})}$

(ii) $|\mathcal{F}_\varepsilon(t,x)| \leq \|\rho_\varepsilon\|_{L^1(I)}\|h_\varepsilon\|_{L^\infty(\mathbb{R})}$

(iii) $|\mathcal{F}_\varepsilon(t,x)| \leq \|\rho_\varepsilon\|_{L^2(\mathbb{R})}\|h_\varepsilon\|_{L^2(\mathbb{R})}$

For positive functions ρ_ε and h_ε, we may also estimate $|\mathcal{F}_\varepsilon(t,x)|$ by extending the integral to a convolution integral, that is,

(iv) $|\mathcal{F}_\varepsilon(t,x)| \leq \|\rho_\varepsilon * h_\varepsilon\|_{L^\infty(\mathbb{R})} \leq \|\rho_\varepsilon\|_{L^p(\mathbb{R})}\|h_\varepsilon\|_{L^q(\mathbb{R})}$

with $\frac{1}{p} + \frac{1}{q} = 1$. Based on these estimates, it is easy to formulate proper asymptotic growth conditions on the nets $(\rho_\varepsilon)_\varepsilon$ and $(h_\varepsilon)_\varepsilon$, ensuring moderateness of $(u_\varepsilon)_\varepsilon$ and thus the existence of a unique generalised solution.

By estimate (ii) it is possible, for instance, to consider $\rho_\varepsilon \to \delta$ $(\varepsilon \to 0)$ and $h_\varepsilon \to H$ $(\varepsilon \to 0)$, representing a coefficient associated with the distribution

$$b_0(t,x) = \delta(t) \otimes H(x).$$

Note that a logarithmic scaling of the delta net $(\rho_\varepsilon)_\varepsilon$ is not necessary. We further specify the coefficients $b_\varepsilon(t,x) = \rho_\varepsilon(t) \otimes h_\varepsilon(x)$ by assuming that

- $(\rho_\varepsilon)_\varepsilon$ is a strict delta net with $\mathrm{supp}(\rho_\varepsilon) \subseteq [-\varepsilon,\varepsilon]$,
- $h_\varepsilon(t) := \int\limits_{-\infty}^{t} \rho_\varepsilon(s)\mathrm{d}s$,
- $g_\varepsilon \to g_0 \in \mathcal{C}^0(\mathbb{R})$ in the L^∞-norm.

Then $\|\mathcal{F}_\varepsilon\|_{L^\infty(I\times\mathbb{R})}$ is uniformly bounded in ε, implying that this is also true for representatives $(u_\varepsilon)_\varepsilon$ of the solution. Next we calculate the distributional limit of the generalised solution. To this end we divide the domain $I\times\mathbb{R}$ into the subdomains I, II, and III, as shown in Figure 2.3.

$\boxed{\text{I}}$ For $x > 0$ and ε_0 small enough, we may write

$$\begin{aligned}\mathcal{F}_\varepsilon(t,x) = \int\limits_x^{1+t+x} \rho_\varepsilon(t+x-s)h_\varepsilon(s)\mathrm{d}s &\overset{\forall\varepsilon\leq\varepsilon_0}{=} \int\limits_x^{1+t+x} \rho_\varepsilon(t+x-s)\mathrm{d}s \\ &\overset{r=t+x-s}{=} \int\limits_{-1}^{t} \rho_\varepsilon(r)\mathrm{d}r = h_\varepsilon(t) - h_\varepsilon(-1) \overset{\forall\varepsilon\leq\varepsilon_0}{=} h_\varepsilon(t).\end{aligned}$$

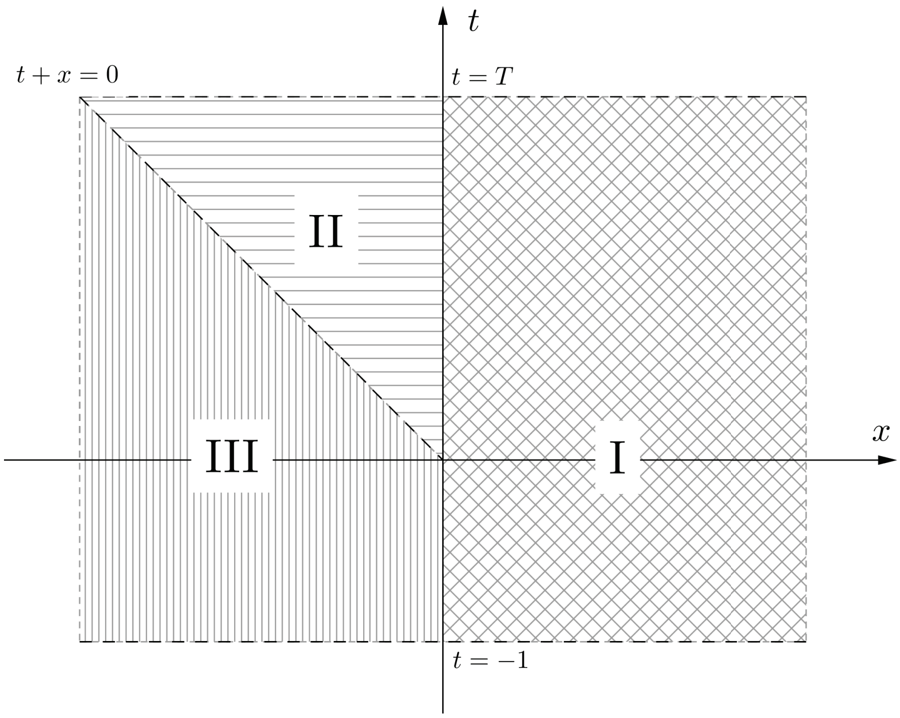

Figure 2.3: Division of the domain $I \times \mathbb{R}$ into the subdomains I, II, and III.

In the first line we have used that $h_\varepsilon(s) = 1$ for all $s > \varepsilon$ and therefore $h_\varepsilon(s) = 1$ in the whole domain of integration if $\varepsilon_0 < x$. Since $u_\varepsilon(t,x) = \exp(-\mathcal{F}_\varepsilon(t,x))g_\varepsilon(1+t+x)$, we obtain the pointwise convergence

$$u_\varepsilon(t,x) \xrightarrow{\varepsilon\to 0} u_0^{\mathrm{I}}(t,x) = \left(1 + \left(\frac{1}{e} - 1\right) H(t)\right) g_0(1+t+x)$$

in subdomain I, that is, for $x > 0$ and $-1 < t < T$.

$\boxed{\mathrm{II}}$ For $x < 0$ and $t + x > 0$, we may write for all $\varepsilon < -x$,

$$\mathcal{F}_\varepsilon(t,x) = \int\limits_{x}^{1+t+x} \rho_\varepsilon(t+x-s)h_\varepsilon(s)\mathrm{d}s = \underbrace{\int\limits_{-\varepsilon}^{\varepsilon} \rho_\varepsilon(t+x-s)h_\varepsilon(s)\mathrm{d}s}_{=:I_1(t+x,\varepsilon)} + \underbrace{\int\limits_{\varepsilon}^{1+t+x} \rho_\varepsilon(t+x-s)\mathrm{d}s}_{=:I_2(t+x,\varepsilon)},$$

since the integrand vanishes vor $x < -\varepsilon$ and $h_\varepsilon(s) = 1$ for $s \geq \varepsilon$. For all $\varepsilon < \varepsilon_0$ small enough, we have

- $I_1(t+x,\varepsilon) = 0$, since the interval of integration $[-\varepsilon,\varepsilon]$ will be disjoint with $\mathrm{supp}(s \mapsto \rho_\varepsilon(t+x-s))$, and
- $I_2(t+x,\varepsilon) = 1$, since $\mathrm{supp}(s \mapsto \rho_\varepsilon(t+x-s)) \subseteq [\varepsilon, 1+t+x]$.

Thus we get $\mathcal{F}_\varepsilon(t,x) = 1$ for $\varepsilon < \varepsilon_0$ small enough and therefore, by (2.31),

$$u_\varepsilon(t,x) \overset{\varepsilon\to 0}{\longrightarrow} u_0^{\mathrm{II}}(t,x) = \frac{1}{e}\, g_0(1+t+x)$$

pointwise in subdomain II, that is, for $x<0$, $-1<t<T$, and $t+x>0$.

$\boxed{\mathrm{III}}$ Now let $x<0$ and $t+x<0$. First note that if $t+x<-1$, the interval $[x,1+t+x]$ will be disjoint with $\mathrm{supp}(h_\varepsilon)$ for all $\varepsilon<\varepsilon_0$ small enough, implying that $\mathcal{F}_\varepsilon(t,x)=0$. If $-1<t+x<0$, we again split the integral for all $\varepsilon<-x$. Then, considering $I_1(t+x,\varepsilon)$ and $I_2(t+x,\varepsilon)$, we observe that for all $\varepsilon<\varepsilon_0$ small enough,

- $I_1(t+x,\varepsilon)=0$, since the interval of integration $[-\varepsilon,\varepsilon]$ will be disjoint with $\mathrm{supp}(s\mapsto\rho_\varepsilon(t+x-s))$, and
- $I_2(t+x,\varepsilon)=0$, since the interval of integration $[\varepsilon,1+t+x]$ will also be disjoint with $\mathrm{supp}(s\mapsto\rho_\varepsilon(t+x-s))$.

Thus, we obtain $\mathcal{F}_\varepsilon(t,x)=0$ for $-1<t+x<0$ as well. To sum up, we obtain the convergene

$$u_\varepsilon(t,x) \overset{\varepsilon\to 0}{\longrightarrow} u_0^{\mathrm{III}}(t,x) = g_0(1+t+x)$$

pointwise in subdomain III, that is, for $x<0$, $-1<t<T$, and $t+x<0$.

Combining these results, the uniform boundedness of $\|u_\varepsilon\|_{L^\infty(I\times\mathbb{R})}$ together with pointwise convergence almost everywhere allows us to write the distributional limit $u_0 \in L^\infty(I\times\mathbb{R})$ of the generalised solution as

$$u_0(t,x) = \begin{cases} \left(1+\left(\frac{1}{e}-1\right)H(t)\right) g_0(1+t+x) & x>0, \\ \left(1+\left(\frac{1}{e}-1\right)H(t+x)\right) g_0(1+t+x) & x<0. \end{cases}$$

In particular, for constant initial data $g_0=1$ we may write this as

$$u_0(t,x) = H(x)\left(1+\left(\frac{1}{e}-1\right)H(t)\right) + (1-H(x))\left(1+\left(\frac{1}{e}-1\right)H(t+x)\right),$$

which is shown in Figure 2.4. The singular support of u_0 is therefore

$$\mathrm{singsupp}(u_0) = \{(t,x) : x\geq 0 \wedge t=0)\} \cup \{(t,x) : x\leq 0 \wedge t+x=0\},$$

whereas the singular support of the coefficient b_0 is given by

$$\mathrm{singsupp}(b_0) = \mathrm{singsupp}(\delta\otimes H) = \{(t,x) : x\geq 0 \wedge t=0)\},$$

implying that

$$\mathrm{singsupp}(u_0)\setminus\mathrm{singsupp}(b_0) = \{(t,x) : x\leq 0 \wedge t+x=0\}.$$

Thus, although the initial data are smooth, the distributional limit of the generalised solution is not only discontinuous at points where the coefficient is singular, but also

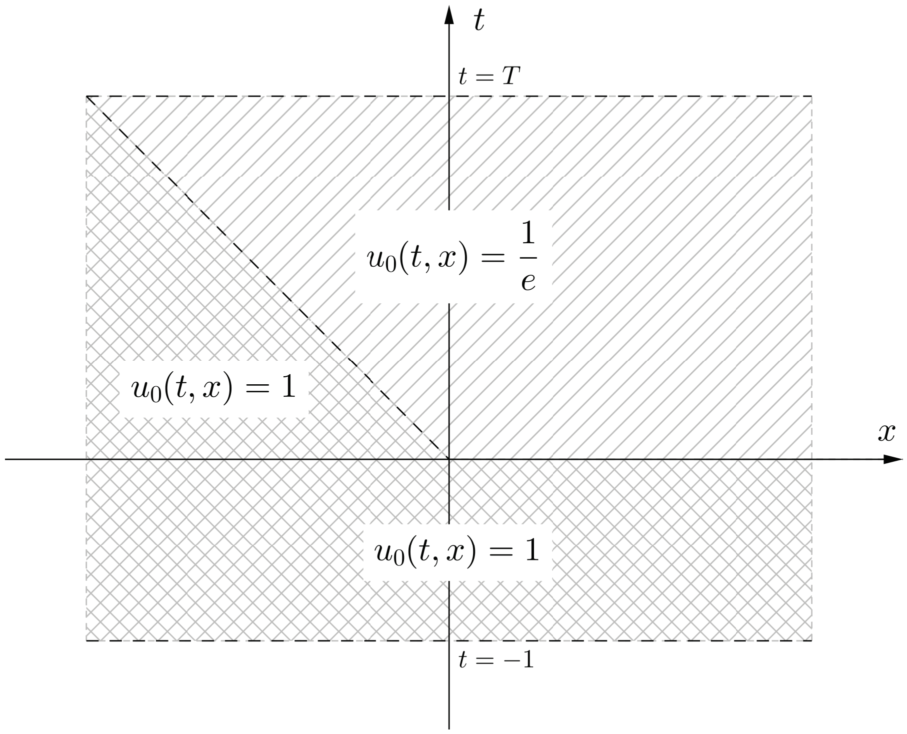

Figure 2.4: Distributional limit of the generalised solution of the initial value problem (2.29–2.30) with $b_0(t,x) = \delta(t) \otimes H(x)$ and $g_0 = 1$.

at points where the coefficient is smooth. A singularity in the coefficient of the PDE propagates along a characteristic curve. The possible visibility of coefficient singularities in the singularity structure of solutions of initial value problems of the type (2.29–2.30) has also been discussed in [HdH01, Section 5].

Note that a coefficient of the form $b_0(t,x) = \delta(t) \otimes H(x)$ is admissible in Proposition 2.36. In fact, the initial value problem (2.29–2.30) can be regarded as a special case of the more general situation covered there.

Inspecting the proof of Proposition 2.36, in particular the splitting of the generalised solution into a continuous factor and a potentially singular part, it is easy to see that propagation of coefficient singularities will occur in much more general circumstances than those we have considered here, with respect to the regularity of the coefficients as well as the dimension of space and the number of equations. For instance, in the initial value problem (2.29–2.30), we may also consider a bounded coefficient of the form $b_0(t,x) = H(t) \otimes H(x)$ or, without time dependence, $b_0(x) = H(x-1)$. Then the distributional limit of the generalised solution would be continuous, but not differentiable along the line $t + x = 0$ (recall that we imposed initial data at time $t = -1$).

However, if we replace the coefficient $b_0(t,x) = \delta(t) \otimes H(x)$ by $\widetilde{b}_0(t,x) = \delta(t) \otimes (1 + H(x))$ and compute the corresponding solution $\widetilde{u}_0$ of the initial value problem (2.29–2.30) with $g_0 = 1$, then its singular support remains the same, that is $\operatorname{singsupp}(\widetilde{u}_0) = \operatorname{singsupp}(u_0)$, although the singular support of the coefficient has changed, that is, $\operatorname{singsupp}(\widetilde{b}_0) \neq \operatorname{singsupp}(b_0)$. On the other hand, we may consider $\overline{b}_0(t,x) = \delta(t)$, for which no propagation of coefficient singularities occurs, that is, $\operatorname{singsupp}(\overline{u}_0) = \operatorname{singsupp}(\overline{b}_0)$ and thus $\operatorname{singsupp}(\overline{u}_0) \neq \operatorname{singsupp}(\widetilde{u}_0)$, although $\operatorname{singsupp}(\overline{b}_0) = \operatorname{singsupp}(\widetilde{b}_0)$. These observations imply, as one would expect, that the notion of singular support is insufficient to fully describe the propagation of coefficient singularities. Microlocal analysis in Colombeau algebras offers a setting that could be capable of providing the means to study coefficient singularity propagation on a mathematically rigorous level.

From a physical viewpoint, the propagation of coefficient singularities is a very realistic scenario, for example in the context of seismic waves. An initially smooth wave may pick up discontinuities by propagating through a medium with singular material parameters. A precise mathematical description of the relation between coefficient singularities and the singularity structure in the solution of an initial value problem would also be desirable for various other applications. Recent progress in this direction of research has been made in [DO16]. The authors derive upper and lower bounds on the $\mathcal{G}^\infty$-singular support of the generalised solution of the one-dimensional wave equation with an x-dependent coefficient exhibiting a jump singularity.

Chapter 3

Second-order wave equations in algebras of generalised functions

Although any system of PDEs can formally be rewritten as a first-order system (cf. [Trè75, Section 18]), the results of Chapter 2 on generalised solutions of the initial value problem for symmetric hyperbolic systems are not directly applicable to second-order hyperbolic PDEs. The transformation from a hyperbolic second-order PDE to a symmetric first-order system requires nonlinear operations of the coefficients, which are not defined for coefficients of low regularity. However, PDEs in physics often present themselves as second-order equations. To utilise the full scope of Theorems 2.7, 2.8, and 2.13 for applications in physics, this chapter is devoted to a reformulation of these results for second-order wave equations. The content of Sections 3.1 and 3.2 is based on [HHSS13]. Section 3.3 briefly discusses the Hadamard parametrix construction for second-order elliptic PDEs.

Wave equations arising via the Laplace-Beltrami operator of a Lorentzian metric of low regularity have been studied in the framework of Colombeau generalised functions in [VW00, May06, GMS09, Han11, HKS12], motivated by potential applications in general relativity. For a wide class of 'weakly singular' space-time metrics, local and global existence and uniqueness of generalised solutions have been established. By translating Cauchy problems for second-order wave equations of low regularity into Cauchy problems for symmetric hyperbolic systems, results obtained for the latter can be made available to the existence and uniqueness theory for the former and, albeit in a less general sense, also vice versa. In view of the broad spectrum of results that have already been achieved in 'both worlds', separated from each other, re-establishing their connection in a low-regularity setting is a reasonable undertaking.

General methods of modern analysis for the reduction of hyperbolic higher-order PDEs to (symmetric) hyperbolic first-order systems rely on pseudodifferential operator techniques (cf. [Tay81, Chapter IV], and [Kg81]). The mathematical operations employed in the reduction algorithm are highly nonlinear and require a certain smoothness of the coefficients (or symbols). In a low-regularity setting, the necessary manipulations on the coefficients are no longer well-defined on the level of distributions. Colombeau theory provides the means to retain the transformation algorithm for rewriting a second-order scalar PDE as a first-order system in the context of low regularities.

Notation

In the following, we will write v_i for the components of a vector v and R_{ij} for the components of a matrix R. The spatial gradient of a scalar function u shall be ∇u and $\mathrm{Hess}(u)$ shall represent the spatial Hessian of u. We denote the i^{th} row or j^{th} column of a matrix R by $R_{i\cdot}$ or $R_{\cdot j}$ respectively. For any symmetric matrix S, we set the divergence $(\operatorname{div} S)_i = \sum_{j=1}^{n} \partial_{x_j} S_{ij}$, meaning that $\operatorname{div} S$ is a vector whose i^{th} entry is the divergence of the i^{th} row of the matrix S. Partial derivatives written as $\partial_j S$ include time derivatives $\partial_t S$ and spatial derivatives $\partial_{x_j} S$.

3.1 Transformation to a first-order system

We want to relate solutions of wave equations to solutions of corresponding symmetric first-order systems. Omitting initial conditions for the moment, we consider a scalar wave equation in $\mathcal{G}(\mathbb{R}^{1+n})$,

$$-\partial_t^2 u + 2\sum_{i=1}^{n} h_i \partial_{x_i}\partial_t u + \sum_{i,j=1}^{n} R_{ij}\partial_{x_i}\partial_{x_j} u + a\partial_t u + \sum_{i=1}^{n} b_i \partial_{x_i} u + cu = f, \tag{3.1}$$

where $R = (R_{ij})$ is a positive definite, symmetric matrix of generalised functions, h and b are vectors with entries in $\mathcal{G}(\mathbb{R}^{1+n})$, and a, c, and f are generalised functions. Thus the principal part of this wave equation is derived from $g = \left(\begin{smallmatrix} -1 & h^{\mathsf{T}} \\ h & R \end{smallmatrix}\right)$, defining a generalised Lorentzian metric[1] (cf. [May08, Section 2.4]). There are several ways to rewrite equation (3.1) as a first-order system (cf. [Eva98, Section 7.3.2]). From the possible algorithms, we employ one that guarantees the symmetry of the arising system. To this end we construct the positive definite matrix $S = R^{\frac{1}{2}}$ via diagonalisation (cf. [HKK14, Proposition 3.25]) and define the vector $w = (u, \partial_t u, S\nabla u)^{\mathsf{T}}$, thereby arriving at the system

$$-\partial_t w + \sum_{j=1}^{n} A^j \partial_{x_j} w + Bw = F \tag{3.2}$$

in $\mathcal{G}(\mathbb{R}^{1+n})^{2+n}$, where

$$\begin{aligned} A^j &= \begin{pmatrix} 0 & 0 & 0_{1\times n} \\ 0 & 2h_j & S_{j\cdot} \\ 0_{n\times 1} & S_{\cdot j} & 0_{n\times n} \end{pmatrix} \qquad F = \begin{pmatrix} 0 \\ f \\ 0_{n\times 1} \end{pmatrix} \\ B &= \begin{pmatrix} 0 & 1 & 0_{1\times n} \\ c & a & (\operatorname{div} S)^{\mathsf{T}} + (b - \operatorname{div} S^2)^{\mathsf{T}} S^{-1} \\ 0_{n\times 1} & 0_{n\times 1} & -(\partial_t S)S^{-1} \end{pmatrix}. \end{aligned} \tag{3.3}$$

[1] In fact, our arguments still hold in the more general case, where the matrix entry g_{00} is a strictly negative generalised function, upon dividing equation (3.1) by g_{00}.

On the level of representatives, the square root matrix S is given by

$$S_\varepsilon(t,x) = U_\varepsilon(t,x)^\mathsf{T} \Lambda_\varepsilon(t,x)^{\frac{1}{2}} U_\varepsilon(t,x)$$

with

$$\Lambda_\varepsilon(t,x)^{\frac{1}{2}} = \operatorname{diag}\Big(\sqrt{\lambda_{1,\varepsilon}(t,x)}, \dots, \sqrt{\lambda_{n,\varepsilon}(t,x)}\Big),$$

and $\lambda_{1,\varepsilon}, \dots, \lambda_{n,\varepsilon}$ are the eigenvalues of a symmetric and positive definite representative $(R_\varepsilon)_\varepsilon$ of R. The nets $(\Lambda_\varepsilon)_\varepsilon$ and $(U_\varepsilon)_\varepsilon$ do not necessarily have smooth entries, yet the product $(U_\varepsilon^\mathsf{T} \Lambda_\varepsilon^{\frac{1}{2}} U_\varepsilon)_\varepsilon$ is smooth by virtue of the following lemma.

Lemma 3.1 (Smoothness of the square root)**.** *Let $\mathbb{S}_n^+(\mathbb{R})$ denote the space of positive definite symmetric matrices in $M_n(\mathbb{R})$. Then the smooth map $f : \mathbb{S}_n^+(\mathbb{R}) \to \mathbb{S}_n^+(\mathbb{R})$ with $f(A) = A^2$ is a diffeomorphism.*

Proof. The map $f : \mathbb{S}_n^+(\mathbb{R}) \to \mathbb{S}_n^+(\mathbb{R})$ is bijective (cf. [Lan02, Proposition 6.8]) and we employ the inverse function theorem to conclude that f is a global diffeomorphism. Let $\mathbb{S}_n(\mathbb{R})$ denote the space of symmetric matrices in $M_n(\mathbb{R})$. Since $\mathbb{S}_n^+(\mathbb{R})$ is an open subset of $\mathbb{S}_n(\mathbb{R})$, we may identify the tangent space $\mathrm{T}_A\mathbb{S}_n^+(\mathbb{R})$ for $A \in \mathbb{S}_n^+(\mathbb{R})$ with $\mathbb{S}_n(\mathbb{R})$ and obtain the derivative $\mathrm{d}f(A)(B) = AB + BA$. To show that $\mathrm{d}f(A)$ is injective, we notice that for any $B \neq 0$ there exists $\lambda \neq 0$ such that $Bv = \lambda v$ for some $v \neq 0$. Now assume that $AB + BA = 0$. Then $BAv = -ABv = -\lambda Av$ and thus Av is an eigenvector of B associated with the eigenvalue $-\lambda$. Since B is symmetric, this means that Av is orthogonal to v, that is $\langle Av, v\rangle = 0$, contradicting positive definiteness of A. $\square$

From the wave equation to a first-order system

Assuming that u is a solution of an initial value problem for the wave equation (3.1), we show that the vector $w = (u, \partial_t u, S\nabla u)^\mathsf{T}$ is indeed a solution of a corresponding initial value problem for the first-order system (3.2). Note that since we do not specify any further assumptions on the coefficients other than being generalised functions, uniqueness of solutions is not guaranteed, but also not relevant at this stage of our considerations.

Lemma 3.2 (Transformation to a system)**.** *Let $u^0, u^1 \in \mathcal{G}(\mathbb{R}^n)$ and consider the second-order equation* (3.1) *with the initial condition*

$$(u, \partial_t u)|_{t=0} = (u^0, u^1). \tag{3.4}$$

Then, if $u \in \mathcal{G}(\mathbb{R}^{1+n})$ is a solution of this initial value problem, the vector $w = (u, \partial_t u, S\nabla u)^\mathsf{T}$ represents a solution of the first-order system (3.2) *with initial condition*

$$w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\mathsf{T}. \tag{3.5}$$

Proof. We basically have to rewrite the wave equation in terms of the three components of w. First we rewrite equation (3.1) in divergence form, that is

$$-\partial_t^2 u + 2\langle h, \partial_t \nabla u\rangle + \operatorname{div}(S^2 \nabla u) + a\partial_t u + \langle b - \operatorname{div} S^2, \nabla u\rangle + cu = f.$$

Introducing new variables $z := \partial_t u$ and $v := S\nabla u$, we may write this as

$$-\partial_t z + 2\langle h, \nabla z\rangle + \operatorname{div}(Sv) + az + \langle b - \operatorname{div} S^2, S^{-1}v\rangle + cu = f.$$

Hence the vector $w = (u, \partial_t u, S\nabla u)^\mathsf{T}$ satisfies the system of equations

$$-\partial_t w + \sum_{i=1}^{n} \begin{pmatrix} 0 & 0 & 0_{1\times n} \\ 0 & 2h_i & S_{i\cdot} \\ 0_{n\times 1} & S_{\cdot i} & 0_{n\times n} \end{pmatrix} \partial_{x_i} w$$
$$+ \begin{pmatrix} 0 & 1 & 0_{1\times n} \\ c & a & (\operatorname{div} S)^\mathsf{T} + (b - \operatorname{div} S^2)^\mathsf{T} S^{-1} \\ 0_{n\times 1} & 0_{n\times 1} & (\partial_t S)S^{-1} \end{pmatrix} w = \begin{pmatrix} 0 \\ f \\ 0_{n\times 1} \end{pmatrix},$$

where we have used that $\operatorname{div}(Sv) = \sum_{i=1}^n S_{i\cdot}\partial_{x_i} v + (\operatorname{div} S)^\mathsf{T} v$. Note that the second equation in the above system is just the original wave equation written in new variables, whereas the other equations represent the transformation of variables. Finally, evaluating w at time $t = 0$ yields $w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\mathsf{T}$. $\square$

From a first-order system to the wave equation

We now look at the converse situation: We start with a system of the form (3.2), where S is a symmetric and invertible n-dimensional matrix (not necessarily positive definite) with entries in $\mathcal{G}(\mathbb{R}^{1+n})$, h and b are vectors with entries in $\mathcal{G}(\mathbb{R}^{1+n})$, and a and c are generalised functions. Then, given a solution of the initial value problem (3.2,3.5), we show that the first component of the solution vector yields a solution of the initial value problem (3.1,3.4).

Lemma 3.3 (Transformation back to the wave equation)**.** *Let u^0 and u^1 belong to $\mathcal{G}(\mathbb{R}^n)$. If $w = (u, z, v)^\mathsf{T} \in \mathcal{G}(\mathbb{R}^{1+n})^{2+n}$ is a solution to the first-order system* (3.2) *with initial condition $w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\mathsf{T}$, then u is a solution to the wave equation* (3.1) *with initial condition $(u, \partial_t u)|_{t=0} = (u^0, u^1)$.*

Proof. Observe that the first equation of our system is just $z = \partial_t u$. The last n equations read $-\partial_t v + S\nabla z + (\partial_t S)S^{-1}v = 0$, which is equivalent to

$$S\nabla z = S\partial_t(S^{-1}v) \tag{3.6}$$

since $S\partial_t(S^{-1}v) = \partial_t(SS^{-1}v) - (\partial_t S)S^{-1}v = \partial_t v - (\partial_t S)S^{-1}v$. Multiplying by S^{-1} from the left in equation (3.6) and substituting $z = \partial_t u$, we find

$$\partial_t(\nabla u - S^{-1}v) = 0. \tag{3.7}$$

By the initial condition we have $v|_{t=0} = S|_{t=0}\nabla u^0$ and therefore $(\nabla u - S^{-1}v)|_{t=0} = 0$. Equation (3.7) then implies $\nabla u - S^{-1}v = 0$ for all t, which is equivalent to $v = S\nabla u$.

Hence, replacing z by $\partial_t u$ and v by $S\nabla u$, the second equation of the system becomes

$$-\partial_t^2 u + 2\sum_{i=1}^n h_i \partial_{x_i}\partial_t u + a\partial_t u + \sum_{i=1}^n b_i \partial_{x_i} u + cu$$
$$+ \sum_{i,j,k=1}^n \big(S_{ij}\partial_{x_i}(S_{jk}\partial_{x_k} u) + (\partial_{x_i} S_{ij}) S_{jk}\partial_{x_k} u - \partial_{x_i}(S_{ij}S_{jk})\partial_{x_k} u\big) = f,$$

which is the same as

$$-\partial_t^2 u + 2\sum_{i=1}^n h_i \partial_{x_i}\partial_t u + \sum_{i,j,k=1}^n S_{ik}S_{kj}\partial_{x_i}\partial_{x_j} u + a\partial_t u + \sum_{i=1}^n b_i \partial_{x_i} u + cu = f.$$

Moreover, the equality $(u, \partial_t u)|_{t=0} = (u^0, u^1)$ is a direct consequence of the initial condition for w. □

Remark 3.4 (Structure of the first-order system)**.** Obviously, the coefficient matrices in the PDE system (3.2) are of a very special type, which simply reflects the fact that in general, a first-order system cannot be transformed into a scalar PDE. However, note that the complete second row of matrix B can be chosen freely, as the structure of the term $(\operatorname{div} S)^\mathsf{T} + (b - \operatorname{div} S^2)^\mathsf{T} S^{-1}$ does not impose any constraints on the corresponding entries of the matrix. Indeed, let $\widetilde{b}$ be an arbitrary element of $\mathcal{G}(\mathbb{R}^{1+n})^n$ and set $\widetilde{b}^\mathsf{T} = (\operatorname{div} S)^\mathsf{T} + (b - \operatorname{div} S^2)^\mathsf{T} S^{-1}$. Multiplying with S from the right gives

$$\widetilde{b}^\mathsf{T} S = (\operatorname{div} S)^\mathsf{T} S + (b - \operatorname{div} S^2)^\mathsf{T}$$

and transposition leads to an equation for b entirely in terms of S and $\widetilde{b}$,

$$b = S^\mathsf{T}\widetilde{b} - S^\mathsf{T}\operatorname{div} S + (\operatorname{div} S^2).$$

Equivalence of initial value problems

In the following theorem, we summarise the content of Lemmas 3.2 and 3.3, showing that the problem of finding a solution of the Cauchy problem for the wave equation (3.1) is equivalent to the problem of finding a solution of the corresponding Cauchy problem for the first-order system (3.2). In addition, uniqueness of solutions is also preserved during the transformation process.

Theorem 3.5 (Equivalence of initial value problems)**.** *Let $u^0, u^1 \in \mathcal{G}(\mathbb{R}^n)$, let $u \in \mathcal{G}(\mathbb{R}^{1+n})$, and define $w = (u, \partial_t u, S\nabla u)^\mathsf{T} \in \mathcal{G}(\mathbb{R}^{1+n})^{2+n}$. Then the following are equivalent:*

(i) *The generalised function u is the unique solution of the wave equation* (3.1) *with initial condition $(u, \partial_t u)|_{t=0} = (u^0, u^1)$.*

(ii) *The vector-valued generalised function w is the unique solution of the first-order system* (3.2) *with initial condition $w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\mathsf{T}$.*

Proof. The translation of solutions between wave equations and first-order systems is an immediate consequence of Lemmas 3.2 and 3.3. Equivalence of unique solvability is more or less obvious, as the following considerations elucidate.

(i)⇒(ii): Let u be the unique solution of the initial value problem (3.1, 3.4) and suppose there were two different solutions w and $\widetilde{w}$ of the initial value problem (3.2,3.5) with $w|_{t=0} = \widetilde{w}|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\mathsf{T}$. Then the first components of w and $\widetilde{w}$ would represent two different solutions of (3.1,3.4). But since the problem (3.1,3.4) has a unique solution, the first component of $\widetilde{w}$ must be equal to the first component of w. From the proof of Lemma 3.3 it is then clear that $\widetilde{z} = z$ and $\widetilde{v} = v$, hence $\widetilde{w} = w$.

(ii)⇒(i): By Lemma 3.2, two different solutions of the initial value problem (3.1,3.4) for the wave equation would give rise to two different solutions of the initial value problem (3.2,3.5), since $u \mapsto w = (u, \partial_t u, S\nabla u)^\mathsf{T}$ is injective. Hence unique solvability of the initial value problem (3.2,3.5) implies unique solvability of the initial value problem (3.1,3.4). □

Theorem 3.5 guarantees that in the context of the differential algebra $\mathcal{G}$, the initial value problem for the second-order wave equation (3.1) is equivalent to the initial value problem for the corresponding first-order system (3.2–3.3). This is true in a very general sense, if only the natural, and merely algebraic, consistency of initial data holds. However, this is not true in spaces of distributions. When the coefficients are of low regularity, the transformation process may fail at various places. For instance, we might end up with differential equations that do not make sense in space of distributions, as we want to illustrate with the help of the following example.

Example 3.6 (Ill-definedness of the transformation algorithm in a distributional setting)**.** We consider a wave equation frequently used as a model in linear acoustics. Let $p : \mathbb{R}^2 \to \mathbb{R}$ and let $c, \rho : \mathbb{R}^2 \to [a, b]$ with $a > 0$. Then the acoustic wave equation can be written as

$$\partial_t^2 p - c^2 \rho\, \partial_x \left(\frac{1}{\rho} \partial_x p \right) = 0, \tag{3.8}$$

where p represents the pressure in the medium, ρ is the density, and c is the propagation speed. All coefficients may depend on both the time variable and the space variable. We define $w = (p, \partial_t p, c\, \partial_x p)^\mathsf{T}$ and formally obtain the symmetric hyperbolic system

$$-\partial_t w_1 + w_2 = 0, \tag{3.9a}$$

$$-\partial_t w_2 + c\partial_x w_3 - (\partial_x c + c\, \partial_x(\ln \rho)) w_3 = 0, \tag{3.9b}$$

$$-\partial_t w_3 + c\partial_x w_2 + (\partial_t c) w_3 = 0. \tag{3.9c}$$

Equation (3.8) can be regarded as an equality in $L^2(\mathbb{R}^2)$, if we have $p \in H^2(\mathbb{R}^2)$, $\rho \in \mathrm{Lip}(\mathbb{R}^2)$ and $c \in L^\infty(\mathbb{R}^2)$. We then have $\partial_t p \in H^1(\mathbb{R}^2)$ and $c\, \partial_x p \in L^2(\mathbb{R}^2)$, since $c \in L^\infty(\mathbb{R}^2)$ and $\partial_x p \in H^1(\mathbb{R}^2)$. Therefore we obtain $w \in H^2(\mathbb{R}^2) \times H^1(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$, but not better in general circumstances. Hence equation (3.9b) will in general not be defined on the level of distributions. For example, if $c(t, x) = 1 + H(x)$, then $(\partial_x c) w_3$ would represent a product of the delta distribution with an L^2-function.

3.2 Existence and uniqueness for the Cauchy problem

Rewriting the wave equation (3.1) as a first-order system (3.2) via Theorem 3.5 allows to apply existence and uniqueness theorems for the latter to prove existence and uniqueness of a solution of the initial value problem for the wave equation. More precisely, let a vector w of generalised functions with representative $(w_\varepsilon)_\varepsilon = ((u_\varepsilon, z_\varepsilon, v_\varepsilon)^\mathsf{T})_\varepsilon$ be given that is the unique solution of the Cauchy problem for the system (3.2–3.3) with initial data $w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\mathsf{T}$, then Theorem 3.5 implies that $[(u_\varepsilon)_\varepsilon]$ will be the unique generalised solution of the Cauchy problem for the wave equation (3.1) with initial data $(u, \partial_t u)|_{t=0} = (u^0, u^1)$. Hence, each of the existence and uniqueness results for hyperbolic first-order systems established in Theorems 2.7, 2.8, and 2.13 gives rise to a corresponding existence and uniqueness result for second-order wave equations. The following theorem gathers these results, formulated for second-order PDEs of the type (3.1). Although we only considered wave equations with real-valued coefficients in Section 3.1, the transformation also works for complex-valued coefficients and a Hermitian principal part (meaning that its spatial part R is a Hermitian positive definite matrix).

Theorem 3.7 (Existence and uniqueness of generalised solutions)**.** *Consider the wave equation*

$$-\partial_t^2 u + 2\sum_{i=1}^{n} h_i \partial_{x_i}\partial_t u + \sum_{i,j=1}^{n} R_{ij}\partial_{x_i}\partial_{x_j} u + a\partial_t u + \sum_{i=1}^{n} b_i \partial_{x_i} u + cu = f$$

on $\Omega_T =]0,T[\,\times\mathbb{R}^n$ *with the initial condition*

$$(u, \partial_t u)|_{t=0} = (u^0, u^1),$$

where a, c, and the entries of R, h, and b belong to $\mathcal{G}_{L^\infty}(\Omega_T)$ and R is symmetric and positive definite. Furthermore, let $S = R^{\frac{1}{2}}$ be the positive definite square root of R, obtained via diagonalisation. Then we have the following three results.

(I) *The Cauchy problem with initial data $u^0, u^1 \in \mathcal{G}(\mathbb{R}^n)$ and right-hand side $f \in \mathcal{G}(\overline{\Omega}_T)$ has a unique solution $u \in \mathcal{G}(\Omega_T)$ if*

  (i) *the coefficients a, c, and the spatial derivatives $\partial_{x_j} h$ are locally of log-type.*

  (ii) *the entries of S, $\partial_{x_j} S$, $(\partial_t S)S^{-1}$, and $(b - \operatorname{div} S^2)^\mathsf{T} S^{-1}$ are locally of log-type.*

  (iii) *there exists $d_{S,h} > 0$ such that $\sup_{(t,x)} |h_\varepsilon(t,x)| = O(1)$ and $\sup_{(t,x)} \|S_\varepsilon(t,x)\|_{\mathrm{op}} = O(1)$ on $]0,T[\,\times\{x \in \mathbb{R}^n : |x| > d_{S,h}\}$ as $\varepsilon \to 0$.*

(II) *The Cauchy problem with initial data $u^0, u^1 \in \mathcal{G}_{L^2}(\mathbb{R}^n)$ and right-hand side $f \in \mathcal{G}_{L^2}(\Omega_T)$ has a unique solution $u \in \mathcal{G}_{L^2}(\Omega_T)$ if*

  (i) *the coefficients a, c, and the spatial derivatives $\partial_{x_j} h$ are of $L^{1,\infty}$-log-type.*

  (ii) *the entries of S, $\partial_{x_j} S$, $(\partial_t S)S^{-1}$, and $(b - \operatorname{div} S^2)^\mathsf{T} S^{-1}$ are of $L^{1,\infty}$-log-type.*

(III) *Let initial data* $u^0, u^1 \in \mathcal{G}_{L^\infty}(\mathbb{R}^n)$ *and right-hand side* $f \in \mathcal{G}_{L^\infty}(\Omega_T)$ *be given. Assume that*

(i) *the coefficients* a*,* c*, and the spatial derivatives* $\partial_{x_j} h$ *are of* L^∞*-log-type.*

(ii) *the entries of* S*,* $\partial_{x_j} S$*,* $(\partial_t S)S^{-1}$*, and* $(b - \operatorname{div} S^2)^\intercal S^{-1}$ *are of* L^∞*-log-type.*

Then there exists a unique solution $u \in \mathcal{G}(\Omega_T)$ *of the wave equation* (3.7) *such that* $(u, \partial_t u)|_{t=0} - (u^0, u^1) \in \mathcal{N}(\mathbb{R}^n) \times \mathcal{N}(\mathbb{R}^n)$*.*

Proof. We rewrite the wave equation (3.7) as the corresponding symmetric hyperbolic system (3.2) with A^j, B and F as in (3.3). Then, considering case (I) first, we observe that all entries of A^j and B belong to $\mathcal{G}_{L^\infty}(\mathbb{R}^{1+n})$. The vector $w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\intercal \in \mathcal{G}(\mathbb{R}^n)^{2+n}$ is admissible as initial data for the system and the right-hand side $F = (0, f, 0)$ belongs to $\mathcal{G}(\Omega_T)^{2+n}$. Hence assumption (1) in Theorem 2.7 is satisfied. Conditions (i) and (ii) ensure that the matrix coefficients A^j and B of the system are locally of L^∞-log-type and thus fulfil assumption (2) in Theorem 2.7. Finally, condition (iii) implies that $\|A^j_\varepsilon(t,x)\|_{\mathrm{op}} = O(1)$ as $\varepsilon \to 0$ uniformly on $]0,T[\, \times \{x \in \mathbb{R}^n : |x| > d_{S,h}\}$ for any representative $(A^j_\varepsilon)_\varepsilon$ of A^j (which is exactly assumption (3) in Theorem 2.7).

Summing up, all conditions in Theorem 2.7 are satisfied and we therefore obtain a unique solution $w \in \mathcal{G}(\Omega_T)^{2+n}$ of the initial value problem for the symmetric hyperbolic system (3.2). Theorem 3.5 then guarantees that the first component $u \in \mathcal{G}(\Omega_T)$ of the vector w is the unique solution of the Cauchy problem for the wave equation.

In case (II), the coefficient matrices $\partial_{x_i} A^j$ and B are of $L^{1,\infty}$-log-type. Moreover, the initial data $w|_{t=0} = (u^0, u^1, S|_{t=0}\nabla u^0)^\intercal$ belong to $\mathcal{G}_{L^2}(\mathbb{R}^n)^{2+n}$, since S has entries in $\mathcal{G}_{L^\infty}(\Omega_T)$. Applying Theorem 2.13, we thus obtain a unique solution $w \in \mathcal{G}_{L^2}(\Omega_T)^{2+n}$ of the initial value problem (3.2,3.5), generating a unique solution $u \in \mathcal{G}_{L^2}(\Omega_T)$ of the corresponding Cauchy problem (3.1,3.4) for the wave equation. The proof of case (III) follows the same pattern and employs Theorem 2.8. □

In Theorem 3.7, all coefficients have to obey certain log-type estimates, either locally as in variant (I) or globally as in variants (II) and (III). Those are the 'minimal assumptions' necessary to ensure that the related first-order system satisfies the corresponding conditions in Theorems 2.7, 2.8, and 2.13. As a result of the underlying transformation algorithm, the conditions labeled (ii) in Theorem 3.7, involving the matrix S (or $R = S^2$ respectively) are not very practical. In the following corollary we formulate more transparent, but stronger conditions on the principal coefficients of the wave equation, similar to the conditions used in [SV09, Definition 4.5].

Corollary 3.8. *In Theorem 3.7* (I)*, replace assumption* (ii) *by the alternative condition*

(ii') $\mathrm{Tr}(R)$ *and* $(\det R)^{-1}$ *are locally uniformly bounded and* b *as well as the derivatives* $\partial_j R$ *(that is,* $\partial_t R$ *and* $\partial_{x_j} R$*) are locally of log-type.*

Then the Cauchy problem with initial data $u^0, u^1 \in \mathcal{G}(\mathbb{R}^n)$ and right-hand side $f \in \mathcal{G}(\overline{\Omega}_T)$ has a unique solution $u \in \mathcal{G}(\Omega_T)$.

Proof. Noting that $\det R$ is invertible since R is positive definite (cf. [HKK14, Proposition 3.20]), condition (ii') implies that the eigenvalues of R are locally uniformly bounded away from zero and bounded from above. Hence R and R^{-1} are locally uniformly bounded in the operator norm and the same is true for the square root S and its inverse S^{-1}, since $\|S\| \leq \|R\|^{1/2}$ and $\|S^{-1}\| \leq \|R^{-1}\|^{1/2}$ by construction. Moreover, the eigenvalues of S are also locally uniformly bounded away from zero. Hence, for each $K \subset\subset \Omega_T$, there exists a constant C_K such that $\|\partial_j S_\varepsilon\|_{L^\infty(K)} \leq C_K \|\partial_j R_\varepsilon\|_{L^\infty(K)}$, where $(R_\varepsilon)_\varepsilon$ is any positive definite representative of R and $(S_\varepsilon)_\varepsilon$ is its square root. Therefore $\partial_j S$ is locally of log-type as well. Summing up, we find that S, $\partial_{x_j} S$, $(\partial_t S)S^{-1}$, and $(b - \operatorname{div} S^2)^\mathsf{T} S^{-1}$ are locally of log-type and hence condition (ii) in Theorem 3.7 (I) is indeed satisfied. $\square$

In cases (II) and (III) of Theorem 3.7, condition (ii) may also be replaced by a simpler but more restrictive version. Instead of local bounds, one has to use global bounds in the respective norms, but otherwise the statements are as in Corollary 3.8 and the proofs follow the same pattern.

Theorem 3.7 demonstrates that transforming results for first-order systems of PDEs into results for higher-order PDEs is straightforward and comparatively simple in a generalised functions setting. Example 3.6 already indicates this is not the case in more classical, weak or distributional solution concepts. The lack of a general equivalence statement of the type of Theorem 3.5 in such frameworks might be part of an explanation for the fact that the existence and uniqueness theory for higher-order PDEs has to a large extent been developed in parallel to the theory for first-order systems, in particular in the context of non-smooth coefficients or symbols. However, in differential algebras of Colombeau generalised functions, transferring results from one world to the other almost reduces to mere algebraic manipulations and to keeping track of their effects on asymptotic estimates for the involved objects.

The asymptotic estimates on the coefficients of the wave equation required in Theorem 3.7 are less restrictive than those supposed in (local) existence results for the initial value problem for the wave equation on 'weakly singular' Lorentzian manifolds, cf. [GMS09, Theorem 3.1] and [Han11, Theorem 3.1].

A Lorentzian metric g on a smooth manifold $\mathcal{M}$ belongs to the Geroch-Traschen class if g and its inverse g^{-1} belong to $H^1_{\text{loc}}(\mathcal{M}) \cap L^\infty_{\text{loc}}(\mathcal{M})$ and, furthermore, if g is non-degenerate in the sense that $|\det g| \geq C > 0$ almost everywhere on compact sets (cf. [SV09, Section 3]). This class of metrics was introduced in [GT87] and is viewed as the 'largest reasonable' class of metrics, for which the Riemann curvature tensor can be defined as a distribution (see also [LM07]).

The conditions of Theorem 3.7 (I) are sufficiently general to cover the Laplace-Beltrami operator $\Box_g$, given by

$$\Box_g u = \sum_{i,j=0}^{n} |\det g|^{-\frac{1}{2}} \partial_i \big(|\det g|^{\frac{1}{2}} g^{ij} \partial_j u \big), \tag{3.10}$$

of a metric g in the Geroch-Traschen class, which was not possible previously. To demonstrate how Theorem 3.7 can be employed to obtain a local existence result for wave equations derived from a Geroch-Traschen metric, we consider a local representative of the metric in a coordinate chart. We cut off this representative outside some ball in $\mathbb{R}^{1+n}$ and extend it to a Lorentzian metric g_0 on $\mathbb{R}^{1+n}$ by setting it constant outside this ball. Then the components of g_0 will belong to $L^\infty(\mathbb{R}^{1+n})$. Regularising g_0 via componentwise convolution with a model delta net, we obtain $g = [(g_\varepsilon)_\varepsilon] \in M_{1+n}(\mathcal{G}_{L^\infty}(\mathbb{R}^{1+n}))$. A detailed discussion of regularisations of Geroch-Traschen metrics can be found in [SV09], including convergence results and more sophisticated ways of smoothing a metric. In order to end up with a generalised metric, as introduced in [KS02b], the metric g_0 has to satisfy a certain stability condition related to its eigenvalues and only a special class of mollifiers are admissible (see [SV09, Proposition 4.6]). Employing a mollifier with a general scaling $\sigma_\varepsilon \to 0$ as in Lemma 2.20, the components of g will belong to $\mathcal{G}_{L^\infty}(\mathbb{R}^{1+n})$ and for any representative $(g_\varepsilon)_\varepsilon$ of g, we have $\|g_\varepsilon\|_{L^\infty(\mathbb{R}^{1+n})} = O(1)$ as well as $\|\partial_j g_\varepsilon\|_{L^\infty(\mathbb{R}^{1+n})} = O(\sigma_\varepsilon^{-1})$. However, unless g_0 is Lipschitz continuous, we will in general not get $\|\partial_j g_\varepsilon\|_{L^\infty(\mathbb{R}^{1+n})} = O(1)$ as $\varepsilon \to 0$ and therefore condition (B) of Theorem 3.1 in [GMS09] is violated, as well as condition (i) of Theorem 3.1 in [Han11]. On the other hand, logarithmically rescaling the mollifier by taking $\sigma_\varepsilon = (\log\frac{1}{\varepsilon})^{-1}$, we may still apply Theorem 3.7 (I), as the following corollary exemplifies.

Corollary 3.9 (Wave equation derived from the Laplace-Beltrami operator of a Lorentzian metric)**.** *Let $g = \begin{pmatrix} -1 & h^\mathsf{T} \\ h & R \end{pmatrix}$ be a generalised Lorentzian metric on $\mathbb{R}^{1+n}$, obtained as the smoothing of a non-degenerate and stable Geroch-Traschen class metric g_0 as described above, using a logarithmically rescaled mollifier. Then the Cauchy problem*

$$\begin{aligned} \Box_g u &= 0 \quad \text{on} \quad \Omega_T =]0,T[\times\mathbb{R}^n \\ (u, \partial_t u)|_{t=0} &= (u^0, u^1) \in \mathcal{G}(\mathbb{R}^n) \times \mathcal{G}(\mathbb{R}^n) \end{aligned}$$

has a unique solution $u \in \mathcal{G}(\Omega_T)$ for arbitrary $T > 0$.

Proof. In local coordinates, the equation $\Box_g u = 0$ takes on the form of equation (3.1) with $c = 0$ and $f = 0$. The components of the generalised metric g arise as special convolution regularisations ensuring that $\det(g_\varepsilon) \to \det(g_0)$ as $\varepsilon \to 0$ in $H^1_{\text{loc}}(\mathbb{R}^{1+n}) \cap L^p_{\text{loc}}(\mathbb{R}^{1+n})$ for all $p < \infty$ and $|\det(g_\varepsilon)| \geq C > 0$ uniformly in ε for any representative of $(g_\varepsilon)_\varepsilon$ of g (see [SV09, Proposition 4.2 and Proposition 4.6]). Moreover, $g(t,x)$ is constant for large $|x|$ as a result of the cut-off procedure. It is then easy to see that all coefficients in equation (3.1) belong to $\mathcal{G}_{L^\infty}(\Omega_T)$. We construct $S = R^{\frac{1}{2}}$ via diagonalisation and check that the conditions of Theorem 3.7 (I) hold (note that $\|S\| \leq \|R\|^{\frac{1}{2}}$). Again thanks to the cut-off, condition (ii) is trivially satisfied, since $h(t,x)$ and $S(t,x)$ are constant for $|x|$ sufficiently large. As for condition (i), S and its inverse S^{-1} are uniformly bounded and the logarithmic rescaling of the mollifier implies that a, all derivatives $\partial_j S$, and the spatial derivatives $\partial_{x_j} h$ are (locally) of L^∞-log-type. Hence $(\partial_t S)S^{-1}$ and $(b - \operatorname{div} S^2)^\mathsf{T} S^{-1}$ are locally of L^∞-log-type as well. □

In Theorem 3.7, the necessary regularity assumptions are naturally weaker for the lower-order coefficients than for the principal part of the PDO. Complementary to Corollary 3.9, the following example discusses an equation with constant principal coefficients, but highly singular lower-order coefficients.

Example 3.10 (Klein-Gordon equation with an impulsive wave potential)**.** The Klein-Gordon equation describing a spinless relativistic particle[2] in an electromagnetic field can be written as

$$-(\partial_t - iA_0)^2\phi + \sum_{j=1}^{3}(\partial_{x_j} - iA_j)^2\phi - m^2\phi = 0, \tag{3.11}$$

where $A = (A_0, A_1, A_2, A_3)^\mathsf{T}$ represents the vector potential describing the electromagnetic field and m is the mass of the particle (cf. [BD64, Chapter 9]). Rewriting equation (3.11) in the form of (3.1), we identify the coefficients therein as $R = I_3$, $h = 0$ (thus the principal part of (3.11) is the d'Alembert operator) and

$$a = 2iA_0, \qquad b_j = 2iA_j, \qquad c = i\partial_t A_0 + A_0^2 - \sum_{j=1}^{3}(i\partial_{x_j}A_j + A_j^2),$$

which could of course be written more elegantly in relativistic notation, but we want to keep the notation and the symbols used in (3.1). As in Example 2.16, we may now consider an impulsive electromagnetic wave travelling in x_1-direction, modelled by a vector potential of the form $A_0(t,x) = A_1(t,x) = \delta(t - x_1)f(x_2, x_3)$, where f satisfies the two-dimensional Laplace equation, and $A_2 = A_3 = 0$. The lower-order coefficients are then formally given by

$$a(t,x) = b_1(t,x) = \delta(t-x_1)f(x_2,x_3), \quad b_2 = b_3 = 0, \quad c(t,x) = 2i\delta'(t-x_1)f(x_2,x_3),$$

hence we encounter δ and even δ' in the coefficients of equation (3.11). Distributional solution concepts will fail in this situation, but in the differential algebra $\mathcal{G}$, the corresponding initial value problem is well-defined and we obtain a unique generalised solution, if the regularisations are chosen appropriately. More precisely, condition (ii) in Theorem 3.7 (II) is trivially satisfied, since $S = I_3$ and $h = 0$. Using a delta net $(\rho_\varepsilon)_\varepsilon$ with $\rho_\varepsilon(x) = \sigma_\varepsilon^{-1}\rho(\sigma_\varepsilon^{-1}x)$ and $\sigma_\varepsilon = (\log(1/\varepsilon))^{-\frac{1}{2}}$ as a model for the delta distribution, all coefficients are locally of L^∞-log-type and thus condition (i) is also satisfied. For the generalised coefficients a, b_1, and c to belong to $\mathcal{G}_{L^\infty}(\Omega_T)$, an additional cut-off is necessary, if $f : \mathbb{R}^2 \to \mathbb{R}$ is not bounded. Imposing initial data $(\phi, \partial_t\phi)|_{t=0} = (\phi^0, \phi^1) \in \mathcal{G}_{L^2}(\mathbb{R}^3) \times \mathcal{G}_{L^2}(\mathbb{R}^3)$, Theorem 3.7 (II) provides a unique solution $\phi \in \mathcal{G}_{L^2}(\Omega_T)$ of the Cauchy problem for the Klein-Gordon equation (3.11) with an impulsive electromagnetic wave potential.

[2]The elementary building blocks of matter are 'fermions' or spin-$\frac{1}{2}$ particles and therefore not correctly described by the Klein-Gordon equation. However, composite particles can have zero spin, for example pions, each consisting of a quark and an antiquark. The Higgs boson, apparently discovered in 2012, would be the first elementary particle described by the Klein-Gordon equation.

Remark 3.11 (Distributional shadows)**.** All results from Chapter 2 on distributional limits of a generalised solution w of a symmetric hyperbolic system can be transformed into corresponding 'extrinsic' regularity results for a generalised solution u of a second-order wave equation. To this end we only have to rephrase the required regularity conditions on the coefficients of the system as conditions for the coefficients of the related second-order wave equation. The regularity of the distributional limit u_0 of the unique generalised solution of the initial value problem for the wave equation is then determined by the regularity of the distributional limit w_0 of the corresponding first-order system via the relation $w = (u, \partial_t u, S\nabla u)^\mathsf{T}$. In particular, $\partial_t u$ and $S\nabla u$ will enjoy the same extrinsic regularity as w. Thus, roughly speaking, the regularity of u_0 will in general be better than that of w_0 by one Sobolev index (recall that the invertibility of S is essential in the transformation algorithm to a first-order system and thus we have $\nabla u = S^{-1}(S\nabla u)$).

3.3 A note on the Hadamard parametrix construction

The general wave equation (3.1) can also be written as

$$\Box_g u + Pu = f,$$

where $g = [(g_\varepsilon)_\varepsilon]$ represents a generalised Lorentzian metric and P is a PDO of first order. If we consider a representative $(g_\varepsilon)_\varepsilon$ of g for fixed ε, we get a PDE with smooth coefficients,

$$\Box_{g_\varepsilon} u + P_\varepsilon u = f_\varepsilon.$$

In fact, for the existence and uniqueness results in Chapter 2, we constructed generalised solutions ε-wise, as nets of smooth solutions of 'smooth initial value problems', with smooth initial data and smooth right-hand side. Subsequently, we investigated regularity properties of generalised solutions mainly in terms of their distributional shadows, that is, by taking any of its representatives $(u_\varepsilon)_\varepsilon$ and considering the limit $\varepsilon \to 0$, in an appropriate sense. However, regularity information about the generalised solution may already be extracted in the process of its construction, in particular with regard to a more detailed analysis of its singularities, in contrast to the essentially global regularity results obtained in Section 2.3.

The singularity structure of the solution of a PDE with smooth coefficients and non-smooth initial data or right-hand side is encoded in its fundamental solution or Green's function. With the help of the topological dual of a Colombeau algebra [Gar05a, Gar05b], it is possible to consider fundamental solutions also in a generalised functions setting [Gar08]. A parametrix represents an 'approximate fundamental solution'. More precisely, a distribution $K \in \mathcal{D}'(\Omega \times \Omega)$ is called a parametrix of the PDO $Q(x, \partial_x)$, where $x \in \Omega \subseteq \mathbb{R}^n$ open, if

$$Q(x, \partial_x) K(x, y) = \delta(x - y) + h(x, y),$$

where $\delta(x - y)$ is the Dirac measure on the diagonal of $\Omega \times \Omega$ and h is a smooth function on $\Omega \times \Omega$ (cf. [Trè67, Section 52]). The Hadamard construction [Hör07, Chapter XVII, Section 17.4] yields an explicit parametrix for elliptic PDOs, giving precise information about the singularities of the fundamental solution. An ε-wise application of Hadamard's method to construct a generalised Hadamard parametrix could provide some new aspects in the regularity theory of generalised solutions, especially concerning the interplay between coefficients, right-hand side, and initial data.

A second-order PDO with smooth coefficients is called elliptic, if its principal symbol is a positive definite quadratic form. If g is a smooth Riemannian metric, the corresponding Laplace-Beltrami operator $\Box_g$ is elliptic (see (3.10)). In the context of wave propagation, elliptic operators typically arise, in a certain sense, as the 'spatial parts' of hyperbolic PDOs. For example, if the principal coefficients of the wave equation (3.1) do not depend on time, then its principal part for fixed ε can be written as $-\partial_t^2 u + Q_\varepsilon(x, \partial_x)u$, where the purely spatial PDO $Q_\varepsilon(x, \partial_x)$ is elliptic. Specifically, the free Klein-Gordon equation, whose principal part is derived from the Minkowski metric[3], is given by

$$- \partial_t^2 \phi + \Delta\phi - m^2\phi = 0 \tag{3.12}$$

with the Laplace operator representing the 'spatial part'.

Hadamard's method uses the fundamental solution of a constant coefficient operator Q_0, obtained by evaluating the coefficients of the elliptic operator Q at a certain point. Based on the constant coefficient case, a parametrix of the operator Q is constructed with the help of the distance function induced by the Riemannian metric associated with the principal symbol of Q. The principal part of the 'frozen coefficients' operator Q_0 is essentially the Laplace operator (up to a rescaling of the coordinates). However, for technical reasons, one considers the operator $-\Delta - z$ where $z \in \mathbb{C} \setminus \overline{\mathbb{R}}_+$ is a fixed complex number, and introduces 'associated Bessel potentials'

$$f_\nu(x) := \mathcal{F}^{-1}_{\xi \to x}\left(\frac{\nu!}{(|\xi|^2 - z)^{\nu+1}}\right) \qquad (\nu \in \mathbb{N}_0), \tag{3.13}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform on the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$. The essential properties of f_ν are

(i) $(-\triangle - z)f_\nu = \nu f_{\nu-1}$ for $\nu \geq 1$ and $(-\triangle - z)f_0 = \delta$,

(ii) $-2\partial_{x_j} f_\nu(x) = x_j f_{\nu-1}(x)$,

which both follow directly from definition of f_ν and basic properties of the Fourier transform. The associated Bessel potentials f_ν are basically fundamental solutions of powers of the operator $-\Delta - z$ and therefore smooth except at the origin (see [Hör03, Theorem 7.1.22]). Moreover, they are also locally integrable at the origin. This we show in the following lemma, using regularisation techniques, basic properties of the Fourier transform, and arguments from distribution theory.

Lemma 3.12 (Integrability of associated Bessel potentials)**.** *Let $\alpha \in \mathbb{N}_0^n$ and $|\alpha| \leq \nu + 1$. Then we have $\partial^\alpha f_\nu \in L^1(\mathbb{R}^n)$.*

[3] In Minkowski space-time, the Laplace-Beltrami operator becomes the d'Alembert operator $\Box = -\partial_t^2 + \Delta$, where $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2$ is the Laplacian.

Proof. We introduce auxiliary distributions $F_\varepsilon \in \mathcal{S}'(\mathbb{R}^n)$, $\varepsilon \in [0,1]$,

$$F_\varepsilon(x) = \mathcal{F}^{-1}_{\xi\to x}\left(\frac{e^{-\varepsilon(|\xi|^2-z)}}{|\xi|^2 - z}\right) \tag{3.14}$$

where $\widehat{F}_\varepsilon(\xi) := \mathcal{F}_{x\to\xi}(F_\varepsilon(x)) = e^{-\varepsilon(|\xi|^2-z)}\widehat{f}_0(\xi)$. Note that $F_0(x) = f_0(x)$ and $F_\varepsilon \in \mathcal{S}(\mathbb{R}^n)$ for $\varepsilon \in]0,1]$. We consider the map $[0,1] \to \mathcal{S}'(\mathbb{R}^n)$, $\varepsilon \mapsto \widehat{F}_\varepsilon$, and differentiate with respect to ε, yielding $\frac{d}{d\varepsilon}\widehat{F}_\varepsilon(\xi) = -e^{-\varepsilon(|\xi|^2-z)}$ and

$$\frac{d}{d\varepsilon}F_\varepsilon(x) = \mathcal{F}^{-1}\left(\frac{d}{d\varepsilon}\widehat{F}_\varepsilon\right) = -(2\pi)^{-\frac{n}{2}}e^{\varepsilon z}(2\varepsilon)^{-\frac{n}{2}}e^{-\frac{1}{4\varepsilon}|x|^2}.$$

Writing $f_0(x) = F_0(x) = F_1(x) - (F_1(x) - F_0(x))$ and applying the fundamental theorem of calculus, we get an integral representation for $f_0(x)$ for all $x \neq 0$ (recall that f_0 is smooth off the origin),

$$f_0(x) = F_1(x) + (2\pi)^{-\frac{n}{2}}\int_0^1 e^{\varepsilon z}(2\varepsilon)^{-\frac{n}{2}}e^{-\frac{1}{4\varepsilon}|x|^2}\mathrm{d}\varepsilon \qquad (x \neq 0). \tag{3.15}$$

For $\alpha \in \mathbb{N}_0^n$ and $|\alpha| \leq 1$ (that is, either $\alpha = 0$ or $\alpha = e_j$, where $\{e_j\}$ are the standard basis vectors), we define h^α as

$$h^0(x) := \begin{cases} f_0(x) & x \neq 0 \\ 0 & x = 0 \end{cases} \qquad h^{e_j}(x) := \begin{cases} \partial_{x_j} f_0(x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

implying that h^α coincides with the distribution $\partial^\alpha f_0$ ($|\alpha| \leq 1$) on $\mathbb{R}^n\backslash\{0\}$. Hence the difference between h^α and $\partial^\alpha f_0$ must be a linear combination of derivatives of the delta distribution, that is

$$h^\alpha - \partial^\alpha f_0 = \sum_{|\gamma|\leq N_\alpha} c^\alpha_\gamma\, \partial^\gamma \delta \tag{3.16}$$

for some constants $N_\alpha \in \mathbb{N}_0$ and $c^\alpha_\gamma \in \mathbb{N}_0^n$ (see [Bha12, Theorem 5.7.2]). Using the integral representation (3.15) of f_0 for $x \neq 0$, it is easy to see that $h^\alpha \in L^1(\mathbb{R}^n)$, where $|\alpha| \leq 1$. Fourier transformation of equation (3.16) leads to $\widehat{h^\alpha}(\xi) - \widehat{\partial^\alpha f_0}(\xi) = \sum_{|\gamma|\leq N_\alpha} c^\alpha_\gamma \xi^\gamma$. From the definition of f_0 in (3.13) it is clear that $\widehat{\partial^\alpha f_0}(\xi) \to 0$ as $|\xi| \to \infty$. But the same is true for $\widehat{h^\alpha}$ by the Riemann-Lebesque lemma (see [Bha12, p. 401]). These observations immediately imply that all coefficients c^α_γ must be zero and therefore $\partial^\alpha f_0 \in L^1(\mathbb{R}^n)$ for $|\alpha| \leq 1$.

Employing the convolution theorem for L^1-functions, we may write

$$f_\nu = \nu!\, \underbrace{f_0 * f_0 * \ldots * f_0}_{\nu+1 \text{ times}},$$

which can also be described in terms of distributional convolutions $\mathcal{E}'(\mathbb{R}^n)\times\mathcal{D}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ by using appropriate cut-off sequences. Distributional derivatives of convolution products can thus be carried out as usual. Since $\partial^\alpha f_0 \in L^1(\mathbb{R}^n)$ for $|\alpha| \leq 1$, we conclude that $\partial^\alpha f_\nu \in L^1(\mathbb{R}^n)$ for $|\alpha| \leq \nu + 1$, in particular these functions are locally integrable, also at the origin, for $|\alpha| \leq 2$, except in the case $\nu = 0$ and $|\alpha| = 2$ (where the derivatives cannot be distributed among convolution factors). □

The functions f_ν defined as

$$f_\nu(x) := \mathcal{F}^{-1}_{\xi \to x}\left(\frac{\nu!}{(|\xi|^2 - z)^{\nu+1}}\right) \qquad (\nu \in \mathbb{N}_0)$$

are essential ingredients of the Hadamard construction. By their definition in (3.13), it is obvious that they are rotationally symmetric, that is, each $f_\nu(x)$ only depends on $|x|$. Thus we may write $f_\nu(x) = \widetilde{f}_\nu(|x|)$. If d_g denotes the Riemannian distance function defined by the principal symbol $\sum_{i,j=1}^n g^{ij}(x)\xi_i\xi_j$ of an elliptic PDO Q (cf. [Hör07, Appendix C.5]), the Hadamard construction yields recursively defined smooth functions U_ν (defined on a neighbourhood of the diagonal in $\Omega \times \Omega$) and a smooth cut-off function χ such that $K_N \in \mathcal{D}'(\Omega \times \Omega)$,

$$K_N(x, y) := \chi(x, y) \sum_{\nu=0}^{N} U_\nu(x, y)\widetilde{f}_\nu(d_g(x, y)) \tag{3.17}$$

satisfies

$$(Q(x, \partial_x) - z)K_N(x, y) = (\det g(y))^{\frac{1}{2}}\delta(x - y) + R(x, y)$$

where $R \in \mathcal{C}^{2N+1-n}(\Omega \times \Omega)$ and can thus be made as smooth as we wish by taking N large enough, that is, by including more and more terms in the expansion (3.17), see [Hör07, Chapter XVII, Section 17.4]. The fact that the distributions $\partial^\alpha f_\nu$ are integrable for $|\alpha| \leq 2$ (except for $\nu = 0$ and $|\alpha| = 2$) is important, in particular for analysing the singularity structure of the Hadamard parametrix K_N.

Chapter 4

Colombeau mollifiers and the Stieltjes moment problem

Distributions are embedded in the special Colombeau algebra by convolution with a mollifier function. For the embedding to be consistent with the constant embedding of smooth functions, recall that the mollifier φ must have all moments of order greater than zero vanishing (cf. [GKOS01, p. 9]), that is

$$\int_{\mathbb{R}^n} \varphi(x)\mathrm{d}x = 1, \qquad \mu_\alpha(\varphi) := \int_{\mathbb{R}^n} x^\alpha \varphi(x)\mathrm{d}x = 0 \quad \text{for all } \alpha \in \mathbb{N}_0^n,\ |\alpha| > 0.$$

A smooth function with the desired properties can be obtained by taking some Schwartz function $\psi \in \mathcal{S}(\mathbb{R}^n)$ with $\psi(0) = 1$ and $\partial^\alpha \psi(0) = 0$ for all $\alpha \in \mathbb{N}_0^n$, $|\alpha| > 0$. Then its Fourier transform $\widehat{\psi} \in \mathcal{S}(\mathbb{R}^n)$ has unit integral (that is, the zeroth moment or mass of $\widehat{\psi}$ is equal to 1) and vanishing moments of nonzero order ($\int_{\mathbb{R}^n} \xi^\alpha \widehat{\psi}(\xi)\mathrm{d}\xi = 0$ for $|\alpha| > 0$). A smooth function satisfying this 'moment condition' cannot belong to $\mathcal{D}(\mathbb{R}^n)$. Indeed, the Fourier transform of a compactly supported smooth function must be an entire function by the Payley-Wiener-Schwartz theorem (cf. [Bha12, Theorem 8.4.1]). If $\varphi \in \mathcal{D}(\mathbb{R}^n)$ has all moments of order $|\alpha| > 0$ vanishing, then $\partial^\alpha \widehat{\varphi}(0) = 0$ for all $|\alpha| > 0$. But since $\widehat{\varphi}$ is an entire function, it must then be constant, which is a contradiction to $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$, if $\varphi \neq 0$. Hence a mollifier cannot have vanishing moments and compact support at the same time.

Now the question arises, how much of $\mathbb{R}^n$ must actually be occupied by the support of a mollifier in order to satisfy the moment condition. Considering the one-dimensional case first, we may therefore ask whether there exists a Schwartz function whose support is bounded from above or from below and which has all moments of nonzero order vanishing. If we assume its support to be contained in the interval $[0, \infty[$, then this question is in fact closely related to the Stieltjes moment problem, posed and solved in its original form by T.J. Stieltjes in 1894 [Sti94]. There one seeks necessary and sufficient conditions on a sequence of real numbers $(\mu_n)_{n=0}^\infty$ to be of the form $\mu_n = \int_0^\infty x^n \mathrm{d}\phi(x)$ for some measure ϕ. Compared to the Hamburger moment problem (cf. [RS75, Chapter X, Example 3]), the domain of integration is only the positive real axis instead of $\mathbb{R}$. In fact, the Hamburger moment problem was posed in [Ham20] as an extension of the Stieltjes moment problem. Over the last century, several variants of the Stieltjes moment problem have been formulated and solved. One of these variants is a direct generalisation of the Colombeau mollifier problem described above: Given an arbitrary sequence of real numbers $(\mu_n)_{n\in\mathbb{N}_0}$, does there exist a function $\varphi \in \mathcal{S}(\mathbb{R})$ with $\mathrm{supp}(\varphi) \subseteq [0,\infty[$ such that its moments coincide with this sequence? Indeed, a function with the desired properties exists, as was shown by A.J. Duran [Dur89], using approximation aspects of the Laguerre polynomials, the Laplace transform and the Hankel transform of order zero. Other proofs were given in [DE94] as well as in [LS08], relying on Fourier analysis, asymptotic power series expansions, and existence results such as the Borel-Ritt theorem (cf. [Was87, Section 9]). Generalisations of the Stieltjes moment problem for rapidly decreasing smooth functions of several variables have been studied in [Est98, GLS01].

In Section 4.1, we show that any function $\mathcal{D}(\mathbb{R})$ can be approximated in $L^p(\mathbb{R})$ for $p \in]\frac{1}{2},\infty]$ by a Schwartz function with prescribed moments and support bounded from below. The presented construction is a modified version of a similar algorithm used in [WB03, Proposition A.1]. We then adapt and generalise the construction to the multidimensional case in Section 4.2, thereby extending the approximation property to $W^{k,p}(\mathbb{R}^n)$ ($k \in \mathbb{N}_0$, $p \in [1,\infty]$). However, as some aspects of the one-dimensional scheme cannot be transferred to $\mathbb{R}^n$, for example the approximation in $L^p(\mathbb{R})$ for $p < 1$, we provide a full proof of the one-dimensional case first. Our results yield an elementary and completely constructive proof of the Stieltjes moment problem for rapidly decreasing smooth functions. They also provide an explicit construction of Colombeau mollifiers in $\mathbb{R}^n$, supported in an arbitrarily small cone and satisfying additional constraints such as approximating a given test function with compact support. Moreover, we obtain generalised mollifiers of the types considered in [OV08, Propositions 4.1 and 4.2] and in [SV09, Lemma 4.3].

4.1 An explicit one-dimensional mollifier construction

The basic idea underlying the mollifier construction presented in this section is the following observation: Consider a test function $\varphi \in \mathcal{D}(\mathbb{R})$ with $\int_{-\infty}^{\infty} \varphi(x)\mathrm{d}x \neq 0$ and set

$$f(x) := \varphi(x) + a\varphi'\left(\frac{x}{b}\right)$$

for some $a \in \mathbb{R}$ and $b > 0$. Then $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_{-\infty}^{\infty} \varphi(x)\mathrm{d}x$, since φ has compact support and thus $\int_{-\infty}^{\infty} \varphi'(x)\mathrm{d}x = 0$. To calculate the first moment of f, we use integration by parts and a transformation of variables to obtain

$$\int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_{-\infty}^{\infty} x\varphi(x)\mathrm{d}x - ab^2 \int_{-\infty}^{\infty} \varphi(x)\mathrm{d}x.$$

It is now easy to see that for all $\varepsilon > 0$ and all $\nu_1 \in \mathbb{R}$, there exist real numbers $a > 0$ and b such that

(i) the first moment of f is equal to ν_1, that is $\int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \nu_1$,

(ii) $\|f - \varphi\|_{L^\infty(\mathbb{R})} < \varepsilon$.

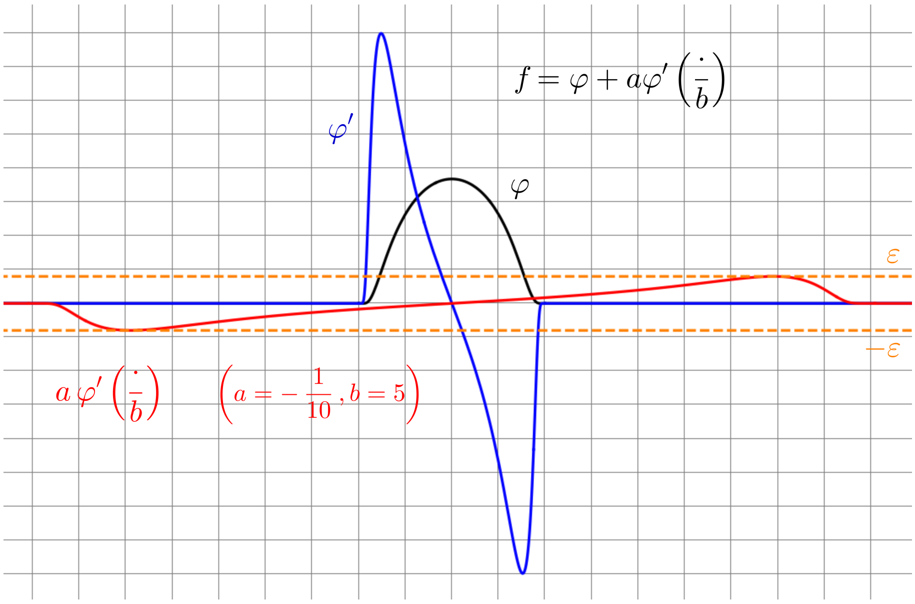

Figure 4.1: The first moment of f vanishes, if the dilation parameter b is chosen large enough.

Condition (ii) can be satisfied by choosing the absolute value of the scaling parameter a small enough. The sign of a and the value of the dilation parameter $b > 0$ are then determined by condition (i). Put differently, the added function $a\varphi'(\cdot/b)$ can be made as small as we want in all L^p-norms, without changing its first moment (see Figure 4.1 for an illustration). This scheme can be extended to approximate φ by a test function with prescribed moments up to some finite order, since adding the m-th derivative of φ to the approximation f will not compromise its moments of order $k < m$.

Preparatory for the general construction to follow, we introduce a sequence of functions $(\varphi_n)_{n\in\mathbb{N}_0}$, derived from a given test function $\varphi_0 \in \mathcal{D}(\mathbb{R})$.

Definition 4.1 (Construction of elementary building blocks)**.** Let $a = (a_n)_{n\in\mathbb{N}_0}$, $a_0 = 1$, and $b = (b_n)_{n\in\mathbb{N}_0}$ be sequences of real numbers such that $(b_n)_{n\in\mathbb{N}_0}$ is strictly increasing. Let $\varphi_0 \in \mathcal{D}(\mathbb{R})$ be a test function with $\operatorname{supp}(\varphi_0) \subseteq [b_0, b_1]$. For any triple (a, b, φ_0) with these properties we define a sequence of test functions $(\varphi_n)_{n\in\mathbb{N}_0}$, $\varphi_n \in \mathcal{D}(\mathbb{R})$,

$$\varphi_n(x) := a_n \varphi_0^{(n)}\left(\frac{x - b_n}{b_{n+1} - b_n}(b_1 - b_0) + b_0\right). \tag{4.1}$$

Basically, the function φ_n is a dilated, translated, and scaled version of the n-th derivative of φ_0. Before collecting some properties of these 'clones' of φ_0, we introduce a notation for a special choice of the antiderivative of a test function. We define the (injective) map $\varphi \mapsto \varphi^{(-1)}$, $\mathcal{D}(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ by setting

$$\varphi^{(-1)}(x) := \int_{-\infty}^{x} \varphi(y)\,\mathrm{d}y \tag{4.2}$$

and write $\varphi^{(-m)}$ for the m-th antiderivative of $\varphi \in \mathcal{D}(\mathbb{R})$. Note that $\varphi^{(-1)} \in \mathcal{C}^\infty(\mathbb{R})$ is again a test function if and only if the mass of φ vanishes.

The functions $\varphi_n \in \mathcal{D}(\mathbb{R})$ can be used as building blocks for the construction of a Schwartz function with prescribed moments which approximates the basic function φ_0 in $L^p(\mathbb{R})$. In the following two lemmas we will prove some properties of φ_n that will be basic ingredients in the proof of Proposition 4.4.

Lemma 4.2 (Antiderivatives of elementary building blocks)**.** *For φ_n as in (4.1), the m-th antiderivative of φ_n, $\varphi_n^{(-m)}$, is given by*

$$\varphi_n^{(-m)}(x) = a_n \left(\frac{b_{n+1}-b_n}{b_1-b_0}\right)^m \varphi_0^{(n-m)}\left(\frac{x-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0\right). \tag{4.3}$$

Proof. For comfort we introduce the affine linear transformation

$$\beta_n(x) := \frac{x-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0$$

with constant derivative $\beta_n' = \frac{b_1-b_0}{b_{n+1}-b_n}$. We then calculate

$$\varphi_n^{(-1)}(x) = \int_{-\infty}^{x} \varphi_n(y)\,\mathrm{d}y = a_n \int_{-\infty}^{x} \varphi_0^{(n)}\left(\beta_n(y)\right)\mathrm{d}y$$

$$\overset{z=\beta_n(y)}{=} a_n \frac{b_{n+1}-b_n}{b_1-b_0} \int_{-\infty}^{\beta_n(x)} \varphi_0^{(n)}(z)\,\mathrm{d}z = a_n \frac{b_{n+1}-b_n}{b_1-b_0} \varphi_0^{(n-1)}\left(\beta_n(x)\right).$$

Thus, after m integrations we arrive at

$$\varphi_n^{(-m)}(x) = a_n \left(\frac{b_{n+1}-b_n}{b_1-b_0}\right)^m \varphi_0^{(n-m)}\left(\frac{x-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0\right).$$

$\square$

Lemma 4.3 (Support and moments of the elementary building blocks)**.** *The test functions φ_n ($n \in \mathbb{N}_0$) as defined in (4.1) have the following properties:*

(i) $\operatorname{supp}(\varphi_n) \subseteq [b_n, b_{n+1}]$ *and* $x \mapsto \sum_{n=0}^{\infty} \varphi_n(x)$ *converges pointwise to a smooth function* φ.

(ii) *Denoting the m-th moment of a function f by $\mu_m(f)$, we have*

$$\mu_m(\varphi_n) = \begin{cases} 0 & (m<n) \\ (-1)^n n! \left(\frac{b_{n+1}-b_n}{b_1-b_0}\right)^{n+1} \int_{-\infty}^{\infty} \varphi_0(x)\mathrm{d}x & (m=n) \\ \frac{(-1)^n m!}{(m-n)!}\, a_n \left(\frac{b_{n+1}-b_n}{b_1-b_0}\right)^n \int_{-\infty}^{\infty} x^{m-n} \varphi_0\left(\frac{(x-b_n)(b_1-b_0)}{b_{n+1}-b_n}+b_0\right)\mathrm{d}x & (m>n) \end{cases}$$

for $m, n \in \mathbb{N}_0$.

Proof. (i): We show that $\operatorname{supp}(\varphi_n) \subseteq [b_n, b_{n+1}]$:

$$\begin{aligned}
x \in \operatorname{supp}(\varphi_n) \quad &\Longrightarrow \quad \frac{x-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0 \in [b_0,b_1] \\
&\Longleftrightarrow \quad 0 \le \frac{x-b_n}{b_{n+1}-b_n}(b_1-b_0) \le b_1-b_0 \\
&\Longleftrightarrow \quad 0 \le \frac{x-b_n}{b_{n+1}-b_n} \le 1 \\
&\Longleftrightarrow \quad 0 \le x-b_n \le b_{n+1}-b_n \\
&\Longleftrightarrow \quad b_n \le x \le b_{n+1}.
\end{aligned}$$

Note also that

- $\varphi_n(b_n) = a_n \varphi_0^{(n)}\left(\frac{b_n-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0\right) = a_n\varphi_0^{(n)}(b_0)$,
- $\varphi_n(b_{n+1}) = a_n \varphi_0^{(n)}\left(\frac{b_{n+1}-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0\right) = a_n\varphi_0^{(n)}(b_1)$.

Pointwise convergence of $\sum_{n=0}^{\infty}\varphi_n(x)$ to a smooth function is guaranteed, since $\operatorname{supp}(\varphi_n) \subseteq [b_n, b_{n+1}]$ and $(b_n)_{n\in\mathbb{N}_0}$ is a strictly increasing sequence by assumption.

(ii): The case $m = 0$ is settled by noting that $\int_{-\infty}^{\infty}\varphi_n(x)\mathrm{d}x = 0$ for all $n > 0$ and verifying that the formula reduces to an identity for $n = m = 0$. For $m, n \ge 1$ we calculate

$$\begin{aligned}
\int_{-\infty}^{\infty} x^m\varphi_n(x)\,\mathrm{d}x \quad &= \quad x^m\varphi_n^{(-1)}(x)\Big|_{-\infty}^{\infty} - m\int_{-\infty}^{\infty} x^{m-1}\varphi_n^{(-1)}(x)\,\mathrm{d}x \\
&= \quad -m\int_{-\infty}^{\infty} x^{m-1}\varphi_n^{(-1)}(x)\,\mathrm{d}x, \qquad (4.4)
\end{aligned}$$

since $\varphi_n^{(-1)}$ has compact support by (4.3). In fact, $\varphi_n^{(-m)} \in \mathcal{D}(\mathbb{R})$ for all $m \le n$ and thus, even after n successive integrations by parts, no boundary terms will survive. If $0 < m < n$, we exploit this fact by performing m integrations by parts, thereby obtaining

$$\int_{-\infty}^{\infty} x^m\varphi_n(x)\,\mathrm{d}x \overset{(4.4)}{=} (-1)^m m! \int_{-\infty}^{\infty}\varphi_n^{(-m)}(x)\,\mathrm{d}x = 0.$$

This proves the first assertion in (ii). If $m \ge n$, we integrate by parts n times, leading us to

$$\begin{aligned}
\int_{-\infty}^{\infty} x^m\varphi_n(x)\,\mathrm{d}x &\overset{(4.4)}{=} \frac{(-1)^n\, m!}{(m-n)!}\int_{-\infty}^{\infty} x^{m-n}\varphi_n^{(-n)}(x)\,\mathrm{d}x \\
&\overset{(4.3)}{=} \frac{(-1)^n\, m!}{(m-n)!}\, a_n\left(\frac{b_{n+1}-b_n}{b_1-b_0}\right)^m \int_{-\infty}^{\infty} x^{m-n}\varphi_0\left(\frac{x-b_n}{b_{n+1}-b_n}(b_1-b_0)+b_0\right)\mathrm{d}x.
\end{aligned}$$

In the case $m = n$, an additional variable transformation as in the proof of Lemma 4.2 gives

$$\int_{-\infty}^{\infty} x^n \varphi_n(x)\,\mathrm{d}x = (-1)^n n!\, a_n \left(\frac{b_{n+1}-b_n}{b_1-b_0}\right)^{n+1} \int_{-\infty}^{\infty} \varphi_0(x)\,\mathrm{d}x.$$

□

We are now prepared to state and prove the main result of this section, showing that any compactly support smooth function on $\mathbb{R}$ with non-vanishing mass can be approximated in L^p ($p > \frac{1}{2}$) by a Schwartz function with prescribed moments of order greater than zero and support bounded from below.

Theorem 4.4 (Schwartz functions with prescribed moments). *Let $\varphi_0 \in \mathcal{D}(\mathbb{R})$ with $\int_{-\infty}^{\infty} \varphi_0(x)\,\mathrm{d}x \neq 0$ and $\operatorname{supp}(\varphi_0) \subseteq [b_0, \infty[$ be given. Let $(\nu_n)_{n\in\mathbb{N}}$ be a sequence of real numbers and $(p_n)_{n\in\mathbb{N}}$ be a sequence of polynomials on $\mathbb{R}$, where p_n is of degree n. Then, for all $q \in]\frac{1}{2}, 1]$ and for all $\varepsilon > 0$ there exists a Schwartz function $\varphi \in \mathcal{S}(\mathbb{R})$ such that*

(i) $\operatorname{supp}(\varphi) \subseteq [b_0, \infty[$,

(ii) $\displaystyle\int_{-\infty}^{\infty} \varphi(x)\mathrm{d}x = \int_{-\infty}^{\infty} \varphi_0(x)\,\mathrm{d}x$ *and* $\displaystyle\int_{-\infty}^{\infty} p_n(x)\varphi(x)\,\mathrm{d}x = \nu_n \ \forall n \in \mathbb{N}$

(iii) $\|\varphi - \varphi_0\|_{L^\infty(\mathbb{R})} < \varepsilon$ *and* $\displaystyle\left(\int_{-\infty}^{\infty} |\varphi(x) - \varphi_0(x)|^p\,\mathrm{d}x\right)^{\frac{1}{p}} < \varepsilon \ \forall p \in [q, \infty[$
(note that the latter is not a norm for $p < 1$).

Proof. We will construct a sequence $(a_n)_{n\in\mathbb{N}_0}$ ($a_0 = 1$) and a strictly increasing sequence $(b_n)_{n\in\mathbb{N}_0}$ ($\operatorname{supp}(\varphi_0) \subseteq [b_0, \infty[$) such that $\varphi = \sum_{n=0}^{\infty} \varphi_n$ with φ_n as defined in (4.1) has the desired properties (i)-(iii). First we assume that

$$\left(\int_{-\infty}^{\infty} x^m \varphi(x)\,\mathrm{d}x = \right) \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} x^m \varphi_n(x)\,\mathrm{d}x = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} x^m \varphi_n(x)\,\mathrm{d}x. \tag{4.5}$$

for all $m \in \mathbb{N}_0$. This assumption will later be justified by a dominated convergence argument. Until then, we may formally calculate

$$\int_{-\infty}^{\infty} \varphi(x)\mathrm{d}x \overset{(4.5)}{=} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \varphi_n(x)\,\mathrm{d}x \overset{4.3(\mathrm{ii})}{=} \int_{-\infty}^{\infty} \varphi_0(x)\,\mathrm{d}x$$

and, using that $p_m(x) = c_m x^m + \sum_{k=0}^{m-1} c_{m,k}\, x^k$ where $c_m \neq 0$, we find for $m \in \mathbb{N}$ via integration by parts that

$$\begin{aligned}
\int_{-\infty}^{\infty} p_m(x)\varphi(x)\,\mathrm{d}x &\overset{(4.5),4.3\text{(ii)}}{=} \underbrace{\sum_{k=0}^{m-1}\int_{-\infty}^{\infty} p_m(x)\varphi_k(x)\,\mathrm{d}x}_{=:J_m} + \int_{-\infty}^{\infty} p_m(x)\varphi_m(x)\,\mathrm{d}x \\
&\overset{4.3\text{(ii)}}{=} J_m + c_m \int_{-\infty}^{\infty} x^m \varphi_m(x)\,\mathrm{d}x \\
&\overset{4.3\text{(ii)}}{=} J_m + c_m(-1)^m m!\, a_m \left(\frac{b_{m+1}-b_m}{b_1-b_0}\right)^{m+1} \underbrace{\int_{-\infty}^{\infty} \varphi_0(x)\,\mathrm{d}x}_{=:J_0},
\end{aligned}$$

where we have introduced the notation

$$J_m := \int_{-\infty}^{\infty} p_m(x) \sum_{k=0}^{m-1} \varphi_k(x)\,\mathrm{d}x \quad \text{for} \quad m \geq 1 \quad \text{and} \quad J_0 := \int_{-\infty}^{\infty} \varphi_0(x)\,\mathrm{d}x. \tag{4.6}$$

In order to satisfy the constraint $\int_{-\infty}^{\infty} p_m(x)\varphi(x)\,\mathrm{d}x = \nu_m$ for all $m \in \mathbb{N}$, we must have

$$\nu_m = J_m + c_m(-1)^m m!\, a_m \left(\frac{b_{m+1}-b_m}{b_1-b_0}\right)^{m+1} J_0,$$

which immediately implies

$$a_m = \frac{(-1)^m}{m!} \frac{\nu_m - J_m}{c_m J_0} \left(\frac{b_1-b_0}{b_{m+1}-b_m}\right)^{m+1} \tag{4.7}$$

since $c_m \neq 0$ and $b_{m+1} - b_m \neq 0$ as well. Setting

$$\lambda_n := \max\left(1,\ \sup_{0\leq k\leq 2n} \left(\|\varphi_0^{(k)}\|_{L^\infty(\mathbb{R})} \big/ \|\varphi_0^{(n)}\|_{L^\infty(\mathbb{R})}\right)\right) \tag{4.8}$$

for all $n \in \mathbb{N}_0$, we impose two growth conditions on the sequence $(b_m)_{m\in\mathbb{N}_0}$, namely

$$b_2 > \max\left(1,\, b_1 + \lambda_1\right), \quad b_{m+1} > \max\left((m!)^{\frac{1}{q(m-1)}},\, b_m + \lambda_m\right) \ \text{ for } m \geq 2 \tag{4.9}$$

and

$$b_{m+1}^{m+\frac{1}{q}-1} \underbrace{\left|\frac{\nu_m - J_m}{m! c_m J_0}\right| \left(\frac{b_1-b_0}{b_{m+1}-b_m}\right)^{m+1}}_{=|a_m|} \|\varphi_0^{(m)}\|_{L^\infty(\mathbb{R})} < \frac{\varepsilon}{e-1}. \tag{4.10}$$

Note that the left-hand side of inequality (4.10) is just

$$b_{m+1}^{m+\frac{1}{q}-1} \|\varphi_m\|_{L^\infty(\mathbb{R})} = b_{m+1}^{m+\frac{1}{q}-1} |a_m| \|\varphi_0^{(m)}\|_{L^\infty(\mathbb{R})}.$$

We will show by induction that sequences $(a_n)_{n\in\mathbb{N}_0}$ and $(b_n)_{n\in\mathbb{N}_0}$ satisfying (4.7), (4.9), and (4.10) can be constructed by an iterative procedure. We put $a_0 = 1$ and pick

$$b_1 > \max\big(0, \sup\{x \in \mathbb{R} : x \in \operatorname{supp}(\varphi_0)\}\big).$$

Now suppose that all distances up to b_m and all coefficients up to a_{m-1} have been picked suitably. Then J_m is uniquely determined according to (4.6). Hence all symbols in (4.9) and (4.10) except b_{m+1} are fixed and we may think of the left-hand side of inequality (4.10) as a function $b_{m+1} \mapsto h(b_{m+1})$. Since $q > \frac{1}{2}$ and therefore $1/q - 1 < 1$, we have $h(b_{m+1}) \to 0$ as $b_{m+1} \to \infty$. Thus conditions (4.9) and (4.10) can always be achieved simultaneously by taking b_{m+1} large enough. The coefficient a_m is then determined by (4.7).

Observe that (4.7) and (4.10) imply the inequality

$$|a_n| < \frac{\varepsilon/(e-1)}{b_{n+1}^{n+\frac{1}{q}-1} \|\varphi_0^{(n)}\|_{L^\infty(\mathbb{R})}}, \tag{4.11}$$

and therefore the amplitudes of the building blocks decay according to the estimate

$$\|\varphi_n\|_{L^\infty(\mathbb{R})} < \frac{\varepsilon/(e-1)}{b_{n+1}^{n+\frac{1}{q}-1}} < \varepsilon\, b_{n+1}^{-n}. \tag{4.12}$$

For the sequences $(a_n)_{n\in\mathbb{N}_0}$ and $(b_n)_{n\in\mathbb{N}_0}$ constructed above, we now define $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ by setting $\varphi(x) := \sum_{n=0}^\infty \varphi_n(x)$. We claim that $\varphi \in \mathcal{S}(\mathbb{R})$. To this end we have to show that all semi-norms $\|x^l\varphi^{(m)}\|_{L^\infty(\mathbb{R})}$ with $l, m \in \mathbb{N}_0$ are finite. We have

$$\begin{aligned}
\|x^l\varphi^{(m)}\|_{L^\infty(\mathbb{R})} &\leq \max\Big(\|x^l\varphi_0^{(m)}\|_{L^\infty(\mathbb{R})}, \sup_{n\in\mathbb{N}} \|x^l\varphi_n^{(m)}\|_{L^\infty(\mathbb{R})}\Big)\\
&\leq \max\left(\|x^l\varphi_0^{(m)}\|_{L^\infty(\mathbb{R})}, \sup_{n\in\mathbb{N}}\left(b_{n+1}^l|a_n|\left(\frac{b_1-b_0}{b_{n+1}-b_n}\right)^m \|\varphi_0^{(m+n)}\|_{L^\infty(\mathbb{R})}\right)\right)\\
&\leq \max\left(\|x^l\varphi_0^{(m)}\|_{L^\infty(\mathbb{R})}, \frac{\varepsilon(b_1-b_0)^m}{e-1}\sup_{n\in\mathbb{N}}\left(b_{n+1}^{l-n-\frac{1}{q}+1}\frac{\|\varphi_0^{(m+n)}\|_{L^\infty(\mathbb{R})}}{\lambda_n^m\|\varphi_0^{(n)}\|_{L^\infty(\mathbb{R})}}\right)\right),
\end{aligned}$$

where the last line follows from (4.11) and (4.9), which we have used to estimate $(b_{n+1}-b_n)^m \geq \lambda_n^m$. The finiteness of the supremum is immediate after substituting for λ_n, more precisely we have for $n \in \mathbb{N}$,

$$\begin{aligned}
b_{n+1}^{l-n-\frac{1}{q}+1}\frac{\|\varphi_0^{(m+n)}\|_{L^\infty(\mathbb{R})}}{\lambda_n^m\|\varphi_0^{(n)}\|_{L^\infty(\mathbb{R})}} &\leq b_{n+1}^{l-n}\frac{\|\varphi_0^{(m+n)}\|_{L^\infty(\mathbb{R})}}{\sup\limits_{0\leq k\leq 2n}\|\varphi_0^{(k)}\|_{L^\infty(\mathbb{R})}}\\
&\leq 1 \text{ for } n \geq \max(l, m).
\end{aligned}$$

The norm $\|x^l\varphi_0^{(m)}\|_{L^\infty(\mathbb{R})}$ is finite, since $\varphi_0 \in \mathcal{D}(\mathbb{R})$. Finally we estimate the L^1-norm of $|\varphi - \varphi_0|^p$ (where $p \geq 1 \geq q > \frac{1}{2}$),

$$\begin{aligned}
\int\limits_{-\infty}^{\infty} |\varphi(x) - \varphi_0(x)|^p \mathrm{d}x &\leq \sum_{n=1}^{\infty} b_{n+1} \|\varphi_n\|_{L^\infty(\mathbb{R})}^p \quad (\mathrm{supp}(\varphi_n) \subseteq [b_n, b_{n+1}]) \\
&\leq \frac{\varepsilon^p}{e-1} \sum_{n=1}^{\infty} b_{n+1}^{1-p(n+\frac{1}{q}-1)} \quad (\text{by } (4.11)) \\
&\leq \frac{\varepsilon^p}{e-1} \sum_{n=1}^{\infty} b_{n+1}^{p(1-n)} \quad (\text{since } 1 - p/q < 0) \\
&\leq \frac{\varepsilon^p}{e-1} \sum_{n=1}^{\infty} (n!)^{-\frac{p}{q}} \quad (\text{by } (4.9)) \\
&\leq \frac{\varepsilon^p}{e-1} \sum_{n=1}^{\infty} \frac{1}{n!} \quad (\text{since } p/q > 1) \\
&= \varepsilon^p
\end{aligned}$$

This proves that

$$\left(\int\limits_{-\infty}^{\infty} |\varphi(x) - \varphi_0(x)|^p \mathrm{d}x \right)^{1/p} < \varepsilon$$

for all $p \geq q$, in particular for all L^p-norms with $1 \leq p < \infty$ (since $q \in]\frac{1}{2}, 1]$), but also for all L^p-metrics with $p \in [q, 1[$. Moreover we find that

$$\begin{aligned}
\|\varphi_0 - \varphi\|_{L^\infty(\mathbb{R})} \leq \sup_{n\in\mathbb{N}} \|\varphi_n\|_{L^\infty(\mathbb{R})} &\leq \sup_{n\in\mathbb{N}} \|a_n \varphi_0^{(n)}\|_{L^\infty(\mathbb{R})} \\
&\leq \frac{\varepsilon}{e-1} \sup_{n\in\mathbb{N}} b_{n+1}^{-n-\frac{1}{q}+1} \quad (\text{by } (4.11)) \\
&\leq \frac{\varepsilon}{e-1} \sup_{n\in\mathbb{N}} b_{n+1}^{-n} < \varepsilon,
\end{aligned}$$

where the last inequality follows from the fact that $-\frac{1}{q} + 1 \leq 0$ and that $b_2 > 1$ by (4.9). It remains to show that φ indeed satisfies the polynomial moment condition (ii), namely

$$\int\limits_{-\infty}^{\infty} p_n(x)\varphi(x)\mathrm{d}x = \nu_n \quad \text{for all} \quad n \in \mathbb{N}.$$

In order to prove this, we only have to show that the formal calculation (4.5) 'was not only formal', that is, we have to check that indeed

$$\int\limits_{-\infty}^{\infty} p_m(x) \sum_{k=0}^{\infty} \varphi_k(x)\mathrm{d}x = \sum_{k=0}^{\infty} \int\limits_{-\infty}^{\infty} p_m(x)\varphi_k(x)\mathrm{d}x.$$

For this purpose we introduce the non-negative functions

$$p_m^+(x) = |c_m x^m| + |c_{m,m-1}x^{m-1}| + ... + |c_{m,0}| \geq p_m(x)$$

and as well as the step functions $g_m : \mathbb{R} \to \mathbb{R}$,

$$g_m(x) = \begin{cases} 0 & \text{for} \quad x \leq b_0 \\ p_m^+(b_{k+1}) \|\varphi_k\|_{L^\infty(\mathbb{R})} & \text{for} \quad b_k \leq x < b_{k+1}, \quad k \in \mathbb{N}_0 \end{cases}$$

whose L^1-norm is finite by the estimate

$$\begin{aligned} \|g_m\|_{L^1(\mathbb{R})} &\leq \sum_{k=0}^{\infty} (b_{k+1} - b_k) p_m^+(b_{k+1}) \|\varphi_k\|_{L^\infty(\mathbb{R})} \\ &\leq \sum_{k=0}^{\infty} b_{k+1} p_m^+(b_{k+1}) \|\varphi_k\|_{L^\infty(\mathbb{R})} \\ &\overset{(4.12)}{\leq} \varepsilon \sum_{k=0}^{\infty} p_m^+(b_{k+1}) b_{k+1}^{1-k} < \infty. \end{aligned}$$

Since $\sum_{k=0}^{n} p_m(x)\varphi_k(x) \leq g_m(x)$ for all $x \in \mathbb{R}$ and this sum converges pointwise to $p_m(x)\varphi(x)$, we can apply dominated convergence to pull the sum through the integral sign. □

The proof of Proposition 4.4 contains an explicit construction of Schwartz functions with support contained in $]0, \infty[$, unit mass and otherwise vanishing moments. Hence they qualify as mollifiers for the embedding of $\mathcal{E}'(]0, \infty[)$ into the special Colombeau $\mathcal{G}(]0, \infty[)$.[1] Moreover, the difference to a given test function of unit integral, measured in L^p-norms, can be made as small as we wish. In fact, Proposition 4.4 can be employed to construct a sequence or net of Schwartz functions with vanishing moments, converging in all $W^{k,p}$-norms to a given test function.

Corollary 4.5 (Nets of Schwartz functions with vanishing moments)**.** *Let* $\varphi \in \mathcal{D}(\mathbb{R})$ *with* $\int_0^\infty \varphi(x)\mathrm{d}x = 1$, $\mathrm{supp}(\varphi) \subseteq]0, \infty[$, *and a net* $(\sigma_\varepsilon)_\varepsilon$ *with* $\sigma_\varepsilon \to 0$ *as* $\varepsilon \to 0$ *be given. Then there exists a net of Schwartz functions* $(\psi_\varepsilon)_\varepsilon$ *such that*

(i) $\mathrm{supp}(\psi_\varepsilon) \subseteq]0, \infty[$ *for all* $\varepsilon > 0$,

(ii) $\displaystyle\int_0^\infty \psi_\varepsilon(x)\mathrm{d}x = 1$ *and* $\displaystyle\int_0^\infty x^n \psi_\varepsilon(x)\mathrm{d}x = 0$ *for all* $\varepsilon > 0$ *and for all* $n \in \mathbb{N}$,

(iii) $\|\psi_\varepsilon - \varphi\|_{W^{k,p}(\mathbb{R})} = O(\sigma_\varepsilon)$ *as* $\varepsilon \to 0$ *for all* $k \in \mathbb{N}_0$ *and* $1 \leq p \leq \infty$.

Proof. The case $k = 0$ is an immediate consequence of Proposition 4.4, as it can be employed to create an approximating sequence (or net) of Schwartz function, converging to the given function at the desired rate. Indeed, we may use Proposition 4.4 for $q = 1$ to construct a net $(\psi_\varepsilon)_\varepsilon$ approximating φ in $L^p(\mathbb{R})$ up to an error proportional to σ_ε. The functions $\psi_\varepsilon \in \mathcal{S}(\mathbb{R})$ are then defined pointwise by

$$\psi_\varepsilon(x) = \sum_{n=1}^{\infty} a_n^\varepsilon \, \varphi^{(n)} \left(\frac{x - b_n^\varepsilon}{b_{n+1}^\varepsilon - b_n^\varepsilon} (b_1^\varepsilon - b_0^\varepsilon) + b_0^\varepsilon \right),$$

[1] This possible application has been suggested by Michael Oberguggenberger.

where $(b_0^\varepsilon)_\varepsilon = (b_0)_\varepsilon$ with $b_0 > 0$ and $(b_1^\varepsilon)_\varepsilon = (b_1)_\varepsilon$, since $\operatorname{supp}(\varphi_0) \subseteq]b_0, b_1[$ is independent of ε. The indexed families of coefficients $a : \mathbb{N}_0 \times]0,1[\to \mathbb{R}$ and distances $b : \mathbb{N}_0 \times]0,1[\to [0,\infty[$ are determined by following the steps in the proof of Proposition 4.4, but with numerator σ_ε in place of ε in growth condition (4.10), implying that

$$|a_n^\varepsilon| \le \frac{\sigma_\varepsilon/(e-1)}{(b_{n+1}^\varepsilon)^n \|\varphi^{(n)}\|_{L^\infty(]0,\infty[)}} \le \sigma_\varepsilon \|\varphi^{(n)}\|^{-1}_{L^\infty(]0,\infty[)}. \tag{4.13}$$

To show convergence in $W^{k,p}(\mathbb{R})$, we note that

$$\begin{aligned}
\|\partial^k(\psi_\varepsilon - \varphi)\|_{L^\infty(]0,\infty[)} &= \sup_{n\in\mathbb{N}} |a_n^\varepsilon| \left(\frac{b_1^\varepsilon - b_0^\varepsilon}{b_{n+1}^\varepsilon - b_n^\varepsilon}\right)^k \|\varphi^{(k+n)}\|_{L^\infty(]0,\infty[)} \\
&\le \sigma_\varepsilon (b_1 - b_0)^k \sup_{n\in\mathbb{N}} \frac{\|\varphi^{(k+n)}\|_{L^\infty(]0,\infty[)}}{\sup\limits_{0\le m\le 2n} \|\varphi^{(m)}\|_{L^\infty(]0,\infty[)}} \\
&\le C_k \sigma_\varepsilon,
\end{aligned}$$

where we have used that $(b_{n+1}^\varepsilon - b_n^\varepsilon)^k \ge \lambda_n^k \ge \lambda_n$ as defined in (4.8), which is independent of ε.

For $p \in [1,\infty[$, we find

$$\begin{aligned}
\|\partial^k(\psi_\varepsilon - \varphi)\|_{L^p(]0,\infty[)} &\le \sum_{n\in\mathbb{N}} (b_{n+1}^\varepsilon)^{\frac{1}{p}} |a_n^\varepsilon| \frac{(b_1 - b_0)^k}{\lambda_n} \|\varphi^{(k+n)}\|_{L^\infty(]0,\infty[)} \\
&\le \sum_{n\in\mathbb{N}} (b_{n+1}^\varepsilon)^{\frac{1}{p}-n} \frac{\sigma_\varepsilon (b_1 - b_0)^k}{e-1} \sup_{n\in\mathbb{N}} \frac{\|\varphi^{(k+n)}\|_{L^\infty(]0,\infty[)}}{\sup\limits_{0\le m\le 2n} \|\varphi^{(m)}\|_{L^\infty(]0,\infty[)}} \\
&\le \frac{\sigma_\varepsilon}{e-1} \widetilde{C}_k \sum_{n\in\mathbb{N}} (b_{n+1}^\varepsilon)^{\frac{1}{p}-n} \\
&\le \widetilde{C}_k \sigma_\varepsilon,
\end{aligned}$$

where we have used the first inequality in (4.13) and that $(b_{n+1}^\varepsilon)^{\frac{1}{p}-n} \le \frac{1}{n!}$ by (4.9). This proves convergence in $W^{k,p}(]0,\infty[)$ for $1 \le p < \infty$. □

Remark 4.6 (Moments and negligibility in $\mathcal{G}$)**.** Corollary 4.5 implies that for any $\varphi \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp}(\varphi) \subseteq]0,\infty[$ and $\int_0^\infty \varphi(x)\mathrm{d}x = 1$, the class of the constant net $(\varphi)_\varepsilon$ in $\mathcal{G}^\infty(\mathbb{R})$ contains a representative $(\psi_\varepsilon)_\varepsilon$ satisfying conditions (i), (ii), and (iii): We obtain a net of the desired type by taking, for example, $\sigma_\varepsilon = e^{-\frac{1}{\varepsilon}}$ in Corollary 4.5. Then we have $\|\psi_\varepsilon\|_{W^{k,\infty}(\mathbb{R})} \le \|\varphi\|_{W^{k,\infty}(\mathbb{R})}$ for all $\varepsilon \le \varepsilon_0$ and $\|\psi_\varepsilon - \varphi\|_{W^{k,\infty}(\mathbb{R})} = O(\varepsilon^m)$ as $\varepsilon \to 0$ for all $m \in \mathbb{N}_0$. This implies $[(\psi_\varepsilon)_\varepsilon] = [(\varphi)_\varepsilon]$.

In fact, the negligible parts of $(\psi_\varepsilon)_\varepsilon$, running off to infinity as $\varepsilon \to 0$, exactly compensate the (nonzero) moments of φ. This may seem counterintuitive at first sight, but it is actually not very surprising. Take, for instance, $\varphi_\varepsilon := \varphi(\,\cdot\, - \varepsilon^{-1})$, yielding $[(\varphi_\varepsilon)_\varepsilon] = 0$ in $\mathcal{G}(\mathbb{R})$, since $(\varphi_\varepsilon)_\varepsilon$ vanishes on any compact set for ε small enough. However, for the componentwise moments of the special representative $(\varphi_\varepsilon)_\varepsilon$, we get $\mu_n(\varphi_\varepsilon) = O(\varepsilon^{-n})$ as $\varepsilon \to 0$. Similarly we may consider $(\widetilde{\varphi}_\varepsilon)_\varepsilon$ with $\widetilde{\varphi}_\varepsilon := e^{-\frac{1}{\varepsilon}} \varphi(e^{-\frac{1}{\varepsilon}}\,\cdot\,)$, which is negligible even in $\mathcal{G}_{L^\infty}(\mathbb{R})$. However, the componentwise moments of this representative are $\mu_n(\widetilde{\varphi}_\varepsilon) = e^{\frac{n}{\varepsilon}} \mu_n(\varphi)$ and thus not even generalised numbers, if $\mu_n(\varphi) \neq 0$ (which must be true for some $n \in \mathbb{N}$). Of course, the problem is that $\widetilde{\varphi}_\varepsilon$ stretches out to infinity as $\varepsilon \to 0$, but there is no inconsistency here, since componentwise integration on the level of representatives differs from integration in $\mathcal{G}$, unless the integrand is a compactly supported generalised function. Integration in Colombeau algebras can be defined in the space of tempered generalised functions $\mathcal{G}_\tau$ (see [GKOS01, Definition 1.2.24]), using damping measures to produce convergence (cf. [Hör99] and [GKOS01, Section 1.2.5]). Both $[(\varphi_\varepsilon)_\varepsilon]$ and $[(\widetilde{\varphi}_\varepsilon)_\varepsilon]$ belong to the null ideal $\mathcal{N}_\tau(\mathbb{R})$ in $\mathcal{G}_\tau(\mathbb{R})$ and, consistently, their moments (in $\mathcal{G}_\tau(\mathbb{R})$) vanish as well (since $[(x \mapsto x^n)_\varepsilon] \in \mathcal{G}_\tau(\mathbb{R})$).

Note that in Proposition 4.4, the mass or zeroth moment of the constructed mollifier is equal to the mass of the given test function, which is required to be nonzero. Yet, this is not the case in the Stieltjes moment problem, where the zeroth moment may be zero as well. The following corollary includes this possibility, as it allows to choose also the zeroth moment of the generated Schwartz function without any restriction. However, this modification of the construction comes at the cost of losing the approximation of the given function in $L^p(\mathbb{R})$ for $p \leq 1$.

Corollary 4.7 (The Stieltjes moment problem for Schwartz functions)**.** *Assume that a function $\varphi \in \mathcal{D}(\mathbb{R})$ with non-vanishing integral $\int_{-\infty}^{\infty} \varphi(x)\,\mathrm{d}x \neq 0$ and* $\operatorname{supp}(\varphi) \subseteq$ *$[b_{-1}, b_0]$, where $b_0 \geq 0$, is given. In addition, let $(\nu_n)_{n\in\mathbb{N}_0}$ be a sequence of real numbers and $(p_n)_{n\in\mathbb{N}_0}$ a sequence of polynomials on $\mathbb{R}$, where p_n is of degree n and $p_0 \neq 0$. Then, for any $q \in]1, 2[$ and for all $\varepsilon > 0$, there exists a Schwartz function $\psi \in \mathcal{S}(\mathbb{R})$ such that*

(i) $\operatorname{supp}(\psi) \subseteq [b_{-1}, \infty[\,$,

(ii) $\displaystyle\int_{-\infty}^{\infty} p_n(x)\psi(x)\,\mathrm{d}x = \nu_n \;\; \forall n \in \mathbb{N}_0,$

(iii) $\|\psi - \varphi\|_{L^p(\mathbb{R})} \leq \varepsilon \quad$ for $q \leq p \leq \infty$.

Proof. For indices $n \in \mathbb{N}_0$ we define the building blocks

$$\psi_n(x) = a_n \varphi^{(n)}\left(\frac{x - b_n}{b_{n+1} - b_n}(b_0 - b_{-1}) + b_{-1}\right) \tag{4.14}$$

and put

$$\psi(x) = \varphi(x) + \sum_{n=0}^{\infty} \psi_n(x). \tag{4.15}$$

From here on, the proof is similar to the proof of Proposition 4.4. Instead of (4.6), we now have

$$J_n := \int_{-\infty}^{\infty} p_n(x)\Big(\varphi(x) + \sum_{k=0}^{n-1} \psi_k(x)\Big)\mathrm{d}x \quad \text{for} \quad n \in \mathbb{N}_0, \tag{4.16}$$

in particular, $J_0 = \int_{-\infty}^{\infty} c_0 \varphi(x)\mathrm{d}x$, where c_n is the highest-order coefficient of the polynomial p_n and $p_0 = c_0 \neq 0$. In order to satisfy $\int_{-\infty}^{\infty} p_n(x)\psi(x)\mathrm{d}x = \nu_n$ for all

$n \in \mathbb{N}_0$, we set

$$a_n = \frac{(-1)^n}{n!} \frac{c_0}{c_n} \frac{\nu_n - J_n}{J_0} \left(\frac{b_0 - b_{-1}}{b_{n+1} - b_n} \right)^{n+1}. \tag{4.17}$$

Analogously to (4.10), we pick the distances b_n such that they fulfil the inequality

$$b_{n+1}^{n+\frac{1}{q}} \|\psi_n\|_{L^\infty(\mathbb{R})} \quad = \quad b_{n+1}^{n+\frac{1}{q}} \left| \frac{c_0(\nu_n - J_n)}{n! c_n J_0} \right| \left(\frac{b_0 - b_{-1}}{b_{n+1} - b_n} \right)^{n+1} \|\varphi^{(n)}\|_{L^\infty(\mathbb{R})} \le \frac{\varepsilon}{e}.$$

The distance $b_0 \ge 0$ has already been chosen. Hence, $b_1 > b_0$ should satisfy

$$b_1^{\frac{1}{q}} \left| \frac{\nu_0 - J_0}{J_0} \right| \frac{b_0 - b_{-1}}{b_1 - b_0} \|\varphi\|_{L^\infty(\mathbb{R})} \le \frac{\varepsilon}{e}.$$

Since $q > 1$ (which is essential here), the left-hand side tends to zero as b_1 tends to infinity, so the inequality will be satisfied for b_1 large enough. The coefficient a_0 is then determined by equation (4.17). The remaining part of the proof follows closely the line of arguments in the proof of Proposition 4.4 and we eventually obtain

$$\|\psi - \varphi\|_{L^p(\mathbb{R})} \le \varepsilon \quad \text{for } q \le p \le \infty.$$

□

Taking $\varphi \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp}(\varphi) \subseteq [0, \infty]$ and $p_n(x) = x^n$ for all $n \in \mathbb{N}_0$, Corollary 4.7 provides a constructive solution to the special variant of the Stieltjes moment problem solved in [Dur89, DE94, LS08] as well as to the ‘strong Hamburger moment problem’ solved in [DE94]. However, our result is stronger in the sense that we additionally require the desired Schwartz function to approximate a given test function.

Remark 4.8 (A counter example)**.** Note that the statement of Corollary 4.7 would be wrong in the case $q = 1$. This can be seen as follows: Consider a non-negative function $\varphi \in \mathcal{D}(\mathbb{R})$ with $\int_{-\infty}^{\infty} \varphi(x) \mathrm{d}x = 1$, and the choices $\nu_0 = 0$ and $c_0 = 1$. Now assume that there exists a function $\psi \in \mathcal{S}(\mathbb{R})$ with the desired properties, in particular

$$0 = \int_{-\infty}^{\infty} \psi(x) \mathrm{d}x = \int_{-\infty}^{\infty} \big(\varphi(x) + \psi_0(x)\big) \mathrm{d}x = 1 + \int_{-\infty}^{\infty} \psi_0(x) \mathrm{d}x,$$

implying that $\|\psi_0\|_{L^1(\mathbb{R})} = 1$, since ψ_0 is either non-negative or non-positive, depending on the sign of a_0.[2] On the other hand,

$$\|\psi - \varphi\|_{L^1(\mathbb{R})} = \left\| \sum_{n=0}^{\infty} \psi_n \right\|_{L^1(\mathbb{R})} = \sum_{n=0}^{\infty} \|\psi_n\|_{L^1(\mathbb{R})} \ge \|\psi_0\|_{L^1(\mathbb{R})} \ge 1,$$

since the functions ψ_n have pairwise disjoint supports. Hence assertion (iii) in Corollary 4.7 is violated, contradicting our assumption.

[2]More generally, for ψ_0 to compensate the unit mass of φ, its L^1-norm cannot be smaller than 1. However, its L^p-norms for $1 < p \le \infty$ can still be made arbitrarily close to zero.

Remark 4.9 (Schwartz functions with vanishing moments are dense in L^p)**.** Corollary 4.7 basically shows that the vector space

$$\mathcal{S}_0(]0,\infty[) := \left\{\varphi \in \mathcal{S}(\mathbb{R}) |\, \forall n \in \mathbb{N}_0 : \int_{-\infty}^{\infty} x^n \varphi(x)\mathrm{d}x = 0 \quad \text{and} \quad \operatorname{supp}(\varphi) \subseteq]0,\infty[\right\}$$

is a dense subspace of $L^p(]0,\infty[)$ for $1 < p < \infty$, since $\mathcal{D}(]0,\infty[)$ is dense (cf. [Bha12, Theorem 6.8.3]). The space $\mathcal{S}_0(]0,\infty[)$ is invariant under dilation $\varphi \mapsto \varphi(\lambda\,\cdot\,)$ with $\lambda > 0$. It is also closed under taking derivatives, since for any $\varphi \in \mathcal{S}_0(]0,\infty[)$ and any $n \in \mathbb{N}$, integration by parts yields

$$\int_{-\infty}^{\infty} x^n \varphi'(x)\mathrm{d}x = -n \int_{-\infty}^{\infty} x^{n-1}\varphi(x)\mathrm{d}x = 0,$$

and the mass of φ' vanishes as well.

4.2 Conical Schwartz functions with vanishing moments

The investigations in this chapter originated from the observation that a function with vanishing moments cannot have compact support, yet its support may not necessarily extend to infinity in all directions. As the one-dimensional case is settled and we even obtained an explicit construction of Schwartz functions with vanishing moments and support contained in $]0,\infty[$, we now turn our attention to the multidimensional case. An obvious possibility to generalise the mollifier construction in the proof of Proposition 4.4 to arbitrary dimensions is to consider n-fold tensor products of one-dimensional mollifiers.

Corollary 4.10 (A trivial generalisation to arbitrary dimensions)**.** *There exists a function $\rho \in \mathcal{S}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \rho(x)\mathrm{d}x = 1$, $\int_{\mathbb{R}^n} x^\alpha \rho(x)\mathrm{d}x = 0$ for all $\alpha \in \mathbb{N}_0^n$, $|\alpha| > 0$, and support contained in the (conical) set $]0,\infty[^{\,n}$.*

Proof. Take $\varphi_0 \in \mathcal{D}(\mathbb{R})$ with $\int_{-\infty}^{\infty} \varphi_0(x)\mathrm{d}x = 1$, $\operatorname{supp}(\varphi_0) \subseteq]0,\infty[$, and apply Proposition 4.4 to construct $\psi \in \mathcal{S}(\mathbb{R})$ with $\operatorname{supp}(\psi) \subseteq]0,\infty[$ and all moments of order $m > 0$ vanishing. Composing a multidimensional mollifier $\rho \in \mathcal{S}(\mathbb{R}^n)$ by taking $\rho(x) = \psi(x_1) \otimes \psi(x_2) \otimes ... \otimes \psi(x_n)$, it is easy to see that ρ has the desired properties. □

Aiming at a generalisation of the one-dimensional construction in Section 4.1, the multidimensional mollifier example in Corollary 4.10 already hints at the possible shape of the required 'minimal' support. This section is devoted to an explicit construction of Schwartz functions on $\mathbb{R}^n$ with vanishing moments and support contained in a conical subset of $\mathbb{R}^n$. It is also possible to create a sequence of Schwartz functions of this type, converging in the space $W^{k,p}(\mathbb{R}^n)$ to a given compactly supported smooth function.

First, we set up the fundamental scheme of the construction by introducing building blocks for the prospective mollifier, derived from a given 'basic function'

$\varphi_0 \in \mathcal{D}(\mathbb{R}^n)$. All derivatives $\partial^\alpha \varphi_0$ of the same order $|\alpha| = m$ are combined to one 'wave packet' φ_m positioned along a ray in the direction of some vector ξ_0. Compared to the one-dimensional case, we allow for independent dilation and translation parameters. This decoupling allows to control the angle covered by the support of each packet.

Lemma 4.11 (Multidimensional building blocks and essential properties)**.** *Let $\xi_0 \in \mathbb{R}^n$ be nonzero, let $a = (a_n)_{n\in\mathbb{N}_0}$ and $b = (b_m)_{m\in\mathbb{N}_0}$ be sequences of real numbers with $a_m \neq 0$, $a_0 = 1$, and $b_0 = 0$. Let $\varphi_0 \in \mathcal{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x \neq 0$ and $c = (c_\alpha)_{\alpha\in\mathbb{N}_0}$ with $c_\alpha \in \mathbb{R}$ (playing the role of a weight) and $c_0 = 1$. For any quadruple (a, b, c, φ_0) with these properties and multi-indices $\alpha \in \mathbb{N}_0^n$, we define a sequence of test functions $(\varphi_m)_{m\in\mathbb{N}_0}$, $\varphi_m \in \mathcal{D}(\mathbb{R}^n)$,*

$$\varphi_m(x) := \sum_{|\alpha|=m} c_\alpha \varphi_\alpha(x), \tag{4.18}$$

where

$$\varphi_\alpha(x) := \partial^\alpha \varphi_0\left(\frac{x}{a_m} - b_m\xi_0\right). \tag{4.19}$$

Then, given $\beta \in \mathbb{N}_0^n$, we have

$$\int\limits_{\mathbb{R}^n} x^\beta \varphi_m(x)\mathrm{d}x = \begin{cases} 0 & \text{for } |\beta| < m, \\ (-1)^{|\beta|} a_m^{|\beta|+n}\, c_\beta \int\limits_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x & \text{for } |\beta| = m. \end{cases} \tag{4.20}$$

Proof. We show that if $|\alpha| = m$, we obtain

$$\int\limits_{\mathbb{R}^n} x^\beta \varphi_\alpha(x)\mathrm{d}x = \begin{cases} 0 & \text{for } |\beta| \leq |\alpha| \text{ and } \beta \neq \alpha, \\ (-a_m)^{|\alpha|+n} \int\limits_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x & \text{for } \beta = \alpha, \end{cases}$$

implying (4.20). If $|\alpha| = m$, $|\beta| \leq m$ and $\beta \neq \alpha$, then there exists at least one component β_j of β which is smaller than the corresponding component of α, that is, $\alpha_j - \beta_j \geq 1$. Rewriting the integral under consideration, using Fubini's theorem and integrating by parts β_j times with respect to x_j yields

$$\begin{aligned} \int\limits_{\mathbb{R}^n} x^\beta \varphi_\alpha(x)\mathrm{d}x &= \int\limits_{\mathbb{R}^{n-1}} x^\beta x_j^{-\beta_j} \int\limits_{\mathbb{R}} x_j^{\beta_j} \varphi_\alpha(x)\mathrm{d}x_j \mathrm{d}x_{[j]} \\ &= \int\limits_{\mathbb{R}^{n-1}} x^\beta x_j^{-\beta_j} \int\limits_{\mathbb{R}} x_j^{\beta_j} \partial^\alpha \varphi_0\left(\frac{x}{a_m} - b_m\xi_0\right) \mathrm{d}x_j \mathrm{d}x_{[j]} \\ &= (-a_m)^{\beta_j} \int\limits_{\mathbb{R}^{n-1}} x^\beta x_j^{-\beta_j} \int\limits_{\mathbb{R}} \partial_{x_j} \partial^\gamma \varphi_0\left(\frac{x}{a_m} - b_m\xi_0\right) \mathrm{d}x_j \mathrm{d}x_{[j]} \\ &= 0. \end{aligned}$$

Here, $\gamma \in \mathbb{N}_0^n$ is some multi-index with $\gamma_j \geq 0$ and $\mathrm{d}x_{[j]}$ is the $(n-1)$-dimensional volume element without $\mathrm{d}x_j$. The inner integral vanishes because the function

$x_j \mapsto \partial^\gamma \varphi_0(x/a_m - b_m\xi_0)$ has compact support. Boundary terms occurring through integration by parts vanish for the same reason. If $\beta = \alpha$, an analogous calculation leads to

$$\int\limits_{\mathbb{R}^n} x^\alpha \varphi_\alpha(x)\mathrm{d}x = (-a_m)^{|\alpha|} \int\limits_{\mathbb{R}^n} \varphi_0\left(\frac{x}{a_m}\right)\mathrm{d}x = (-1)^{|\alpha|} a_m^{|\alpha|+n} \int\limits_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x.$$

□

The functions φ_m as defined in Lemma 4.11 will be the basic elements for the construction of a mollifier function with vanishing moments on $\mathbb{R}^n$ in Theorem 4.12. To obtain a rapidly decreasing function φ which is close to its ancestor φ_0 in the $W^{k,p}$-norm, the involved sequences of dilations and translation have to be constructed carefully. The essential steps in the proof are similar to the one-dimensional case, yet some of the estimates are slightly more delicate and the recurrence relations have to be modified to achieve the cone property and to include the approximation of φ_0 in $W^{k,p}(\mathbb{R}^n)$ instead of $L^p(\mathbb{R}^n)$ right from the start.

Theorem 4.12 (Conical Schwartz functions with vanishing moments)**.** *Let $\varphi_0 \in \mathcal{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi_0(x)\,\mathrm{d}x \neq 0$ and $0 \notin \operatorname{supp}(\varphi_0)$ be given. Choose $\xi_0 \in \mathbb{R}^n$ and $r_0 > 0$, such that $\operatorname{supp}(\varphi_0) \subseteq \xi_0 + B_{r_0}$, where $|\xi_0| > r_0 > 0$ and B_{r_0} is the open ball of radius r_0, centred at 0. Let Γ be the smallest open cone (that is, for all $\xi \in \Gamma$ and $\lambda > 0$, $\lambda\xi \in \Gamma$), containing the set $\xi_0 + B_{r_0}$. Then, for all $k \in \mathbb{N}_0$, for all $1 \leq p \leq \infty$ and for all $\varepsilon > 0$, there exists a Schwartz function $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with*

(i) $\operatorname{supp}(\varphi) \subseteq \Gamma$,

(ii) $\int\limits_{\mathbb{R}^n} \varphi(x)\mathrm{d}x = \int\limits_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x$ *and* $\int\limits_{\mathbb{R}^n} x^\alpha \varphi(x)\mathrm{d}x = 0$ *for all* $\alpha \in \mathbb{N}_0^n$, $|\alpha| > 0$,

(iii) $\|\varphi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} < \varepsilon$.

Proof. We will construct φ as the infinite sum of building blocks φ_m as defined in Lemma 4.11, that is,

$$\varphi(x) := \sum_{m=0}^{\infty} \varphi_m(x) = \sum_{m=0}^{\infty} \sum_{|\alpha|=m} c_\alpha \partial^\alpha \varphi_0\left(\frac{x}{a_m} - b_m\xi_0\right), \tag{4.21}$$

where a, b, c have to be defined appropriately. We put $a_0 = 1$, $b_0 = 0$ and require the sequences $(a_n)_{n\in\mathbb{N}_0}$, $(b_n)_{n\in\mathbb{N}_0}$ to be strictly increasing. In addition, we assume that the sum in (4.21) converges pointwise. This will be justified later on. By setting

$$b_{m+1} = \max\left(a_{m+1}, b_m + \frac{r_0}{|\xi_0|}(a_m + a_{m+1})\right) \tag{4.22}$$

for all $m \in \mathbb{N}_0$, we automatically have $\operatorname{supp}(\varphi_m) \subseteq \Gamma$ (for all $m \in \mathbb{N}_0$) and $\operatorname{supp}(\varphi_m) \cap \operatorname{supp}(\varphi_n) = \emptyset$ if $m \neq n$ (see Figure 4.2).

From (4.21) it follows that $\operatorname{supp}(\varphi_m) \subseteq b_m\xi_0 + B_{a_m r_0}$, that is, φ_m has its support in a ball of radius $a_m r_0$ centred at $b_m\xi_0$. A simple scaling argument then reveals

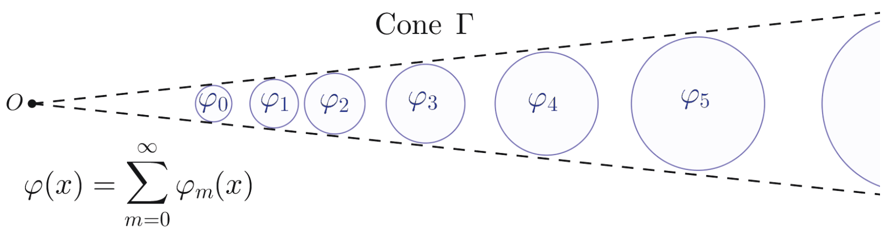

Figure 4.2: The 'wave packets' φ_m have pairwise disjoint supports and are contained in the cone Γ.

that the cone Γ contains all sets of the form $\mu\xi_0 + B_{\lambda r_0}$ with $\mu \geq \lambda > 0$. Hence, the ball $b_m\xi_0 + B_{a_m r_0}$ will be contained in Γ if only $b_m \geq a_m$. Moreover,

$$\begin{aligned}
\operatorname{supp}(\varphi_{m+1}) \cap \operatorname{supp}(\varphi_m) = \emptyset \quad &\iff \quad (b_{m+1}\xi_0 + B_{a_{m+1}r_0}) \cap (b_m\xi_0 + B_{a_m r_0}) = \emptyset \\
&\iff \quad |b_{m+1}\xi_0 - b_m\xi_0| \geq a_m r_0 + a_{m+1} r_0 \\
&\iff \quad b_{m+1} \geq b_m + \frac{r_0}{|\xi_0|}(a_m + a_{m+1}),
\end{aligned}$$

which is also guaranteed by (4.22). Finally, since the supports of the functions φ_m are lined up along the ray $\{\lambda\xi_0 : \lambda > 0\}$, this also ensures that $\operatorname{supp}(\varphi_m) \cap \operatorname{supp}(\varphi_n) = \emptyset$ for any $m \neq n$.

Concerning the dilation parameters a_m, for the time being we only impose that

$$a_{m+1} \geq \max(a_m, b_m). \tag{4.23}$$

Note that this together with (4.22) implies the estimate

$$a_m \leq b_m \leq a_m \left(1 + 2\frac{r_0}{|\xi_0|}\right) \leq 3a_m. \tag{4.24}$$

Since the building blocks $(\varphi_m)_{m\in\mathbb{N}_0}$ have pairwise disjoint supports, the infinite sum in (4.21) converges pointwise to a smooth function. Most importantly, we want φ to have vanishing moments. To achieve this, we need for any $\beta \in \mathbb{N}_0^n$ with $|\beta| > 0$ that

$$\begin{aligned}
\int_{\mathbb{R}^n} x^\beta \varphi(x)\mathrm{d}x = 0 \quad &\iff \quad \underbrace{\sum_{k=0}^{|\beta|-1} \int_{\mathbb{R}^n} x^\beta \varphi_k(x)\mathrm{d}x}_{=:J_\beta \ (|\beta|>0)} + \sum_{k=|\beta|}^{\infty} \int_{\mathbb{R}^n} x^\beta \varphi_k(x)\mathrm{d}x = 0 \\
&\overset{(4.20)}{\iff} \quad J_\beta + (-1)^{|\beta|}(a_{|\beta|})^{|\beta|+n}\, c_\beta \underbrace{\int_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x}_{=:J_0} = 0 \\
&\iff \quad c_\beta = \frac{(-1)^{|\beta|+1} J_\beta}{a_{|\beta|}^{|\beta|+n} J_0}
\end{aligned}$$

where J_β depends on all a_m, b_m, and c_α with $m < |\beta|$ and $|\alpha| < |\beta|$. Hence the coefficients c_β will be determined by the equation

$$c_\beta = \frac{(-1)^{|\beta|+1} J_\beta}{a_{|\beta|}^{|\beta|+n} J_0}, \tag{4.25}$$

making sure that the moment of order β ($|\beta| > 0$) will be zero. Before we proceed, we estimate $\|\varphi - \varphi_0\|^p_{L^p(\mathbb{R}^n)}$ for $1 \le p < \infty$, assuming that all limits exist and the infinite sum commutes with the integral, leading to

$$\begin{aligned}\int\limits_{\mathbb{R}^n} |\varphi(x) - \varphi_0(x)|^p \mathrm{d}x &\le \sum_{m=1}^{\infty} \int\limits_{\mathbb{R}^n} |\varphi_m(x)|^p \mathrm{d}x \\ &\le \frac{\pi^{\frac{n}{2}} r_0^n}{\Gamma(n/2+1)} \sum_{m=1}^{\infty} a_m^n \|\varphi_m\|^p_{L^\infty(\mathbb{R}^n)},\end{aligned}$$

since the support of φ_m is contained in an n-dimensional ball of radius $a_m r_0$. Here and in the following equations, $\Gamma(n/2+1)$ denotes the Gamma function, evaluated at $\frac{n}{2}+1$. Similarly we estimate $\|\varphi - \varphi_0\|^p_{W^{k,p}(\mathbb{R}^n)}$, more precisely

$$\begin{aligned}\sum_{|\alpha|\le k} \int\limits_{\mathbb{R}^n} |\partial^\alpha \varphi(x) - \partial^\alpha \varphi_0(x)|^p \mathrm{d}x &\le \sum_{|\alpha|\le k} \sum_{m=1}^{\infty} \int\limits_{\mathbb{R}^n} |\partial^\alpha \varphi_m(x)|^p \mathrm{d}x \\ &\le \frac{\pi^{\frac{n}{2}} r_0^n}{\Gamma(\frac{n}{2}+1)} \sum_{m=1}^{\infty} \sum_{|\alpha|\le k} a_m^n \|\partial^\alpha \varphi_m\|^p_{L^\infty(\mathbb{R}^n)} \\ &\le \frac{\pi^{\frac{n}{2}} r_0^n}{\Gamma(\frac{n}{2}+1)} \sum_{m=1}^{\infty} \sum_{|\alpha|\le k} a_m^n \parallel a^{-|\alpha|} \sum_{|\beta|=m} c_\beta \partial^{\alpha+\beta} \varphi_0 \parallel^p_{L^\infty(\mathbb{R}^n)} \\ &\le \frac{\pi^{\frac{n}{2}} r_0^n}{\Gamma(\frac{n}{2}+1)} \sum_{m=1}^{\infty} \sum_{|\alpha|\le k} a_m^{n-|\alpha|p} \parallel a^{-m-n} \sum_{|\beta|=m} \left|\frac{J_\beta}{J_0}\right| \partial^{\alpha+\beta} \varphi_0 \parallel^p_{L^\infty(\mathbb{R}^n)} \\ &\overset{(4.25)}{\le} \frac{\pi^{\frac{n}{2}} r_0^n}{\Gamma(\frac{n}{2}+1)} \sum_{m=1}^{\infty} \sum_{|\alpha|\le k} a_m^{-|\alpha|p-mp-n(p-1)} \parallel \sum_{|\beta|=m} \left|\frac{J_\beta}{J_0}\right| \partial^{\alpha+\beta} \varphi_0 \parallel^p_{L^\infty(\mathbb{R}^n)} \\ &\le \frac{\pi^{\frac{n}{2}} r_0^n}{\Gamma(\frac{n}{2}+1)} \sum_{m=1}^{\infty} a_m^{-mp-n(p-1)} k n^k \max_{|\alpha|\le k} \parallel \sum_{|\beta|=m} \left|\frac{J_\beta}{J_0}\right| \partial^{\alpha+\beta} \varphi_0 \parallel^p_{L^\infty(\mathbb{R}^n)} .\end{aligned}$$

Hence we get $\|\varphi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} < \varepsilon$, if $a_m \ge N_m^{k,p,\varepsilon}$, where the sequence $(N_m^{k,p,\varepsilon})_{m\in\mathbb{N}_0}$ is defined by

$$N_m^{k,p,\varepsilon} := \left(\frac{2^m}{\varepsilon^p} \frac{\pi^{\frac{n}{2}} r_0^n k n^k}{\Gamma(\frac{n}{2}+1)} \max_{|\alpha|\le k} \parallel \sum_{|\beta|=m} \left|\frac{J_\beta}{J_0}\right| \partial^{\alpha+\beta} \varphi_0 \parallel^p_{L^\infty(\mathbb{R}^n)} \right)^{\frac{1}{mp+n(p-1)}} . \tag{4.26}$$

for $1 \le p < \infty$. The required growth $a_m \ge N_m^{k,p,\varepsilon}$ can be accomplished for all $m \ge 1$, since for $|\beta| = m$, J_β only depends on $(a_k)_{0\le k\le m-1}$. In addition we obtain

$$\begin{aligned}
\sum_{|\alpha|\le k} \|\partial^\alpha\varphi - \partial^\alpha\varphi_0\|_{L^\infty(\mathbb{R}^n)} &\le \sum_{|\alpha|\le k}\sum_{m=1}^{\infty} \|\partial^\alpha\varphi_m\|_{L^\infty(\mathbb{R}^n)} \\
&\le \sum_{m=1}^{\infty}\sum_{|\alpha|\le k}\sum_{|\beta|=m} a_m^{-|\alpha|}\|c_\beta\partial^{\alpha+\beta}\varphi_0\|_{L^\infty(\mathbb{R}^n)} \\
&\le \sum_{m=1}^{\infty}\sum_{|\alpha|\le k} a_m^{-|\alpha|-m-n}\sum_{|\beta|=m}\left|\frac{J_\beta}{J_0}\right| \|\partial^{\alpha+\beta}\varphi_0\|_{L^\infty(\mathbb{R}^n)} \\
&\le \sum_{m=1}^{\infty} a_m^{-m-n} kn^k \max_{|\alpha|\le k}\Big(\sum_{|\beta|=m}\left|\frac{J_\beta}{J_0}\right| \|\partial^{\alpha+\beta}\varphi_0\|_{L^\infty(\mathbb{R}^n)}\Big),
\end{aligned}$$

implying that $\|\varphi-\varphi_0\|_{W^{k,\infty}(\mathbb{R}^n)} < \varepsilon$, if $a_m \ge N_m^{k,\infty,\varepsilon}$, where

$$N_m^{k,\infty,\varepsilon} := \left(\frac{2^m kn^k}{\varepsilon}\max_{|\alpha|\le k}\Big(\sum_{|\beta|=m}\left|\frac{J_\beta}{J_0}\right| \|\partial^{\alpha+\beta}\varphi_0\|_{L^\infty(\mathbb{R}^n)}\Big)\right)^{\frac{1}{m+n}}. \tag{4.27}$$

Since φ shall be a Schwartz function, the supremum of the functions φ_m must decrease sufficiently fast as $m\to\infty$. Aiming at a corresponding condition on the parameters a_m and b_m, we estimate

$$\begin{aligned}
\sup_{x\in\mathbb{R}^n}\left|x^\alpha\partial^\beta\varphi(x)\right| &\le \sum_{m=0}^{\infty}\sup_{x\in\mathbb{R}^n}\left|x^\alpha\partial^\beta\varphi_m(x)\right| \\
&\overset{(4.19)}{\le} \sum_{m=0}^{\infty}\sum_{|\gamma|=m} c_\gamma a_m^{-|\beta|}\sup_{x\in\mathbb{R}^n}\left|x^\alpha(\partial^{\beta+\gamma}\varphi_0)\left(a_m^{-1}x - b_m\xi_0\right)\right| \\
&\overset{(4.25)}{\le} \sum_{m=0}^{\infty} a_m^{-|\beta|-m-n}\sum_{|\gamma|=m}\left|\frac{J_\gamma}{J_0}\right|\sup_{x\in\mathbb{R}^n}\left|x^\alpha(\partial^{\beta+\gamma}\varphi_0)\left(a_m^{-1}x - b_m\xi_0\right)\right| \\
&\le \sum_{m=0}^{\infty} a_m^{-|\beta|-m-n}(b_m|\xi_0| + a_m r_0)^{|\alpha|}\sum_{|\gamma|=m}\left|\frac{J_\gamma}{J_0}\right| \|\partial^{\beta+\gamma}\varphi_0\|_{L^\infty(\mathbb{R}^n)} \\
&\overset{(4.24)}{\le} \sum_{m=0}^{\infty} a_m^{-|\beta|-m-n}(a_m(|\xi_0| + 2r_0) + a_m r_0)^{|\alpha|}\sum_{|\gamma|=m}\left|\frac{J_\gamma}{J_0}\right| \|\partial^{\beta+\gamma}\varphi_0\|_{L^\infty(\mathbb{R}^n)} \\
&\le \sum_{m=0}^{\infty} a_m^{|\alpha|-|\beta|-m-n}(|\xi_0| + 3r_0)^{|\alpha|}\sum_{|\gamma|=m}\left|\frac{J_\gamma}{J_0}\right| \|\partial^{\beta+\gamma}\varphi_0\|_{L^\infty(\mathbb{R}^n)}
\end{aligned}$$

from which we conclude that in order to achieve $\varphi\in\mathcal{S}(\mathbb{R}^n)$, it would be sufficient to have

$$a_m^{|\alpha|-|\beta|-m-n}\sum_{|\gamma|=m}|J_\gamma|\,\|\partial^{\beta+\gamma}\varphi_0\|_{L^\infty(\mathbb{R}^n)} \le \frac{C_{\alpha,\beta}}{2^m} \quad \text{for} \quad m\ge m_0. \tag{4.28}$$

where m_0 may depend on α and β. We have to make sure that for all $\alpha,\beta\in\mathbb{N}_0^n$, there exists $m_0 \in \mathbb{N}_0$ such that (4.28) holds. Therefore we set

$$\lambda_m := 1 + \sup_{|\beta| \le 2m} \sum_{|\gamma|=m} |J_\gamma| \|\partial^{\beta+\gamma} \varphi_0\|_{L^\infty(\mathbb{R}^n)}, \tag{4.29}$$

noting that λ_m is uniquely determined, once the coefficients c_γ for $|\gamma| < m$ are fixed. If

$$a_m \ge 2^m \lambda_m, \tag{4.30}$$

then, given $\alpha, \beta \in \mathbb{N}_0^n$, condition (4.28) will be satisfied for all numbers $m \ge \max(|\alpha| - |\beta| - n, |\beta|/2)$, in particular for all $m \ge m_0 = \max(|\alpha|, |\beta|)$.

For convenience of the reader, we briefly summarise the first part of the proof. Attempting to construct a function $\varphi \in \mathcal{S}(\mathbb{R}^n)$ which satisfies the conditions $\mathrm{supp}(\varphi) \subseteq \Gamma$ and $\|\varphi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} < \varepsilon$, we formally defined

$$\varphi(x) = \sum_{m=0}^{\infty} \sum_{|\alpha|=m} c_\alpha \partial^\alpha \varphi_0 \left(\frac{x}{a_m} - b_m \xi_0 \right),$$

in (4.21) and subsequently derived several growth conditions on the sequence scaling parameters $(a_m)_{m \in \mathbb{N}_0}$, with $a_m r_0$ representing the radius of the ball containing $\mathrm{supp}(\varphi_m)$, see the figure below.

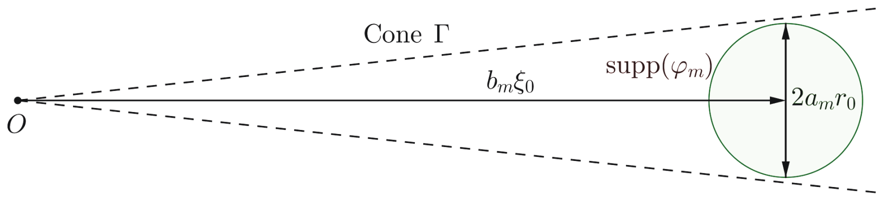

Figure 4.3: The translation along ξ_0 is determined by b_m, the dilation by a_m.

All of the conditions arising from (4.23), (4.26), (4.27) and (4.30) are satisfied, if we put

$$a_m := \max\left(a_{m-1}, b_{m-1}, N_m^{k,p,\varepsilon}, 2^m \lambda_m\right) \tag{4.31}$$

for $m \ge 1$ and $a_0 = 1$. Here, $a_m \ge \max(a_{m-1}, b_{m-1})$ was a technical condition necessary to control the distances $b_m |\xi_0|$ in terms of the radii $a_m r_0$, whereas the condition $a_m \ge N_m^{k,p,\varepsilon}$ ($1 \le p \le \infty$, see (4.26–4.27)) implies $\|\varphi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} < \varepsilon$, and finally, $a_m \ge 2^m \lambda_m$ (see (4.29)) settles the rapid decay property of φ, that is, $\varphi \in \mathcal{S}(\mathbb{R}^n)$. The other parameters were also defined recursively by

$$b_m = \max\left(a_m, b_{m-1} + \frac{r_0}{|\xi_0|} (a_{m-1} + a_m) \right)$$

for $m \ge 1$ ($b_0 = 0$) in (4.22) and by

$$c_\beta = \frac{(-1)^{|\beta|+1}}{a_{|\beta|}^{|\beta|+n} J_0} \sum_{k=0}^{|\beta|-1} \int_{\mathbb{R}^n} x^\beta \varphi_k(x) \mathrm{d}x$$

for $|\beta| \geq 1$ ($c_0 = 1$) in (4.25). So far, all calculations were performed under the assumption that the infinite sum in (4.21) converges in $L^1(\mathbb{R}^n)$ to $\varphi \in \mathcal{S}(\mathbb{R}^n)$. It remains to show that $(a_m)_{m\in\mathbb{N}_0}$, $(b_m)_{m\in\mathbb{N}_0}$, and $(c_\alpha)_{\alpha\in\mathbb{N}_0^n}$, satisfying (4.22), (4.25), and (4.31), can indeed be constructed and that the resulting function $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$ has all the desired properties. We show by induction that for all $m \in \mathbb{N}$,

$$\int_{\mathbb{R}^n} x^\alpha \sum_{k=0}^{m} \varphi_k(x)\mathrm{d}x = \delta_{|\alpha|,0}\, J_0 \quad \text{for} \quad 0 \leq |\alpha| < m, \tag{4.32}$$

where $\delta_{|\alpha|,0}$ represents the Kronecker delta. First, we construct φ_1 explicitly and consider $m = 1$ as the base case. Note that $a_0 = 1$, $b_0 = 0$ and $c_0 = 1$. We then have

$$\varphi_1(x) = c_{e_j} (\partial_j \varphi_0) \left(a_1^{-1}x - b_1\xi_0\right)$$

and put

$$a_1 := \max\left(a_0, b_0, N_1^{k,p,\varepsilon}, 2\lambda_1\right),$$

according to (4.31), where $N_1^{k,p,\varepsilon}$ and $2\lambda_1$ are already determined by φ_0. We set

$$b_1 = \max\left(a_1, b_0 + \frac{r_0}{|\xi_0|}\,(a_0 + a_1)\right) \tag{4.33}$$

as in (4.22) and compute the coefficients c_{e_j} by

$$c_{e_j} = \frac{J_{e_j}}{a_1^{1+n} J_0} = \frac{1}{a_1^{1+n} J_0} \int_{\mathbb{R}^n} x_j \varphi_0(x)\mathrm{d}x$$

as in (4.25). Then

$$\int_{\mathbb{R}^n} (\varphi_0(x) + \varphi_1(x))\mathrm{d}x \quad = \quad J_0 + \int_{\mathbb{R}^n} \varphi_1(x)\mathrm{d}x = J_0$$

by (4.20). Now, regarding (4.32) as the induction hypothesis, we construct φ_{m+1} by computing a_{m+1}, b_{m+1} and c_α for $|\alpha| = m+1$, which are all determined by their direct predecessors and by all previously constructed wave packets $(\varphi_k)_{k=0}^m$. We then consider

$$\begin{aligned}
\int_{\mathbb{R}^n} x^\alpha \sum_{k=0}^{m+1} \varphi_k(x)\mathrm{d}x \quad &= \quad \int_{\mathbb{R}^n} x^\alpha \sum_{k=0}^{m} \varphi_k(x)\mathrm{d}x + \int_{\mathbb{R}^n} x^\alpha \varphi_{m+1}(x)\mathrm{d}x \\
&\overset{(4.32)}{=} \quad \delta_{|\alpha|,0} J_0 + \int_{\mathbb{R}^n} x^\alpha \varphi_{m+1}(x)\mathrm{d}x \\
&\overset{(4.20)}{=} \quad \delta_{|\alpha|,0} J_0 \quad \text{for} \quad |\alpha| \leq m.
\end{aligned}$$

This completes the induction and also proves that φ with $\varphi(x) = \sum_{m=0}^\infty \varphi_m(x)$ is indeed rapidly decreasing, as well as all its derivatives, since we obtain a sequence $(\varphi_m)_{m\in\mathbb{N}_0}$ satisfying (4.28).

Finally, we show that $(x \mapsto x^\alpha \sum_{k=0}^m \varphi_k(x))_{m\in\mathbb{N}_0}$ converges in $L^1(\mathbb{R}^n)$ as $m \to \infty$ for every $\alpha \in \mathbb{N}_0^n$. Therefore we define $\psi_\alpha : \mathbb{R}^n \to [0,\infty[$,

$$\psi_\alpha(x) := \begin{cases} (b_m|\xi_0| + a_m r_0)^{|\alpha|}\|\varphi_m\|_{L^\infty(\mathbb{R}^n)} & \text{for} \quad x \in b_m\xi_0 + B_{a_m r_0} \\ 0 & \text{else,} \end{cases}$$

implying that $|x^\alpha \varphi_m(x)| \le \psi_\alpha(x)$ for all $x \in \mathbb{R}^n$. We have

$$\begin{aligned}
\|\psi_\alpha\|_{L^1(\mathbb{R}^n)} &\le \sum_{m=0}^\infty (b_m|\xi_0| + a_m r_0)^{|\alpha|} \frac{\pi^{\frac{n}{2}} r_0^n a_m^n}{\Gamma(\frac{n}{2}+1)} \sum_{|\beta|=m} \|c_\beta \partial^\beta \varphi_0\|_{L^\infty(\mathbb{R}^n)} \\
&\overset{(4.24)}{\le} \frac{4^{|\alpha|}(|\xi_0| + r_0)^{|\alpha|}\pi^{\frac{n}{2}} r_0^n}{|J_0|\Gamma(\frac{n}{2}+1)} \sum_{m=0}^\infty a_m^{|\alpha|-m} \sum_{|\beta|=m} \|J_\beta \partial^\beta \varphi_0\|_{L^\infty(\mathbb{R}^n)} \\
&\overset{(4.29)}{\le} C \sum_{m=0}^\infty a_m^{|\alpha|-m}\lambda_m \\
&\le C \sum_{m=0}^{|\alpha|} a_m^{|\alpha|-m}\lambda_m + C \sum_{m=|\alpha|+1}^\infty a_m^{-1}\lambda_m \\
&\overset{(4.30)}{\le} C \sum_{m=0}^{|\alpha|} a_m^{|\alpha|-m}\lambda_m + C \sum_{m=|\alpha|+1}^\infty \frac{1}{2^m}
\end{aligned}$$

and thus $\|\psi_\alpha\|_{L^1(\mathbb{R}^n)}$ is finite. Since $x \mapsto x^\alpha \sum_{k=0}^m \varphi_k(x)$ converges pointwise as $m \to \infty$ and is dominated by $\psi_\alpha \in L^1(\mathbb{R}^n)$, taking the limit $m \to \infty$ in (4.32) commutes with integration and

$$\int_{\mathbb{R}^n} x^\alpha \sum_{k=0}^\infty \varphi_k(x)\mathrm{d}x = \sum_{k=0}^\infty \int_{\mathbb{R}^n} x^\alpha \varphi_k(x)\mathrm{d}x = \delta_{|\alpha|,0} J_0 \quad \text{for} \quad \alpha \in \mathbb{N}_0^n.$$

In particular, the formal estimates showing that $\|\varphi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} < \varepsilon$ are valid. □

While in Theorem 4.12, the conic set Γ ist determined by the support of φ_0, we may also take any cone Γ, not related to $\mathrm{supp}(\varphi_0)$, as this does not affect the construction in any way. More precisely, we are free to choose an arbitrary cone Γ (which may be as narrow as we wish) and obtain $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with $\mathrm{supp}(\varphi) \subseteq \mathrm{supp}(\varphi_0) \cup \Gamma$ and satisfying conditions (ii) and (iii). Here it is not required that $0 \notin \mathrm{supp}(\varphi_0)$.

Corollary 4.13 (Mollifiers with vanishing moments and 'minimal' support)**.** *Let Γ be a cone in $\mathbb{R}^n$, $R > 0$ and $B_R = \{x \in \mathbb{R}^n : |x| < R\}$). Then there exists $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with $\mathrm{supp}(\varphi) \subseteq \Gamma \setminus B_R$, $\int_{\mathbb{R}^n} \varphi(x)\mathrm{d}x = 1$ and $\int_{\mathbb{R}^n} x^\alpha\varphi(x)\mathrm{d}x = 0$ for all $\alpha \in \mathbb{N}_0^n$, $|\alpha| > 0$.*

Proof. The statement follows immediately from Theorem 4.12 by choosing any $\varphi_0 \in \mathcal{D}(\mathbb{R}^n)$ with $\mathrm{supp}(\varphi_0) \subseteq \Gamma \setminus B_R$ and $\int_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x = 1$. □

The construction in the proof of Theorem 4.12 can easily be modified to provide a solution of the Stieltjes moment problem in arbitrary dimensions, formulated on a cone Γ in $\mathbb{R}^n$. To this end we take any $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with $\operatorname{supp}(\varphi) \subseteq \Gamma$, $\int_{\mathbb{R}^n} \varphi(x)\mathrm{d}x \neq 0$ and set $\psi(x) := \varphi(x) + \sum_{m=0}^{\infty} \varphi_m(x)$ with φ_m defined as in (4.18–4.19), in particular we have $\varphi_0(x) = c_0\varphi(x/a_0 - b_0\xi_0)$. Note that compared to definition (4.21), the sum representing the function ψ additionally includes the original function φ itself. Whereas we had put $a_0 = c_0 = 1$ and $b_0 = 0$ in Lemma 4.11, we now determine $a_0 > 0$, $b_0 > 0$, and $c_0 \in \mathbb{R}$ by demanding that the conditions $\int_{\mathbb{R}^n}(\varphi(x)+\varphi_0(x))\mathrm{d}x = 0$ and $\operatorname{supp}(\varphi_0) \subseteq \Gamma$ are satisfied. We then proceed as in the proof of Theorem 4.12, but replace J_α by $J_\alpha - \nu_\alpha$ in the recurrence relation (4.25). By this we obtain $\psi \in \mathcal{S}(\mathbb{R}^n)$ with $\operatorname{supp}(\psi) \subseteq \Gamma$ and $\int_{\mathbb{R}^n} x^\alpha \psi(x)\mathrm{d}x = \nu_\alpha$ for all $\alpha \in \mathbb{N}_0^n$ and thus a solution to a variant of the 'vector moment problem for functions of several variables' considered in [Est98] and in [GLS01]. Moreover, we can make $\|\psi - \varphi\|_{L^p(\mathbb{R}^n)}$ as small as we wish for any fixed $p > 1$. This also implies that

$$\mathcal{S}_0(\Gamma) := \left\{ \psi \in \mathcal{S}(\mathbb{R}^n) \,|\, \forall \alpha \in \mathbb{N}_0^n : \int\limits_{\mathbb{R}^n} x^\alpha \psi(x)\mathrm{d}x = 0 \quad \text{and} \quad \operatorname{supp}(\psi) \subseteq \Gamma \right\}$$

is a dense subspace of $L^p(\Gamma)$ for $1 < p < \infty$, since $\mathcal{D}(\Gamma)$ is dense (cf. [Bha12, Theorem 6.8.3]). As already discussed for the one-dimensional case (see Remark 4.8), the approximation fails for $p = 1$, since the condition $\|\psi - \varphi\|_{L^1(\mathbb{R}^n)} < \varepsilon$ will be in contradiction to $\int_{\mathbb{R}^n} \psi(x)\mathrm{d}x = 0$ for small ε. Note that the statements in Theorem 4.12 remain true, if we replace the monomials $\{x^\alpha\}_{\alpha\in\mathbb{N}_0^n}$ in the moment condition $\int_{\mathbb{R}^n} x^\alpha \varphi(x)\mathrm{d}x = \nu_\alpha$ by polynomials $\{p_\alpha\}_{\alpha\in\mathbb{N}_0^n}$, where $p_\alpha(x) = \sum_{|\beta|\leq|\alpha|} \widetilde{c}_{\alpha,\beta} x^\beta$, as this modification does not compromise the algorithm employed in the proof.

Corollary 4.14 (The multidimensional Stieltjes moment problem)**.** *Let real numbers* $(\nu_\alpha)_{\alpha\in\mathbb{N}_0^n}$*, a cone* $\Gamma \subseteq \mathbb{R}^n$*, and a function* $\varphi \in \mathcal{D}(\mathbb{R}^n)$ *with* $\operatorname{supp}(\varphi) \subseteq \Gamma$ *and* $\int_{\mathbb{R}^n} \varphi(x)\,\mathrm{d}x \neq 0$ *be given. Let* $\varepsilon > 0$ *and* $1 < p \leq \infty$*. Then there exists a function* $\psi \in \mathcal{S}(\mathbb{R})$ *such that*

(i) $\operatorname{supp}(\psi) \subseteq \Gamma$,

(ii) $\displaystyle\int\limits_{\mathbb{R}^n} x^\alpha \psi(x)\,\mathrm{d}x = \nu_\alpha \;\; \forall \alpha \in \mathbb{N}_0^n$,

(iii) $\|\psi - \varphi\|_{L^p(\mathbb{R}^n)} \leq \varepsilon$.

Proof. The proof follows along the lines of the proof of Theorem 4.12, with minor adaptions as described above. □

In the embedding of distributions into the special Colombeau algebra, technical difficulties arise from the fact that a mollifier $\rho \in \mathcal{S}(\mathbb{R}^n)$ with vanishing moments cannot have compact support and therefore $\rho_\varepsilon(\cdot) = \frac{1}{\varepsilon^n}\rho\left(\frac{\cdot}{\varepsilon}\right)$ cannot be convolved with every element of $\mathcal{D}'(\mathbb{R}^n)$. To overcome this restriction, nonstandard or generalised mollifiers $\rho \in \mathcal{G}(\mathbb{R}^n)$ with compact support have been introduced in [OT98] and [OV08]. Another approach is to use smooth cut-off functions $\chi_\varepsilon \in \mathcal{D}(\mathbb{R}^n)$ converging to 1 at a certain rate and create a delta net $\left(\frac{1}{\varepsilon^n}\rho\left(\frac{\cdot}{\varepsilon}\right)\chi_\varepsilon\right)_\varepsilon$ whose moments differ from those of ρ_ε only by an element of $\mathcal{N}(\mathbb{R}^n)$, as discussed in [NPS98] and in [Del05]. The purpose of the following corollary is to put our results in relation to these constructions and to offer a different perspective on Theorem 4.12.

Corollary 4.15 (Approximating sequences with vanishing moments)**.** *Let* $\Gamma \subseteq \mathbb{R}^n$ *be an open cone,* $\varphi_0 \in \mathcal{D}(\mathbb{R}^n)$*,* $\operatorname{supp}(\varphi_0) \subseteq \Gamma$*,* $k \in \mathbb{N}_0$*,* $1 \leq p \leq \infty$*, and* $d > 0$*. Then there exists a sequence* $(\psi_m)_{m\in\mathbb{N}_0}$ *in* $\mathcal{S}(\mathbb{R}^n)$*,* $\operatorname{supp}(\varphi_m) \subseteq \Gamma$*, and a function* $\psi \in \mathcal{S}(\mathbb{R}^n)$ *with* $\operatorname{supp}(\psi) \subseteq \Gamma$*, such that*

(i) $\lim\limits_{m\to\infty} \|\psi_m - \psi\|_{W^{k,p}(\mathbb{R}^n)} = 0$,

(ii) $\int\limits_{\mathbb{R}^n} \psi(x)\mathrm{d}x = \int\limits_{\mathbb{R}^n} \psi_m(x)\mathrm{d}x = 1$ *for all* $m \in \mathbb{N}_0$,

(iii) $\int\limits_{\mathbb{R}^n} x^\alpha \psi(x)\mathrm{d}x = 0$ *for* $|\alpha| > 0$ *and* $\int\limits_{\mathbb{R}^n} x^\alpha \psi_m(x)\mathrm{d}x = 0$ *for* $0 < |\alpha| < m$,

(iv) $\|\psi_m - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} \leq \|\psi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} < d$ *for all* $m \in \mathbb{N}_0$.

Proof. Assertions (i)-(iii) follow immediately from Theorem 4.12 by defining ψ_m as the m-th partial sum in (4.21). The inequalities in (iv) are also clear, since

$$\|\psi_m - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)} \leq \|\psi_m - \psi\|_{W^{k,p}(\mathbb{R}^n)} + \|\psi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)}$$

and $\|\psi - \varphi_0\|_{W^{k,p}(\mathbb{R}^n)}$ can be made as small as we wish. □

Notice that if we assume that $\varphi_0(x) \geq 0$, we automatically get $\psi(x) + d \geq 0$ for all $x \in \mathbb{R}^n$, that is, we also obtain 'almost non-negative' mollifiers. The following Proposition 4.16 demonstrates how Corollary 4.15 can be employed to construct a generalised function $\rho \in \mathcal{G}(\mathbb{R}^n)$, sharing the properties of generalised mollifiers with compact support as established in [OV08, Propositions 4.1 and 4.2]. An alternative proof of this result can be found in [SV09, Lemma A.1].

Proposition 4.16 (Generalised mollifiers with support conditions)**.** *Let* $\Gamma \subseteq \mathbb{R}^n$ *be an open cone,* $B_1 = \{x \in \mathbb{R}^n : |x| < 1\}$*, and* $d > 0$*. Then there exists an element* $\rho = [(\rho_\varepsilon)_\varepsilon] \in \mathcal{G}(\mathbb{R}^n)$ *such that*

(i) $\int\limits_{\mathbb{R}^n} \rho(x)\mathrm{d}x = 1$,

(ii) $\int\limits_{\mathbb{R}^n} x^\alpha \rho(x)\mathrm{d}x = 0$ *for all* $\alpha \in \mathbb{N}_0^n$, $|\alpha| > 0$,

(iii) $\operatorname{supp}(\rho) \subseteq \overline{B}_1 \cap \overline{\Gamma}$*, more precisely*
$\operatorname{supp}(\rho_\varepsilon) \subseteq B_1 \cap \Gamma$ *for all* $\varepsilon > 0$,

(iv) $\int\limits_{\mathbb{R}^n} |\rho(x)|\mathrm{d}x \leq 1 + d$.

Proof. First we choose any non-negative function $\varphi_0 \in \mathcal{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x = 1$ and $\operatorname{supp}(\varphi_0) \subseteq B_1 \cap \Gamma$. We construct a sequence $(\psi_m)_m$ converging to $\psi \in \mathcal{S}(\mathbb{R}^n)$ as in Corollary 4.15 such that $\|\psi - \varphi_0\|_{L^1(\mathbb{R}^n)} < d$. We may then define

$$N(\varepsilon) := \max\left\{m \in \mathbb{N}_0 : \operatorname{supp}\left(\psi_m\left(\frac{\cdot}{\varepsilon}\right)\right) \subseteq B_1\right\}$$

and set $\rho_\varepsilon(x) := \frac{1}{\varepsilon^n}\psi_{N(\varepsilon)}\big(\frac{x}{\varepsilon}\big)$, representing a finite sum of functions with support contained in $B_1 \cap \Gamma$. Since $\|\partial^\alpha \rho_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \le C_\alpha \varepsilon^{-|\alpha|-n}$, we obtain a generalised function $\rho = [(\rho_\varepsilon)_\varepsilon] \in \mathcal{G}(\mathbb{R}^n)$. Condition (iii) is automatically satisfied, implying that the integrals in (i) and (ii) can be carried out componentwise on representatives. Since $\int_{\mathbb{R}^n} \psi_{N(\varepsilon)}\big(\frac{x}{\varepsilon}\big)\,\mathrm{d}x = \int_{\mathbb{R}^n} \varphi_0(x)\mathrm{d}x = 1$ for all $\varepsilon > 0$ by Corollary 4.15 (ii), we immediately get (i). Moreover we have $\int_{\mathbb{R}^n} x^\alpha \psi_{N(\varepsilon)}\big(\frac{x}{\varepsilon}\big)\,\mathrm{d}x = 0$ for all $\alpha \in \mathbb{N}_0^n$, $0 < |\alpha| < N(\varepsilon)$, as a consequence of Corollary 4.15 (iii). This proves assertion (ii), since $N(\varepsilon) \to \infty$ as $\varepsilon \to 0$. Finally, we observe that

$$\|\rho_\varepsilon\|_{L^1(\mathbb{R}^n)} = \| \varepsilon^{-n}\psi_{N(\varepsilon)}(\cdot/\varepsilon) \|_{L^1(\mathbb{R}^n)} = \|\psi_{N(\varepsilon)}\|_{L^1(\mathbb{R}^n)} \le \|\psi\|_{L^1(\mathbb{R}^n)} < 1 + d$$

since $\|\psi\|_{L^1(\mathbb{R}^n)} \le \|\varphi_0\|_{L^1(\mathbb{R}^n)} + \|\psi - \varphi_0\|_{L^1(\mathbb{R}^n)} < 1 + d$. □

Remark 4.17 (Dependence on ε and rescaling)**.** The net $(\psi_{N(\varepsilon)})_\varepsilon$ employed in the proof of Proposition 4.16 does not incorporate any ε-dependent scaling or dilation of the underlying sequence $(\psi_m)_m$. The role of ε is merely that of a cut-off parameter, determining how many of the previously constructed wave packets are included in $\psi_{N(\varepsilon)}$. Using a more general scaling net $\sigma_\varepsilon \to 0$, $\sigma_\varepsilon^{-1} = O(\varepsilon^{-q})$ as $\varepsilon \to 0$ for some $q \in \mathbb{N}$, to eventually define $\rho_\varepsilon(x) := \sigma_\varepsilon^{-n}\psi_{N(\sigma_\varepsilon)}\big(\sigma_\varepsilon^{-1}x\big)$, allows us to control the rate of convergence of $(\rho_\varepsilon)_\varepsilon$, for instance to satisfy additional asymptotic conditions such as L^∞-log-type estimates. In particular, taking $\sigma_\varepsilon = 1/\log(1/\varepsilon)$, the mollifier $\rho = [(\rho_\varepsilon)_\varepsilon]$ will be a regular generalised function, that is $\rho \in \mathcal{G}^\infty(\mathbb{R}^n)$, as in [OV08, Propositions 4.1 and 4.2]. Further adaptation is possible by choosing a specific cone $\overline{\Gamma}$ containing the support of ρ.

The generalised function $\rho \in \mathcal{G}(\mathbb{R}^n)$ obtained in Proposition 4.16 is associated with the delta distribution. In fact, the constructed family $(\rho_\varepsilon)_\varepsilon$ is a strict delta net satisfying the conditions used in [SV09, Lemma 4.3] to specify admissible mollifiers for regularisations of Geroch-Traschen class metrics in order to end up with a generalised metric. An application of mollifiers whose L^1-norm can be made arbitrarily close to 1 can also be found in [Hör15, Section 3].

The explicit construction of Schwartz functions with prescribed moments presented in this chapter has some advantages compared to less explicit solutions of the Stieltjes moment problem for rapidly decreasing smooth functions ([Dur89, DE94]). Moreover, the underlying method may also be applied to generalisations and other variants of the Stieltjes moment problem discussed in [Est98, GLS01, LS08]. Our construction is flexible in the sense that there is a lot of freedom in the choice of the basic function as well as the positions of the wave packets composed of higher-order derivatives of the basic function. Mollifiers satisfying additional support and stability conditions as in [Del05, OV08, SV09] are easily obtained by exploiting this freedom. The need for tailor-made mollifiers with special properties may arise in future applications of Colombeau theory.

References

[AB59] Y. Aharonov and D. Bohm. Significance of Electromagnetic Potentials in the Quantum Theory. *Physical Review*, 115:485–491, August 1959.

[AF03] R.A. Adams and J.J.F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[AFJ09] J. Aragona, R. Fernandez, and S.O. Juriaans. Natural topologies on Colombeau algebras. *Topol. Methods Nonlinear Anal.*, 34(1):161–180, 2009.

[AJ01] J. Aragona and S.O. Juriaans. Some structural properties of the topological ring of Colombeau's generalized numbers. *Comm. in Algebra*, 29:2201–2230, 2001.

[AJOS08] J. Aragona, S.O. Juriaans, O.R.B. Oliveira, and D. Scarpalezos. Algebraic and geometric theory of the topological ring of Colombeau generalized functions. *Proc. Edinb. Math. Soc. (2)*, 51(3):545–564, 2008.

[AS71] P.C. Aichelburg and R.U. Sexl. On the gravitational field of a massless particle. *General Relativity and Gravitation*, 2(4):303–312, 1971.

[Bal97] H. Balasin. Geodesics for impulsive gravitational waves and the multiplication of distributions. *Classical Quantum Gravity*, 14(2):455–462, 1997.

[BD64] J.D. Bjorken and S.D. Drell. *Relativistic quantum mechanics*. McGraw-Hill Book Co., New York-Toronto-London, 1964.

[BGS07] S. Benzoni-Gavage and D. Serre. *Multidimensional hyperbolic partial differential equations*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007.

[Bha12] P.K. Bhattacharyya. *Distributions*. De Gruyter Textbook. Walter de Gruyter & Co., Berlin, 2012.

[BK12] A. Burtscher and M. Kunzinger. Algebras of generalized functions with smooth parameter dependence. *Proc. Edinb. Math. Soc. (2)*, 55(1):105–124, 2012.

[Cau27] A.-L. Cauchy. *Théorie de la propagation des ondes a la surface d'un fluide pesant d'une profondeur indéfinie. Œuvres complètes. Series 1. Volume 1.* Cambridge Library Collection. Cambridge University Press, Cambridge, 2009 (first published 1827).

[CL88] J.-F. Colombeau and A.Y. LeRoux. Multiplications of distributions in elasticity and hydrodynamics. *J. Math. Phys.*, 29(2):315–319, 1988.

[CO90] J.-F. Colombeau and M. Oberguggenberger. On a hyperbolic system with a compatible quadratic term: generalized solutions, delta waves, and multiplication of distributions. *Comm. Partial Differential Equations*, 15(7):905–938, 1990.

[Col83] J.-F. Colombeau. Une multiplication générale des distributions. *C. R. Acad. Sci. Paris Sér. I Math.*, 296(8):357–360, 1983.

[Col84] J.-F. Colombeau. *New generalized functions and multiplication of distributions*, volume 84 of *North-Holland Mathematics Studies.* North-Holland Publishing Co., Amsterdam, 1984.

[Col85] J.-F. Colombeau. *Elementary introduction to new generalized functions*, volume 113 of *North-Holland Mathematics Studies.* North-Holland Publishing Co., Amsterdam, 1985.

[Col92] J.-F. Colombeau. *Multiplication of distributions*, volume 1532 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1992.

[CVW96] C.J.S. Clarke, J.A. Vickers, and J.P. Wilson. Generalized functions and distributional curvature of cosmic strings. *Classical Quantum Gravity*, 13(9):2485–2498, 1996.

[DE94] A.L. Durán and R. Estrada. Strong moment problems for rapidly decreasing smooth functions. *Proc. Amer. Math. Soc.*, 120(2):529–534, 1994.

[Deg11] H. Deguchi. A linear first-order hyperbolic equation with a discontinuous coefficient: distributional shadows and propagation of singularities. *Electron. J. Differential Equations*, 2011.

[Del05] A. Delcroix. Remarks on the embedding of spaces of distributions into spaces of Colombeau generalized functions. *Novi Sad J. Math.*, 35(2):27–40, 2005.

[dHHO08] M. de Hoop, G. Hörmann, and M. Oberguggenberger. Evolution systems for paraxial wave equations of Schrödinger-type with non-smooth coefficients. *J. Differential Equations*, 245(6):1413–1432, 2008.

[DHO13] H. Deguchi, G. Hörmann, and M. Oberguggenberger. The wave equation with a discontinuous coefficient depending on time only: generalized solutions and propagation of singularities. In *Pseudo-differential operators, generalized functions and asymptotics*, volume 231 of *Oper. Theory Adv. Appl.*, pages 323–339. Birkhäuser/Springer Basel AG, Basel, 2013.

[Dir27] P.A.M. Dirac. The physical interpretation of the quantum dynamics. *Proceedings of the Royal Society of London. Series A*, 113(765):621–641, 1927.

[Dir30] P.A.M. Dirac. *The Principles of Quantum Mechanics (International Series of Monographs on Physics).* Oxford University Press, USA, 1958 (1st edition 1930).

[Dir78] P.A.M. Dirac. The mathematical foundation of quantum theory. In A.R. Marlow, editor, *Mathematical Foundations of Quantum Theory*, pages 1–8. Academic Press, 1978.

[DL92] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 5.* Springer-Verlag, Berlin, 1992.

[DO16] H. Deguchi and M. Oberguggenberger. Propagation of singularities for generalized solutions to wave equations with discontinuous coefficients. *SIAM J. Math. Anal.*, 48(1):397–442, 2016.

[Dur89] A.J. Durán. The Stieltjes moments problem for rapidly decreasing functions. *Proc. Amer. Math. Soc.*, 107(3):731–741, 1989.

[EI38] A. Einstein and L. Infeld. *The Evolution of Physics.* Camb.Lib.of Mod.Science. Scientific Book Club, 1938.

[ES49] W. Ehrenberg and R.E. Siday. The Refractive Index in Electron Optics and the Principles of Dynamics. *Proceedings of the Physical Society B*, 62:8–21, January 1949.

[Est98] R. Estrada. Vector moment problems for rapidly decreasing smooth functions of several variables. *Proc. Amer. Math. Soc.*, 126(3):761–768, 1998.

[Eva85] J.-C. Evard. On matrix functions which commute with their derivative. *Linear Algebra Appl.*, 68:145–178, 1985.

[Eva98] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[Fol99] G.B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, second edition, 1999.

[Fou09] J.-B.J. Fourier. *Théorie analytique de la chaleur.* Cambridge Library Collection. Cambridge University Press, Cambridge, 2009. Reprint of the 1822 original, Previously published by Éditions Jacques Gabay, Paris, 1988.

[Fri75] F.G. Friedlander. *The wave equation on a curved space-time.* Cambridge Monographs on Mathematical Physics, No. 2. Cambridge University Press, Cambridge, 1975.

[Gar05a] C. Garetto. Topological structures in Colombeau algebras: investigation of the duals of $\mathcal{G}_c(\Omega)$, $\mathcal{G}(\Omega)$ and $\mathcal{G}_{\mathcal{S}}(\mathbb{R}^n)$. *Monatsh. Math.*, 146(3):203–226, 2005.

[Gar05b] C. Garetto. Topological structures in Colombeau algebras: topological $\widetilde{\mathbb{C}}$-modules and duality theory. *Acta Appl. Math.*, 88(1):81–123, 2005.

[Gar08] C. Garetto. Fundamental solutions in the Colombeau framework: applications to solvability and regularity theory. *Acta Appl. Math.*, 102(2-3):281–318, 2008.

[Gar15] C. Garetto. On hyperbolic equations and systems with non-regular time dependent coefficients. *J. Differential Equations*, 259(11):5846–5874, 2015.

[GH05] C. Garetto and G. Hörmann. Microlocal analysis of generalized functions: pseudodifferential techniques and propagation of singularities. *Proc. Edinb. Math. Soc. (2)*, 48(3):603–629, 2005.

[GKOS01] M. Grosser, M. Kunzinger, M. Oberguggenberger, and R. Steinbauer. *Geometric theory of generalized functions with applications to general relativity*, volume 537 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.

[GLS01] F. Galindo, F. López, and J. Sanz. A generalized moment problem for rapidly decreasing smooth vector functions of several variables. *J. Math. Anal. Appl.*, 263(2):655–665, 2001.

[GMS09] J.D.E. Grant, E. Mayerhofer, and R. Steinbauer. The wave equation on singular space-times. *Comm. Math. Phys.*, 285(2):399–420, 2009.

[GO14] C. Garetto and M. Oberguggenberger. Generalized Fourier integral operator methods for hyperbolic equations with singularities. *Proc. Edinb. Math. Soc. (2)*, 57(2):423–463, 2014.

[GO15] C. Garetto and M. Oberguggenberger. Symmetrisers and generalised solutions for strictly hyperbolic systems with singular coefficients. *Math. Nachr.*, 288(2-3):185–205, 2015.

[GOT00] H. Grosse, M. Oberguggenberger, and I.T. Todorov. Generalized functions for quantum fields obeying quadratic exchange relations. *Tr. Mat. Inst. Steklova*, 228(Probl. Sovrem. Mat. Fiz.):90–100, 2000.

[GR15] C. Garetto and M. Ruzhansky. Hyperbolic second order equations with non-regular time dependent coefficients. *Arch. Ration. Mech. Anal.*, 217(1):113–154, 2015.

[Gra94] T. Gramchev. Classical solutions to singular hyperbolic systems modelling acoustic wave propagation. In *Qualitative aspects and applications of nonlinear evolution equations (Trieste, 1993)*, pages 143–148. World Sci. Publ., River Edge, NJ, 1994.

[GT87] R. Geroch and J. Traschen. Strings and other distributional sources in general relativity. *Phys. Rev. D (3)*, 36(4):1017–1031, 1987.

[Ham20] H. Hamburger. Über eine Erweiterung des Stieltjesschen Momentenproblems. *Math. Ann.*, 81(2-4):235–319, 1920.

[Han11] C. Hanel. Wave-type equations of low regularity. *Appl. Anal.*, 90(11):1691–1705, 2011.

[HdH01] G. Hörmann and M.V. de Hoop. Microlocal analysis and global solutions of some hyperbolic equations with discontinuous coefficients. *Acta Appl. Math.*, 67(2):173–224, 2001.

[HE73] S.W. Hawking and G.F.R. Ellis. *The large scale structure of space-time.* Cambridge University Press, London-New York, 1973.

[HH08] S. Haller and G. Hörmann. Comparison of some solution concepts for linear first-order hyperbolic differential equations with non-smooth coefficients. *Publ. Inst. Math. (Beograd) (N.S.)*, 84(98):123–157, 2008.

[HHSS13] C. Hanel, G. Hörmann, C. Spreitzer, and R. Steinbauer. Wave equations and symmetric first-order systems in case of low regularity. In *Pseudo-differential operators, generalized functions and asymptotics*, volume 231 of *Oper. Theory Adv. Appl.*, pages 283–296. Birkhäuser/Springer Basel AG, Basel, 2013.

[HK01] G. Hörmann and M. Kunzinger. Microlocal properties of basic operations in Colombeau algebras. *J. Math. Anal. Appl.*, 261(1):254–270, 2001.

[HKK14] G. Hörmann, S. Konjik, and M. Kunzinger. Symplectic modules over Colombeau-generalized numbers. *Comm. Algebra*, 42(8):3558–3577, 2014.

[HKS12] G. Hörmann, M. Kunzinger, and R. Steinbauer. Wave equations on non-smooth space-times. In *Evolution equations of hyperbolic and Schrödinger type*, volume 301 of *Progr. Math.*, pages 163–186. Birkhäuser/Springer Basel AG, Basel, 2012.

[HO04] G. Hörmann and M. Oberguggenberger. Elliptic regularity and solvability for partial differential equations with Colombeau coefficients. *Electron. J. Differential Equations*, (14):1–30, 2004.

[HOP06] G. Hörmann, M. Oberguggenberger, and S. Pilipović. Microlocal hypoellipticity of linear partial differential operators with generalized functions as coefficients. *Trans. Amer. Math. Soc.*, 358(8):3363–3383, 2006.

[Hör99] G. Hörmann. Integration and microlocal analysis in Colombeau algebras of generalized functions. *J. Math. Anal. Appl.*, 239(2):332–348, 1999.

[Hör03] L. Hörmander. *The analysis of linear partial differential operators. I.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[Hör04] G. Hörmann. First-order hyperbolic pseudodifferential equations with generalized symbols. *J. Math. Anal. Appl.*, 293(1):40–56, 2004.

[Hör07] L. Hörmander. *The analysis of linear partial differential operators. III.* Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.

[Hör11] G. Hörmann. The Cauchy problem for Schrödinger-type partial differential operators with generalized functions in the principal part and as data. *Monatsh. Math.*, 163(4):445–460, 2011.

[Hör15] G. Hörmann. Conical spacetimes and global hyperbolicity. *Novi Sad J. Math.*, 45(1):215–229, 2015.

[HS12] G. Hörmann and C. Spreitzer. Symmetric hyperbolic systems in algebras of generalized functions and distributional limits. *J. Math. Anal. Appl.*, 388(2):1166–1179, 2012.

[Kg81] H. Kumano-go. *Pseudodifferential operators.* MIT Press, Cambridge, Mass.-London, 1981. Translated from the Japanese by the author, Rémi Vaillancourt and Michihiro Nagase.

[KH01] I. Kmit and G. Hörmann. Semilinear hyperbolic systems with singular non-local boundary conditions: reflection of singularities and delta waves. *Z. Anal. Anwendungen*, 20(3):637–659, 2001.

[Kmi06] I. Kmit. Generalized solutions to hyperbolic systems with non-linear conditions and strongly singular data. *Integral Transforms Spec. Funct.*, 17(2-3):177–183, 2006.

[Kmi07] I. Kmit. A distributional solution to a hyperbolic problem arising in population dynamics. *Electron. J. Differential Equations*, pages No. 132, 23, 2007.

[KS02a] M. Kunzinger and R. Steinbauer. Foundations of a nonlinear distributional geometry. *Acta Appl. Math.*, 71(2):179–206, 2002.

[KS02b] M. Kunzinger and R. Steinbauer. Generalized pseudo-Riemannian geometry. *Trans. Amer. Math. Soc.*, 354(10):4179–4199, 2002.

[KSV05] M. Kunzinger, R. Steinbauer, and J.A. Vickers. Generalised connections and curvature. *Math. Proc. Cambridge Philos. Soc.*, 139(3):497–521, 2005.

[KT13] M.G. Katz and D. Tall. A Cauchy-Dirac delta function. *Found. Sci.*, 18(1):107–123, 2013.

[Lan02] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, third edition, 2002.

[Lau89] D. Laugwitz. Definite values of infinite sums: aspects of the foundations of infinitesimal analysis around 1820. *Arch. Hist. Exact Sci.*, 39(3):195–245, 1989.

[Lau92] D. Laugwitz. Early delta functions and the use of infinitesimals in research. *Rev. Histoire Sci.*, 45(1):115–128, 1992. Études sur Cauchy (1789–1857).

[LM07] P.G. LeFloch and C. Mardare. Definition and stability of Lorentzian manifolds with distributional curvature. *Port. Math. (N.S.)*, 64(4):535–573, 2007.

[LO91] F. Lafon and M. Oberguggenberger. Generalized solutions to symmetric hyperbolic systems with discontinuous coefficients: the multidimensional case. *J. Math. Anal. Appl.*, 160(1):93–106, 1991.

[Loj57] S. Lojasiewicz. Sur la valeur et la limite d'une distribution en un point. *Studia Math.*, 16:1–36, 1957.

[LS08] A. Lastra and J. Sanz. Linear continuous operators for the Stieltjes moment problem in Gelfand-Shilov spaces. *J. Math. Anal. Appl.*, 340(2):968–981, 2008.

[Lüt82] J. Lützen. *The prehistory of the theory of distributions*, volume 7 of *Studies in the History of Mathematics and Physical Sciences.* Springer-Verlag, New York-Berlin, 1982.

[May06] E. Mayerhofer. *The wave equation on singular space-times.* Ph.D. thesis, University of Vienna, 2006.

[May08] E. Mayerhofer. On Lorentz geometry in algebras of generalized functions. *Proc. Roy. Soc. Edinburgh Sect. A*, 138(4):843–871, 2008.

[McM02] E. McMullin. The origins of the field concept in physics. *Physics in Perspective*, 4(1):13–39, 2002.

[Mik48] J.G. Mikusiński. Sur la méthode de généralisation de M. Laurent Schwartz et sur la convergence faible. *Fund. Math.*, 35:235–239, 1948.

[Nig13] E.A. Nigsch. Point value characterizations and related results in the full Colombeau algebras $\mathcal{G}^e(\Omega)$ and $\mathcal{G}^d(\Omega)$. *Math. Nachr.*, 286(10):1007–1021, 2013.

[NO12] M. Nedeljkov and M. Oberguggenberger. Ordinary differential equations with delta function terms. *Publ. Inst. Math. (Beograd) (N.S.)*, 91(105):125–135, 2012.

[NOP05] M. Nedeljkov, M. Oberguggenberger, and S. Pilipović. Generalized solutions to a semilinear wave equation. *Nonlinear Anal.*, 61(3):461–475, 2005.

[NPS98] M. Nedeljkov, S. Pilipović, and D. Scarpalézos. *The linear theory of Colombeau generalized functions*, volume 385 of *Pitman Research Notes in Mathematics Series.* Longman, Harlow, 1998.

[Obe88] M. Oberguggenberger. Hyperbolic systems with discontinuous coefficients: examples. In *Generalized functions, convergence structures, and their applications (Dubrovnik, 1987)*, pages 257–266. Plenum, New York, 1988.

[Obe89] M. Oberguggenberger. Hyperbolic systems with discontinuous coefficients: generalized solutions and a transmission problem in acoustics. *J. Math. Anal. Appl.*, 142(2):452–467, 1989.

[Obe92] M. Oberguggenberger. *Multiplication of distributions and applications to partial differential equations*, volume 259 of *Pitman Research Notes in Mathematics Series.* Longman Scientific & Technical, Harlow, 1992.

[Obe09] M. Oberguggenberger. Hyperbolic systems with discontinuous coefficients: generalized wavefront sets. In *New developments in pseudo-differential operators*, volume 189 of *Oper. Theory Adv. Appl.*, pages 117–136. Birkhäuser, Basel, 2009.

[OK99] M. Oberguggenberger and M. Kunzinger. Characterization of Colombeau generalized functions by their pointvalues. *Math. Nachr.*, 203:147–157, 1999.

[OT98] M. Oberguggenberger and T. Todorov. An embedding of Schwartz distributions in the algebra of asymptotic functions. *Internat. J. Math. Math. Sci.*, 21(3):417–428, 1998.

[OV08] M. Oberguggenberger and H. Vernaeve. Internal sets and internal functions in Colombeau theory. *J. Math. Anal. Appl.*, 341(1):649–659, 2008.

[Pac98] B.G. Pachpatte. *Inequalities for differential and integral equations*, volume 197 of *Mathematics in Science and Engineering.* Academic Press, Inc., San Diego, CA, 1998.

[RR04] M. Renardy and R.C. Rogers. *An introduction to partial differential equations*, volume 13 of *Texts in Applied Mathematics.* Springer-Verlag, New York, second edition, 2004.

[RS75] M. Reed and B. Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[Rud87] W. Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[Sca98] D. Scarpalézos. Some remarks on functoriality of Colombeau's construction; topological and microlocal aspects and applications. *Integral Transform. Spec. Funct.*, 6(1-4):295–307, 1998.

[Sch54] L. Schwartz. Sur l'impossibilité de la multiplication des distributions. *C. R. Acad. Sci. Paris*, 239:847–848, 1954.

[Sti94] T.-J. Stieltjes. Recherches sur les fractions continues. *Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys.*, 8(4):J1–J122, 1894.

[SV06] R. Steinbauer and J.A. Vickers. The use of generalized functions and distributions in general relativity. *Classical Quantum Gravity*, 23(10):R91–R114, 2006.

[SV09] R. Steinbauer and J.A. Vickers. On the Geroch-Traschen class of metrics. *Classical Quantum Gravity*, 26(6):065001, 19, 2009.

[Tay81] M.E. Taylor. *Pseudodifferential operators*, volume 34 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1981.

[Tay11] M.E. Taylor. *Partial differential equations I. Basic theory*, volume 115 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2011.

[Tem53] G. Temple. Theories and applications of generalized functions. *J. London Math. Soc.*, 28:134–148, 1953.

[Trè67] F. Trèves. *Topological vector spaces, distributions and kernels*. Dover Publications, Inc., Mineola, NY, 2006 (first published 1967).

[Trè75] F. Trèves. *Basic linear partial differential equations*. Academic Press, New York-London, 1975.

[Ver10] H. Vernaeve. Ideals in the ring of Colombeau generalized numbers. *Comm. Algebra*, 38(6):2199–2228, 2010.

[Vic99] J.A. Vickers. Nonlinear generalised functions in general relativity. In *Nonlinear theory of generalized functions (Vienna, 1997)*, volume 401 of *Chapman & Hall/CRC Res. Notes Math.*, pages 275–290. Chapman & Hall/CRC, Boca Raton, FL, 1999.

[Vic12] J.A. Vickers. Distributional geometry in general relativity. *J. Geom. Phys.*, 62(3):692–705, 2012.

[VW00] J.A. Vickers and J.P. Wilson. Generalized hyperbolicity in conical spacetimes. *Classical Quantum Gravity*, 17(6):1333–1360, 2000.

[Was87] W. Wasow. *Asymptotic expansions for ordinary differential equations*. Dover Publications, Inc., New York, 1987. Reprint of the 1976 edition.

[WB03] S. Wickramasekara and A. Bohm. Representation of semigroups in rigged Hilbert spaces: Subsemigroups of the Weyl-Heisenberg group. *Journal of Mathematical Physics*, 44(2):930–942, 2003.

[WF98] J.A. Wheeler and K.W. Ford. Geons, Black Holes, and Quantum Foam: A Life in Physics, 1998.

Zusammenfassung

In dieser Arbeit wird das Cauchy-Problem für symmetrisch hyperbolische Systeme partieller Differentialgleichungen in Colombeau-Algebren verallgemeinerter Funktionen untersucht. Dieser Rahmen erlaubt es, Gleichungen mit hochsingulären Koeffizienten zu betrachten, für welche distributionelle Lösungskonzepte nicht anwendbar sind. Insbesondere können auch Koeffizienten, die die Delta-Distribution oder ihre Ableitungen enthalten, in mathematisch rigoroser Weise behandelt werden. Differentialgleichungen mit Koeffizienten dieses Typs treten in der Physik in verschiedensten Zusammenhängen auf. Die wesentlichen Methoden für die Existenz- und Eindeutigkeitsresultate dieser Arbeit basieren auf Energieabschätzungen auf linsenförmigen Teilgebieten der Raumzeit. Darüber hinaus wird die Regularität verallgemeinerter Lösungen analysiert, wenn die Koeffizienten als Regularisierungen nichtglatter Sobolev-Funktionen definiert sind. Es wird gezeigt, dass distributionelle Limiten verallgemeinerter Lösungen auch im Falle von Koeffizienten existieren, deren Regularitäten schwächer sind als jene, die in klassischen schwachen Löungskonzepten verlangt werden. Weiters wird die Transformation einer skalaren Wellengleichung mit verallgemeinerten Funktionen als Koeffizienten in ein symmetrisch hyperbolisches System erster Ordnung untersucht. Die dafür benötigten mathematischen Operationen sind für Koeffizienten schwacher Regularität in einem distributionellen Rahmen nicht anwendbar. Als Beispiele aus der Physik werden die Dirac-Gleichung und die Klein-Gordon-Gleichung mit einer Delta-Distribution als Potential diskutiert. Auch die durch den Laplace-Beltrami-Operator einer nichtglatten Lorentz-Metrik definierte Wellengleichung wird behandelt. Ausgehend von der wichtigen Rolle von Faltungsregularisierungen bei der Einbettung von Distributionen in Colombeau-Algebren wird eine explizite Konstruktion von Schwartz-Funktionen mit vorgegebenen Momenten und einem in einer kegelförmigen Menge enthaltenen Träger entwickelt. Diese Konstruktion liefert maßgeschneiderte Glättungsfunktionen, die adaptierbare Konvergenz- und Trägerbedingungen erfüllen. Insbesondere ergeben sich daraus explizite Lösungen für Varianten des Stieltjes-Momentenproblems für rasch fallende glatte Funktionen.