





Hyperbolic systems with non-diagonalisable principal part and variable multiplicities, I: well-posedness

Claudia Garetto¹ · Christian Jäh¹  ·
Michael Ruzhansky² 

Received: 10 January 2018 / Revised: 8 March 2018 / Published online: 22 March 2018
© The Author(s) 2018

Abstract In this paper we analyse the well-posedness of the Cauchy problem for a rather general class of hyperbolic systems with space-time dependent coefficients and with multiple characteristics of variable multiplicity. First, we establish a well-posedness result in anisotropic Sobolev spaces for systems with upper triangular principal part under interesting natural conditions on the orders of lower order terms below the diagonal. Namely, the terms below the diagonal at a distance k to it must be of order $-k$. This setting also allows for the Jordan block structure in the system. Second, we give conditions for the Schur type triangularisation of general systems with variable coefficients for reducing them to the form with an upper triangular principal part for which the first result can be applied. We give explicit details for the appearing

Communicated by Y. Giga.

Michael Ruzhansky was supported in parts by EPSRC Grant EP/R003025/1 and by the Leverhulme Grant RPG-2017-151. No new data was collected or generated during the course of research.

Inspired by our colleague and friend Todor Gramchev (1956–2015).

✉ Michael Ruzhansky
m.ruzhansky@imperial.ac.uk

Claudia Garetto
c.garetto@lboro.ac.uk

Christian Jäh
c.jaeh@lboro.ac.uk

¹ Department of Mathematical Sciences, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

² Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK

conditions and constructions for 2×2 and 3×3 systems, complemented by several examples.

Mathematics Subject Classification 35L45 (primary) · 46E35 (secondary)

Contents

- 1 Introduction 1598
- 2 Well-posedness in anisotropic Sobolev spaces 1603
 - 2.1 Auxiliary remarks 1604
 - 2.2 The case $m = 2$ 1605
 - 2.3 The case $m = 3$ 1607
 - 2.4 The general case 1610
- 3 Schur decomposition of $m \times m$ matrices 1614
 - 3.1 First step or Schur step 1614
 - 3.2 The triangularisation procedure 1616
 - 3.3 The case $m = 2$ 1622
 - 3.4 The case $m = 3$ 1624
- References 1627

1 Introduction

The main aim of this paper is to consider the Cauchy problem for hyperbolic systems

$$\begin{cases} D_t u = A(t, x, D_x)u + B(t, x, D_x)u + f(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\ u|_{t=0} = u_0, & x \in \mathbb{R}^n, \end{cases} \quad (1)$$

with the usual notation $D_t = -i\partial_t$ and $D_x = -i\partial_x$. Here, we assume that $A(t, x, D_x) = [a_{ij}(t, x, D_x)]_{i,j=1}^m$ is an $m \times m$ matrix of pseudo-differential operators of order 1, i.e. $a_{ij} \in C([0, T], \Psi_{1,0}^1(\mathbb{R}^n))$ with possibly complex valued symbols. In the first part of the paper we will also assume that

$$A(t, x, D_x) = \Lambda(t, x, D_x) + N(t, x, D_x), \quad (2)$$

with real-valued symbols in

$$\Lambda(t, x, D_x) = \text{diag}(\lambda_1(t, x, D_x), \lambda_2(t, x, D_x), \dots, \lambda_m(t, x, D_x)),$$

and

$$N(t, x, D_x) = \begin{bmatrix} 0 & a_{12}(t, x, D_x) & a_{13}(t, x, D_x) & \cdots & a_{1m}(t, x, D_x) \\ 0 & 0 & a_{23}(t, x, D_x) & \cdots & a_{2m}(t, x, D_x) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{m-1m}(t, x, D_x) \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Finally, we assume that

$$B(t, x, D_x) = [b_{ij}(t, x, D_x)]_{i,j=1}^m, \quad b_{ij} \in C([0, T], \Psi_{1,0}^0(\mathbb{R}^n)),$$

is an $m \times m$ matrix of pseudo-differential operators of order 0 with possibly complex valued symbols. We can take any $n \geq 1$ and we can assume that $m \geq 2$ since in the case $m = 1$ there are no multiplicities and thus much more is known. It is also well-known that even if all the coefficients in A and B depend only on time, due to multiplicities, the best one can hope for is the well-posedness of the Cauchy problem (1) in suitable classes of Gevrey spaces. Thus, the main questions that we address in this paper are:

- (Q1) Under what structural conditions on the zero order part $B(t, x, D_x)$ is the Cauchy problem (1) well-posed in C^∞ or, even better, in suitable scales of Sobolev spaces?
- (Q2) Under what conditions on the general matrix $A(t, x, D_x)$ of first order pseudo-differential operators can we reduce it (microlocally) to another system with A satisfying the upper triangular condition (2)?

Note that this paper is part of a wider analysis of hyperbolic systems with multiplicities. Here we investigate the well-posedness of these systems. In the second part of this paper we plan to carry out the microlocal analysis of their solutions.

In the case of 2×2 systems the questions above have been analysed with the answer to (Q1) given by the following theorem:

Theorem A ([27, Theorem 7.2]) *Let $m = 2$. Suppose that the pseudo-differential operator b_{21} is of order not greater than -1 . Then the Cauchy problem (1) is well-posed in C^∞ . Moreover, it is well-posed in the anisotropic Sobolev space $\begin{bmatrix} H^{s_1}(\mathbb{R}^n) \\ H^{s_2}(\mathbb{R}^n) \end{bmatrix}$ provided $s_2 - s_1 \geq 1$. In that case the solution satisfies the following estimates:*

$$\|u_1(t, \cdot)\|_{H^s} + \|u_2(t, \cdot)\|_{H^{s+1}} \leq ce^{ct} \left(\|u_1^0\|_{H^s} + \|u_2^0\|_{H^{s+1}} \right), \quad 0 \leq t \leq T,$$

for $u_j^0 \in H_{comp}^{s+j-1}(\mathbb{R}^n)$, $j = 1, 2$, with $c > 0$ depending on s, T , and the support of the initial data.

The case of systems of general size but for coefficients depending only on t and for $n = 1$ was also considered. More precisely, in [26] the authors considered the Cauchy problem

$$\begin{cases} D_t u = A(t)D_x u + B(t)D_x u + f(t, x), & (t, x) \in [0, T] \times \mathbb{R}, \\ u|_{t=0} = 0, & x \in \mathbb{R}, \end{cases} \quad (3)$$

with $A(t) = [a_{ij}(t)]_{i,j=1}^m \in C([0, T])^{m \times m}$ in the form

$$A(t) = \Lambda(t) + N(t),$$

similar to (2). They showed the following result in the absence of lower order terms and for zero Cauchy data:

Theorem B ([26, Proposition 1]) *Let $B(t) \equiv 0$ and let $s \in \mathbb{R}$. Then the Cauchy problem (3) is C^∞ -well-posed. Moreover, there exist $r_1, \dots, r_{m-1} \in [0, 1]$ such that for every $f \in C(\mathbb{R}, (H^s(\mathbb{R}))^m)$ identically 0 at $t = 0$ it admits a unique solution $u \in C(\mathbb{R}, (\mathcal{S}'(\mathbb{R}))^m)$ satisfying*

$$u_m \in C(\mathbb{R}, H^s(\mathbb{R})), \quad u_{m-j} \in C(\mathbb{R}, H^{s-r_1-\dots-r_{j-1}}(\mathbb{R})),$$

for $j = 1, \dots, m-1$, and identically 0 at $t = 0$. In particular, if $\lambda_j(t) \neq \lambda_k(t)$, $t \in \mathbb{R}$, $1 \leq j < k \leq m$, no loss of anisotropic regularity appears.

The case of (microlocally) *diagonalisable* systems of any order with fully variable coefficients was considered by Rozenblum [41] under the condition of transversality of the intersecting characteristics. Also allowing the variable multiplicities, this transversality condition was later removed in [32,33] with sharp L^p -estimates for solutions, with further applications to the spectral asymptotics of the corresponding elliptic systems.

Before stating our main results and collecting some necessary basic notions we give a brief overview of the state of the art for hyperbolic equations and systems. We have a complete understanding of strictly hyperbolic systems, i.e., systems without multiplicities, with C^∞ -coefficients. This starts with the groundbreaking work of Lax [35] and Hörmander [28] and heavily relies on the modern theory of Fourier integral operators (FIO). Well-posedness is here obtained in the space of distributions \mathcal{D}' . There are also well-posedness results for less regular coefficients with respect to t . For instance, well-posedness with loss of derivatives has been obtained by Colombini and Lerner [9] for second order strictly hyperbolic equations with Log-Lipschitz coefficients with respect to t and smooth in x . It is possible to further drop the regularity in t (for instance Hölder), however, this has to be balanced by stronger regularity in x (Gevrey) and leads to more specific (Gevrey) well-posedness results (see [3,31] and references therein). Paradifferential techniques have been recently used for this kind of strictly hyperbolic equations by Colombini et al. [6,7].

The analysis of hyperbolic equations with multiplicities (weakly hyperbolic) has started with the seminal paper by Colombini et al. [5] in the case of coefficients depending only on time. Profound difficulties in such analysis have been exhibited by Colombini et al. [4,8] showing that even the second order wave equation in \mathbb{R} with smooth time-dependent propagation speed (but with multiplicity) and smooth Cauchy data need not be well-posed in \mathcal{D}' . However, they turn out to be well-posed in suitable Gevrey classes or spaces of ultradistributions. In the last decades many results were obtained for weakly hyperbolic equations with t -dependent coefficients ([3,11,16,18–20,34], to quote only very few). More recently, advances in the theory of weakly hyperbolic systems with t -dependent coefficients have been obtained for systems of any size in presence of multiplicities with regular or low regular (Hölder) coefficients [16,22,23]. In addition, in [17] precise conditions on the lower order terms (Levi conditions) have been formulated to guarantee Gevrey and ultradistributional well-posedness. Previously very few results were known in the field for systems of a

certain size ($2 \times 2, 3 \times 3$) [12,13] or of a certain form (for instance without lower order terms or with principal part of a certain form) [44].

Weakly hyperbolic equations with x -dependent coefficients were considered for the first time in the celebrated paper by Bronshtein [2]. As shown already in some earlier works by Ivrii, the corresponding Cauchy problem is well-posed under “almost analytic regularity”, namely, if the coefficients and initial data are in suitable Gevrey classes. Bronshtein’s result was extended to (t, x) -dependent scalar equations by Ohya and Tarama [38] and to systems by Kajitani and Yuzawa [31]. The regularity assumptions are always quite strong with respect to x (Gevrey) and not below Hölder in t . See also [10,37]. Geometrical and microlocal analytic approaches are known for equations or systems under specific assumptions on the characteristics and/or lower order terms. See [29,30,33,36,39], to quote only a few. Time-dependent coefficients of low regularity (distributional) have been considered in [21].

In this paper we will be interested in the case of coefficients depending on both t and x and we will make use of the usual definitions of symbol classes. We say that a (possibly) complex valued function $a = a(x, \xi) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ belongs to $S^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n)$ if there exist constants $C_{\alpha,\beta} > 0$ such that

$$\forall \alpha, \beta \in \mathbb{N}^n : |\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \leq C_{\alpha,\beta} \langle \xi \rangle^{m-|\beta|} \quad \forall (x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n.$$

The set of pseudo-differential operators associated to the symbols in $S^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n)$ is denoted by $\Psi^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n)$.

If there is no question about the domain under consideration, we will abbreviate the symbol- and operator-classes by $S^m_{1,0}$ and $\Psi^m_{1,0}$, respectively, or simply by S^m and Ψ^m .

We also denote by $C([0, T], S^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n))$ the space of all symbols $a(t, x, \xi) \in S^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n)$ which are continuous with respect to t . The set of operators associated to the symbols in $C([0, T], S^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n))$ is denoted by $C([0, T], \Psi^m_{1,0}(\mathbb{R}^n \times \mathbb{R}^n))$.

Again, if there is no question about the domain under consideration, we will abbreviate the symbol- and operator-classes by $CS^m_{1,0}$ and $C\Psi^m_{1,0}$, respectively, or simply by CS^m and $C\Psi^m$.

Let us give our main result concerning the first question (Q1) for the systems with the principal part A satisfying the upper triangular condition (2). Here, f_k, u_k and u_k^0 , for $k = 1, \dots, m$, stand for the components of the vectors f, u and u_0 , respectively.

Theorem 1 *Let $n \geq 1, m \geq 2$, and let*

$$\begin{cases} D_t u = A(t, x, D_x)u + B(t, x, D_x)u + f(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\ u|_{t=0} = u_0(x), & x \in \mathbb{R}^n, \end{cases} \quad (4)$$

where $A(t, x, D_x) \in (CS^1)^{m \times m}$ is an upper-triangular matrix of pseudo-differential operators of order 1 in the form (2), and $B(t, x, D_x) \in (CS^0)^{m \times m}$ is a matrix of pseudo-differential operators of order 0, continuous with respect to t . Hence, if

the lower order terms b_{ij} belong to $C([0, T], \Psi^{j-i})$ for $i > j$,

$u_k^0 \in H^{s+k-1}(\mathbb{R}^n)$ and $f_k \in C([0, T], H^{s+k-1})$ for $k = 1, \dots, m$, then (4) has a unique anisotropic Sobolev solution u , i.e., $u_k \in C([0, T], H^{s+k-1})$ for $k = 1, \dots, m$.

Remark 1 As stated earlier, we allow A and B to have complex valued symbols as long as the symbols of Λ in (2), i.e. the eigenvalues of $A(t, x, \xi)$, are real valued.

The main condition of Theorem 1 for the Sobolev well-posedness is that the pseudo-differential operator b_{ij} below the diagonal (i.e. for $i > j$) must be of order $j - i$. In other words, the terms below the diagonal at a distance k to it must be of order $-k$.

In solving the Cauchy problem (4) we will make use of Fourier integral operators depending on the parameter $t \in [0, T]$. Namely, we will work with operators of the type

$$\int_0^t \int_{\mathbb{R}^n} e^{i\varphi(t,s,x,\xi)} a(t, s, x, \xi) \widehat{g}(s, \xi) d\xi ds$$

where φ is the solution of a certain eikonal equation and the symbol a is determined via asymptotic expansion and transport equations. In Sect. 2.1 we will recall some well-known Sobolev estimates for this type of operators.

In Sect. 2 we will prove Theorem 1 after we explain its idea in the cases of $m = 2$ and $m = 3$.

Consequently, in Sect. 3 we give an answer to the second question (Q2) above in the form of a suitable variable coefficients extension of the Schur triangularisation. For constant matrices such a procedure is well known (see e.g. [1, Theorem 5.4.1]).

Theorem C (Schur’s triangularisation theorem) *Given a (constant) $m \times m$ matrix A with eigenvalues $\lambda_1, \dots, \lambda_m$ in any prescribed order, there is a unitary $m \times m$ matrix T such that $R = T^{-1}AT$ is upper triangular with the diagonal elements $r_{ii} = \lambda_i$. Furthermore, if the entries of A and its eigenvalues are all real, T may be chosen to be real orthogonal.*

It follows that R can be written as $D + N$, where $D = \text{diag}(\lambda_1, \dots, \lambda_m)$ and N is a nilpotent upper triangular matrix.

If the matrix A depends on one or several parameters, namely $A = A(t, x, \xi)$, the situation becomes less clear and it is difficult to give a complete description, in particular together with a prescribed regularity of the involved transformation matrices. The regularity of the matrix A and the desire to maintain it through the transformation puts already constraints on the matrix as, in general, the eigenvalues can only be expected to be Lipschitz continuous in the parameters even if all the entries depend smoothly on the parameters (see, e.g., [2, 40] and the references therein). In the sequel, we will present some sufficient conditions to ensure the existence of an upper triangularisation for $A(t, x, \xi)$ which respects its regularity. For example, it will apply to the case when A is a matrix of first order symbols continuous with respect to t , i.e., $A(t, x, \xi) \in (CS^1)^{m \times m}$.

Our main result for this part of the problem is the following theorem.

Theorem 2 Let $A(t, x, \xi) \in (CS^1)^{m \times m}$, be a $m \times m$ -matrix with eigenvalues $\lambda_1, \dots, \lambda_m \in CS^1$, and let $h_1, \dots, h_{m-1} \in (CS^0)^m$ be the corresponding eigenvectors. Suppose that for $e_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^{m-i+1}$ the condition

$$\left\langle h^{(i)}(t, x, \xi) | e_1 \right\rangle \neq 0, \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \tag{5}$$

holds for all $i = 1, \dots, m - 1$, with the notation for $h^{(i)}$ explained in (37). Then, there exists a matrix-valued symbol $T(t, x, \xi) \in (CS^0)^{m \times m}$, invertible for $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$ with $T^{-1}(t, x, \xi) \in (CS^0)^{m \times m}$, such that

$$T^{-1}(t, x, \xi)A(t, x, \xi)T(t, x, \xi) = \Lambda(t, x, \xi) + N(t, x, \xi)$$

for all $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$, where

$$\Lambda(t, x, \xi) = \text{diag}(\lambda_1(t, x, \xi), \lambda_2(t, x, \xi), \dots, \lambda_m(t, x, \xi))$$

and

$$N(t, x, \xi) = \begin{bmatrix} 0 & N_{12}(t, x, \xi) & N_{13}(t, x, \xi) & \dots & N_{1m}(t, x, \xi) \\ 0 & 0 & N_{23}(t, x, \xi) & \dots & N_{2m}(t, x, \xi) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & N_{m-1m}(t, x, \xi) \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

and N is a nilpotent matrix with entries in CS^1 .

Furthermore, there is an expression for the matrix symbol T which will be given in Theorem 6. Also, the assumption (5) can be relaxed, see Remark 6. In Sect. 3 we will prove this result as well as describe the procedure how to obtain the desired upper triangular form. Moreover, we work out in detail the cases of $m = 2$ and $m = 3$ clarifying this Schur triangualisation procedure and give a number of examples.

The results and techniques of this paper are a natural outgrowth of the paper [27] where the case $m = 2$ was considered and to which the results of the present paper reduce in the case of 2×2 systems. It is with great sorrow that we remember the untimely departure of our colleague and friend **Todor Gramchev** who was the inspiration for both [27] and the present paper.

2 Well-posedness in anisotropic Sobolev spaces

This section is devoted to proving the well-posedness of the Cauchy problem (1). For the reader’s convenience we first give a detailed proof in the cases $m = 2$ and $m = 3$. This will inspire us in proving Theorem 1. We note that the case $m = 2$ has been studied in [27] and we will briefly review its derivartion. However, first we collect a few results about Fourier integral operators that we will need in the sequel.

2.1 Auxiliary remarks

In solving the Cauchy problem (1), we will deal with solutions of certain scalar pseudo-differential equations. For each characteristic λ_j of A , we will be denoting by $G_j^0\theta$ the solution to

$$\begin{cases} D_t w = \lambda_j(t, x, D_x)w + b_{jj}(t, x, D_x)w, \\ w(0, x) = \theta(x), \end{cases}$$

and by $G_j g$ the solution to

$$\begin{cases} D_t w = \lambda_j(t, x, D_x)w + b_{jj}(t, x, D_x)w + g(t, x), \\ w(0, x) = 0. \end{cases}$$

The operators G_j^0 and G_j can be microlocally represented by Fourier integral operators

$$G_j^0\theta(t, x) = \int_{\mathbb{R}^n} e^{i\varphi_j(t,x,\xi)} a_j(t, x, \xi) \widehat{\theta}(\xi) d\xi \tag{6}$$

and

$$G_j g(t, x) = \int_0^t \int_{\mathbb{R}^n} e^{i\varphi_j(t,s,x,\xi)} A_j(t, s, x, \xi) \widehat{g}(s, \xi) d\xi ds,$$

with $\varphi_j(t, s, x, \xi)$ solving the eikonal equation

$$\begin{cases} \partial_t \varphi_j = \lambda_j(t, x, \nabla_x \varphi_j), \\ \varphi_j(s, s, x, \xi) = x \cdot \xi, \end{cases}$$

and with the notation

$$\varphi_j(t, x, \xi) = \varphi_j(t, 0, x, \xi).$$

Here we also have the amplitudes $A_{j,-k}(t, s, x, \xi)$ of order $-k$, $k \in \mathbb{N}$, giving $A_j \sim \sum_{k=0}^\infty A_{j,-k}$, and they satisfy the transport equations with initial data at $t = s$, and we have $a_j(t, x, \xi) = A_j(t, 0, x, \xi)$.

If $a_j \in S^m$, i.e. if the amplitude a_j in (6) is a symbol of order m , we will write $G_j^0 \in I_{1,0}^m$. However, in the above construction of propagators for hyperbolic equations, we have $a_j \in S^0$, so that $G_j^0 \in I_{1,0}^0$.

By $I_{1,0}^m$, we denote the class of Fourier integral operators with amplitudes in $S_{1,0}^m$. For further information, the reader may consult [15, 42, 43] and the references therein.

With that, we can record the following estimate:

Lemma 1 For any $\sigma \in \mathbb{R}$, for sufficiently small t , we have

$$\|G_j^0 \theta(t)\|_{H^\sigma} \leq C_{A,\sigma,u_0} \|\theta\|_{H^\sigma}, \quad \|G_j g(t)\|_{H^\sigma} \leq C_{A,\sigma} t \|g\|_{L_x^\infty H_x^\sigma}.$$

This statement follows from the continuity of $\lambda_j, \varphi_j, a_j, A_j$ with respect to t and from the H^σ -boundedness of non-degenerate Fourier integral operators, see e.g. [15] (there are also surveys on such questions [42,43]). It is important to note that the constant for the estimate for G_j does not depend on the initial data of the Cauchy problem; see also Remark 2.

2.2 The case $m = 2$

To motivate the higher order cases, here we review the construction for 2×2 systems adapting it for the subsequent higher order arguments. Hence, in this subsection we follow the proof in [27]. Thus, we consider the system

$$\begin{cases} D_t u = A(t, x, D_x)u + B(t, x, D_x)u + f(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\ u|_{t=0} = u_0, & x \in \mathbb{R}^n, \end{cases} \tag{7}$$

where $u_0(x) = [u_1^0(x), u_2^0(x)]^T$, $f(t, x) = [f_1(t, x), f_2(t, x)]^T$, and with the operators $A(t, x, D_x)$ and $B(t, x, D_x)$ given by

$$A(t, x, D_x) = \begin{bmatrix} \lambda_1(t, x, D_x) & a_{12}(t, x, D_x) \\ 0 & \lambda_2(t, x, D_x) \end{bmatrix} \tag{8}$$

and

$$B(t, x, D_x) = \begin{bmatrix} b_{11}(t, x, D_x) & b_{12}(t, x, D_x) \\ b_{21}(t, x, D_x) & b_{22}(t, x, D_x) \end{bmatrix}.$$

We suppose that all entries of $A(t, x, D_x)$ belong to $C\Psi_{1,0}^1$ and all entries of $B(t, x, D_x)$ belong to $C\Psi_{1,0}^0$. By using the operators G_j^0 and G_j introduced in Sect. 2.1, we can reformulate the Eq. (7) as

$$u_1 = U_1^0 + G_1((a_{12} + b_{12})u_2), \tag{9}$$

$$u_2 = U_2^0 + G_2(b_{21}u_1), \tag{10}$$

where

$$U_j^0 = G_j^0 u_j^0 + G_j(f_j), \quad j = 1, 2. \tag{11}$$

Plugging (10) in (9), we obtain

$$u_1 = \tilde{U}_1^0 + G_1(a_{12}G_2(b_{21}u_1)) + G_1(b_{12}G_2(b_{21}u_1)), \tag{12}$$

where

$$\tilde{U}_1^0 = G_1^0 u_1^0 + G_1(f_1) + G_1((a_{12} + b_{12})U_2^0). \tag{13}$$

Using the rules of composition of Fourier integral operators, see e.g. [15], and by Lemma 1, we get that the operator $G_1 \circ a_{12} \circ G_2 \circ b_{21}$ in (12) acts continuously on H^s if it is of order 0. Since $a_{12} \in C\Psi_{1,0}^1$ we therefore need to assume that $b_{21} \in C\Psi_{1,0}^{-1}$.

The operator $G_1 \circ b_{12} \circ G_2 \circ b_{21}$ belongs to $CI_{1,0}^{-1}$ since $b_{21} \in C\Psi_{1,0}^{-1}$ and $b_{12} \in C\Psi_{1,0}^0$.

We now introduce the following scale of Banach spaces $X^s(t) := C([0, t], H^s)$, $t \in [0, T]$, equipped with the norm

$$\|u\|_{X^s(t)} = \sup_{\tau \in [0,t]} \|u(\tau, \cdot)\|_{H^s}.$$

Let

$$\mathcal{G}_1^0 u_1 := G_1(a_{12}G_2(b_{21}u_1)) + G_1(b_{12}G_2(b_{21}u_1)).$$

It follows that (12) can be written as

$$u_1 = \tilde{U}_1^0 + \mathcal{G}_1^0 u_1.$$

By composition of Fourier integral operators and Lemma 1 we have that the 0-order Fourier integral operator \mathcal{G}_1^0 maps $C([0, T], H^s)$ continuously into itself and for small time interval it is a contraction, in the sense that there exists $T^* \in [0, T]$ such that

$$\|\mathcal{G}_1^0(u - v)\|_{X^s(T^*)} \leq C_{A,s} T^* \|u - v\|_{X^s(T^*)},$$

with $C_{A,s} T^* < 1$. Banach’s fixed point theorem ensures the existence of a unique fixed point u_1 for the map \mathcal{G}_1^0 . Hence, by assuming that the initial data \tilde{U}_1^0 belongs to $C([0, T^*], H^s)$ we conclude that there exists a unique $u_1 \in C([0, T^*], H^s)$ solving (12). Note that the same argument proves that the operator $I - \mathcal{G}_1^0$ is invertible on a sufficiently small interval in t since $\mathcal{G}_1^0 = I$ at $t = 0$. From formula (13) it is clear that in order to get \tilde{U}_1^0 to belong to $C([0, T^*], H^s)$ we need to assume that $U_2^0 \in H^{s+1}$. Finally, we get u_2 by substitution of u_1 in (10).

Remark 2 Note that the constant T^* depends only on A and s . Thus, the argument above can be iterated by taking $u(T^*, x)$ as new initial data. In this way one can cover an arbitrary finite interval $[0, T]$ and obtain a solution in $C([0, T], H^s) \times C([0, T], H^{s+1})$.

Remark 3 Since $a_{12}(t, x, D_x)$ is a first order operator combining (11) with (13) we easily see that in order to get Sobolev well-posedness of order s we need to take initial data u_1^0 and u_2^0 in H^s and H^{s+1} , respectively, and right hand-side functions f_1 and f_2 in $C([0, T], H^s)$ and $C([0, T], H^{s+1})$, respectively.

We have therefore proved the following theorem stated for the first time in [27, Theorem 7.2].

Theorem 3 Consider the Cauchy problem (7), with the 2×2 matrices

$$A(t, x, D_x) \in (CS^1)^{2 \times 2} \quad \text{and} \quad B(t, x, D_x) \in (CS^0)^{2 \times 2},$$

where A is of the form (2). Assume that $b_{21} \in C([0, T], \Psi_{1,0}^{-1})$, the right hand-side functions f_1 and f_2 belong to $C([0, T], H^s)$ and $C([0, T], H^{s+1})$, respectively, and the initial data u_1^0 and u_2^0 belong to H^s and H^{s+1} , respectively. Then, (7) has a unique solution in $C([0, T], H^s) \times C([0, T], H^{s+1})$. More generally it is well-posed in the anisotropic Sobolev space $C([0, T], H^{s_1}) \times C([0, T], H^{s_2})$, provided $s_2 - s_1 = 1$.

Remark 4 It was also shown in [27] that the solution u satisfies the estimate

$$\|u_1(t, \cdot)\|_{H^s} + \|u_2(t, \cdot)\|_{H^{s+1}} \leq ce^{ct} \left(\|u_1^0\|_{H^s} + \|u_2^0\|_{H^{s+1}} \right), \quad 0 \leq t \leq T,$$

for $u_j^0 \in H_{comp}^{s+j-1}$, $j = 1, 2$ with $c > 0$ depending on s, T , and the support of the initial data. Since well-posedness is obtained for any Sobolev order s it follows that the Cauchy problem (7) is also C^∞ well-posed.

2.3 The case $m = 3$

In this section we will extend the construction to the case of 3×3 systems. In the argument there is an additional substitution and a fixed point argument step compared to the case $m = 2$. The advantage of giving the case of $m = 3$ here is that we can make the argument more concrete compared to the more abstract construction in the general case that will be given in the following section. Thus, let

$$\begin{cases} D_t u = A(t, x, D_x)u + B(t, x, D_x)u + f(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\ u|_{t=0} = u_0, & x \in \mathbb{R}^n, \end{cases} \quad (14)$$

where $u_0(x) = [u_1^0(x), u_2^0(x), u_3^0(x)]^T$, $f(t, x) = [f_1(t, x), f_2(t, x), f_3(t, x)]^T$, $A(t, x, D_x)$ is defined by the matrix

$$\begin{bmatrix} \lambda_1(t, x, D_x) & a_{12}(t, x, D_x) & a_{13}(t, x, D_x) \\ 0 & \lambda_2(t, x, D_x) & a_{23}(t, x, D_x) \\ 0 & 0 & \lambda_3(t, x, D_x) \end{bmatrix}, \quad (15)$$

and

$$B(t, x, D_x) = \begin{bmatrix} b_{11}(t, x, D_x) & b_{12}(t, x, D_x) & b_{13}(t, x, D_x) \\ b_{21}(t, x, D_x) & b_{22}(t, x, D_x) & b_{23}(t, x, D_x) \\ b_{31}(t, x, D_x) & b_{32}(t, x, D_x) & b_{33}(t, x, D_x) \end{bmatrix}.$$

We assume that all the entries of $A(t, x, D_x)$ and $B(t, x, \xi)$ belong to $C\Psi_{1,0}^1$ and $C\Psi_{1,0}^0$, respectively. Using the notations introduced earlier, we can write

$$\begin{aligned} u_3(t, x) &= U_3^0 + G_3(b_{31}u_1) + G_3(b_{32}u_2), \\ u_2(t, x) &= U_2^0 + G_2((a_{23} + b_{23})u_3) + G_2(b_{21}u_1), \\ u_1(t, x) &= U_1^0 + G_1((a_{12} + b_{12})u_2) + G_1((a_{13} + b_{13})u_3), \end{aligned} \quad (16)$$

where

$$U_j^0(t, x) = G_j^0(u_j^0) + G_j(f_j), \quad j = 1, 2, 3. \quad (17)$$

Now, we plug u_3 into u_1 and u_2 in formula (16) and, thus, obtain

$$\begin{aligned} u_2(t, x) &= \tilde{U}_2^0 + G_2(b_{21}u_1) + G_2((a_{23} + b_{23})G_3(b_{31}u_1)) \\ &\quad + G_2((a_{23} + b_{23})G_3(b_{32}u_2)), \\ u_1(t, x) &= \tilde{U}_1^0 + G_1((a_{13} + b_{13})G_3(b_{31}u_1)) \\ &\quad + G_1((a_{13} + b_{13})G_3(b_{32}u_2)) + G_1((a_{12} + b_{12})u_2), \end{aligned} \quad (18)$$

where

$$\tilde{U}_j^0 = U_j^0 + G_j((a_{j3} + b_{j3})(t, x, D_x)U_3^0), \quad j = 1, 2.$$

We introduce the operator \mathcal{G}_2^0 by setting

$$\mathcal{G}_2^0 u_2 := G_2((a_{23} + b_{23})G_3(b_{32}u_2)) \quad (19)$$

and in analogy with the case $m = 2$ we define

$$L_2 u_2 := u_2 - \mathcal{G}_2^0 u_2.$$

By Lemma 1 we have that for any s , \mathcal{G}_2^0 has the operator norm in H^s strictly less than 1 on a sufficiently small interval $[0, T^*]$, so L_2 is a perturbation of the identity operator. By the Neumann series it follows that L_2 is invertible as a continuous operator from $C([0, T^*], H^s)$ to $C([0, T^*], H^s)$. Noting now that

$$u_2 - \mathcal{G}_2^0 u_2 = L_2 u_2 = \tilde{U}_2^0 + G_2(b_{21}u_1) + G_2((a_{23} + b_{23})G_3(b_{31}u_1)),$$

we have that

$$u_2(t, x) = L_2^{-1} \tilde{U}_2^0 + L_2^{-1} G_2((a_{23} + b_{23})G_3(b_{31}u_1)) + L_2^{-1} G_2(b_{21}u_1).$$

Since this expression depends only on u_1 , we can plug it into the formula for u_1 in (18) and obtain

$$\begin{aligned}
 u_1(t, x) &= \tilde{U}_1^0 + G_1((a_{13} + b_{13})G_3(b_{31}u_1)) + \\
 &\quad + G_1((a_{13} + b_{13})G_3(b_{32}u_2)) + G_1((a_{12} + b_{12})u_2) \\
 &= \tilde{U}_1^0 + G_1((a_{13} + b_{13})G_3(b_{31}u_1)) + \\
 &\quad + G_1((a_{13} + b_{13})G_3(b_{32}(L_2^{-1}\tilde{U}_2^0)) \\
 &\quad + G_1((a_{13} + b_{13})G_3(b_{32}L_2^{-1}G_2((a_{23} + b_{23})G_3(b_{31}u_1)))) \\
 &\quad + G_1((a_{13} + b_{13})G_3(b_{32}(L_2^{-1}G_2(b_{21}u_1))) \\
 &\quad + G_1((a_{12} + b_{12})L_2^{-1}\tilde{U}_2^0) \\
 &\quad + G_1((a_{12} + b_{12})L_2^{-1}G_2((a_{23} + b_{23})G_3(b_{31}u_1))) \\
 &\quad + G_1((a_{12} + b_{12})L_2^{-1}G_2(b_{21}u_1)).
 \end{aligned}$$

By collecting now the terms with order ≤ 0 we can simplify the previous formula as follows:

$$\begin{aligned}
 u_1(t, x) &= \tilde{U}_1^0 + G_1(a_{13}G_3(b_{31}u_1)) + G_1(a_{13}G_3(b_{32}(L_2^{-1}\tilde{U}_2^0))) \\
 &\quad + G_1(a_{13}G_3b_{32}L_2^{-1}G_2(a_{23}G_3(b_{31}u_1))) \\
 &\quad + G_1(a_{13}G_3b_{32}L_2^{-1}G_2(b_{23}G_3(b_{31}u_1))) \\
 &\quad + G_1(a_{13}G_3b_{32}(L_2^{-1}G_2(b_{21}u_1))) \\
 &\quad + G_1(b_{13}G_3(b_{32}L_2^{-1}G_2(a_{23}G_3(b_{31}u_1)))) \\
 &\quad + G_1(a_{12}L_2^{-1}\tilde{U}_2^0) \\
 &\quad + G_1(a_{12}L_2^{-1}G_2(a_{23}G_3(b_{31}u_1))) \\
 &\quad + G_1(a_{12}L_2^{-1}G_2(b_{23})G_3(b_{31}u_1))) \\
 &\quad + G_1(b_{12}L_2^{-1}G_2((a_{23}G_3(b_{31}u_1)))) \\
 &\quad + G_1(a_{12}L_2^{-1}G_2(b_{21}u_1)) + \text{l.o.t.}
 \end{aligned}$$

Looking at the terms

$$\begin{aligned}
 &G_1(a_{13}G_3(b_{32}(L_2^{-1}\tilde{U}_2^0))), \\
 &G_1(a_{12}L_2^{-1}G_2(b_{21}u_1)), \\
 &G_1(a_{12}L_2^{-1}G_2(a_{23}G_3(b_{31}u_1)))
 \end{aligned}$$

and keeping in mind that in order to get the right Sobolev regularity we need to have operators of order 0, we deduce that b_{21} and b_{32} must have order -1 while b_{31} must have order -2 . Considering now the initial data

$$\tilde{U}_j^0 = U_j^0 + G_j((a_{j3} + b_{j3})U_3^0), \quad j = 1, 2,$$

by using (17) we obtain

$$\tilde{U}_j^0 = U_j^0 + G_j((a_{j3} + b_{j3})(G_3^0(u_3^0) + G_3(f_3))), \quad j = 1, 2.$$

Combining these formulas with an analysis of the term $G_1(a_{12}L_2^{-1}\tilde{U}_2^0)$ we deduce that \tilde{U}_2^0 must belong to H^{s+1} . This implies $U_2^0 \in H^{s+1}$ and $U_3^0 \in \dot{H}^{s+2}$. Concluding, similarly to the case $m = 2$, that is by the Banach fixed point theorem argument on u_1 and substitution in u_2 and u_3 , we get anisotropic Sobolev well-posedness by assuming u_1^0 and f_1 in H^s , u_2^0 and f_2 in H^{s+1} , and u_3^0 and f_3 in H^{s+2} . This well-posedness is obtained by means of one invertible operator L_2 , and in analogy with case $m = 2$ the well-posedness can be extended to the whole interval $[0, T]$ by an iterated argument. This proves Theorem 1 in the case $m = 3$.

2.4 The general case

We are now ready to prove the main result of our paper in the general case of an upper-triangular $m \times m$ matrix, i.e, a matrix A of the type

$$\begin{bmatrix} \lambda_1(t, x, D_x) & a_{12}(t, x, D_x) & \cdots & a_{1m}(t, x, D_x) \\ 0 & \lambda_2(t, x, D_x) & \cdots & a_{2m}(t, x, D_x) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \lambda_{m-1}(t, x, D_x) & a_{m-1m}(t, x, D_x) \\ 0 & 0 & \cdots & \lambda_m(t, x, D_x) \end{bmatrix}.$$

For the convenience of the reader we recall here the statement of Theorem 1.

Theorem 1 *Let*

$$\begin{cases} D_t u = A(t, x, D_x)u + B(t, x, D_x)u + f(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\ u|_{t=0} = u_0(x), & x \in \mathbb{R}^n, \end{cases} \quad (20)$$

where $A(t, x, D_x)$ is an upper-triangular matrix of pseudo-differential operators of order 1 in the form (2), and $B(t, x, D_x)$ is a matrix of pseudo-differential operators of order 0, continuous with respect to t . Hence, if the lower order terms b_{ij} belong to $C([0, T], \Psi^{j-i})$ for $i > j$, $u_k^0 \in H^{s+k-1}$ and $f_k \in C([0, T], H^{s+k-1})$ for $k = 1, \dots, m$ then (20) has a unique anisotropic Sobolev solution u , i.e., $u_k \in C([0, T], H^{s+k-1})$ for $k = 1, \dots, m$.

Proof Making use of the notations introduced earlier we can write the components of the solution u as

$$\begin{aligned}
 u_i(t, x) &= U_i^0 + G_i \left(\sum_{j>i}^m a_{ij}(t, x, D_x)u_j \right) + G_i \left(\sum_{\substack{j=1 \\ j\neq i}}^m b_{ij}(t, x, D_x)u_j \right) \\
 &= U_i^0 + \sum_{j<i} G_i(b_{ij}(t, x, D_x)u_j) + \sum_{i<j\leq m} G_i((a_{ij} + b_{ij})(t, x, D_x)u_j),
 \end{aligned}
 \tag{21}$$

where

$$U_i^0 = G_i^0 u_j^0 + G_i(f_i),$$

and G_i, G_i^0 are Fourier integral operator of order 0 for $i = 1, \dots, m$. Note that from the fact that b_{ij} is a symbol of order 0 for every i, j and, in particular, of order $j - i$ for $j < i$ we obtain that the operator $G_i(b_{ij})$ is of order $j - i$ for $j < i$, while $G_i(a_{ij} + b_{ij})$ is, in general, of order 1. To simplify the argument we introduce the notations $G_{i,j}^{j-i}$ and $G_{i,j}^1$ for the operators $G_i(b_{ij})$ and $G_i(a_{ij} + b_{ij})$, respectively. Here the superscript stands to remind us of the order of the operator. Hence,

$$u_i = U_i^0 + \sum_{j<i} G_{i,j}^{j-i}(u_j) + \sum_{i<j\leq m} G_{i,j}^1(u_j),$$

for $i = 1, \dots, m$. By begin by substituting

$$u_m = U_m^0 + \sum_{j<m} G_{m,j}^{j-m}(u_j),$$

into

$$u_{m-1} = U_{m-1}^0 + \sum_{j<m-1} G_{m-1,j}^{j-m+1}(u_j) + G_{m-1,m}^1(u_m).$$

We get

$$\begin{aligned}
 u_{m-1} &= U_{m-1}^0 + \sum_{j<m-1} G_{m-1,j}^{j-m+1}(u_j) + G_{m-1,m}^1 U_m^0 + \sum_{j<m} G_{m-1,m}^1 G_{m,j}^{j-m}(u_j) \\
 &= (U_{m-1}^0 + G_{m-1,m}^1 U_m^0) + \sum_{j<m-1} (G_{m-1,j}^{j-m+1}(u_j) + G_{m-1,m}^1 G_{m,j}^{j-m}(u_j)) \\
 &\quad + G_{m-1,m}^1 G_{m,m-1}^{-1} u_{m-1}.
 \end{aligned}$$

Note that it is enough to assume $U_m^0 \in H^{s+1}$ and $U_{m-1}^0 \in H^s$ to obtain $U_{m-1}^0 + G_{m-1,m}^1 U_m^0 \in H^s$. Since all the operators above are of order ≤ 0 we conclude that the operator

$$L_{m-1} = I - G_{m-1,m}^1 G_{m,m-1}^{-1} := I - \mathcal{G}_{m-1}^0$$

is invertible on a sufficiently small interval $[0, T]$ and, therefore,

$$\begin{aligned}
 u_{m-1} - G_{m-1,m}^1 G_{m,m-1}^{-1} u_{m-1} &= (U_{m-1}^0 + G_{m-1,m}^1 U_m^0) \\
 &\quad + \sum_{j < m-1} (G_{m-1,j}^{j-m+1}(u_j) + G_{m-1,m}^1 G_{m,j}^{j-m}(u_j)),
 \end{aligned}
 \tag{22}$$

yields

$$u_{m-1} = L_{m-1}^{-1} \tilde{U}_{m-1}^0 + L_{m-1}^{-1} \sum_{j < m-1} \tilde{G}_{m-1}^{j-m+1} u_j, \tag{23}$$

with \tilde{U}_{m-1}^0 and \tilde{G}_{m-1}^{j-m+1} defined by the right-hand side of (22). We now substitute u_m and u_{m-1} into u_{m-2} making use of (23). We obtain

$$\begin{aligned}
 u_{m-2} &= U_{m-2}^0 + \sum_{j < m-2} G_{m-2,j}^{j-m+2}(u_j) \\
 &\quad + G_{m-2,m-1}^1(u_{m-1}) + G_{m-2,m}^1(u_m) \\
 &= U_{m-2}^0 + \sum_{j < m-2} G_{m-2,j}^{j-m+2}(u_j) + G_{m-2,m-1}^1 L_{m-1}^{-1} \tilde{U}_{m-1}^0 \\
 &\quad + G_{m-2,m-1}^1 L_{m-1}^{-1} \sum_{j < m-2} \tilde{G}_{m-1}^{j-m+1} u_j \\
 &\quad + G_{m-2,m-1}^1 L_{m-1}^{-1} \tilde{G}_{m-1}^{-1} u_{m-2} + G_{m-2,m}^1 U_m^0 \\
 &\quad + G_{m-2,m}^1 \sum_{j < m-2} G_{m,j}^{j-m}(u_j) + G_{m-2,m}^1 G_{m,m-2}^{-2} u_{m-2} \\
 &\quad + G_{m-2,m}^1 G_{m,m-1}^{-1} L_{m-1}^{-1} \tilde{U}_{m-1}^0 \\
 &\quad + G_{m-2,m}^1 G_{m,m-1}^{-1} L_{m-1}^{-1} \sum_{j < m-2} \tilde{G}_{m-1}^{j-m+1} u_j \\
 &\quad + G_{m-2,m}^1 G_{m,m-1}^{-1} L_{m-1}^{-1} \tilde{G}_{m-1}^{-1} u_{m-2}.
 \end{aligned}
 \tag{24}$$

We set

$$\begin{aligned}
 \tilde{U}_{m-2}^0 &= U_{m-2}^0 + G_{m-2,m-1}^1 L_{m-1}^{-1} \tilde{U}_{m-1}^0 \\
 &\quad + G_{m-2,m}^1 U_m^0 + G_{m-2,m}^1 G_{m,m-1}^{-1} L_{m-1}^{-1} \tilde{U}_{m-1}^0.
 \end{aligned}
 \tag{25}$$

The operators $G_{m-2,m-1}^1 L_{m-1}^{-1}$ and $G_{m-2,m}^1$ in (25) are of order 1. Keeping in mind that we already assumed $U_m^0 \in H^{s+1}$ and $U_{m-1}^0 \in H^s$, in order to obtain Sobolev

order s the initial data U_m^0, U_{m-1}^0 and U_{m-2}^0 must belong to H^{s+2}, H^{s+1} and H^s , respectively. Thus,

$$u_{m-2} = \tilde{U}_{m-2}^0 + \mathcal{G}_{m-2}^0 u_{m-2} + \sum_{j < m-2} \tilde{G}_{m-2}^{j-m+2} u_j, \tag{26}$$

where \mathcal{G}_{m-2}^0 is a zero order operator defined by

$$\begin{aligned} \mathcal{G}_{m-2}^0 u_{m-2} &= G_{m-2,m-1}^1 L_{m-1}^{-1} \tilde{G}_{m-1}^{-1} u_{m-2} + G_{m-2,m}^1 G_{m,m-2}^{-2} u_{m-2} \\ &\quad + G_{m-2,m}^1 G_{m,m-1}^{-1} L_{m-1}^{-1} \tilde{G}_{m-1}^{-1} u_{m-2}, \end{aligned}$$

and the last summand in (26) is obtained by collecting all the operators acting on u_j with $j < m - 2$ in (24). Since the norm of \mathcal{G}_{m-2}^0 can be taken strictly less than one in a sufficiently small interval $[0, T]$ we have that the operator

$$L_{m-2} = I - \mathcal{G}_{m-2}^0$$

is invertible and, therefore,

$$u_{m-2} = L_{m-2}^{-1} \tilde{U}_{m-2}^0 + \sum_{j < m-2} L_{m-2}^{-1} \tilde{G}_{m-2}^{j-m+2} u_j. \tag{27}$$

Note that $\tilde{U}_{m-2}^0 \in H^s$ if $U_m^0 \in H^{s+2}, U_{m-1}^0 \in H^{s+1}$ and $U_{m-2}^0 \in H^s$. By iterating the same procedure we deduce that

$$u_k = \tilde{U}_k^0 + \mathcal{G}_k^0 u_k + \sum_{j < k} \tilde{G}_k^{j-k} u_j, \tag{28}$$

where \tilde{U}_k^0 depends on U_k^0, U_j^0 and \tilde{U}_j^0 with $j > k$ and \mathcal{G}_k^0 is a zero order operator defined by using invertible operators $L_{m-1}, L_{m-2}, \dots, L_k$. In addition, we obtain $\tilde{U}_k^0 \in H^s$ since $U_m^0 \in H^{s+m-k}, U_{m-1}^0 \in H^{s+m-k-1}, \dots, U_k^0 \in H^s$. It follows that for $k = 2$ we have

$$u_2 = \tilde{U}_2^0 + \mathcal{G}_2^0 u_2 + \tilde{G}_2^{-1} u_1,$$

where the operator \mathcal{G}_2^0 is of zero operator and defined by invertible operators $L_{m-1}, L_{m-2}, \dots, L_2, \tilde{G}_2^{-1}$ is of order -1 , and $\tilde{U}_2^0 \in H^s$ since $U_m^0 \in H^{s+m-2}, U_{m-1}^0 \in H^{s+m-3}, \dots, U_2^0 \in H^s$. Hence, by inverting the operator $L_2 = I - \mathcal{G}_2^0$ on a sufficiently small interval $[0, T]$ we have

$$u_2 = L_2^{-1} \tilde{U}_2^0 + L_2^{-1} \tilde{G}_2^{-1} u_1.$$

Now by substitution of u_2, u_3, \dots, u_m in the equation of u_1 we arrive at the formula (28) with $k = 1$, i.e.,

$$u_1 = \tilde{U}_1^0 + \mathcal{G}_1^0 u_1,$$

where $\tilde{U}_1^0 \in H^s$ since $U_m^0 \in H^{s+m-1}, U_{m-1}^0 \in H^{s+m-2}, \dots, U_2^0 \in H^{s+1}, U_1^0 \in H^s$. Concluding, by the Banach fix point argument we prove that there exists a unique $u_1 \in C([0, T], H^s)$ solving the equation above with the given initial conditions. By substitution in the equations for u_2, \dots, u_{m-1}, u_m we arrive at the desired Sobolev well-posedness with $u_k \in C([0, T], H^{s+k-1})$ for $k = 2, \dots, m$. Note that, since the sufficiently small interval $[0, T]$ where we get well-posedness does not depend on the initial data, by a standard iteration argument we can achieve well-posedness on any bounded interval $[0, T]$ as stated in the theorem. \square

3 Schur decomposition of $m \times m$ matrices

In this section we investigate how to reduce an $m \times m$ matrix to the upper triangular form. We recall that such decomposition is well-known for constant matrices and goes under the name of Schur’s triangularisation, with its statement given in Theorem C.

One of the difficulties when dealing with variable multiplicities is the loss of regularity in the parameters at the points of multiplicities. In the following, we will assume that A is a matrix of (possibly) complex valued first order symbols, continuous with respect to t , i.e., $A(t, x, \xi) \in (CS^1)^{m \times m}$.

We will now develop a parameter dependent extension of the Schur triangularisation procedure and we will describe it step by step. Then we will give an example for it for the systems of low sizes, namely, for $m = 2$ and $m = 3$.

In the case of $m = 2$ the construction below was introduced in [27] and now we give its general version for systems of any size.

Normal forms of matrices depending on several parameters have a long history and are notoriously involved; for some remarks and related works, we refer the reader to [14, 24, 25, 45].

3.1 First step or Schur step

The first step in our triangularisation follows the construction in the constant case except that we will not get a unitary transformation matrix. For this reason we talk of a Schur step. Throughout this paper e_i denotes the i -th vector of the standard basis of \mathbb{R}^n with an appropriate dimension n .

Proposition 1 (Schur step) *Let the $m \times m$ matrix valued symbol $A(t, x, \xi) \in (CS^1)^{m \times m}$, have a real eigenvalue $\lambda \in CS^1$ and a corresponding eigenvector $h \in (CS^1)^m$ such that there exists $j \in \{1, \dots, m\}$ with*

$$\langle h(t, x, \xi) | e_j \rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}, \tag{29}$$

for a sufficiently large $M > 0$. Then there exist an $m \times m$ matrix valued symbol $T(t, x, \xi) \in (CS^0)^{m \times m}$, invertible for $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$ with $T^{-1} \in$

$(CS^0)^{m \times m}$, and an $(m - 1) \times (m - 1)$ matrix valued symbol $E(t, x, \xi) \in (CS^1)^{m \times m}$, such that

$$T^{-1}(t, x, \xi)A(t, x, \xi)T(t, x, \xi) = \begin{bmatrix} \lambda & a_{12} & \dots & a_{1m} \\ 0 & & & \\ \vdots & & E(t, x, \xi) & \\ 0 & & & \end{bmatrix}$$

for all $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$.

Proof First let us note that we can assume that $j = 1$ in (29). If that is not the case, we can exchange the rows 1 and j as well as columns 1 and j to move the j th component of the eigenvector to the first component.

We define the rescaled eigenvector μ componentwise by

$$\mu_i(t, x, \xi) = \frac{\langle h(t, x, \xi) | e_i \rangle}{\langle h(t, x, \xi) | e_1 \rangle} \quad \forall i = 1, \dots, m.$$

Now we set

$$T(t, x, \xi) = \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ \mu_2 & & & \\ \vdots & & I_{m-1} & \\ \mu_m & & & \end{bmatrix}.$$

Since $\mu_1 \equiv 1$ it follows that

$$T^{-1}(t, x, \xi) = \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ -\mu_2 & & & \\ \vdots & & I_{m-1} & \\ -\mu_m & & & \end{bmatrix},$$

where I_{m-1} is the $(m - 1) \times (m - 1)$ identity matrix. By direct computations we get

$$AT = \begin{bmatrix} \sum_{j=1}^m a_{1j} \mu_j & & & \\ \vdots & A_{(2)} & \dots & A_{(m)} \\ \sum_{j=1}^m a_{mj} \mu_j & & & \end{bmatrix} = \begin{bmatrix} \lambda \mu_1 & & & \\ \vdots & A_{(2)} & \dots & A_{(m)} \\ \lambda \mu_m & & & \end{bmatrix},$$

where we used that

$$\sum_{j=1}^m a_{ij} \mu_j = \lambda \mu_i, \quad i = 1, \dots, m, \tag{30}$$

and denoted the i th column of A by $A_{(i)}$. The equations in (30) are given by the eigenvalue equation $A\mu = \lambda\mu$. Further, from $\mu_1 \equiv 1$ we obtain

$$\begin{aligned}
 T^{-1}AT &= \begin{bmatrix} \lambda\mu_1^2 & a_{12}\mu_1 & \dots & a_{1m}\mu_1 \\ -\mu_2\mu_1\lambda + \mu_2\lambda & & & \\ \vdots & & E & \\ -\mu_m\mu_1\lambda + \mu_m\lambda & & & \end{bmatrix} \\
 &= \begin{bmatrix} \lambda & a_{12} & \dots & a_{1m} \\ 0 & & & \\ \vdots & & E & \\ 0 & & & \end{bmatrix}, \tag{31}
 \end{aligned}$$

which concludes the proof. Note that by construction the matrix E has entries in CS^1 which depend on A . In particular its eigenvalues are the eigenvalues of A excluding λ (counted as many times as they occur). \square

Applying Proposition 1 repeatedly for $m - 2$ times to E , we obtain a full Schur transformation of A , that is a full reduction to an upper triangular form. In the next subsection we describe this iteration in detail. This triangularisation procedure is summarised in Theorem 6 where sufficient conditions on the eigenvectors of A are given.

3.2 The triangularisation procedure

The reduction to an upper triangular form or the Schur transformation of A is possible under certain conditions on its eigenvectors. More precisely, let

$$h_1(t, x, \xi), \dots, h_{m-1}(t, x, \xi) \in (CS^0)^m$$

be $m - 1$ eigenvectors of $A(t, x, \xi) = [a_{ij}(t, x, \xi)]_{i,j=1}^m$, $a_{ij} \in CS^1$, corresponding to the eigenvalues $\lambda_1(t, x, \xi), \dots, \lambda_{m-1}(t, x, \xi) \in CS^1$. To formulate the sufficient conditions for the existence of such Schur transformation, we introduce a set of auxiliary vectors $h^{(i)}$, $i = 1, \dots, m - 1$, which depend only on h_i and the previous vectors $h^{(j)} \in CS^0$, $j = 1, \dots, i - 1$. When $i = 1$ we set $h^{(1)} = h_1$.

As in Proposition 1 we begin by assuming

$$\langle h^{(1)}(t, x, \xi) | e_1 \rangle \neq 0 \tag{32}$$

for $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$.

Remark 5 As noted in the proof of Proposition 1, we could have that

$$\langle h^{(1)}(t, x, \xi) | e_j \rangle \neq 0$$

for another arbitrary $j \in \{1, \dots, m\}$. Then, we could transform the matrix $A(t, x, \xi)$ by a constant permutation matrix P such that $P^{-1}h^{(1)}$ is eigenvector of $P^{-1}AP$ corresponding to λ_1 which satisfies $\langle P^{-1}h^{(1)}(t, x, \xi)|e_1 \rangle \neq 0$. For this reason we state (32) with $h^{(1)}$ and e_1 .

Step 1 By Proposition 1 there exists a matrix T_1 such that

$$T_1^{-1}AT_1 = \begin{bmatrix} \lambda_1 & a_{12} & \dots & a_{1m} \\ 0 & & & \\ \vdots & & E_{m-1} & \\ 0 & & & \end{bmatrix}.$$

The matrix T_1 is given by

$$T_1 = [\omega_1 \quad e_2 \quad \dots \quad e_m], \quad \omega_1 = [\omega_{11} \quad \dots \quad \omega_{1m}]^T$$

with

$$\omega_{1j} = \frac{\langle h^{(1)}(t, x, \xi)|e_j \rangle}{\langle h^{(1)}(t, x, \xi)|e_1 \rangle}.$$

In the sequel we make use of the projector $\Pi_k : \mathbb{R}^m \rightarrow \mathbb{R}^{m-k}, 0 \leq k \leq m - 1$, defined by

$$\Pi_k \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} x_{k+1} \\ \vdots \\ x_m \end{bmatrix}.$$

Note that Π_0 is the identity map $I_m : \mathbb{R}^m \rightarrow \mathbb{R}^m$.

Step 2 Since h_2 is an eigenvector of A with eigenvalue λ_2 we get that $T_1^{-1}h_2$ is an eigenvector of $T_1^{-1}AT_1$ with eigenvalue λ_2 as well. By the structure of $T_1^{-1}AT_1$ we easily see that $h^{(2)} := \Pi_1 T_1^{-1}h_2$ is an eigenvector of E_{m-1} , corresponding to λ_2 .

Arguing as in Remark 5 we assume that

$$\langle \Pi_1 T_1^{-1}h_2|e_1 \rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}, \tag{33}$$

to be able to apply Proposition 1 to E_{m-1} . We get that there exists an $(m - 1) \times (m - 1)$ matrix \tilde{T}_2 such that $\tilde{T}_2^{-1}E_{m-1}\tilde{T}_2$ is of form

$$\begin{bmatrix} \lambda_2 & * & \dots & * \\ 0 & & & \\ \vdots & & E_{m-2} & \\ 0 & & & \end{bmatrix},$$

where in the first row the first row of E_{m-1} appears. Thus, setting

$$T_2 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & \tilde{T}_2 & & \\ 0 & & & \end{bmatrix},$$

we obtain

$$T_2^{-1}T_1^{-1}AT_1T_2 = \begin{bmatrix} \lambda_1 & * & * & \dots & * \\ 0 & \lambda_2 & * & \dots & * \\ 0 & 0 & & & \\ \vdots & \vdots & & E_{m-2} & \\ 0 & 0 & & & \end{bmatrix}. \tag{34}$$

Note that in (34) we write explicitly only the entries most relevant to our triangularisation. To compute the matrix \tilde{T}_2 , we set

$$\omega_2 = [\omega_{22} \dots \omega_{2m}]^T,$$

where

$$\omega_{2j}(t, x, \xi) := \frac{\langle h^{(2)}(t, x, \xi) | e_j \rangle}{\langle h^{(2)}(t, x, \xi) | e_1 \rangle}, \quad j = 2, \dots, m,$$

and then

$$\tilde{T}_2 = [\omega_2 \ e_2 \ \dots \ e_{m-1}].$$

It is clear that T_2 has the same structure as T_1 , i.e., it is defined via a rescaled eigenvector as the first column and an identity matrix (I_{m-1} for T_1 and I_{m-2} for T_2).

Step K By iterating the method $k - 1$ times we can find $k - 1$ matrices T_1, T_2, \dots, T_{k-1} of size $m \times m$ such that

$$T_{k-1}^{-1} \cdot \dots \cdot T_1^{-1}AT_1 \cdot \dots \cdot T_{k-1} = \begin{bmatrix} \lambda_1 & * & * & \dots & \dots & * \\ 0 & \ddots & * & \dots & \dots & * \\ 0 & 0 & \lambda_{k-1} & * & \dots & * \\ 0 & 0 & 0 & & & \\ \vdots & \vdots & \vdots & & E_{m-k+1} & \\ 0 & 0 & 0 & & & \end{bmatrix},$$

where E_{m-k+1} is a $(m - k + 1) \times (m - k + 1)$ matrix and the equality is true on $[0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$. Since h_k is an eigenvector of A corresponding to λ_k , the vector

$$T_{k-1}^{-1} T_{k-2}^{-1} \cdots T_1^{-1} h_k$$

is an eigenvector of

$$T_{k-1}^{-1} T_{k-2}^{-1} \cdots T_1^{-1} A T_1 T_2 \cdots T_{k-1}$$

and

$$h^{(k)} := \Pi_{k-1} T_{k-1}^{-1} T_{k-2}^{-1} \cdots T_1^{-1} h_k \in (CS^0)^{m-k+1}$$

an eigenvector of E_{m-k+1} corresponding to λ_k . Thus, to satisfy the assumptions of Proposition 1 and keeping in mind Remark 5, we require that

$$\left\langle h^{(k)}(t, x, \xi) | e_1 \right\rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}. \tag{35}$$

It follows that there exists an $(m - k + 1) \times (m - k + 1)$ transformation matrix \tilde{T}_k such that $\tilde{T}_k^{-1} \dots \tilde{T}_1^{-1} A \tilde{T}_1 \dots \tilde{T}_k$ is of the form

$$\begin{bmatrix} \lambda_k & * & \dots & * \\ 0 & & & \\ \vdots & & E_{m-k} & \\ 0 & & & \end{bmatrix}.$$

and set

$$T_k = \begin{bmatrix} I_{k-1} & \mathbf{0} \\ \mathbf{0} & \tilde{T}_k \end{bmatrix}.$$

The matrix \tilde{T}_k is defined by

$$\tilde{T}_k = [\omega_k \quad e_2 \quad \dots \quad e_{m-k+1}], \quad \omega_k = [\omega_{kk} \quad \dots \quad \omega_{km}]^T,$$

where

$$\omega_{kj} = \frac{\langle h^{(k)}(t, x, \xi) | e_j \rangle}{\langle h^{(k)}(t, x, \xi) | e_1 \rangle}, \quad j = k, \dots, m.$$

Step m-1 This is the last step as E_2 is a 2×2 matrix. We have that

$$h^{(m-1)} = \Pi_{m-2} T_{m-2}^{-1} \cdots T_1^{-1} h_{m-1} \in (CS^0)^2$$

is an eigenvector of E_2 corresponding to λ_{m-1} and that \tilde{T}_{m-1} exists as before if

$$\langle h^{(m-1)}(t, x, \xi) | e_1 \rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}. \quad (36)$$

The matrix \tilde{T}_{m-1} is given by

$$\tilde{T}_{m-1} = [\omega_{m-1} \ e_2] = \begin{bmatrix} \omega_{m-1,m-1} & 0 \\ \omega_{m-1,m} & 1 \end{bmatrix},$$

where

$$\omega_{m-1,j} = \frac{\langle h^{(m-1)}(t, x, \xi) | e_j \rangle}{\langle h^{(m-1)}(t, x, \xi) | e_1 \rangle}, \quad j = m - 1, m,$$

and then

$$T_{m-1} = \begin{bmatrix} I_{m-2} & \mathbf{0} \\ \mathbf{0} & \tilde{T}_{m-1} \end{bmatrix}.$$

We are now ready to state Theorem 6 which summarises the triangularisation procedure explained above. For the convenience of the reader we recall the notations introduced so far:

- h_1, \dots, h_{m-1} are the eigenvectors of the matrix A corresponding to the eigenvalues $\lambda_1, \dots, \lambda_{m-1}$.
- $h^{(1)} = h_1$ and

$$h^{(i)} = \Pi_{i-1} T_{i-1}^{-1} T_{i-2}^{-1} \cdot \dots \cdot T_1^{-1} h_i \in (C S^0)^{m-k+1}, \quad (37)$$

for $i = 2, \dots, m - 1$.

- the matrices T_k are inductively defined as follows: $T_0 = I_m$ and

$$T_k = \begin{bmatrix} I_{k-1} & \mathbf{0} \\ \mathbf{0} & \tilde{T}_k \end{bmatrix}, \quad \tilde{T}_k = [\omega_k \ e_2 \ \dots \ e_{m-k}], \quad e_i \in \mathbb{R}^{m-k},$$

where

$$\omega_{kj} = \frac{\langle h^{(k)}(t, x, \xi) | e_j \rangle}{\langle h^{(k)}(t, x, \xi) | e_1 \rangle}, \quad j = k, \dots, m.$$

Finally, we note that $h^{(k)}$ depends only on T_{k-1}, \dots, T_1 and, thus, only on the eigenvectors $h^{(k-1)}, \dots, h^{(1)}$.

Summarising, we can formulate a more precise version of Theorem 2.

Theorem 6 (Schur Decomposition) *Let $A(t, x, \xi) \in (CS^1)^{m \times m}$ be a matrix with eigenvalues $\lambda_1, \dots, \lambda_m \in CS^1$, and let $h_1, \dots, h_{m-1} \in (CS^0)^m$ be the corresponding eigenvectors. Suppose that for $e_1 \in \mathbb{R}^{m-i+1}$ the condition*

$$\left\langle h^{(i)}(t, x, \xi) | e_1 \right\rangle \neq 0, \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \tag{38}$$

holds for all $i = 1, \dots, m - 1$, with the notation explained above. Then, there exists a matrix-valued symbol $T(t, x, \xi) \in (CS^0)^{m \times m}$, invertible for $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$, $T^{-1}(t, x, \xi) \in (CS^0)^{m \times m}$, such that

$$T^{-1}(t, x, \xi)A(t, x, \xi)T(t, x, \xi) = \Lambda(t, x, \xi) + N(t, x, \xi)$$

for all $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$, where

$$\Lambda(t, x, \xi) = \text{diag}(\lambda_1(t, x, \xi), \lambda_2(t, x, \xi), \dots, \lambda_m(t, x, \xi))$$

and

$$N(t, x, \xi) = \begin{bmatrix} 0 & N_{12}(t, x, \xi) & N_{13}(t, x, \xi) & \dots & N_{1m}(t, x, \xi) \\ 0 & 0 & N_{23}(t, x, \xi) & \dots & N_{2m}(t, x, \xi) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & N_{m-1m}(t, x, \xi) \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

and N is a nilpotent matrix with entries in CS^1 . Furthermore, the matrix symbol T is given by

$$T(t, x, \xi) = T_1 T_2 \cdots T_{m-1},$$

with the notation explained above.

Remark 6 Taking into account Remark 5, let us stress that condition (38) is not restrictive as it can be replaced by the following: suppose that there exist $m - 1$ numbers $j_i \in \{1, \dots, m - i + 1\}$, $i = 1, \dots, m - 1$, such that for all $i = 1, \dots, m - 1$

$$\left\langle h^{(i)}(t, x, \xi) | e_{j_i} \right\rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\} \tag{39}$$

holds.

Remark 7 If $A(t, x, \xi)$ has complex symbols (as allowed in Theorem 1, see also Remark 1) and real eigenvalues, the eigenvalues of the Schur transformed system clearly remain real. The upper triangular entries may still be complex valued symbols.

Remark 8 Theorem 6 is quite general in the sense that the functions a_{ij} could be complex-valued. In this paper, we are concerned with hyperbolic matrices, i.e. we

assume that the eigenvalues $\lambda_1, \dots, \lambda_m$ are real. We stress that the Schur transform does not change the hyperbolicity of the matrix as the eigenvalues of $T^{-1}AT$ are also $\lambda_1, \dots, \lambda_m$.

Remark 9 For our applications in this and future work it is important that the transform T in Theorem 6 keeps the regularity of the original matrix A , i.e. that the elements of the Schur transform $T^{-1}AT$ are in the same class as the elements of A . Here, we stated everything with CS^1 and CS^0 as that is the regularity considered in this paper. Note that one could replace C with C^k or C^∞ and find a matrix T such that the transformed matrix $T^{-1}AT$ inherits the same regularity with respect to t . In addition, one could also drop the regularity in t to L^∞ and the triangularisation procedure would still work preserving the boundedness in t through every step.

For the sake of simplicity and the reader’s convenience, in the next subsections we analyse Theorem 6 in the special cases of $m = 2$ and $m = 3$.

3.3 The case $m = 2$

We now formulate Theorem 6 in the special case $m = 2$. In this way we recover the formulation given in [27].

Theorem 7 ([27, Theorem 7.1]) *Suppose that $A(t, x, \xi) \in (CS^1)^{2 \times 2}$ admits eigenvalues $\lambda_j(t, x, \xi) \in CS^1, j = 1, 2$, and an eigenvector $h(t, x, \xi) \in (CS^0)^2$ satisfying*

$$\langle h(t, x, \xi) | e_j \rangle \neq 0, \quad (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}, \tag{40}$$

for $j = 1$ or $j = 2$. Then, we can find a 2×2 matrix valued symbol $T(t, x, \xi) \in (CS^0)^{2 \times 2}$, invertible for $\{|\xi| \geq M\}$ with $T^{-1}(t, x, \xi) \in (CS^0)^{2 \times 2}$, such that

$$T^{-1}(t, x, \xi)A(t, x, \xi)T(t, x, \xi) = \begin{bmatrix} \lambda_1(t, x, \xi) & a_{12}(t, x, \xi) \\ 0 & \lambda_2(t, x, \xi) \end{bmatrix}$$

for all $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$.

Proof For 2×2 matrices the triangularisation procedure described in the previous subsection can stop at Step 1. By Remark 5, we may assume that (40) holds for the eigenvector h corresponding to λ_1 and for $j = 1$. We set $h = h_1$ and $h^{(1)} = h_1$. The vector

$$\omega_1 = \begin{bmatrix} \omega_{11}(t, x, \xi) \\ \omega_{12}(t, x, \xi) \end{bmatrix}, \quad \omega_{1j}(t, x, \xi) = \frac{\langle h^{(1)}(t, x, \xi) | e_j \rangle}{\langle h^{(1)}(t, x, \xi) | e_1 \rangle},$$

belongs to CS^0 and is an eigenvector of A associated to λ_1 . We then set

$$T_1(t, x, \xi) = [\omega_1 \quad e_2] = \begin{bmatrix} \omega_{11}(t, x, \xi) & 0 \\ \omega_{12}(t, x, \xi) & 1 \end{bmatrix}.$$

With that, we obtain

$$A(t, x, \xi)T_1(t, x, \xi) = \begin{bmatrix} a_{11}\omega_{11} + a_{12}\omega_{12} & a_{12} \\ a_{21}\omega_{11} + a_{22}\omega_{12} & a_{22} \end{bmatrix}$$

and finally, with

$$T_1^{-1}(t, x, \xi) = \begin{bmatrix} \omega_{11}(t, x, \xi) & 0 \\ -\omega_{12}(t, x, \xi) & 1 \end{bmatrix},$$

we obtain

$$\begin{aligned} & T_1^{-1}(t, x, \xi)A(t, x, \xi)T_1(t, x, \xi) \\ &= \begin{bmatrix} a_{11}\omega_{11}^2 + a_{12}\omega_{12}\omega_{11} & a_{12}\omega_{11} \\ -a_{11}\omega_{12}\omega_{11} - a_{12}\omega_{12}^2 + a_{21}\omega_{11} + a_{22}\omega_{12} & -\omega_{12}a_{12} + a_{22} \end{bmatrix} \end{aligned}$$

By construction, we have

$$\begin{aligned} a_{11}\omega_{11} + a_{12}\omega_{12} &= \lambda_1\omega_{11}, \\ a_{21}\omega_{11} + a_{22}\omega_{12} &= \lambda_1\omega_{12}, \end{aligned}$$

and $\omega_1 = 1$. This yields $a_{11}\omega_{11} + a_{12}\omega_{12} = \lambda_1\omega_{11} = \lambda_1$ and

$$\begin{aligned} & -a_{11}\omega_{12}\omega_{11} - a_{12}\omega_{12}^2 + a_{21}\omega_{11} + a_{22}\omega_{12} \\ &= -\omega_{12}(a_{11}\omega_{11} + a_{12}\omega_{12}) + a_{21}\omega_{11} + a_{22}\omega_{12} = -\lambda_1\omega_{12} + \lambda_1\omega_{12} = 0. \end{aligned}$$

Using $a_{11} + a_{22} = \lambda_1 + \lambda_2$, we obtain

$$-\omega_{12}a_{12} + a_{22} = -\omega_{12}a_{12} + a_{22} + a_{11}\omega_{11} - a_{11}\omega_{11} = \lambda_2.$$

Thus, we get that

$$T_1^{-1}(t, x, \xi)A(t, x, \xi)T_1(t, x, \xi) = \begin{bmatrix} \lambda_1(t, x, \xi) & a_{12}(t, x, \xi) \\ 0 & \lambda_2(t, x, \xi) \end{bmatrix}$$

for $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$. This concludes the proof. \square

3.3.1 Example

(i) By direct computations we can easily see that if $h_1 = [h_{11} \ h_{12}]^T = e_1$ then the matrix A is automatically in the upper triangular form. Indeed,

$$a_{21}h_{11} + a_{22}h_{12} = \lambda_1h_{12}$$

implies $a_{21} = 0$. A typical example (already discussed in [27]) is the Jordan block matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

where $\lambda_1 = 0$ is an eigenvalue with eigenvector $h_1 = e_1$.

(ii) Condition (40) is trivially fulfilled when $\det A \equiv 0$ and A is of the form

$$\begin{bmatrix} a & a \\ -a & -a \end{bmatrix},$$

for $a = a(t, x, \xi)$. Indeed, also in this case one can take 0 as an eigenvalue with eigenvector $h_1 = [1 \ 1]^T$.

3.4 The case $m = 3$

With the notation introduced in Sect. 3.2, we assume that the 3×3 matrix $A(t, x, \xi) \in (CS^1)^{3 \times 3}$ admits three eigenvalues $\lambda_i(t, x, \xi) \in CS^1, i = 1, 2, 3$, and two corresponding eigenvectors $h_i(t, x, \xi) \in (CS^0)^3, i = 1, 2$. Then, we set $h^{(1)} := h_1$ and, as in Remark 6 we suppose that there is a $j_1 \in \{1, 2, 3\}$ with

$$\langle h^{(1)}(t, x, \xi) | e_{j_1} \rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}.$$

Thus, we can set

$$\omega_{1j}(t, x, \xi) = \frac{\langle h^{(1)}(t, x, \xi) | e_j \rangle}{\langle h^{(1)}(t, x, \xi) | e_{j_1} \rangle}.$$

Now, we rearrange the matrix A such that the first component of ω_1 becomes identically equal to 1. Then, with $j_2, j_3 \in \{1, 2, 3\} \setminus \{j_1\}$, we can write

$$T_1^{-1} = [\omega_1 \ e_2 \ e_3]^{-1} = \begin{bmatrix} \omega_{1j_1} & 0 & 0 \\ -\omega_{1j_2} & 1 & 0 \\ -\omega_{1j_3} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\omega_{1j_2} & 1 & 0 \\ -\omega_{1j_3} & 0 & 1 \end{bmatrix},$$

which leads to

$$T_1^{-1}h_2 = \begin{bmatrix} \omega_{1j_1} & 0 & 0 \\ -\omega_{1j_2} & 1 & 0 \\ -\omega_{1j_3} & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{2j_1} \\ h_{2j_2} \\ h_{2j_3} \end{bmatrix} = \begin{bmatrix} h_{2j_1} \\ -\omega_{1j_2}h_{2j_1} + h_{2j_2} \\ -\omega_{1j_3}h_{2j_1} + h_{2j_3} \end{bmatrix}.$$

We then get

$$h^{(2)} = \Pi_1 T_1^{-1}h_2 = \begin{bmatrix} -\omega_{1j_2}h_{2j_1} + h_{2j_2} \\ -\omega_{1j_3}h_{2j_1} + h_{2j_3} \end{bmatrix}$$

and the condition (38) that there exists $j \in \{1, 2\}$ such that

$$\langle h^{(2)}(t, x, \xi) | e_j \rangle \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$$

translates to: either

$$-\omega_{1j_2} h_{2j_1} + h_{2j_2} \neq 0 \quad \Rightarrow \quad h_{2j_2} h_{1j_1} - h_{1j_2} h_{2j_1} \neq 0 \tag{41}$$

or

$$-\omega_{1j_3} h_{2j_1} + h_{2j_3} \neq 0 \quad \Rightarrow \quad h_{2j_3} h_{1j_1} - h_{1j_3} h_{2j_1} \neq 0 \tag{42}$$

holds. Thus, assuming that (41) holds, the matrix \tilde{T}_2 is given by

$$\begin{bmatrix} \omega_{21} & 0 \\ \omega_{21} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{-\omega_{1j_3} h_{2j_1} + h_{2j_3}}{-\omega_{1j_2} h_{2j_1} + h_{2j_2}} & 1 \end{bmatrix}, \quad \omega_{2j} = \frac{\langle h^{(2)}(t, x, \xi) | e_j \rangle}{\langle h^{(2)}(t, x, \xi) | e_{j_2} \rangle}, \quad j = 1, 2,$$

and the matrix T_2 by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega_{21} & 0 \\ 0 & \omega_{22} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{-\omega_{1j_3} h_{2j_1} + h_{2j_3}}{-\omega_{1j_2} h_{2j_1} + h_{2j_2}} & 1 \end{bmatrix}.$$

Thus, we obtain

$$T(t, x, \xi) = T_1 T_2 = \begin{bmatrix} 1 & 0 & 0 \\ \omega_{1j_2} & 1 & 0 \\ \omega_{1j_3} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{-\omega_{1j_3} h_{2j_1} + h_{2j_3}}{-\omega_{1j_2} h_{2j_1} + h_{2j_2}} & 1 \end{bmatrix}. \tag{43}$$

If we have (42) instead of (41), then we would need a permutation matrix

$$P_{j_2 \leftrightarrow j_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

in (43), i.e.

$$T(t, x, \xi) = T_1(t, x, \xi) P_{j_2 \leftrightarrow j_3} T_2(t, x, \xi)$$

and

$$T_2(t, x, \xi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{-\omega_{1j_2}h_{2j_1}+h_{2j_2}}{-\omega_{1j_3}h_{2j_1}+h_{2j_3}} & 1 \end{bmatrix}.$$

See also Remark 6.

Thus, we can state

Theorem 8 *Suppose that $A(t, x, \xi) \in (CS^1)^{3 \times 3}$ admits three eigenvalues $\lambda_i \in CS^1$, $i = 1, 2, 3$, and two corresponding eigenvectors $h_i(t, x, \xi) \in (CS^1)^3$, $i = 1, 2$. Suppose that there exists a $j_1 \in \{1, 2, 3\}$ such that*

$$h_{1j_1}(t, x, \xi) \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}. \quad (44)$$

Further suppose that there exists $j_2 \in \{1, 2, 3\} \setminus \{j_1\}$ such that

$$h_{2j_2}h_{1j_1} - h_{1j_2}h_{2j_1} \neq 0 \quad \forall (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}. \quad (45)$$

Then, there exists a matrix-valued symbol $T(t, x, \xi) \in (CS^0)^{3 \times 3}$, invertible for all $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$ with $T^{-1}(t, x, \xi) \in (CS^0)^{3 \times 3}$, such that

$$T^{-1}(t, x, \xi)A(t, x, \xi)T(t, x, \xi) = \Lambda(t, x, \xi) + N(t, x, \xi)$$

holds for all $(t, x, \xi) \in [0, T] \times \mathbb{R}^n \times \{|\xi| \geq M\}$, where $\Lambda(t, x, \xi) = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ and

$$N(t, x, \xi) = \begin{bmatrix} 0 & N_{13}(t, x, \xi) & N_{13}(t, x, \xi) \\ 0 & 0 & N_{23}(t, x, \xi) \\ 0 & 0 & 0 \end{bmatrix}.$$

We end this subsection by discussing some examples of 3×3 matrices fulfilling the assumptions above on their eigenvalues.

3.4.1 Examples

(i) If the matrix A has eigenvectors

$$h_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad h_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

then conditions (44) and (45) are easily fulfilled with $j_1 = 1$ and $j_2 = 2$. Indeed, $h_{11} = 1$ and

$$h_{22}h_{11} - h_{12}h_{21} = h_{22}h_{11} = 1.$$

More in general to satisfy (44) and (45) it would be enough to have two eigenvectors

$$h_1 = \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \end{bmatrix} \quad \text{and} \quad h_2 = \begin{bmatrix} h_{21} \\ h_{22} \\ h_{23} \end{bmatrix}$$

with $h_{11} \neq 0$, $h_{22} \neq 0$ and $h_{12} = 0$.

(ii) A matrix with eigenvectors

$$h_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad h_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

has a special form. Indeed, for λ_1 and λ_2 eigenvalues corresponding to h_1 and h_2 , respectively, by using the eigenvector equations we obtain

$$\begin{aligned} a_{13} &= \lambda_1 - a_{11}, \\ a_{23} &= -a_{21}, \\ a_{33} &= \lambda_1 - a_{31}, \end{aligned}$$

and

$$\begin{aligned} a_{12} &= \lambda_2 - a_{11}, \\ a_{22} &= \lambda_2 - a_{21}, \\ a_{32} &= -a_{31}. \end{aligned}$$

Hence

$$A = \begin{bmatrix} a_{11} & \lambda_2 - a_{11} & \lambda_1 - a_{11} \\ a_{21} & \lambda_2 - a_{21} & -a_{21} \\ a_{31} & -a_{31} & \lambda_1 - a_{31} \end{bmatrix}.$$

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Bernstein, D.S.: *Matrix Mathematics—Theory, Facts, and Formulas*, 2nd edn. Princeton University Press, Princeton (2009)
2. Bronshtein, M.D.: Smoothness of roots of polynomials depending on parameters. *Sibirsk. Mat. Zh.*, 20(3), 493–501, (1979). *Sib. Math. J.* **20**, 347–352 (1980). (English transl)
3. Colombini, F., Kinoshita, T.: On the Gevrey well posedness of the Cauchy problem for weakly hyperbolic equations of higher order. *J. Differ. Equ.* **186**, 394–419 (2002). [https://doi.org/10.1016/S0022-0396\(02\)00009-8](https://doi.org/10.1016/S0022-0396(02)00009-8)

4. Colombini, F., Spagnolo, S.: An example of a weakly hyperbolic Cauchy problem not well posed in C^∞ . *Acta Math.* **148**, 243–253 (1982). <https://doi.org/10.1007/BF02392730>
5. Colombini, F., De Giorgi, E., Spagnolo, S.: Sur les équations hyperboliques avec des coefficients qui ne dépendent que du temps. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **6**, 511–559 (1979)
6. Colombini, F., Del Santo, D., Fanelli, F., Métivier, G.: Time-dependent loss of derivatives for hyperbolic operators with non regular coefficients. *Commun. Partial Differ. Equ.* **38**(10), 1791–1817 (2013). <https://doi.org/10.1080/03605302.2013.795968>
7. Colombini, F., Del Santo, D., Fanelli, F., Métivier, G.: A well-posedness result for hyperbolic operators with Zygmund coefficients. *J. Math. Pures Appl.* **9**(100), 455–475 (2013). <https://doi.org/10.1016/j.matpur.2013.01.009>
8. Colombini, F., Jannelli, E., Spagnolo, S.: Nonuniqueness in hyperbolic Cauchy problems. *Ann. Math.* **126**, 495–524 (1987). <https://doi.org/10.2307/1971359>
9. Colombini, F., Lerner, N.: Hyperbolic operators with non-Lipschitz coefficients. *Duke Math. J.* **77**(3), 657–698 (1995). <https://doi.org/10.1215/S0012-7094-95-07721-7>
10. Colombini, F., Nishitani, T.: Second order weakly hyperbolic operators with coefficients sum of powers of functions. *Osaka J. Math.* **44**(1), 121–137 (2007)
11. D’Ancona, P., Kinoshita, T.: On the wellposedness of the Cauchy problem for weakly hyperbolic equations of higher order. *Math. Nachr.* **278**, 1147–1162 (2005). <https://doi.org/10.1002/mana.200310299>
12. D’Ancona, P., Kinoshita, T., Spagnolo, S.: Weakly hyperbolic systems with Hölder continuous coefficients. *J. Differ. Equ.* **203**(1), 64–81 (2004). <https://doi.org/10.1016/j.jde.2004.03.016>
13. D’Ancona, P., Kinoshita, T., Spagnolo, S.: On the 2 by 2 weakly hyperbolic systems. *Osaka J. Math.* **45**(4), 921–939 (2008)
14. Dieci, L., Eirola, T.: On smooth decompositions of matrices. *SIAM J. Matrix Anal. Appl.* **20**(3), 800–819 (1999). <https://doi.org/10.1137/S0895479897330182>
15. Duistermaat, J.J.: *Fourier Intergal Operators Progress in Mathematics*, vol. 130. Birkhäuser Boston, Inc, Boston (1996)
16. Garetto, C.: On hyperbolic equations and systems with non-regular time dependent coefficients. *J. Differ. Equ.* **259**(11), 5846–5874 (2015). <https://doi.org/10.1016/j.jde.2015.07.011>
17. Garetto, C., Jäh, C.: Well-posedness of hyperbolic systems with multiplicities and smooth coefficients. *Math. Ann.* **369**(1–2), 441–485 (2017). <https://doi.org/10.1007/s00208-016-1436-8>
18. Garetto, C., Ruzhansky, M.: Well-posedness of weakly hyperbolic equations with time dependent coefficients. *J. Differ. Equ.* **253**(5), 1317–1340 (2012). <https://doi.org/10.1016/j.jde.2012.05.001>
19. Garetto, C., Ruzhansky, M.: Weakly hyperbolic equations with non-analytic coefficients and lower order terms. *Math. Ann.* **357**(2), 401–440 (2013). <https://doi.org/10.1007/s00208-013-0910-9>
20. Garetto, C., Ruzhansky, M.: A note on weakly hyperbolic equations with analytic principal part. *J. Math. Anal. Appl.* **412**(1), 1–14 (2014). <https://doi.org/10.1016/j.jmaa.2013.09.011>
21. Garetto, C., Ruzhansky, M.: Hyperbolic second order equations with non-regular time dependent coefficients. *Arch. Ration. Mech. Anal.* **217**(1), 113–154 (2015). <https://doi.org/10.1007/s00205-014-0830-1>
22. Garetto, C., Ruzhansky, M.: On hyperbolic systems with time dependent Hölder characteristics. *Ann. Mat. Pura Appl.* **196**(1), 155–164 (2017). <https://doi.org/10.1007/s10231-016-0567-6>
23. Garetto, C., Ruzhansky, M.: On C^∞ well-posedness of hyperbolic systems with multiplicities. *Ann. Mat. Pura Appl.* **196**(5), 1819–1834 (2017). <https://doi.org/10.1007/s10231-017-0639-2>
24. Gingold, H.: On continuous triangularization of matrix functions. *SIAM J. Math. Anal.* **10**(4), 709–720 (1979). <https://doi.org/10.1137/0510065>
25. Gingold, H., Hsieh, P.-F.: Globally analytic triangularization of a matrix function. *Linear Algebra Appl.* **169**, 75–101 (1992). [https://doi.org/10.1016/0024-3795\(92\)90172-7](https://doi.org/10.1016/0024-3795(92)90172-7)
26. Gramchev, T., Orrù, N.: Cauchy problem for a class of nondiagonalizable hyperbolic systems. *Discret. Contin. Dyn. Syst.*, 533–542 (2011). <https://doi.org/10.3934/proc.2011.2011.533>
27. Gramchev, T., Ruzhansky, M.: Cauchy problem for 2×2 hyperbolic systems of pseudo-differential equations with nondiagonalisable principal part. *Studies in phase space analysis with applications to PDEs. Progr. Nonlinear Differ. Equ. Appl.* **84**, 129–144 (2013)
28. Hörmander, L.: *The Analysis of Linear Partial Differential Operators*, vol. I–IV. Springer, Heidelberg (1985)
29. Hörmander, L.: Hyperbolic systems with double characteristics. *Comm. Pure Appl. Math.* **46**, 261–301 (1993). <https://doi.org/10.1002/cpa.3160460207>

30. Ivrii, V.Ya., Petkov, V.M.: Necessary conditions for the correctness of the Cauchy problem for non-strictly hyperbolic equations. (Russian). *Russ. Math. Surv.* **29**, 3–70 (1974). <https://doi.org/10.1070/RM1974v029n05ABEH001295>
31. Kajitani, K., Yuzawa, Y.: The Cauchy problem for hyperbolic systems with Hölder continuous coefficients with respect to the time variable. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **5**(4), 465–482 (2006)
32. Kamotski, I., Ruzhansky, M.: Estimates and spectral asymptotics for systems with multiplicities. *Funct. Anal. Appl.* **39**, 308–310 (2005). <https://doi.org/10.1007/s10688-005-0052-2>
33. Kamotski, I., Ruzhansky, M.: Regularity properties, representation of solutions and spectral asymptotics of systems with multiplicities. *Comm. Partial Differ. Equ.* **32**, 1–35 (2007). <https://doi.org/10.1080/03605300600856816>
34. Kinoshita, T., Spagnolo, S.: Hyperbolic equations with non-analytic coefficients. *Math. Ann.* **336**, 551–569 (2006). <https://doi.org/10.1007/s00208-006-0009-7>
35. Lax, P.: Asymptotic solutions of oscillatory initial value problems. *Duke Math. J.* **24**, 627–646 (1957). <https://doi.org/10.1215/S0012-7094-57-02471-7>
36. Melrose, R.B., Uhlmann, G.A.: Microlocal structure of involutiveconical refraction. *Duke Math. J.* **46**, 571–582 (1982). <https://doi.org/10.1215/S0012-7094-79-04630-1>
37. Nishitani, T.: On the Cauchy problem for $D_t^2 - D_x a(t, x)^n D_x$. *Ann. Univ. Ferrara Sez. VII Sci. Mat.* **52**(2), 395–430 (2006)
38. Ohya, Y., Tarama, S.: The Cauchy Problem with multiple characteristics in the Gevery class–Hölder coefficients in t . *Hyperbolic Equations and Related Topics. Kataka/Kioto*, pp. 273–306 (1984)
39. Parenti, C., Parmeggiani, A.: On the Cauchy problem for hyperbolic operators with double characteristics. *Commun. Partial Differ. Equ.* **34**, 837–888 (2009). <https://doi.org/10.1080/03605300902892360>
40. Parusinski, A., Rainer, A.: Regularity of roots of polynomials. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **16**(2), 481–517 (2016)
41. Rozenblum, G.: Spectral asymptotic behaviour of elliptic systems (Russian). *Zap. LOMI* **96**, 255–271 (1980)
42. Ruzhansky, M.: Singularities of affine fibrations in the theory of regularity of Fourier integral operators. *Russ. Math. Surv.* **55**(1), 93–161 (2000). <https://doi.org/10.1070/RM2000v055n01ABEH000250>
43. Ruzhansky, M.: Regularity theory of Fourier integral operators with complex phases and singularities of affine fibrations. *CWI Tract*, 131. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam (2001)
44. Yuzawa, Y.: The Cauchy problem for hyperbolic systems with Hölder continuous coefficients with respect to time. *J. Differ. Equ.* **219**(2), 363–374 (2005). <https://doi.org/10.1016/j.jde.2004.12.006>
45. Wasow, W.: On holomorphically similar matrices. *J. Math. Anal. Appl.* **4**(2), 202–206 (1962). [https://doi.org/10.1016/0022-247X\(62\)90050-1](https://doi.org/10.1016/0022-247X(62)90050-1)