Essentials of pseudodifferential operators*

Gabriel Nagy

March 27, 2004

1 Introduction

Pseudodifferential operators are a generalization of differential operators. The idea is to think of a differential operator acting upon a function as the inverse Fourier transform of a polynomial in the Fourier variable times the Fourier transform of the function. This integral representation leads to a generalization of differential operators, which correspond to functions other than polynomials in the Fourier variable, as far as the integral converges.

In other words, given a smooth, complex valued function $\underline{p}(x,y)$ from $\mathbb{R}^n \times \mathbb{R}^n$ with a given asymptotic behavior at infinity, associate with it an operator $p(x,\partial_x) : \mathcal{S} \to \mathcal{S}$, where $\mathcal{S}$ is the Schwartz space, that is, the set of complex valued smooth functions in $\mathbb{R}^n$, such that the function and every derivative decay faster than any polynomial at infinity. The association $\underline{p} \to p$, that is, functions into differential operators, is not unique. This is known to anyone acquainted with quantum mechanics. Different maps from functions $\underline{p}(x,y)$ into operators $p(x,\partial_x)$ give rise to different theories of pseudodifferential calculus. Every generalization must coincide in the following: The polynomial $\underline{p}(x,y) = \sum_{|\alpha|\leq m} a_\alpha(x)(iy)^\alpha$ where α is a multi-index in $\mathbb{R}^n$ must be associated with the operator $p(x,\partial_x) = \sum_{|\alpha|\leq m} a_\alpha(x)\partial_x^\alpha$, that is, with a differential operator of order m. The map $\underline{p} \to p$ used in these notes is introduced in Sec. 3. It is the most used definition of pseudodifferential operators in the literature, and the one most studied.

The Fourier transform is used to rewrite the differential operator because it maps derivatives into multiplication, that is, $[\partial_x u(x)]\hat{} = iy\hat{u}(y)$. This property is used to solve constant coefficient partial differential equations (PDEs) by transforming the whole equation into an algebraic equation. This technique is not useful on variable coefficient PDEs, because of the inverse property, that is, $[xu(x)]\hat{} = i\partial_y\hat{u}(y)$. For example, one has $[\partial_x u + xu]\hat{} = i[\partial_y\hat{u} + y\hat{u}]$, and nothing has been simplified by the Fourier transform. That is why one looks for

*These notes are included as an appendix in the work: G. Nagy. O. Ortiz, O. Reula, *Strongly hyperbolic second order Einstein's evolution equations*, Phys. Rev. D, **70**, 044012 (2004). G.N. thanks Omar Ortiz, Oscar Reula, and Bruce Driver for nice discussions, and for their help to improve these notes.

other ways of rewriting differential operators. The generalization to pseudodifferential operators is an additional consequence. Other transforms can be used to define different generalizations of differential operators. For example, Mellin transforms are used in [14].

The functions $\underline{p}(x, y)$ are called symbols. Differential operators correspond to polynomial symbols in y. They contain the main equations from physics. Even strongly hyperbolic PDEs have polynomial symbols. Why should one consider more general symbols? Because the generalization is evident, and it has proved worth doing it. The Atiyah and Singer index theorem is proved using pseudodifferential operators with smooth symbols, which are more suitable for studying homotopy invariants than polynomial symbols, [17]. Techniques to prove the well posedness of the Cauchy problem for a strongly hyperbolic system require one to mollify polynomial symbols into smooth non-polynomial ones, [15]. The main application in these notes is simple: to reduce a second order partial differential equation to a first order system without adding new characteristics into the system. This is done by introducing the operator, a square root of the Laplacian, which is a first order pseudodifferential, but not differential, operator. The main idea for this type of reduction was introduced in [1].

How far should this generalization be carried? In other words, how is the set of symbols that define the pseudodifferential operators determined? The answer depends on which properties of differential operators one wants to be preserved by the general operators, and which additional properties one wants the latter to have. There are different spaces of symbols defined in the literature. Essentially all of them agree that the associated space of pseudodifferential operators is closed under taking the inverse. The inverse of a pseudodifferential operator is another pseudodifferential operator. This statement is not true for differential operators. The algebra developed in studying pseudodifferential operators is useful to compute their inverses. This is important from a physical point of view, because the behavior of solutions of PDEs can be inferred from the inverse operator. One could even say that pseudodifferential operators were created in the middle 1950s from the procedure to compute parametrices to elliptic equations. [A parametrix is a function that differs from a solution of the equation $p(u) = \delta$ by a smooth function, where p is a differential operator, and δ is Dirac's delta distribution.] To know the parametrix is essentially the same as to have the inverse operator.

Spaces of pseudodifferential operators are usually defined to be closed under composition and transpose, and to act on distribution spaces and on Sobolev spaces. They can be invariant under diffeomorphism, so they can be defined on a manifold. This definition for pseudodifferential operators is not so simple as for differential operators, because the latter are local operators and the former are not. Pseudodifferential operators are pseudolocal. An operator p acting on a distribution u is local if $p(u)$ is smooth in the same set where u is smooth. Pseudolocal means that the set where $p(u)$ is smooth includes the set where u is smooth. This means that p could smooth out a nonsmoothness of u. Mollifiers are an example of this kind of smoothing operator. They are integral operators, which justifies the name of pseudolocal. Differential operators with smooth coefficients are an example of local operators. The proofs of all these properties of pseudodifferential operators are essentially algebraic calculations on the symbols. One could say that the main practical advantage of pseudodifferential calculus is, precisely, turning differential problems into algebraic ones.

2 Function spaces

The study of the existence and uniqueness of solutions to PDEs, as well as the qualitative behavior of these solutions, is at the core of mathematical physics. Function spaces are the basic ground for carrying on this study. The mathematical structure needed is that of the Hilbert space, or Banach space, or Fréchet space, which are complete vector spaces having, respectively, an inner product, a norm, and a particular metric constructed with a family of seminorms. Every Hilbert space is a Banach space, and every Banach space is a Fréchet space. The main examples of Hilbert spaces are the space of square integrable functions L^2, and the Sobolev spaces H^k, with k a positive integer, which consist of functions whose k derivatives belong to L^2. The Fourier transform makes it possible to extend Sobolev spaces to real indices. This generalization in the idea of the derivative is essentially the same as one uses to construct pseudodifferential operators. Examples of Banach spaces are L^p, spaces of p-power integrable functions, where L^2 is the particular case $p = 2$. The main examples of Fréchet spaces are $C^\infty(\Omega)$, the set of smooth functions in any open set $\Omega \subset \mathbb{R}^n$, with a particular metric on it, (the case $\Omega = \mathbb{R}^n$ is denoted C^∞), the Schwartz spaces of smooth functions of rapid decrease, and its dual, which is a space of distributions.

This section presents only Sobolev spaces, first with non-negative integer index, and the generalization to a real index. The Fourier transform is needed to generalize the Sobolev spaces. Therefore, Schwartz spaces are introduced to present the Fourier transform, and to extend it to L^2. The next section is dedicated to introduce pseudodifferential operators.

Let L^2 be the vector space of complex valued, square integrable functions on $\mathbb{R}^n$, that is functions such that $\|u\| < \infty$, where $\|u\| := \sqrt{(u,u)}$ and

$$(u, v) := \int_{\mathbb{R}^n} \bar{u}(x)v(x)dx,$$

with $\bar{u}$ the complex conjugate of u. This set is a Hilbert space, that is a complete vector space with inner product, where the inner product is given by $(\ ,\ )$ and is complete with respect to the associated norm $\|\ \|$.

The Sobolev spaces H^k, for k a non-negative integer, are the elements of L^2 such that

$$\|u\|_k^2 := \sum_{|\alpha| \leq k} \|\partial^\alpha u\|^2 < \infty,$$

where α is a multi-index and the notation $|\alpha|$ when α is a multi-index means the positive integer $|\alpha| = \sum_{i=1}^n \alpha_i$. The inner product in L^2 defines an inner product in H^k given by

$$(u,v)_k := \sum_{|\alpha|\leq k} (\partial^\alpha u, \partial^\alpha v).$$

Let $\mathcal{S}$ be the space of functions of rapid decrease, also called the Schwartz space, that is, the set of complex valued, smooth functions on $\mathbb{R}^n$, satisfying

$$|u|_{k,\alpha} := \sup_{x\in\mathbb{R}^n} |(1+|x|^2)^{k/2}\partial^\alpha u| < \infty$$

for every multi-index α, and all $k \in \mathbb{N}$ natural, with $|x|$ the Euclidean length in $\mathbb{R}^n$. The Schwartz space is useful in several contexts. It is the appropriate space to introduce the Fourier transform. It is simple to check that the Fourier transform is well defined on elements in that space, in other words, the integral converges. It is also simple to check the main properties of the transformed function. More important is that the Fourier transform is an isomorphism between Schwartz spaces. As mentioned earlier, the Schwartz space provided with an appropriate metric is an example of a Fréchet space, a slight generalization of a Banach space. Its dual space is the set of distributions, which generalizes the usual concept of functions.

The Fourier transform of any function $u \in \mathcal{S}$ is given by

$$\mathcal{F}[u](x) = \hat{u}(x) := \int_{\mathbb{R}^n} e^{-ix\cdot y} u(y) đy,$$

where $đy = dy/(2\pi)^{n/2}$, while dy and $x\cdot y = \delta_{ij}x^i y^j$ are the Euclidean volume element and scalar product in $\mathbb{R}^n$, respectively. The map $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is an isomorphism. The inverse map is given by

$$\mathcal{F}^{-1}[u](x) = \check{u}(x) := \int_{\mathbb{R}^n} e^{ix\cdot y} u(y) đy.$$

An important property of the Fourier transform useful in PDE theory is the following: $[\partial_x^\alpha u(x)]\hat{} = i^{|\alpha|} y^\alpha \hat{u}(y)$, and $[x^\alpha u(x)]\hat{} = i^{|\alpha|}\partial_y^\alpha \hat{u}(y)$, that is, it converts smoothness of the function into decay properties of the transformed function, and vice versa. The Fourier transform is extended to an isomorphism $\mathcal{F} : L^2 \to L^2$, first proving Parseval's theorem, that is, $(u,v) = (\hat{u},\hat{v})$ for all u, $v \in \mathcal{S}$ (which gives Plancherel's formula for norms, $\|u\| = \|\hat{u}\|$, in the case that the norm comes from an inner product, with $u = v$) and second recalling that $\mathcal{S}$ is dense in L^2.

The definition of Sobolev spaces H^s for s real is based on Parseval's theorem. First recall that every $u \in H^k$ with non-negative integer k satisfies $\partial^\alpha u \in L^2$ for $|\alpha| \leq k$, so Parseval's theorem implies $|y|^k\hat{u}(y) \in L^2$. Second, notice that there exists a positive constant c such that $(1/c)\langle y\rangle \leq (1+|y|) \leq c\langle y\rangle$, where $\langle y\rangle = (1+|y|^2)^{1/2}$. Therefore, one arrives at the following definition. The Sobolev space H^s for any $s \in \mathbb{R}$ consists of locally square integrable functions in $\mathbb{R}^n$ such that $\langle y\rangle^s\hat{u} \in L^2$. This space is a Hilbert space with the inner product

$$(u,v)_s := \int_{\mathbb{R}^n} \langle y\rangle^{2s}\bar{\hat{u}}(y)\hat{v}(y)dy,$$

and the associated norm is denoted by

$$\|u\|_s^2 := \int_{\mathbb{R}^n} \langle y\rangle^{2s} |\hat{u}(y)|^2 dy.$$

The definition is based on a property of the Fourier Transform, that converts smoothness of the function into decay at infinity of the transformed function. One can check that $H^s \subset H^{s'}$ whenever $s' \leq s$. Notice that negative indices are allowed. The elements of those spaces are distributions. Furthermore, the Hilbert space H^{-s} is the dual of H^s. Finally, two more spaces are needed later on, $H^{-\infty} := \cup_{s\in\mathbb{R}} H^s$ and $H^{\infty} := \cap_{s\in\mathbb{R}} H^s$. These spaces are, with appropriate metrics on them, Fréchet spaces. A closer picture of the kind of element these spaces may contain is given by the following observations. The Sobolev embedding lemma implies that $H^\infty \subset C^\infty$, while the opposite inclusion is not true. Also notice that $\mathcal{S} \subset H^\infty$, therefore $H^{-\infty} \subset \mathcal{S}'$, so the elements of $H^{-\infty}$ are tempered distributions.

3 Pseudodifferential operators

Let S^m with $m \in \mathbb{R}$, be the set of complex valued, smooth functions $\underline{p}(x,y)$, from $\mathbb{R}^n \times \mathbb{R}^n$ such that

$$|\partial_x^\beta \partial_y^\alpha \underline{p}(x,y)| \leq C_\alpha \langle y\rangle^{m-|\alpha|}, \tag{1}$$

with C_α a constant depending on the multi-index α, and $\langle y\rangle = (1+|y|^2)^{1/2}$. This is the space of functions whose elements are associated with operators. It is called the space of symbols, and its elements $\underline{p}(x,y)$ symbols. There is no asymptotic behavior needed in the x variable, because Fourier integrals are thought to be carried out in the y variable. And the asymptotic behavior of this variable is related to the order of the associated differential operator, as one can shortly see in the definition of the map that associates functions $\underline{p}(x,y)$ with operators $p(x,\partial_x)$. One can check that $S^{m'} \subset S^m$ whenever $m' \leq m$. Two more spaces are needed later on, $S^\infty := \cup_{m\in\mathbb{R}} S^m$ and $S^{-\infty} := \cap_{m\in\mathbb{R}} S^m$.

Given any $\underline{p}(x,y) \in S^m$, then the associated operator $p(x,\partial_x) : \mathcal{S} \to \mathcal{S}$ is said to belong to ψ^m and is determined by

$$p(x,\partial_x)(u) = \int_{\mathbb{R}^n} e^{ix\cdot y} \underline{p}(x,y) \hat{u}(y) đy, \tag{2}$$

for all $u \in \mathcal{S}$. The constant m is called the order of the operator. It is clear that $u \in \mathcal{S}$ implies $p(x,\partial_x)(u) \in C^\infty$; however, the proof that $p(u) \in \mathcal{S}$ is more involved. One has to show that $p(u)$ and its derivatives decay faster than any polynomial in x. The idea is to multiply Eq. (2) by x^α and recall the relation $i^{|\alpha|} x^\alpha e^{iy\cdot x} = \partial_y^\alpha e^{iy\cdot x}$. Integration by parts and the inequality (1) imply that the resulting integral converge and is bounded in x. This gives the decay.

The polynomial symbols $\underline{p}(x,y) = \sum_{|\alpha|=0}^m a_\alpha(x)(iy)^\alpha$ with non-negative integer m correspond to differential operators $p(x,\partial_x) = \sum_{|\alpha|=0}^m a_\alpha(x)\partial_x^\alpha$ of order m. An example of a pseudodifferential operator that is not differential is given by the symbol $\underline{p}(y) = \chi(y)|y|^k \sin[\ln(|y|)]$, where k is a real constant and $\chi(y)$ is a cut function at $|y| = 1/2$, that is a smooth function that vanishes for $|y| \leq 1/2$ and is identically 1 for $|y| \geq 1$. The cut function is needed to have a smooth function at $y = 0$. This symbol belongs to S^k. The function $\underline{p}(y) = \chi(y)\ln(|y|)$ is not a symbol, because $|\underline{p}(y)| \leq c_0\langle y\rangle^\epsilon$, for every $\epsilon > 0$, but $|\partial_y \underline{p}(y)| \leq c_1\langle y\rangle^{-1}$, and the change in the decay is bigger than 1, which is the value of $|\alpha|$ in this case. Another useful example to understand the symbol spaces is $\underline{p}(y) = \chi(y)|y|^k \ln(|y|)$, with k a real constant. This function is not a symbol for k natural or zero, for the same reason as in the previous example. However, it is a symbol for the remaining cases, belonging to $S^{k+\epsilon}$, for every $\epsilon > 0$.

A very useful operator is $\Lambda^s : \mathcal{S} \to \mathcal{S}$ given by

$$\Lambda^s(u) := \int_{\mathbb{R}^n} e^{iy\cdot x}\langle y\rangle^s \hat{u}(y) đy,$$

where s is any real constant. This is a pseudodifferential operator that is not differential. Its symbol is $\underline{\Lambda}^s = \langle y\rangle^s$, which belongs to S^s, and then one says $\Lambda^s \in \psi^s$. It is usually denoted as $\Lambda^s = (1-\Delta)^{s/2}$. It can be extended to Sobolev spaces, that is, to an operator $\Lambda^s : H^s \to L^2$. This is done by noticing the bound $\|\Lambda^s(u)\| = \|u\|_s$ for all $u \in \mathcal{S}$, and recalling that $\mathcal{S}$ is a dense subset of L^2. This operator gives a picture of what is meant by an s derivative, for s real. One can also rewrite the definition of H^s, saying that $u \in H^s$ if and only if $\Lambda^s(u) \in L^2$.

Pseudodifferential operators can be extended to operators acting on Sobolev spaces. Given $p \in \psi^m$, it defines an operator $p(x, \partial_x) : H^{s+m} \to H^s$. This is the reason to call m the order of the operator. The main idea of the proof is again to translate the basic estimate (1) in to the symbol to an L^2-type estimate for the operator, and then use the density of $\mathcal{S}$ in L^2. The translation is more complicated for a general pseudodifferential operator than for Λ^s, because symbols can depend on x. Intermediate steps are needed, involving estimates on an integral representation of the symbol, called the kernel of the pseudodifferential operator. Pseudodifferential operators can also be extended to act on distribution spaces $\mathcal{S}'$, the dual of Schwartz spaces $\mathcal{S}$.

An operator $p : H^{-\infty} \to H^{-\infty}$ is called a smoothing operator if $p(H^{-\infty}) \subset C^\infty$. That means $p(u)$ is smooth regardless of u being smooth. One can check that a pseudodifferential operator whose symbol belongs to $S^{-\infty}$ is a smoothing operator. For example, $\underline{p}(y) = e^{-|y|^2} \in S^{-\infty}$. However, not every smoothing operator is pseudodifferential. For example, $\underline{p}(y) = \rho(y)$, with $\rho \in H^s$ for some s and having compact support, is a smoothing operator which is not pseudodifferential unless ρ is smooth. Friedrichs' mollifiers, J_ϵ for $\epsilon \in (0,1]$, are a useful family of smoothing operators, which satisfy $J_\epsilon(u) \to u$ in the L^2 sense, in the limit $\epsilon \to 0$, for each $u \in L^2$.

Consider one more example, the operator $\lambda : \mathcal{S} \to \mathcal{S}$ given by

$$\lambda(u) := \int_{\mathbb{R}^n} e^{iy\cdot x} i|y|\chi(y)\hat{u}(y) đy,$$

where $\chi(y)$ is again a cut function at $|y| = 1/2$. The symbol is $\underline{\lambda}(y) = i|y|\chi(y)$. The cut function χ makes $\underline{\lambda}$ smooth at $y = 0$. The operator without the cut function is $\ell : \mathcal{S} \to L^2$ given by

$$\ell(u) := \int_{\mathbb{R}^n} e^{iy\cdot x} i|y|\hat{u}(y) đy.$$

Its symbol $\underline{\ell}(y) = i|y|$ does not belong to any S^m because it is not smooth at $y = 0$. Both operators λ, ℓ can be extended to maps $H^1 \to L^2$. What is more important, their extensions are essentially the same, because they differ in a smoothing, although not pseudodifferential, operator.

The asymptotic expansion of symbols is maybe the most useful notion related to pseudodifferential calculus. Consider a decreasing sequence $\{m_j\}_{j=1}^{\infty}$, with $\lim_{j\to\infty} m_j = -\infty$. Let $\{\underline{p}_j\}_{j=1}^{\infty}$ be a sequence of symbols $\underline{p}_j(x, y) \in S^{m_j}$. Assume that these symbols are asymptotically homogeneous in y of degree m_j, that is, they satisfy $\underline{p}_j(x, ty) = t^{m_j}\underline{p}_j(x, y)$ for $|y| \geq 1$. Then, a symbol $\underline{p} \in S^{m_1}$ has the asymptotic expansion $\sum_j \underline{p}_j$ if and only if

$$\left(\underline{p} - \sum_{j=1}^{k} \underline{p}_j\right) \in S^{m_{(k+1)}}, \quad \forall\, k \geq 1, \tag{3}$$

and it is denoted by $\underline{p} \sim \sum_j \underline{p}_j$. The first order term in the expansion, $\underline{p}_1$, is called the principal symbol. Notice that m_j are real constants, not necessarily integers. Every asymptotic expansion defines a symbol, that is, every function of the form $\sum_j \underline{p}_j$ belongs to some symbol space S^{m_1}. However, not every symbol $\underline{p} \in S^m$ has an asymptotic expansion. Consider the example $\underline{p}(y) = \chi(y)|y|^{1/2}\ln(|y|)$. The set of symbols that admit an asymptotic expansion of the form (3) is called classical, it is denoted by S^m_{cl}, and the corresponding operators are said to belong to ψ^m_{cl}. One then has $S^m_{\text{cl}} \subset S^m$. Notice that if two symbols $\underline{p}$ and $\underline{q}$ have the same asymptotic expansion $\sum_j \underline{p}_j$, then they differ in a pseudodifferential smoothing operator, because

$$\underline{p} - \underline{q} = \left(\underline{p} - \sum_{j=1}^{k} \underline{p}_j\right) - \left(\underline{q} - \sum_{j=1}^{k} \underline{p}_j\right) \in S^{m_{(k+1)}}$$

for all k, and $\lim_{k\to\infty} m_j = -\infty$, so $(\underline{p} - \underline{q}) \in S^{-\infty}$. This is the precise meaning for the rough sentence, "what really matters is the asymptotic expansion."

There is in the literature a more general concept of asymptotic expansion. It does not require that the $\underline{p}_j$ to be asymptotically homogeneous. We do not consider this generalization in these notes.

Most of the calculus of pseudodifferential operators consists of performing calculations with the highest order term in the asymptotic expansion and keeping careful track of the lower order terms. The symbol of a product of pseudodifferential operators is not the product of the individual symbols. Moreover, the former is difficult to compute. However, an asymptotic expansion can be explicitly written, and one can check that the principal symbol of the product is equal to the product of the individual principal symbols. More precisely, given $p \in \psi^r$ and $q \in \psi^s$, then the product is a well defined operator $pq \in \psi^{r+s}$ and the asymptotic expansion of its symbol is

$$\underline{pq} \sim \sum_{|\alpha|\geq 0} \frac{1}{i^{|\alpha|}\alpha!} \left[\partial_y^\alpha \underline{p}(x,y)\right] \left[\partial_x^\alpha \underline{q}(x,y)\right].$$

Notice that the first term in the asymptotic expansion of a commutator $[p,q] = pq - qp$, that is its principal symbol, is precisely $1/i$ times the Poisson bracket of their respective symbols, $\{\underline{p},\underline{q}\} = \sum_j \left(\partial_{y^j}\underline{p}\partial_{x^j}\underline{q} - \partial_{x^j}\underline{p}\partial_{y^j}\underline{q}\right)$. Similarly, the symbol of the adjoint pseudodifferential operator is not the adjoint of the original symbol. However, this is true for the principal symbols. The proof is based in an asymptotic expansion of the following equation:

$$\underline{(p^*)}(x,y) = \int\int_{\mathbb{R}^n} e^{-i(x-x')\cdot(y-y')}\,(\underline{p})^*(x',y')\,đx đy.$$

There are three main generalizations of the theory of pseudodifferential operators present in the literature. First, the operators act on vector valued functions instead of on scalar functions. While this is straightforward, the other generalizations are more involved. Second, the space of symbols is enlarged, first done in [6]. It is denoted as $S^m_{\rho\delta}$, and its elements satisfy $|\partial_x^\beta\partial_y^\alpha\underline{p}(x,y)| \leq C_{\alpha,\beta}\langle y\rangle^{m-\rho|\alpha|+\delta|\beta|}$, with $C_{\alpha,\beta}$ a constant depending on the multi-indices α and β. The extra indices have been tuned to balance two opposite tendencies; on the one hand, to preserve some properties of differential operators; on the other hand, to maximize the amount of new objects in the generalization. These symbol spaces contain functions like $\underline{p}(x,y) = \langle y\rangle^{a(x)}$, which belongs to $S^m_{1,\delta}$, where $\delta > 0$ and $m = \max_{x\in\mathbb{R}^n} a(x)$. Third, the domain of the functions $\underline{p}(x,y)$ is changed from $\mathbb{R}^n \times \mathbb{R}^n$ to $\Omega \times \mathbb{R}^n$, with $\Omega \subset \mathbb{R}^n$ any open set. A consequence in the change of the domain is that $p : C_0^\infty(\Omega) \to C^\infty(\Omega)$, so the domain and range of p are not the same, which makes it more difficult to define the product of pseudodifferential operators. These notes are intended to be applied to hyperbolic PDEs on $\mathbb{R}^n$, which are going to be converted to pseudodifferential operators in S^1, so there is no need to consider the last two generalizations.

4 Strongly hyperbolic systems

Consider the Cauchy problem for a linear first order pseudodifferential system

$$\partial_t u = p(t,x,\partial_x)u, \quad u|_{t=0} = f, \tag{4}$$

where u, f are m-dimensional vector valued functions, $m \geq 1$, and x represents Cartesian coordinates in $\mathbb{R}^n$, $n \geq 1$. Here $p(t,x,\partial_x)$ is a smooth family of pseudodifferential operators in ψ^1_{cl}, parametrized by $t \in \mathbb{R}^+$. Let $\underline{p}(t,x,y)$ and $\underline{p}_1(t,x,y)$ be their symbols and principal symbols, respectively.

If p is a differential operator with analytic coefficients, then the Cauchy-Kowalewski Theorem asserts that there exists a unique solution for every analytic data f. However, solutions corresponding to smooth data behave very different depending on the type of operator p. For example, write the Laplace equation in $\mathbb{R}^{n+1}$ and the wave equation in $\mathbb{R}^{n+1}$ as a first order system in the form $\partial_t u = A^i \partial_i u$. Matrices A^i are skew-symmetric for the Laplacian, and symmetric for the wave operator. Therefore the eigenvalues are pure imaginary for the former and real for the latter. Write the equations in the base of their respective eigenvectors. Try solutions of the form $u(t,x) = \hat{u}(t)e^{iy\cdot x}$. The solutions are $\hat{u}(t) = e^{(\lambda\cdot y)t}\hat{u}(0)$ and $\hat{u}(t) = e^{i(\lambda\cdot y)t}\hat{u}(0)$, respectively, for some real constants λ, and for all $y \in \mathbb{R}^n$. Therefore, the high frequency solutions of the Cauchy problem for the Laplace equation diverge in the limit $|y| \to \infty$, while the solutions of the wave equation do not diverge in that limit.

An explicit example in $\mathbb{R}^2$, presented by Hadamard in [4], may clarify this. Consider the functions

$$v(t,x) = \sin(nt)\sin(nx)/n^{p+1},$$

$$w(t,x) = \sinh(nt)\sin(nx)/n^{p+1},$$

with $p \geq 1$, n, constants, defined on $t \geq 0$, $x \in [0,1]$. They are solutions of the Cauchy problem for wave equation and the Laplace equation, respectively, with precisely the same Cauchy data on $t = 0$, that is

$$v_{tt} - v_{xx} = 0, \quad v|_{t=0} = 0, \quad v_t|_{t=0} = \sin(nx)/n^p,$$

$$w_{tt} + w_{xx} = 0, \quad w|_{t=0} = 0, \quad w_t|_{t=0} = \sin(nx)/n^p.$$

As $n \to \infty$, the Cauchy data converges to zero in $C^{p-1}([0,1])$. In this limit, the solution of the wave equation converges to zero, while the solution of the Laplace equation diverges. The concept of well posedness is introduced in order to capture this behavior of the wave equation's solution under high frequency perturbations on its Cauchy data.

The Cauchy problem (4) is well posed in a norm $\|\ \|$ if given the data $f(x)$ there exists a unique solution $u(t,x)$ in $[0,T) \times \mathbb{R}^n$ for some $T > 0$; and given any number $\epsilon > 0$ there exists $\delta > 0$ such that, for every data $\tilde{f}(x)$ satisfying $\|\tilde{f} - f\| < \delta$ there exists a unique solution $\tilde{u}(t,x)$ in $[0,\tilde{T}) \times \mathbb{R}^n$ for some $\tilde{T} > 0$ with $|\tilde{T} - T| < \epsilon$, and satisfying $\|\tilde{u}(t) - u(t)\| < \epsilon$, for all $t \in [0, \min(\tilde{T}, T))$. This means that the solution depends continuously on the data in the norm $\|\ \|$.

As pointed out above, well posedness is essentially a statement about the behavior of the solutions of a Cauchy problem under high frequency perturbations on the initial data. Here is where pseudodifferential calculus is most useful to study solutions of the Cauchy problem. The high frequency part of the solution can be determined studying the higher order terms in the asymptotic expansion of symbols.

A wide class of operators with well posed Cauchy problem is called strongly hyperbolic. A first order pseudodifferential system (4) is strongly hyperbolic if $p \in \psi^1_{\rm cl}$ and the principal symbol is symmetrizable. This means that there exists a positive definite, Hermitian operator, $H(t,x,y)$, homogeneous of degree zero in y, smooth in all its arguments for $y \neq 0$, such that

$$(H\underline{p}_1 + \underline{p}_1^* H) \in S^0,$$

where $\underline{p}_1^*$ is the adjoint of the principal symbol $\underline{p}_1$.

Consider first order differential systems of the form $\partial_t u = A^i(t,x)\partial_i u + B(t,x)u$. The symbol is $\underline{p}(t,x,y) = iA^j(t,x)y_j + B(t,x)$, and the principal symbol is $\underline{p}_1(t,x,y) = iA^j y_j$. If the matrices A^i are all symmetric, then the system is called symmetric hyperbolic. The symmetrizer H is the identity, and $\underline{p}_1 + \underline{p}_1^* = 0$. The wave equation on a fixed background, written as a first order system is an example of a symmetric hyperbolic system. Well posedness for symmetric hyperbolic systems can be shown without pseudodifferential calculus. The basic energy estimate can be obtained by integration by parts in space-time.

If the matrices A^i are symmetrizable, then the differential system is called strongly hyperbolic. The symmetrizer $H = H(t,x,y)$ is assumed to depend smoothly on ω. Every symmetric hyperbolic system is strongly hyperbolic. pseudodifferential calculus must be used to show well posedness for variable coefficient strongly hyperbolic systems that are not symmetric hyperbolic [15]. The definition given two paragraphs above is more general because the symbol does not need to be a polynomial in y. The definition given above includes first order pseudodifferential reductions of second order differential systems. These type of reductions are performed with operators like Λ, λ, or ℓ.

In the particular case of constant coefficient systems there exists in the literature a more general definition of strong hyperbolicity [3, 8]. The principal symbol $\underline{p}_1$ must have only imaginary eigenvalues, and a complete set of linearly independent eigenvectors. The latter must be uniformly linear independent in $y \neq 0$ over the whole integration region. Kreiss's matrix theorem (see Sec. 2.3 in [8]) says that this definition is equivalent to the existence of a symmetrizer H. Nothing is known about the smoothness of H with respect to t, x, and y. The existence of this symmetrizer is equivalent to well posedness for constant coefficient systems. However, the proof of well posedness for variable coefficient and quasilinear systems does require the smoothness of the symmetrizer. There are examples showing that this smoothness does not follow from the previous hypothesis on eigenvalues and eigenvectors of $\underline{p}_1$. Because it is not known what additional hypothesis on the latter could imply this smoothness, one has to include it into the definition of strong hyperbolicity for nonconstant coefficient systems.

A more fragile notion of hyperbolicity is called weak hyperbolicity, where the operator $\underline{p}_1$ has imaginary eigenvalues, but nothing is required of its eigenvectors. Quasilinear weakly hyperbolic systems are not well posed. The following example gives an idea of the problem. The 2×2 system $\partial_t u = A\partial_x u$ with

$t, x \in \mathbb{R}$ and

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

is weakly hyperbolic. Plane wave solutions of the form $u(t,x) = \hat{u}(t)e^{iy\cdot x}$ satisfy $|\hat{u}(t)| \leq |\hat{u}(0)|(1+|y|t)$. Therefore, plane wave solutions to a weakly hyperbolic system do not diverge exponentially in the high frequency limit (as in the case of Cauchy problem for the Laplace equation) but only polynomically. This divergence causes solutions to variable coefficient weak hyperbolic systems to be unstable under perturbations in the lower order terms of the operator, as well as in the initial data.

The main theorem about well posedness for strongly hyperbolic systems is the following. *The Cauchy problem (4) for a strongly hyperbolic system is well posed with respect to the Sobolev norm* $\| \ \|_s$ *with* $s > n/2 + 1$. *The solution belongs to* $C([0,T), H^s)$, *and* $T > 0$ *depends only on* $\|f\|_s$.

In the case of strongly hyperbolic differential systems, this is Theorem 5.2.D in [16]. The proof for pseudodifferential strongly hyperbolic systems is essentially the same. One builds an estimate for the solution in a norm, defined using the symmetrizer, equivalent to the Sobolev norm H^s. Then the argument follows the standard proof for differential systems. The construction of the symmetrizer is basically the one carried out in [9].

5 Further reading

There is no main reference followed in these notes; however, a good place to start is [15]. Notice that the notation is not precisely the one in that reference. The introduction is good, and the definitions are clear. The proofs are difficult to follow. More extended proofs can be found in [13], together with some historical remarks. The whole subject is clearly written in [12]. It is not the most general theory of pseudodifferential operators, but it is close to these notes. A slightly different approach can be found in [2], and detailed calculations to find parametrices are given in [18]. The introduction of [10] is very instructive. The first order reduction using Λ is due to Calderón in [1], and a clear summary of this reduction is given in [12].

The field of pseudodifferential operators grew out of a special class of integral operators called singular integral operators. Mikhlin in 1936, and Calderón and Zygmund in the begining of 1950s carried out the first investigations. The field started to develop really fast after a suggestion by Peter Lax in 1963 [11], who introduced the Fourier transform to represent singular integral operators in a different way. Finally, the work of Kohn and Nirenberg [7] presented the pseudodifferential operators as they are known today, and they proved their main properties. They showed that singular integral operators are the particular case of pseudodifferential operators of order zero. Further enlargements of the theory were due to Lars Hörmander [5, 6].

References

[1] A. Calderón. Existence and uniqueness theorems for systems of partial differential equations. In *Proc. Sympos. Fluid Dynamics and Appl. Math.*, volume 27. Gordon and Breach, New York, 1962.

[2] K. O. Friedrichs. *Pseudo-differential operators, an introduction.* Courant Institute of Mathematical Sciences, New York, 1968.

[3] B. Gustafsson, H. Kreiss, and J. Oliger. *Time dependent methods and difference methods.* John Wiley & Sons, New York, 1995.

[4] Jacques Hadamard. *Lectures on Cauchy's problem in linear partial differential equations.* Dover Publications, New York, 1952. Unabridged reprint from Lectures on Cauchy's problem, Yale University Press, 1923.

[5] L. Hörmander. Pseudodifferential operators. *Commun. Pure Appl. Math.*, 18:501–517, 1965.

[6] L. Hörmander. Pseudodifferential operators and hypoelliptic equations. *Symp. Pure Math.*, 10:138–183, 1967.

[7] J. Kohn and L. Nirenberg. An algebra of pseudo-differential operators. *Commun. Pure Appl. Math.*, 18:269–305, 1965.

[8] H. Kreiss and J. Lorentz. *Initial-boundary value problem and the Navier-Stokes equations.* Academic Press, Boston, 1989.

[9] H. Kreiss, O. E. Ortiz, and O. Reula. Stability of quasi-linear hyperbolic dissipative systems. *J. Diff. Eqs.*, 142:78–96, 1998.

[10] H. Kumano-go. *Pseudo-differential operators.* The MIT Press, Cambridge, Massachusetts, London, 1981.

[11] P. Lax. *The L_2 operator calculus of Mikhlin, Calderón and Zygmund.* Courant Inst. Math. Sci., New York University, 1963. Lectures Notes.

[12] L. Nirenberg. Lectures on linear partial differential equations. In B. Cockburn, G. Karniadakis, and C.W. Shu, editors, *Regional conference series in mathematics*, volume 17, pages 1–58. AMS, Providence, Rode Island, 1973.

[13] B. E. Petersen. *An introduction to the Fourier Transform and pseudo-differential operators.* Pitman Advanced Publishing Program, Boston, London, Melbourne, 1983.

[14] B. Schulze. *Pseudo-differential operators on manifolds with singularities.* North Holland, Amsterdam, 1991.

[15] M. E. Taylor. *Pseudodifferential operators.* Princeton University Press, Princeton, New Jersey, 1981.

[16] M. E. Taylor. Pseudodifferential operators and nonlinear PDE. In *Progress in Mathematics*, volume 100. Birkhäuser, Boston-Basel-Berlin, 1991. Second printing 1993.

[17] M. E. Taylor. *Partial Differential Equations II.* Springer, 1996.

[18] F. Treves. *Introduction to pseudodifferential and Fourier integral operators, I, II.* Plenum Press, New York and London, 1980.