Nonlinear Geometric Optics for Short Pulses

Deborah ALTERMAN *

Department of Applied Mathematics, University of Colorado, Boulder 80309 CO

Jeffrey RAUCH †

Department of Mathematics, University of Michigan, Ann Arbor 48109 MI

Abstract

This paper studies the propagation of pulse like solutions of semilinear hyperbolic equations in the limit of short wavelength. The pulses are located at a wavefront $\Sigma := \{\phi = 0\}$ where ϕ satisfies the eikonal equation and $d\phi$ lies on a regular sheet of the characteristic variety. The approximate solutions are $u^\epsilon_{\rm approx} = U(t,x,\phi(t,x)/\epsilon)$ where $U(t,x,r)$ is a smooth function with compact support in r. When U satisfies a familiar nonlinear transport equation from geometric optics it is proved that there is a family of exact solutions $u^\epsilon_{\rm exact}$ such that $u^\epsilon_{\rm approx}$ has relative error $O(\epsilon)$ as $\epsilon \to 0$. While the transport equation is familiar, the construction of correctors and justification of the approximation are different from the analogous problems concerning the propagation of wave trains with slowly varying envelope.

1 Introduction.

The methods of nonlinear geometric optics construct approximate solutions of partial differential equations of hyperbolic type. The solutions are accurate as a parameter ϵ measuring the wavelength tends to zero in units chosen so that the natural unit length for the problem is $O(1)$. The usual science article derivations are valid for wavetrains and go under the name of the slowly varying envelope approximation. The derivations suppose that the amplitude of the waves changes little over a distance of one wavelength. This article is concerned with the diametrically opposite case of pulse like solutions, in particular pulses whose length may be comparable to ϵ. These violate the slowly varying amplitude assumption.

The research is motivated by ultrashort laser pulses which may contain few wavelengths and for which the shortcomings inherent in the slowly varying amplitude assumption have been recognized for a long time (see e.g. [R]). We

*Research partially supported by U.S. National Science Foundation grant NSF-DMS 9810751

†Research partially supported by U.S. National Science Foundation grants NSF-DMS-9803296 and NSF-INT-9314095 .

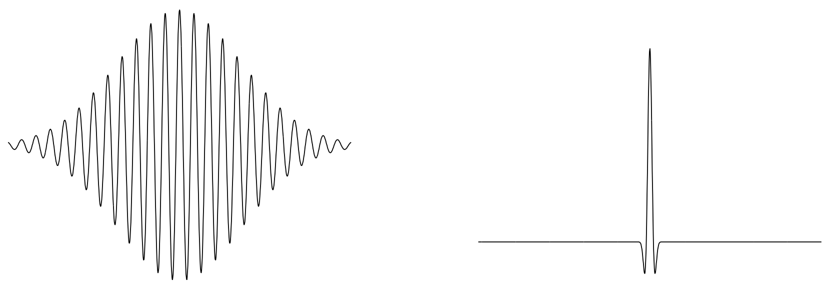


Figure 1: A wavetrain and a short pulse.

present semilinear results. The quasilinear analogues do not present serious additional difficulties.

In the present article we study pulses whose amplitudes are so large that nonlinear effects are important before diffractive effects. This scaling is called the scaling of geometric optics as opposed to diffractive geometric optics. For the scaling of geometric optics, an accurate asymptotic solution is constucted and the leading amplitude satisfies essentially the same nonlinear transport equation that is appropriate for slowly varying wavetrains and also for the transport of jump discontinuities. To our knowledge, the result may even be new in the linear case. The key corrector construction from §2 would also be needed to justify a variable coefficient linear result. Related but weaker results than those in this paper are contained in Chapter 3 of [A]. When amplitudes are smaller so that nonlinear effects are not important until the longer time scale of diffractive geometric optics, we will show in a second paper that the asymptotic equations for pulses on that scale differ from the familiar Shrödinger equations which describe wavetrains.

A snapshot at a fixed time of a wavetrain solution with slowly varying amplitude is suggested in Figure 1. Actually this figure violates the rule of thumb that the amplitude should change no more than 10% over a distance of one wavelength, but the idea should be clear. An analytic expression for an example is

$$a(x_1, x_2)\, e^{ix_1/\epsilon}, \quad \epsilon << 1, \qquad a \in \mathcal{S}(\mathbb{R}^2)\,.$$

In the same spirit, an analytic expression for a snapshot of a short pulse is

$$b(x_2)\, f(x_1/\epsilon), \quad \epsilon << 1, \qquad b\,, f \;\in \mathcal{S}(\mathbb{R})\,.$$

The Fourier transforms of these snapshots are given by

$$\hat{a}(\xi_1 - \frac{1}{\epsilon}, \xi_2)\,, \qquad \text{and} \qquad \hat{b}(\xi_2)\, \epsilon \hat{f}(\epsilon\, \xi_1)$$

repectively. In both cases the solutions are high frequency in the sense that the Fourier tranform is concentrated in a region $|\xi| > C/\epsilon$. However the concentration is stronger in the wavetrain case where the spectrum is essentially contained in an $O(1)$ neighborhood of the point $\xi = (1,0)/\epsilon$. The wavetrain has the form of a rapidly varying exponential prefactor times a function which is not rapidly varying. This is a key to the slowly varying envelope approximation. In contrast, the pulse has Fourier transform is spread over a region of size $O(1)$ in ξ_2 while it resembles a broad plateau of size $O(1/\epsilon)$ in ξ_1. In particular, no exponential prefactor leaves a slowly varying quotient. In both cases if one truncates frequencies from a bounded set by multiplying by $1 - \chi(\xi)$ with $\chi \in C_0^\infty$ the relative error tends to zero with epsilon. For wavetrains it is $O(\epsilon^\infty)$ while for pulses it is $O(\epsilon)$. These relatively larger low frequency contributions make the short wavelength asymptotic analysis for pulses harder than the corresponding results for wavetrains. In particular, it is harder to construct correctors to the leading approximations. The broad spectrum of pulses parallels exactly the laboratory strategy of spectral broadening which is followed in the production of short laser pulses.

In the important work of Majda and Rosales ([Ma], [MaR]), both pulse and wavetrain solutions were envisaged. Their formal analysis includes the construction of correctors like ours at least in the case of one space dimension. Using the Haar inequalities in one space dimension and the corrector they constructed one arrives at results resembling ours in the 1-d case. They concentrated on the concept of resonance, which only takes place for wave trains because pulses with different speeds do not overlap on large enough regions of space-time. They also concentrated on quasilinear as opposed to semilinear equations. In the class of smooth solutions the quasilinear and semilinear cases are similar. Yoshikawa also addressed pulse like solutions in a series of papers [Y1, Y2] and references therein. He constructed formal solutions making additional assumptions of vanishing moments for the profile U. The additional hypotheses guaranteed that correctors without the new terms in the present paper could be constructed. Considering only profiles with with vanishing moments is not reasonable for the practical problems. For example there are half cycle laser pulses for which the amplitude is overwhelmingly of one sign and therefore has nonvanishing integral. Our results provide a natural setting and a far reaching generalization of these earlier attacks.

We consider semilinear systems of partial differential equations of the form

$$L(y, \partial_y)\, u + F(y, u) = 0\,, \qquad L(y, \partial_y) := \sum_{\mu=0}^{d} A_\mu(y) \frac{\partial}{\partial y_\mu}\,. \tag{1}$$

Here

$$u = (u_1, \dots, u_N) \in \mathbb{C}^N\,, \quad \text{and} \quad y = (y_0, y_1, \dots, y_d) := (t, x) \in \mathbb{R} \times \mathbb{R}^d\,. \tag{2}$$

Assumption 1. *The operator L is symmetric hyperbolic on a neighborhood of a domain of determinacy $\overline{\Omega}_{\underline{T}}$ given by*

$$\overline{\Omega}_{\underline{T}} := \Big\{ (t, x) \;:\; x \in \overline{\mathcal{O}}\,, \quad \text{and} \quad 0 \le t \le \min\{\ell(x), \underline{T}\} \Big\}\,, \tag{3}$$

where $\mathcal{O}$ is a connected bounded open subset of $\mathbb{R}^d$ lying on one side of its smooth boundary. The symmetry hypothesis means that the A_μ are C^∞ and hermitian on a neigborhood of $\overline{\Omega}_{\underline{T}}$ and $A_0 = I$. To guarantee that $\overline{\Omega}_{\underline{T}}$ is a domain of determinacy assume that $\ell \in C^1(\overline{\mathcal{O}})$ vanishes at $\partial\mathcal{O}$, is strictly positive on $\mathcal{O}$, and whenever (t,x) belongs to the **lateral surface**

$$\Gamma_{\underline{T}} \ := \ \Big\{(t,x) \ : \ x \in \mathcal{O}\,, \ t = \ell(x) < \underline{T}\Big\}\,, \tag{4}$$

one has

$$A_0(t,x) - \sum_{j=1}^{d} \frac{\partial \ell(x)}{\partial x_j}\, A_j(t,x) \geq 0\,. \tag{5}$$

Assumption 2. *The nonlinear function F is infinitely differentiable (in the real sense) from $\overline{\Omega}_{\underline{T}} \times \mathbb{C}^N$ to $\mathbb{C}^N$.*

The pulses will be located near the wavefront which is given as a nondegenerate level set of a smooth function ϕ,

$$\Sigma_{\underline{T}} := \Big\{y \in \overline{\Omega}_{\underline{T}} \ : \ \phi(y) = 0\Big\}\,. \tag{6}$$

Assumption 3. *The defining function ϕ is smooth on a neighborhood of $\overline{\Omega}_{\underline{T}}$, the wavefront $\Sigma_{\underline{T}}$ is nonempty, and $d\phi(y) \neq 0$ wherever $\phi(y) = 0$. The defining function ϕ satisfies the eikonal equation*

$$\det L(y, d\phi(y)) = 0 \tag{7}$$

on a neighborhood of $\Sigma_{\underline{T}}$. This implies that $\Sigma_{\underline{T}}$ and the nearby level surfaces of ϕ are characteristic. The surface $\Sigma_{\underline{T}}$ is assumed to be transverse to the lateral boundary $\overline{\Gamma}_{\underline{T}}$.

Examples. Two motivating examples are sketched in Figure 2. They concern a system whose characteristic variety is given by $4\tau^2 = |\xi|^2$ so the corresponding speed of propagation is equal to $1/2$. The domain of determinacy is $\Omega_{\underline{T}} := \{|x| \leq R - t\,,\ 0 < t < \underline{T}\}$ where $R > \underline{T}$.

In the example on the left, Σ is the hyperplane $\{t = -2x_1 + c\}$ which correponds to a pulse with planar wave front. A different geometry is given by the example on the right with $\Sigma = \{|x| = \underline{R} - t/2\}$ where $\underline{T} < 2\underline{R} < R$. In this case the pulse is focussing with spherical wavefronts. Note that the transversality condition is satisfied in the second example because Σ does not meet the lateral boundary.

Assumption 4. *There is an open conic neighborhood $\mathcal{N}$ of the conormal variety*

$$N^*(\Sigma_{\underline{T}}) \ = \ \Big\{(y, s\, d\phi(y)) \ : \ y \in \overline{\Sigma}_{\underline{T}}\,, \ \ 0 \neq s \in \mathbb{R}\Big\}$$

on which the characteristic variety is a graph of a smooth function $\omega(y,\xi)$. That is $(y;\tau,\xi) \in \mathrm{Char}\, L(y,\partial)$ if and only if $\tau = \omega(y,\xi)$ with ω a smooth function homogeneous of degree one in ξ.

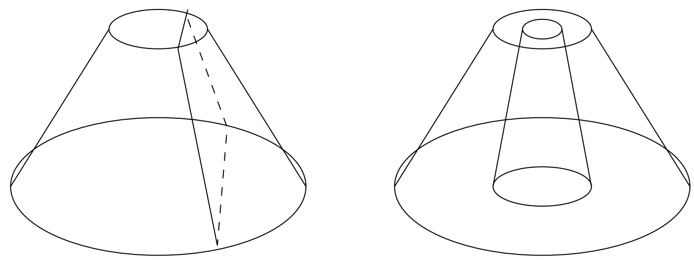


Figure 2: A planar wavefront and a focussing spherical wavefront.

It follows that near $\Sigma_{\underline{T}}$ the defining function satisfies the reduced eikonal equation

$$\phi_t = \omega(y, \nabla_x \phi)\,. \tag{8}$$

On the conic neighborhood from Assumption 4, the characteristic variety is given by the equation $\tau - \omega(y, \xi) = 0$ so the hamiltonian vector field with hamiltonian $\tau - \omega(y, \xi)$,

$$\frac{\partial}{\partial t} - \sum_{j=1}^{d} \frac{\partial \omega}{\partial \xi_j} \frac{\partial}{\partial x_j} - \sum_{\mu=0}^{d} \frac{\partial \omega}{\partial y_\mu} \frac{\partial}{\partial \eta_\mu}\,,$$

is tangent to the variety.

Definitions. *On a neighborhood of $\Sigma_{\underline{T}}$ define the smooth funtion $\pi(y)$ to be the orthogonal projector of $\mathbb{C}^N$ onto the kernel of $L(y, d\phi(y))$. Define the group velocity field $\mathbf{v}.\partial_x$ from the spatial projection of the hamiltonian field*

$$\mathbf{v}(y).\partial_x := -\sum_{j=1}^{d} \frac{\partial \omega}{\partial \xi_j}\Big|_{(y, d\phi(y))} \frac{\partial}{\partial x_j}\,. \tag{9}$$

The approximate solutions that we construct have the form

$$u^\epsilon_{\text{approx}} := U(y, \phi(y)/\epsilon)\,, \qquad \text{with} \qquad \lim_{r \to \pm\infty} U(y, r) = 0\,.$$

For ϵ small, the approximate solution is concentrated in an ϵ neighborhood of $\Sigma_{\underline{T}}$ so that up to an error $O(\epsilon)$ only the values of $U(y, r)$ with $y \in \Sigma_{\underline{T}}$ are needed. The graph of U as r changes describes the cross section of the pulse as one moves away from $\Sigma_{\underline{T}}$. If U decays at least as fast as $1/r$ then $U = \chi(r)U + V(y, r)/r$ where $\chi(r)$ has compact support and V is bounded. Inserting $r = \phi/\epsilon$ shows that replacing U by χU changes $u^\epsilon_{\text{approx}}$ by $O(\epsilon)$. These two remarks show that it is natural to seek the leading profile, $\mathtt{U}(y, r)$ as a function on $\Sigma_{\underline{T}} \times \mathbb{R}$ compactly supported in r.

In order that the approximation be good, the profile $\mathtt{U}$ is chosen to satisfy two equations. The first is the polarization identity

$$\forall y \in \Sigma_{\underline{T}}\,, \; r \in \mathbb{R}, \quad \pi(y)\,\mathtt{U}(y,r) = \mathtt{U}(y,r)\,. \tag{10}$$

The second is a transport equation along the rays which are integral curves of the vector field $\partial_t + \mathbf{v}.\partial_x$. It reads

$$\Big(\partial_t + \mathbf{v}.\partial_x\Big)\,\mathtt{U} + \pi(y)\sum_{\mu=1}^{d} A_\mu(y)\frac{\partial \pi(y)}{\partial y_\mu}\,\mathtt{U} + \pi(y)F(y,\mathtt{U}) \;=\; 0\,. \tag{11}$$

Since $\pi(y)\,\mathtt{U} = \mathtt{U}$, the sum of the first two terms is equal to

$$\pi(y)\Big(\partial_t + \mathbf{v}.\partial_x\Big)\,\mathtt{U}\,.$$

If F were an affine function of u, (11) would be exactly the transport equation of linear geometric optics. In this sense it is familiar. In particular, for conservative linear operators the symmetric part of the zeroth order term yields growth or decay of amplitudes corresponding to converging or diverging rays. The transport equation (11) is a compact family of ordinary differential equations parametrized by $\Sigma_{\underline{T}} \cap \{t=0\}$. Since Ω_T is a domain of determinacy for $L(y,\partial_y)$ it follows that the vector field $\partial_t + \mathbf{v}.\partial_x$ is outgoing at the lateral boundary of Ω. Therefore, Picard's existence theorem for ordinary differential equations implies that the profile $\mathtt{U}$ exists locally and for T small is uniquely determined in Σ_T by its initial values.

Proposition 1.1 *Suppose that* $\underline{\mathtt{U}}(x,r) \in C^\infty\big((\Sigma_{\underline{T}} \cap \{t=0\}) \times \mathbb{R}\big)$ *is supported in* $|r| \le \overline{r}$ *and satisfies* $\pi(0,x)\,\underline{\mathtt{U}}(x,r) = \underline{\mathtt{U}}(x,r)$. *Then there is a* $0 < T \le \underline{T}$ *and a unique solution* $\mathtt{U}(y,r) \in C^\infty(\Sigma_T \times \mathbb{R})$ *to equations* $(10-11)$ *also supported in* $|r| \le \overline{r}$ *which satisfies the initial condition*

$$\mathtt{U}(0,x,r) = \underline{\mathtt{U}}(x,r)\,, \qquad x \in \Sigma_{\underline{T}} \;\cap\; \{t=0\}\,. \tag{12}$$

Our main theorem is a refinement of the following result which asserts that if the approximate solution exists for $0 \le t \le T$ then smooth exact solutions within $O(\epsilon)$ also exist on that time interval.

Theorem 1.2 *Suppose that* $0 < T \le \underline{T}$, *and that* $\mathtt{U} \in C^\infty(\Sigma_T \times \mathbb{R})$ *is a solution of* (11) *with support in* $|r| \le \overline{r} < \infty$. *Suppose that* $U_0 = \pi(y)U_0 \in C^\infty(\overline{\Omega}_T \times \mathbb{R})$ *is a smooth extension of* $\mathtt{U}$ *supported in a small neighborhood of* $\overline{\Sigma}_T \times \{|r| \le \overline{r}\}$. *Define a family of approximate solutions in* $\overline{\Omega}_T$ *by*

$$u^\epsilon_{\rm approx}(y) \;:=\; U_0(y, \phi(y)/\epsilon)\,. \tag{13}$$

The approximate solutions are asymptotically accurate in the sense that there is an $\epsilon_0 > 0$ *and a family of exact solutions* $v^\epsilon \in C^\infty(\overline{\Omega}_T)$ *of the equation*

$$L(y,\partial_y)\,v^\epsilon + F(v^\epsilon) = 0\,, \qquad 0 < \epsilon < \epsilon_0 \tag{14}$$

satisfying

$$\left\| Z_1 Z_2 \dots Z_M \left(u^\epsilon_{\text{approx}} - v^\epsilon\right)\right\|_{L^\infty(\overline{\Omega}_T)} = O(\epsilon) \quad \text{as} \quad \epsilon \to 0\,, \tag{15}$$

for any finite family of smooth vector fields $Z_1, \dots, Z_M$ on $\overline{\Omega}_T$ each tangent to $\Sigma_{\underline{T}}$.

The proof has four steps; correction, extension-truncation, linear conormal estimates, and a perturbation argument. The first and most important step is the construction of a corrector so that the residual $R^\epsilon := L\, u^\epsilon_{\text{approx}} + F(y, u^\epsilon_{\text{approx}})$ becomes $O(\epsilon)$. Then the family of corrected approximate solutions is extended to $t \geq -a$ with $a > 0$ preserving the essential geometric features of the set $\overline{\Omega}_T$ and the smallness of the residual. The exact solution v^ϵ is the solution to $L\, v^\epsilon + F(y, v^\epsilon) = \chi(t)\, R^\epsilon$ where the smooth cutoff function χ vanishes for $t > -a/3$ and is equal to 1 for $t < -2ta/3$. The solution v^ϵ is chosen equal to $u^\epsilon_{\text{approx}}$ for $t < -2a/3$. Finally a perturbation argument using linear conormal estimates shows that for ϵ small $v^\epsilon = u^\epsilon_{\text{approx}} + w^\epsilon$ with $w^\epsilon = O(\epsilon)$.

2 The approximate solution and first corrector.

As is usually the case in geometric optics expansions, the leading term alone when inserted in the equation does not have a small enough residual to yield a useful error estimate. One needs correctors. One of the striking facets of the problem of pulse propagation is that the corrector strategy which works in the case of wavetrains fails for pulses. Based on experience with wavetrains, it is natural to take as an *ansatz*

$$u^\epsilon := U_0(y, \phi(y)/\epsilon) + \epsilon\, U_1(y, \phi(y)/\epsilon)\,. \tag{16}$$

with $U_j(y, r)$ tending to zero as $|r| \to \infty$. The *ansatz* which works for wave trains is exactly of this form but with U_j periodic in r. Plugging (16) into the equation yields

$$L(y, \partial_y)(u^\epsilon) = L\big(y, \partial_y + \frac{d\phi(y)}{\epsilon}\,\partial_r\big)\,\Big(U_0(y,r) + \epsilon U_1(y,r)\Big)\Big|_{r=\phi(y)/\epsilon}\,,$$

$$F(y, u^\epsilon) \;\;=\;\; F\big(y, U_0(y,r) + \epsilon U_1(y,r)\big)\Big|_{r=\phi(y)/\epsilon}\,,$$

and

$$F\big(y, U_0(y,r) + \epsilon U_1(y,r)\big) = F(y, U_0(y,r)) + \epsilon H\big(\epsilon, y, U_0(y,r), \epsilon U_1(y,r)\big) \cdot U_1(y,r)\,.$$

The last assertion uses Taylor's Theorem with remainder to provide the smooth function H.

Combining these three equations one has the important residual formula

$$L(u^\epsilon) + F(u^\epsilon) = \Big(\frac{1}{\epsilon} W_{-1}(y,r) + W_0(y,r) + \epsilon W_1(\epsilon, y, r)\Big)\Big|_{r=\phi(y)/\epsilon}, \tag{17}$$

where

$$\begin{aligned} W_{-1}(y,r) &= L(y, d\phi(y))\, \partial_r U_0\,, \\ W_0(y,r) &= L(y, d\phi(y)\,) \partial_r U_1 + L(y, \partial_y)\, U_0 + F(y, U_0)\,, \\ W_1(\epsilon, y, r) &= L(y, \partial_y)\, U_1(y,r) + H\big(\epsilon, y, U_0(y,r), \epsilon U_1(y,r)\big) \cdot U_1(y,r)\,. \end{aligned} \tag{18}$$

The strategy is to choose U_0 and U_1 so that $W_{-1} = W_0 = 0$. In general, this turns out to be impossible. The problem occurs in the W_0 equation. Integrating (18) from $r = -\infty$ to $r = \infty$ shows that

$$L(y, d\phi(y))\, \big(U_1(y, \infty) - U_1(y, -\infty)\big) = \int_{-\infty}^{\infty} L(y, \partial_r)\, U_0(y,r) + F(y, U_0(y,r))\; dr\,.$$

If U_1 tends to zero as $r \to \pm\infty$ then the left hand side must vanish. However, for generic U_0, the right hand side is nonzero and it is therefore not possible to find such U_1. The solution is to allow the corrector to have different values at $r = \pm\infty$. Choose a function

$$g(r) \in C_0^\infty(\mathbb{R})\,, \qquad \text{with} \qquad \int_{-\infty}^{\infty} g(r)\; dr = 1\,. \tag{19}$$

Define

$$G(r) := \int_{-\infty}^{r} g(r)\; dr\,. \tag{20}$$

The new *ansatz* is

$$u^\epsilon(y) \;:=\; U_0(y, \phi(y)/\epsilon) + \epsilon \Big(U_1(y, \phi(y)/\epsilon) + c(y)\, G(\phi(y)/\epsilon)\Big)\,. \tag{21}$$

Computing as above shows that (17) is valid with

$$W_{-1}(y,r) = L(y, d\phi(y))\, \partial_r U_0 \tag{22}$$

$$W_0(y,r) = L(y, d\phi(y))\big(\partial_r U_1 + c(y) g(r)\big) + L(y, \partial_y) U_0 + F(y, U_0) \tag{23}$$

$$\begin{aligned} W_1(\epsilon, y, r) &= L(y, \partial_y)\, \big(U_1(y,r) + c(y) G(r)\big) + \\ &\quad H\big(\epsilon, y, U_0, \epsilon(U_1 + c(y) G(r))\,\big) \cdot \big(U_1 + c(y)\, G(r)\big)\,. \end{aligned} \tag{24}$$

To guarantee that $W_{-1} = 0$, we choose U_0 satisfying the polarization condition

$$\pi\, U_0 = U_0\,. \tag{25}$$

To analyse equation (23), introduce the partial inverse $Q(y) \in C^\infty(\overline{\Omega}_{\underline{T}})$ uniquely defined by the conditions

$$Q(y)\, \pi(y) = 0\,, \qquad \text{and} \qquad Q(y)\, L(y, d\phi(y)) = (I - \pi(y))\,. \tag{26}$$

The equation $W_0 = 0$ is equivalent to the pair of equations

$$\pi(y)\, W_0 = 0\,, \qquad \text{and} \qquad Q(y)\, W_0 = 0\,.$$

The first of these equations involves only U_0 and thanks to (10) is equivalent to

$$\pi(y)\, L(y, \partial_y)\, \pi(y)\, U_0 + \pi(y)\, F(y, U_0) \;=\; 0\,. \tag{27}$$

Expanding the operator in the first term yields

$$\pi(y)\, L(y, \partial_y)\, \pi(y) \;=\; \sum_{\mu=0}^{d} \pi(y)\, A_\mu(y)\, \pi(y)\, \frac{\partial}{\partial y_\mu} + \pi(y) \sum_{\mu=0}^{d} A_\mu(y) \frac{\partial \pi(y)}{\partial y_\mu}\,. \tag{28}$$

Thanks to Assumptions 3 and 4 and the fact that $A_0 = I$ one has (see [DJMR])

$$\pi(y)\, A_j(y)\, \pi(y) = \mathbf{v}_j(y)\, \pi(y)\,, \tag{29}$$

where $\mathbf{v}$ is the group velocity defined in equation (9). Inject this in the previous formula to find

$$\pi(y)\, L(y, \partial_y)\, \pi(y) \;=\; \pi(y) \left(\partial_t + \mathbf{v}.\partial_x\right) + \pi(y) \sum_{\mu=0}^{d} A_\mu(y) \frac{\partial \pi(y)}{\partial y_\mu}\,. \tag{30}$$

This together with (27) yields the important equation (11) of the last section. It also shows that (27) is a set of transport equations along rays which can be used to determine U_0 from its initial data.

When (25) as well as (27) are satisfied, one has $W_{-1} = 0 = \pi(y)\, W_0$. Then the leading term in the residual is $(1 - \pi(y)) W_0(y, \phi(y)/\epsilon)$ and is comparable in size to the approximate solution u^ϵ. This is typical of two scale asymptotic methods. The leading term in the expansion is not sufficient to create a residual which is small compared to u^ϵ.

The next goal is to construct a corrector to eliminate this leading term in the residual. Multiply (23) by $Q(y)$ to find

$$(1 - \pi(y)) W_0 = (I - \pi(y)) \Big(\partial_r U_1 + c(y) g(r)\Big) + Q(y) \Big(L(y, \partial_y)\, U_0 + F(y, U_0)\Big). \tag{31}$$

Thus W_0 vanishes exactly when the corrector satisfies

$$\big((I - \pi(y)\big) \Big(\partial_r U_1 + c(y) g(r)\Big) \;=\; -Q(y) \Big(L(y, \partial_y)\, U_0 + F(y, U_0)\Big)\,. \tag{32}$$

Integrating this equation and using the fact that $U_1(y, \pm\infty) = 0$ shows that

$$\big(I - \pi(y)\big) c(y) \;=\; -\int_{-\infty}^{\infty} Q(y) \Big(L(y, \partial_y)\, U_0 + F(y, U_0)\Big)\, dr\,. \tag{33}$$

When $c(y)$ satisfies this equation, there is an $\overline{r} < \infty$ so that the function

$$\big(I - \pi(y)\big) c(y) g(r) - Q(y) \Big(L(y, \partial_y)\, U_0 + F(y, U_0)\Big) \tag{34}$$

is smooth and compactly supported in $|r| \leq \overline{r}$ with vanishing r integral so (33) uniquely determines

$$(I - \pi(y))U_1(y,r) = \tag{35}$$
$$-\int_{-\infty}^{r} (I-\pi(y))c(y)g(y) + Q(y)\Big(L(y,\partial_y)U_0(y,r) + F(U_0(y,r))\Big)\, dr.$$

Neither $\pi(y)\, c(y)$ nor $\pi(y)U_1$ affect the value of W_0 so for simplicity we complete the specification of $c(y)$ and U_1 by

$$\pi(y)\, c(y) \;=\; 0 \;=\; \pi(y)\, U_1(y,r)\,. \tag{36}$$

Proposition 2.1 *Suppose that* $\mathtt{U} \in C^\infty(\Sigma_T \times \mathbb{R})$ *is a solution of* (11) *as in Proposition 1.1. Suppose that* $U_0 \in C^\infty(\overline{\Omega}_T \times \mathbb{R})$ *is an extension of* $\mathtt{U}$ *as in Theorem 1.2. Define* $c(y)$ *amd* U_1 *by formulas* (33), (35), *and* (36). *Then*
1. $c(y) \in C^\infty(\overline{\Omega}_T)$, $U_1 \in C^\infty(\overline{\Omega}_T \times \mathbb{R})$ *and* U_1 *all vanish for* $|r| > \overline{r}$.
2. *The approximate solution* u^ϵ *defined by the ansatz* (21) *satisfies* (17) *with* $W_{-1} = 0$ *and* $W_0\big|_{\Sigma_T} = 0$.
3. *For any finite family* $Z_1, \ldots, Z_M$ *of smooth vector fields on* Ω_T *each of which is tangent to* Σ_T *one has*

$$\sup_{0<\epsilon<1} \Big\| Z_1 Z_2 \ldots Z_M\, u^\epsilon \Big\|_{L^\infty(\Omega_T)} \;<\; \infty\,, \tag{37}$$

and

$$\Big\|\, Z_1 Z_2 \ldots Z_M \Big(L(y,\partial_y)\, u^\epsilon + F(y,u^\epsilon)\Big)\,\Big\|_{L^\infty(\Omega_T)} \;=\; O(\epsilon)\,, \qquad \text{as} \quad \epsilon \to 0\,. \tag{38}$$

Proof. The first two parts are immediate from the construction.

The leading term of the residual is equal to $W_0(y,\phi(y)/\epsilon)$. Since $W_0(y,r) = 0$ when $\phi(y) = 0$ and $d\phi \neq 0$ at such points, Taylor's Theorem implies that on a neighborhood of $\Sigma_T \times \mathbb{R}$, $W_0(y,r) = \phi(y)\, V(y,r)$ with $V \in C^\infty(\overline{\Omega}_T \times \mathbb{R})$ supported in $|r| \leq \overline{r}$. The leading term is then equal to $\epsilon\, \mathrm{K}(y,\phi(y)/\epsilon)$ with $\mathrm{K}(y,r) := r\, V(y,r)$. This together with the formula for W_1 shows that the residual is of the form $\epsilon\, K(\epsilon,y,\phi(y)/\epsilon)$ where $K \in C^\infty([0,1] \times \overline{\Omega}_T \times \mathbb{R})$ with $\partial_r K$ supported in a compact interval of r uniformly in ϵ, y.

To prove the third part it suffices to show that for such $K(\epsilon,y,r)$,

$$\sup_{0<\epsilon<1} \|\, Z_1 Z_2 \ldots Z_M K(\epsilon,y,\phi(y)/\epsilon)\, \|_{L^\infty(\Omega_T)} \;<\; \infty\,. \tag{39}$$

This is proved by showing that $Z_1 Z_2 \ldots Z_M K(\epsilon,y,\phi(y)/\epsilon)$ is a finite sum of terms $K_j(\epsilon,y,\phi(y)/\epsilon)$ with K_j having the same properties as K.

By an inductive argument it suffices to show that $Z\, K(\epsilon,y,\phi(y)/\epsilon)$ is a finite sum of this type.

The chain rule implies that

$$Z\big(\, K(\epsilon,y,\phi(y)/\epsilon)\,\big) = (ZK)(\epsilon,y,\phi(y)/\epsilon) + K_r(\epsilon,y,\phi(y)/\epsilon)\, Z\phi(y)/\epsilon\,. \tag{40}$$

Since Z is tangent to Σ it follows that $Z\phi = 0$ when $\phi = 0$. Since the zeroes of ϕ are nondegenerate, Taylor's Theorem implies that there is a smooth function $a(y)$ on a neighborhood of $\overline{\Sigma}_T$ so that $Z\phi = a(y)\phi(y)$. Thus, the right hand side of (40) is equal to $\mathbf{K}(\epsilon, y, \phi/\epsilon)$ with $\mathbf{K}(\epsilon, y, r) := (ZK)(\epsilon, y, r) + K_r(\epsilon, y, r)a(y)r$. This completes the proof. □

3 Extension to negative times.

The proof of the main theorem begins by extending the approximate solution to a domain of determination which reaches into the past. The extension is done in three stages, extend ϕ, then extend $\mathtt{U}$, then extend Ω_T.

First step, In $\{t = 0\}$ extend $\phi(0, x)$ to be smooth on a neighborhood of $\Sigma_{\underline{T}} \cap \{t = 0\}$. Then solve the reduced eikonal equation (8) to construct an extension of ϕ to a neighborhood of $\Sigma_{\underline{T}} \cap \{t = 0\}$. In doing this note that the domain of determinacy part of Assumption 1 implies that the ray direction $\partial_t + \mathbf{v}.\partial_x$ is tangent or outgoing at the lateral boundary of Ω_T, which permits this extension without altering the values of ϕ inside Ω_T.

Having extended ϕ, the set $\Sigma_{\underline{T}} \cap \{t = 0\}$ extends across the boundary of $\mathcal{O}$ wherever the two surfaces meet, thanks to the transversality condition in Assumption 3.

Extend the initial values $\mathtt{U}(0, x)$ across $\partial\mathcal{O}$ as a smooth function satisfying the polarization identity (10). Then solve the intial value problem defined by (11) to extend $\mathtt{U}$ to the past on a neighborhood of $\Sigma_{\underline{T}} \cap \{t = 0\}$.

Finally as t moves negative, move the boundary of Ω outward at a speed which slightly exceeds the maximal (in absolute value) speed of propagation for the system $L(y, \partial)$ at the point. The intersection of the domain swept out for $\{t > -a\}$ is denoted $\Omega_{a,T}$.

This argument proves the following lemmas. The first extends Assumptions 3 and 4 to $\Omega_{a,T}$ while the second extends Proposition 2.1.

Lemma 3.1 *If $a > 0$ is sufficiently small the above procedure yields a domain of determinacy $\Omega_{a,T}$ with C^1 lateral boundary, a characteristic surface $\Sigma_{a,T}$ which is transverse to the boundary with conormal $\overline{\mathcal{N}^*}(\Sigma_{a,T})$ belonging to a smooth sheet of the characteristic variety, and, a solution $\mathtt{U} \in C^\infty(\overline{\Sigma}_{a,T} \cap \overline{\Omega}_{a,T})$ to the polarization and profile equations* (10) *and* (11).

Lemma 3.2 *The analogue of Proposition 2.1 in $\Omega_{a,T}$ as opposed to Ω_T is true. That is, suppose that U_0 extends $\mathtt{U}$ from $\overline{\Sigma}_{a,T}$ to $\overline{\Omega}_{a,T}$, satisfies the polarization identity, and is supported in* $\{\text{a small neighborhood of } \overline{\Sigma}_{a,T}\} \times [-r, r]$. *Then defining an approximate solution on $\Omega_{a,T}$ by* (33), (35), (36), *and* (21), *the residuals satisfy* (37) *and* (38) *with Ω_T replaced by $\Omega_{a,T}$.*

4 Linear conormal estimates.

In this section we recall two types of conormal estimates associated with $\Sigma_{a,T} \subset \Omega_{a,T}$. The first family are L^2 estimates for derivatives tangent to $\Sigma_{a,T}$ and the second family are pointwise estimates coming ultimately from integration along characteristics. These ideas were perfected in the eighties in the study of the nonlinear propagation of singularities for hyperbolic equations and systems. To our knowlege the conormal category was introduced in these problems by Bony in [Bo1]. The present treatment is a semiglobal version of [RR]. References for further developements can be found in the book of Beals [Be].

The motivating problem is that the exact solution v^ϵ of the main theorem is constructed as a perturbation of the approximate solution u^ϵ. To analyse this process one needs control on the family of operators linearized at the approximate solutions u^ϵ,

$$\Big[L(y,\partial) + D_u F(y,u^\epsilon)\Big]^{-1}.$$

The zeroth order term, $D_u F(y,u^\epsilon)$ has bounded derivatives tangent to $\Sigma_{a,T}$. The derivative transverse to $\Sigma_{a,T}$ can grow like $1/\epsilon$.

Definition. *Denote by $\mathcal{Z}$ the set of smooth vector fields on $\overline{\Omega}_{a,T}$ which are tangent to $\overline{\Sigma}_{a,T}$*

$\mathcal{Z}$ is a $C^\infty(\overline{\Omega}_{a,T})$ module, that is it is closed under multiplication by functions in $C^\infty(\overline{\Omega}_{a,T})$. It is closed under Lie bracket in the sense that if Z_1 and Z_2 belong to $\mathcal{Z}$ then so does the commutator $[Z_1, Z_2]$. In local coordinates so that $\Sigma = \{y_d = 0\}$, $\mathcal{Z}$ is generated by $\partial_1, \dots, \partial_{d-1}, y_d\partial_d$ in the sense that elements of $\mathcal{Z}$ are linear combinations of these fields coefficients in $C^\infty(\overline{\Omega}_{a,T})$. A finite partition of unity for $\overline{\Omega}_{a,T}$ shows that $\mathcal{Z}$ is finitely generated as a module, that is, there is a finite set of elements $Z_1, \dots, Z_K \in \mathcal{Z}$ so that an arbitrary member of $\mathcal{Z}$ is a linear combination of the Z_j with coefficients in $C^\infty(\overline{\Omega}_{a,T})$.

With the linearized operators as a model we study estimates uniform in ϵ for the family of operators

$$L^\epsilon(y,\partial) \ := \ L(y,\partial) + B^\epsilon(y)\,, \tag{41}$$

where $B^\epsilon \in C^\infty(\overline{\Omega}_{a,T})$ is defined for $0 < \epsilon \le \epsilon_0$ and satisfies the following uniform bounds. If $M < \infty$ is any positive integer and $Z_1\,, \dots, Z_M \in \mathcal{Z}$, then

$$\sup_{0<\epsilon\le\epsilon_0} \big\| \, Z_1 \, \dots, Z_M \, B^\epsilon \, \big\|_{L^\infty(\overline{\Omega}_{a,T})} \ < \ \infty\,. \tag{42}$$

This section adds some ideas to earlier work to achieve global estimates throughout $\Omega_{a,T}$ and to treat the characteristic surface whose constant multiplicity may be greater than one. It also introduces some innovations borrowed from progress made in the last years in the study of nonlinear geometric optics. This is particularly true for the presentation of the transport equations.

Definition. *For $0 \leq s \in \mathbb{N}$, the space $H^s_{\mathcal{Z}}$ is the set of $u \in L^2(\Omega_{a,T})$ so that for any $\alpha \in \mathbb{N}^K$ with $|\alpha| \leq s$ one has $(Z_1, Z_2, \ldots, Z_K)^\alpha u \in L^2(\Omega_{a,T})$ where $Z_1, \ldots, Z_K$ is a generating set for $\mathcal{Z}$. The linear space $H^s_{\mathcal{Z}}$ is a Hilbert space with a family of pairwise equivalent norms defined by*

$$\|u\|^2_{s,\lambda} := \sum_{|\alpha| \leq s} \left\| e^{-\lambda t} \left(Z_1, \ldots, Z_K\right)^\alpha u \right\|^2_{L^2(\Omega_{a,T})}. \tag{43}$$

The symbol $\| \ \|_{H^s_{\mathcal{Z}}}$ when used without further clarification is meant to denote one of the family of norms.

The main result of the section is the following.

Theorem 4.1 *Suppose that $(3+d)/2 < s \in \mathbb{N}$, $0 \leq k \in \mathbb{N}$, and that for all $|\alpha| \leq k$,*

$$\left(Z_1, \ldots, Z_K\right)^\alpha f \in L^\infty(\Omega_{a,T}) \ \cap \ H^s_{\mathcal{Z}}$$

and vanishes on a neighborhood of $\{t = -a\}$. Then the unique $u^\epsilon \in L^2(\Omega_{a,T})$ solving the initial value problem

$$L^\epsilon u^\epsilon = f \ \text{ in } \ \Omega_{a,T}, \qquad u^\epsilon\big|_{t=-a} = 0$$

satisfies for all $|\alpha| \leq k$,

$$\left(Z_1, \ldots, Z_K\right)^\alpha u^\epsilon \in L^\infty(\Omega_{a,T})) \ \cap \ H^s_{\mathcal{Z}},$$

and for all $|\beta| \leq k+1$

$$\left(Z_1, \ldots, Z_K\right)^\beta (I - \pi(y)) u^\epsilon \in L^\infty(\Omega_{a,T}).$$

There is a constant C independent of f and ϵ so that

$$\begin{aligned} &\sum_{|\alpha| \leq k} \|(Z_1, \ldots, Z_K)^\alpha u^\epsilon\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}} + \\ &\quad \sum_{|\beta| \leq k+1} \|(Z_1, \ldots, Z_K)^\beta \big((1 - \pi(y)) u^\epsilon\big)\|_{L^\infty(\Omega_{a,T})} \\ &\qquad \leq C \sum_{|\alpha| \leq k} \|\left(Z_1, \ldots, Z_K\right)^\alpha f\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}}. \end{aligned} \tag{44}$$

To prove the theorem the key is to prove (44) as an *a priori* estimate. This is broken into two parts, L^2 estimates and then L^∞ estimates which are proved in the next two subsections. Readers familiar with such conormal estimates may want to skip directly to §5.

4.1 $\mathbf{L^2}$ estimates.

The L^2 estimates are implied by the next elementary L^2 estimate proposition followed by a nontrivial commutation argument.

Proposition 4.2 *For each $0 \leq s \in \mathbb{N}$ there are constants λ_s and $C_s > 0$ so that for all $u \in C^{s+1}(\overline{\Omega}_{a,T})$ with $(Z_1, \ldots, Z_K)^\alpha u\big|_{t=-a} = 0$ for all $|\alpha| \leq s$, one has*

$$\forall \lambda > \lambda_s\,, \qquad (\lambda - \lambda_s)\big\|u\big\|_{s,\lambda} \;\leq\; C_s\big\|L^\epsilon\, u\big\|_{s,\lambda}\,. \tag{45}$$

This estimate descends from the following basic estimate.

Lemma 4.3 *There is a $\lambda_0 > 0$ so that for all $u \in C^1(\overline{\Omega}_{a,T}))$ with $u\big|_{t=-a} = 0$ one has*

$$\mathrm{Re}\,\big(e^{-\lambda t}\, L^\epsilon u\,,\, e^{-\lambda t} u\big)_{\Omega_{a,T}} \;\geq\; (\lambda - \lambda_0)\;\mathrm{Re}\,\big(e^{-\lambda t}\, u\,,\, e^{-\lambda t} u\big)_{\Omega_{a,T}}\,, \tag{46}$$

where the scalar product is that in $L^2(\Omega_{a,T})$.

Proof of Lemma. Begin with the identity

$$\mathrm{Re}\,\big(e^{-\lambda t}\, L^\epsilon\, u\,,\, e^{-\lambda t} u\big)_{\Omega_{a,T}} \;=\; \mathrm{Re}\,\big((L^\epsilon + \lambda I)\, e^{-\lambda t} u\,,\, e^{-\lambda t} u\big)_{\Omega_{a,T}}\,.$$

Integration by parts in the L term on the right yields

$$\mathrm{Re}\,\big(L^\epsilon\, e^{-\lambda t}\, u, e^{-\lambda t} u\big)_{\Omega_{a,T}} = \int_{\partial\Omega_{a,T}} \Big\langle \sum_{\mu\geq 0} \nu_\mu A_\mu\, e^{-\lambda t} u\,,\, e^{-\lambda t} u\Big\rangle\, d\sigma$$
$$+ \int_{\Omega_{a,T}} \Big\langle\Big(\frac{B^\epsilon + B^{\epsilon *}}{2} + \sum_{\mu\geq 0} \frac{\partial A_\mu}{\partial y_\mu}\Big) e^{-\lambda t} u, e^{-\lambda t} u\Big\rangle\, dy\,,$$

where ν is the unit outward normal and $d\sigma$ the element of surface area.

Since u vanishes at $t = -a$ the integral over that part of the boundary vanishes. The outward normal to the boundary $\{t = T\}$ is $(1, 0, \ldots, 0)$ and the normal at the lateral boundary is $\nabla_{t,x}(t - \ell(x)) = (1, -\nabla_x \ell)$. Therefore, Assumption 1 guarantees that on these parts of the boundary the matrix $\sum \nu_\mu A_\mu$ is nonegative so that the boundary integral is nonnegative and the lemma follows since B^ϵ and $\partial_\mu A_\mu$ are uniformly bounded. □

The proof of the proposition proceeds by a commutation argument in suitable coordinates together with a patching using a partition of unity. The coordinates for both dependent and independent variables are chosen to achieve the commutation relation (52) between L^ϵ and $\mathcal{Z}$. To prepare for that we begin with a moderately general discussion.

Definition. *Denote by $\mathcal{M}$ the family of smooth $N \times N$ matrix valued functions on $\overline{\Omega}_{a,T}$. If $P(y,\partial)$ is an $N \times N$ system of linear partial differential operators of order one whose coefficients are smooth on $\overline{\Omega}_{a,T}$, we say that*

$$P \subset \mathcal{M}\,L + \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,, \tag{47}$$

when there are smooth matrix valued functions $M_L(y), M_1(y), \ldots, M_K(y)$, and $M_0(y)$ on $\overline{\Omega}_{a,T}$ so that

$$P = M_L\,L + \sum_{j=1}^{K} M_j\,Z_j + M_0\,.$$

The definition of

$$P \subset \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,, \tag{48}$$

is analgous but without the $M_L\,L$ term.

When condition (48) holds P is sometimes called *tangential* and sometimes *totally characteristic*. The first part of the next lemma is both well known and easy. We learned it from the work of Melrose. The second part of the lemma is closely related to ideas investigated in [MeR], [Bo1, Bo2], and [RR].

Since the principal symbol $P(y,\eta)$ of a differential operator is homogeneous of degree one in the the fiber variable η and the fiber in $N^*(\Sigma)$ has dimension equal to one, it follows that for $y \in \Sigma$ and $(y,\eta) \in N^*(\Sigma)$, $\ker \sigma(P)(y,\eta)$ and $\operatorname{range} \sigma(P)(y,\eta)$ do not depend on the choice of $\eta \neq 0$.

Lemma 4.4 i. (48) *holds if and only if the principal symbol $\sigma(P)$ satisfies $\sigma(P)|_{N^*(\Sigma)} = 0$.*
ii. *Suppose that L is an $N \times N$ system of first order partial differential operators with smooth coefficients on $\overline{\Omega}_{a,T}$, not necessarily hermitian, and that* $\ker \sigma(L)|_{N^*(\Sigma)}$ *and* $\operatorname{range} \sigma(L)|_{N^*(\Sigma)}$ *are smooth vector bundles on Σ satisfying*

$$\operatorname{kernel} \sigma(L)|_{N^*(\Sigma)} \;\cap\; \operatorname{range} \sigma(L)|_{N^*(\Sigma)} \;=\; \{0\}\,. \tag{49}$$

Then (47) *holds if and only if*

$$\sigma(P)|_{\operatorname{kernel} \sigma(L)|_{N^*(\Sigma)}} = 0\,. \tag{50}$$

Remark. The hypothesis (49) is equivalent to

$$\mathbb{C}^N = \operatorname{kernel} \sigma(L)|_{N^*(\Sigma)} \;\oplus\; \operatorname{range} \sigma(L)|_{N^*(\Sigma)}\,, \tag{51}$$

and also to

$$\operatorname{kernel} \sigma(L)|_{N^*(\Sigma)} = \operatorname{kernel} \sigma(L)^2|_{N^*(\Sigma)}\,.$$

These equivalent hypotheses are automatically satisfied when the A_μ are hermitian.

Proof. The *only if* parts of each of assertions **i.** and **ii.** are immediate.

The assertions are local and independent of the choice of y coordinates so it suffices to reason in coordinates so that $\Sigma = \{y_d = 0\}$. In such coordinates one has

$$P = \sum_{\mu=0}^{d} C_\mu(y)\,\partial_\mu + \text{l.o.t.} \;\equiv\; C_d(y)\,\partial_d \quad \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}$$

Using Taylor's theorem write

$$C_d(y_0, \ldots, y_{d-1}, y_d) = C_d(y_0, \ldots, y_{d-1}, 0) + y_d\,\tilde{C}_d(y)\,.$$

Then since $y_d\,\partial_d \in \mathcal{Z}$ one has

$$P \;\equiv\; C_d(y_0, \ldots, y_{d-1}, 0)\,\partial_d \quad \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,.$$

On the other hand $\sigma(P)|_{N^*(\Sigma)} = C_d(y_0, \ldots, y_{d-1}, 0)\eta_d$, so if $\sigma(P)|_{N^*(\Sigma)} = 0$ then (48) holds which proves the *if* part of **i.**.

Next turn to **ii.**. Analogous computations for L show that

$$M\,L \;\equiv\; M(y_0, \ldots, y_{d-1}, 0)\,A_d(y_0, \ldots, y_{d-1}, 0)\,\partial_d \quad \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,.$$

In these coordinates, (50) is equivalent to the assertion

$$\text{kernel}\,A_d(y_0, \ldots, y_{d-1}, 0) \;\subset\; \text{kernel}\,C_d(y_0, \ldots, y_{d-1}, 0)\,.$$

This together with (51) implies that there is a unique smooth matrix valued function $M(y_0, \ldots, y_{d-1})$ so that

$$\text{kernel}\,A_d(y_0, \ldots, y_{d-1}, 0) \subset \text{kernel}\,M(y_0, \ldots, y_{d-1})\,,$$

and

$$M(y_0, \ldots, y_{d-1})\,A_d(y_0, \ldots, y_{d-1}, 0) \;=\; C_d(y_0, \ldots, y_{d-1}, 0)\,.$$

It follows that $P - ML \subset \mathcal{M}\mathcal{Z} + \mathcal{M}$ proving the *if* part of **ii.**. □

The key to conormal estimates are commutation relations of the form

$$[\mathcal{Z}, L] \;\subset\; \mathcal{M}\,L + \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,. \tag{52}$$

It is clear that such a relation is invariant under a change of independent variable $y' = y'(y)$. In addition, (52) is invariant under multiplying L by a smooth invertible matrix valued function. However, linear systems are also invariant under linear changes in the dependent variable $u' = M(y)u(y)$. One of the subtleties of the commutation relation (52) is that they are not invariant under such changes of variables. They depend on the choice of basis for the dependent variable, and that choice may need to depend on y as the next example illustrates. This is in sharp contrast to the identities (47) and (48) which hold independent of the choice of basis.

Example. Consider $d = N = 2$ and $\Sigma = \{y_2 = 0\}$. Suppose that $\mathbf{r}(y_1)$ a smooth unit vector valued function of y_1 with $\mathbf{r}' \neq 0$. Consider operators with

$$L \;\equiv\; |\mathbf{r}(y_1)\rangle\langle\mathbf{r}(y_1)|\,\partial_2 \quad \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,. \tag{53}$$

Then

$$[\partial_1, L] \;\equiv\; \Big(|\mathbf{r}'(y_1)\rangle\langle\mathbf{r}(y_1)| + |\mathbf{r}(y_1)\rangle\langle\mathbf{r}'(y_1)|\Big)\,\partial_2 \quad \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,. \tag{54}$$

Since the coefficient of ∂_2 in (53) has rank one, and the coefficient in the commutator has rank two the operator $[\partial_1, L]$ does not satisfy (50) and therefore the commutator identity (52) does not hold.

On the other hand, if one makes a smooth unitary change of variables in $\mathbb{C}^2$

$$u = U(y_1)^* v \quad \text{with} \quad U(y_1)\,\mathbf{r}(y_1) = (1, 0)$$

and then multiplies by U, the operator L is transformed to a symmetric hyperbolic operator $\tilde{L}$ with leading coefficients $U\,A_\mu\,U^*$. Therefore

$$\tilde{L} = U\,|\mathbf{r}(y_1)\rangle\langle\mathbf{r}(y_1)|\,U^*\,\partial_2 \quad \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,.$$

However, by construction,

$$U\,|\mathbf{r}(y_1)\rangle\langle\mathbf{r}(y_1)|\,U^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and it follows that all the commutators $[Z, \tilde{L}]$ satisfy (50).

Lemma 4.5 *When assumptions 1,3, and 4 are satisfied and $\underline{y} \in \overline{\Sigma}_{a,T}$ there is a neighborhood of $\underline{y}$ and a smooth unitary matrix valued function $U(y)$ so that the operator $\tilde{L}v := U\,L\,(U^* v)$ satisfies*

$$[\mathcal{Z}, \tilde{L}] \;\subset\; \mathcal{M}\,\tilde{L} + \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,. \tag{55}$$

Proof. The commutator identity holds or not independent of the choice of coordinates y. On a neighborhood of $\underline{y}$ introduce local coordinates preserving the time coordinate and so that $\Sigma = \{y_d = 0\}$.

Lemma 3.1 implies that $\ker L(y, d\phi)$ is a smooth vector bundle on $\overline{\Sigma}_{a,T}$. Denote by $k \geq 1$ the dimension of the fibers on a neighborhood of $\underline{y}$. Write $\mathbb{C}^N = \mathbb{C}^k \times \mathbb{C}^{N-k}$ so that $\mathbb{C}^k$ (resp $\mathbb{C}^{N-k}$) is the k (resp. $N-k$) dimensional subspace of vectors whose last (resp. first) components vanish. On a neighborhood of $\underline{y}$ one can choose a smooth unitary matrix $U(y)$ so that

$$\ker L(y, d\phi(y)) \;=\; U(y)^*\,\Big[\mathbb{C}^k\Big]\,.$$

Then

$$\tilde{L} := U\,L\,U^* \;\equiv\; \begin{pmatrix} 0_{k\times k} & 0 \\ 0 & \text{Invertible}_{(N-k)\times(N-k)} \end{pmatrix} \partial_d \ \ \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,,$$

and

$$[Z, \tilde{L}] \;\equiv\; \begin{pmatrix} 0_{k\times k} & 0 \\ 0 & * \end{pmatrix} \partial_d \ \ \text{mod}\ \mathcal{M}\,\mathcal{Z} + \mathcal{M}\,.$$

The criterion of Lemma 4.4.**ii.** then implies that (55) holds. $\square$

Proof of Proposition 4.2. For each $y \in \overline{\Sigma}_{a,T}$ choose a neighborhood and unitary map as in Lemma 4.5. For each $y \in \overline{\Omega}_{a,T} \setminus \overline{\Sigma}_{a,T}$ choose a relatively open neighborhood in $\overline{\Omega}_{a,T}$ which does not meet $\overline{\Sigma}_{a,T}$. To the neighboorhoods disjoint from Σ associate the unitary matrix valued function which is identically equal to I.

Choose a finite cover of $\overline{\Omega}_{a,T}$ by neighborhoods from the preceding paragraph, and a smooth partition of unity $\{\psi_i \in C^\infty(\overline{\Omega}_{a,T})\}_{1\le i\le \kappa}$ subordinate to this cover. For each i denote by U_i the associated unitary matrix valued function.

We now use the big system trick to derive the estimate (45). Define a long vector of functions on $\Omega_{a,T}$ by

$$\mathbf{u} \ := \ \Big(U_1(y)\psi_1(y)u\,,\, U_2(y)\,\psi_2(y)u\,,\,\dots\,,\, U_\kappa(y)\psi_\kappa(y)u\Big)\,. \tag{56}$$

Define a diagonal operator

$$\mathbf{L} \ := \ \operatorname{diag}\big\{\, U_i\, L\, U_i^* \,\big\}\,. \tag{57}$$

The equation for u^ϵ then implies an equation

$$\Big(\mathbf{L} + \mathbf{B}^\epsilon\Big)\, \mathbf{u} = \mathbf{f}\,, \tag{58}$$

with

$$\big\|\mathbf{f}\big\|_{s,\lambda} \ \le \ C\Big(\big\|f\big\|_{s,\lambda} + \big\|u\big\|_{s,\lambda}\Big)\,, \qquad \text{for all } \lambda, \tag{59}$$

and

$$\sup_\epsilon \big\|(Z_1,\dots,Z_K)^\alpha \mathbf{B}^\epsilon\big\|_{L^\infty(\Omega_{a,T})} < \infty\,.$$

Note that the diagonal entries of $\mathbf{L}$ are not everywhere defined in $\Omega_{a,T}$ but that they are applied in (58) only to functions with compact support within their domain of definition. The interest of the bold equation is that one has the estimate

$$(\lambda - \lambda_0)\,\big\|\mathbf{u}\big\|_{0,\lambda} \ \le \ \big\|(\mathbf{L} + \mathbf{B}^\epsilon)\mathbf{u}\big\|_{0,\lambda}$$

which is the case $s = 0$ of the proposition, *and* one has the commutation relation

$$[\mathcal{Z}, \mathbf{L}] \subset \mathcal{M}\mathbf{L} + \mathcal{M}\mathcal{Z} + \mathcal{M}\,. \tag{60}$$

The proof is by induction. Assuming the case $s-1$, we prove the case s. Compute

$$\begin{aligned}
(\mathbf{L} + \mathbf{B}^\epsilon)\, Z\mathbf{u} \ &= \ Z\,(\mathbf{L} + \mathbf{B}^\epsilon)\mathbf{u} + [\mathbf{L}, Z]\mathbf{u} - (Z\mathbf{B}^\epsilon)\mathbf{u} \\
&= \ Z\,(\mathbf{L} + \mathbf{B}^\epsilon)\mathbf{u} + \mathcal{M}\mathbf{L}\mathbf{u} + \mathcal{M}\mathcal{Z}\mathbf{u} + \mathcal{M}\mathbf{u} - (Z\mathbf{B}^\epsilon)\mathbf{u} \\
&= \ Z\,(\mathbf{L} + \mathbf{B}^\epsilon)\mathbf{u} + \mathcal{M}(\mathbf{L} + \mathbf{B}^\epsilon)\mathbf{u} + \mathcal{M}\mathcal{Z}\mathbf{u} + \mathcal{M}^\epsilon\mathbf{u}\,,
\end{aligned} \tag{61}$$

where the family $\mathcal{M}^\epsilon$ contains terms from the commutator and $[\mathcal{Z}, \mathbf{L}]$, terms of the form $\mathcal{M}\mathbf{B}^\epsilon u$ and terms of the form $(\mathcal{Z}\mathbf{B}^\epsilon)u$. Therefore $\mathcal{M}^\epsilon$ satisfies uniform bounds like $\mathbf{B}^\epsilon$.

The inductive hypothesis implies that

$$(\lambda-\lambda_{s-1})\|Z\mathbf{u}\|_{s-1,\lambda} \le C_{s-1}\|(\mathbf{L}+\mathbf{B}^\epsilon)\, Z\mathbf{u}\|_{s-1,\lambda} \le C\left(\|(\mathbf{L}+\mathbf{B}^\epsilon)\mathbf{u}\|_{s,\lambda}+\|\mathbf{u}\|_{s,\lambda}\right),$$

where (61) is used in the last estimate. Summing over $Z \in \{Z_1 \ldots, Z_K\}$ and adding the estimate from the case $s-1$ yields

$$(\lambda-\lambda_{s-1})\Big(\|\mathbf{u}\|_{s-1,\lambda}+\sum_1^K \|Z_j\mathbf{u}\|_{s-1,\lambda}\Big) \;\le\; C\,\Big(\|\,(\mathbf{L}+\mathbf{B}^\epsilon)\mathbf{u}\|_{s,\lambda}+\|\mathbf{u}\|_{s,\lambda}\Big)\,.$$

Absorbing the last term on the right in the left hand side yields

$$(\lambda-\lambda_s)\|\mathbf{u}\|_{s,\lambda} \;\le\; C_s\,\|\,(\mathbf{L}+\mathbf{B}^\epsilon)\,\mathbf{u}\|_{s,\lambda}\,,$$

which is equivalent to the case s of the proposition. □

4.2 $\mathbf{L}^\infty$ estimates.

The next Sobolev type lemma gives two easy L^∞ estimates.

Lemma 4.6 *If $s > (1+d)/2$ and $u \in H^s_{\mathcal{Z}}$ and $\nabla u \in H^{s-1}_{\mathcal{Z}}$, then $u \in C(\overline{\Omega}_{a,T})$ and there is a constant $C = C(s, \Omega_{a,T})$ so that*

$$\big\|u\big\|_{L^\infty(\Omega_{a,T})} \;\le\; C\,\Big(\,\big\|\nabla u\big\|_{H^{s-1}_{\mathcal{Z}}}+\big\|u\big\|_{H^s_{\mathcal{Z}}}\,\Big)\,.$$

Proof. At points away from $\overline{\Sigma}$ the standard Sobolev imbedding theorem applies.

If u satisfies the hypotheses then so does $\psi\, u$ for any $\psi \in C^\infty(\overline{\Omega}_{a,T})$ so it suffices to consider functions u with support in small local coordinate neighborhoods.

On a neighborhood of points of $\overline{\Sigma} \setminus \partial\overline{\Omega}$ choose local coordinates so that $\overline{\Sigma} = \{y_d = 0\}$. Then the hypotheses on u and ∇u imply that

$$\psi u \in L^2(y_d\,;\,H^s(\mathbb{R}^d)) \quad \text{and} \quad \partial_d(\psi u) \in L^2(y_d\,;\,H^{s-1}(\mathbb{R}^d))\,.$$

Standard trace theorems imply that ψu is a continuous function of y_d with values in $H^{s-1/2}(\mathbb{R}^d) \subset (L^\infty \cap C)(\mathbb{R}^d)$.

At points of $\overline{\Sigma}$ at which $t = -a$ or $t = T$ but not on the lateral boundary of Ω the local coordinates as above can be chosen respecting the time variable. Repeating the argument of the last paragraph shows that ψu is a continuous function of y_d with values in $H^{s-1/2}(\mathbb{R}^d_+) \subset (L^\infty \cap C)(\mathbb{R}^d_+)$.

Finally where $\overline{\Sigma}$, $\partial\Omega$ and $t = -a$ or $t = T$ meet, the transversality in Assumption 3 implies that $d \ge 2$ and the surfaces are in general position. One can choose local coordinates so that $\overline{\Sigma} = \{y_d = 0\}$ and Ω is the quadrant $\{y_0 > 0\} \cap \{y_1 > 0\}$. Then, repeating as above one finds that ψu is a continuous function of y_d with values in $H^{s-1/2}(\mathbb{R}^2_+ \times \mathbb{R}^{d-2}) \subset (L^\infty \cap C)(\mathbb{R}^2_+ \times \mathbb{R}^{d-2})$. □

Lemma 4.7 *If $s > (1+d)/2$ and $u \in H^s_{\mathcal{Z}}$ and ϕ is the defining function of $\Sigma_{a,T}$, then $\phi u \in C(\overline{\Omega}_{a,T})$ and there is a constant C so that*

$$\left\| \phi(y) u \right\|_{L^\infty(\Omega_{a,T})} \ \le\ C \left\| u \right\|_{H^s_{\mathcal{Z}}} .$$

Proof. Choose a smooth vector field X on Σ which is transverse to Σ. Since $X(\phi u) = (\phi X)u + (X\phi)u$ and $\phi(y)X$ is tangential it follows that $X(\phi u) \in H^{s-1}_{\mathcal{Z}}$ and

$$\| X(\phi u) \|_{H^{s-1}_{\mathcal{Z}}} \ \le\ C \, \| \phi u \|_{H^s_{\mathcal{Z}}} .$$

Therefore

$$\| \nabla (\phi u) \|_{H^{s-1}_{\mathcal{Z}}} \ \le\ C \left(\| X(\phi u) \|_{H^{s-1}_{\mathcal{Z}}} + \| \phi u \|_{H^{s-1}_{\mathcal{Z}}} \right) \ \le\ C' \, \| \phi u \|_{H^{s-1}_{\mathcal{Z}}} .$$

An application of the preceding lemma completes the proof. □

For elements of $H^s_{\mathcal{Z}}$ one therefore has L^∞ control except near Σ. Using the differential equation one has L^∞ control for $(I-\pi)\,u$ as the next result shows. Estimates like these at a noncharacteristic hypersurface Σ go under the name partial hypoellipticity. Here the surface is characteristic and one might call the estimate *partial partial hypoellipticity.*

Lemma 4.8 *Denote by $\pi(y) \in C^\infty(\overline{\Omega}_{a,T})$ a smooth extension, from a neighborhood of $\overline{\Sigma}_{a,T}$ to all of $\overline{\Omega}_{a,T}$, of the orthogonal projector on* $\ker L(y, d\phi)$. *There is a constant C so that for all $u \in C^1(\overline{\Omega}_{a,T})$ and $\epsilon \in]0, \epsilon_0]$*

$$\left\| \, \nabla \, (I - \pi(y)) \, u \, \right\|_{L^2(\Omega_{a,T})} \ \le\ C \left(\left\| u \right\|_{H^1_{\mathcal{Z}}} + \left\| L^\epsilon u \right\|_{L^2(\Omega_{a,T})} \right) .$$

Proof of Lemma. There is a uniquely determined smooth partial inverse $Q(y)$ to $L(y, d\phi(y))$ defined on a neighborhood of $\overline{\Sigma}_{a,T}$ by the conditions

$$Q\,\pi = 0\,, \quad \text{and} \quad Q(y) L(y, d\phi(y)) = I - \pi(y) \,.$$

In local coordinates so that $\Sigma = \{y_d = 0\}$,

$$Q\, L(y, \partial) \ \equiv\ (I - \pi(y))\, \partial_d \quad \mathrm{mod} \quad \mathcal{M}\mathcal{Z} + \mathcal{M} \,,$$

so

$$\partial_d \, (I - \pi(y))\, u \ \equiv\ Q\, Lu \quad \mathrm{mod} \quad \mathcal{M}\mathcal{Z}u + \mathcal{M}u \,,$$

whence

$$\left\| \, \partial_d \, (I - \pi(y))\, u \, \right\|_{L^2} \ \le\ C \left(\left\| \, L^\epsilon \, u \, \right\|_{L^2} + \left\| \, u \, \right\|_{H^1_{\mathcal{Z}}} \right) .$$

This together with the tangential derivatives shows that $\nabla(I-\pi)u \in L^2$ and the proof is complete. □

Combining Lemmas 4.8 and 4.6 proves the following corollary.

Corollary 4.9 *If $s > (1+d)/2$ then there is a constant C so that for all u as in the previous lemma*

$$\big\| \,(I - \pi(y))\, u\, \big\|_{L^\infty(\Omega_{a,T})} \ \leq\ C\,\Big(\, \|u\|_{H^s_{\mathcal{Z}}} + \|L^\epsilon u\|_{H^{s-1}_{\mathcal{Z}}}\Big)\,. \tag{62}$$

The more interesting sup norm estimates are those for $\pi(y)\, u$, which are proved by the method of characteristics.

Lemma 4.10 *There is a neighborhood $\mathbb{O}$ of $\overline{\Sigma}_{a,T}$ in $\overline{\Omega}_{a,T}$ and a constant C so that for all $u \in C^1(\overline{\Omega}_{a,T})$ with $u\big|_{t=-a} = 0$,*

$$\big\|\, \pi(y)\, u\, \big\|_{L^\infty(\mathbb{O})} \ \leq \tag{63}$$
$$C\Big(\, \big\|L^\epsilon u\big\|_{L^\infty(\Omega_{a,T})} + \big\|\big(I - \pi(y)\big)u\big\|_{L^\infty(\Omega_{a,T})} + \big\|\mathcal{Z}\,\big(I - \pi(y)\big)u\big\|_{L^\infty(\Omega_{a,T})}\Big)\,.$$

Proof. Begin with the identity

$$\pi\, L\, \pi\, u \ =\ \pi\, Lu - \pi\, L\,(I - \pi)u\,.$$

The $\pi(y)$ sandwich identity (29) implies that

$$\pi\, L\, \pi \ \subset\ \big(\partial_t + \mathbf{v}.\partial_x\big)\pi + \mathcal{M}\pi(y)\,.$$

Lemma 4.4.**i**. implies that $\pi\, L\,(I - \pi) \subset \mathcal{M}\mathcal{Z} + \mathcal{M}$. Multiplying on the right by $(I - \pi)$ yields

$$\pi\, L\,(I - \pi) \ \subset\ \mathcal{M}\mathcal{Z}\,(I - \pi) + \mathcal{M}(I - \pi)\,.$$

The last three displayed identities show that

$$\big(\partial_t + \mathbf{v}.\partial_x\big)\pi\, u + \mathcal{M}\pi\, u \ \subset\ \pi\, Lu + \mathcal{M}\mathcal{Z}\,(I - \pi)u + \mathcal{M}(I - \pi)u\,. \tag{64}$$

Therefore

$$\Big\|\big(\partial_t + \mathbf{v}.\partial_x\big)\pi\, u + \mathcal{M}\pi\, u\Big\|_{L^\infty} \ \leq$$
$$C\,\Big(\, \big\|L^\epsilon u\big\|_{L^\infty} + \big\|\big(I - \pi(y)\big)u\big\|_{L^\infty(\Omega_{a,T})} + \big\|\mathcal{Z}\,\big(I - \pi(y)\big)u\big\|_{L^\infty(\Omega_{a,T})}\Big)\,.$$

The fact that $\Omega_{a,T}$ is a domain of determinacy in the sense that (5) is satisfied on the lateral boundary implies that the backward integral curves of the vector field $\partial_t + \mathbf{v}.\nabla_x$ beginning at a point of $\Omega_{a,T}$ reach $\{t = -a\}$ before leaving $\Omega_{a,T}$. Thus integrating the preceding inequality along such integral curves proves the lemma. ☐

Estimates (62) and (63) are the key sup norm estimates but they are just short of sufficient to estimate the sup norm of u since the second estimate requires a sup norm estimate for the tangential derivatives of $(I - \pi)u$. Now, $Z\,(I - \pi)u = \big(Z(I - \pi)\big)\, u + (I - \pi)Zu$. The idea to estimate the second term would be to differentiate to find $L^\epsilon\, Zu = Z\, L^\epsilon u + [L^\epsilon, Z]\, u$. For this strategy to work requires good commutation between L^ϵ and Z and that in turn may require a change of dependent variable. Fortunately the machinery has already been set up in the derivation of the L^2 estimates.

Proof of Theorem 4.1. The main step is to prove the estimate (44) for smooth u. The $H^s_{\mathcal{Z}}$ part of the conclusion follows from (45). What is new are the sup norm estimates. The strategy is to use the even bigger system trick which comes from differentiating (58) tangent to Σ while taking advantage of the commutation relation (60).

Introduce the very long vector

$$\underline{\mathbf{u}} \; := \; \Big\{ (Z_1, \ldots, Z_K)^\alpha \mathbf{u} \Big\}_{|\alpha| \le k} .$$

Applying tangential derivatives to (58) and commuting yields a very big system

$$\big(\mathbb{L} + \mathbb{B}^\epsilon(y)\big)\, \underline{\mathbf{u}} \; = \; \underline{\mathbf{f}} \in L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}} \,,$$

with

$$\mathbb{L} := \operatorname{diag}\{\mathbf{L}\}\,, \qquad \text{and} \qquad \forall \beta, \;\; \sup_\epsilon \big\| (Z_1, \ldots, Z_K)^\beta \mathbb{B}^\epsilon(y) \big\|_{L^\infty(\Omega_{a,T})} < \infty \,.$$

Note that in these special coordinates $Z\pi(y) = 0$ so Z_j commutes with $(I - \pi)$. Then, the proofs of Corollary 4.9 and Lemma 4.10 yield the estimate

$$\begin{aligned} \big\| \underline{\mathbf{u}} \big\|_{L^\infty(\Omega_{a,T})} + \sum_{j=1}^K \big\| Z_j (I - \pi) \underline{\mathbf{u}} \big\|_{L^\infty(\Omega_{a,T})} & \\ \le \; C\Big(\|(\mathbb{L} + \mathbb{B}^\epsilon) \underline{\mathbf{u}}\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}} + \|\underline{\mathbf{u}}\|_{H^s_{\mathcal{Z}}} \Big) . & \end{aligned}$$

This estimate is equivalent to the L^∞ part of (44) which completes the proof of the latter estimate.

Having proven the *a priori* estimate, choose $f_n \in C^\infty(\overline{\Omega}_{a,T})$ so that for all $|\alpha| \le k$ one has the weak star convergence

$$(Z_1, \ldots, Z_K)^\alpha f_n \rightharpoonup (Z_1, \ldots, Z_K)^\alpha f \qquad \text{in} \quad L^\infty(\Omega_{a,T}) \cap \, H^s_{\mathcal{Z}} \,,$$

and

$$\|(Z_1, \ldots, Z_K)^\alpha f_n\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}} \; \le \; 2\, \|(Z_1, \ldots, Z_K)^\alpha f\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}} \,.$$

The resulting solutions u_n^ϵ are smooth on $\overline{\Omega}_{a,T}$ so the *a priori* estimate holds for u_n^ϵ. Thus

$$\begin{aligned} \sum_{|\beta| \le k+1} \|(Z_1, \ldots, Z_M)^\beta (I - \pi) u_n^\epsilon\|_{L^\infty(\Omega_{a,T})} + & \\ \sum_{|\alpha| \le k} \|(Z_1, \ldots, Z_M)^\alpha u_n^\epsilon\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}} & \\ \le C \sum_{|\alpha| \le k} \|(Z_1, \ldots, Z_M)^\alpha f_n\|_{L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}} . & \end{aligned}$$

Passing to the limit $n \to \infty$ proves the Theorem. $\square$

5 Perturbation proof of the Main Theorem 1.2.

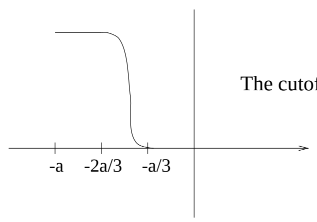
Proof. Recall that there are two approximate solutions in play, u^ϵ from Proposition 2.1 which is a corrected version of $u^\epsilon_{\rm approx}$ from Theorem 1.2. They differ by $O(\epsilon)$ in the sense that

$$\forall\ \beta, \quad \big\|\, (Z_1, \dots, Z_K)^\beta \big(u^\epsilon_{\rm approx} - u^\epsilon\big)\,\big\|_{L^\infty(\Omega_{a,T})} \ =\ O(\epsilon)\,.$$

Thus to prove the theorem is suffices to find exact solutions v^ϵ which differ from u^ϵ by $O(\epsilon)$.

The exact solutions v^ϵ are constructed as follows. Choose a cutoff function $\chi(t) \in C^\infty([-a, \infty[)$ as sketched in the figure, so that

$$\chi(t) \ =\ 1 \quad \text{if } t \le -2a/3 \qquad \chi(t) \ =\ 0 \quad \text{if } t \ge -a/3\,.$$



The cutoff function χ.

Let

$$r^\epsilon(y) \ :=\ L(y, \partial)\, u^\epsilon + F(y, u^\epsilon) \tag{65}$$

so that $r^\epsilon = O(\epsilon)$ in the sense of (38). The exact solution v^ϵ is chosen to satisfy the nonlinear initial value problem

$$L(y, \partial)\, v^\epsilon + F(y, v^\epsilon) = \chi(t)\, r^\epsilon\,, \qquad v^\epsilon = u^\epsilon \ \text{ for } \ t \le -2a/3\,. \tag{66}$$

This problem automatically has a local solution on $-a < t < T(\epsilon)$ with $T(\epsilon) > -2a/3$. To prove the theorem we must show that v^ϵ exists throughout $\Omega_{a,T}$ and that $u^\epsilon - v^\epsilon = O(\epsilon)$. With that in mind define w^ϵ by

$$v^\epsilon := u^\epsilon + \epsilon\, w^\epsilon\,.$$

Subtracting (65) from (66) shows that the initial value problem for v^ϵ is equivalent to the following initial value problem for w^ϵ,

$$L\, w^\epsilon + \frac{F(y, u^\epsilon + \epsilon w^\epsilon) - F(y, u^\epsilon)}{\epsilon} = \frac{(1 - \chi(t)) r^\epsilon}{\epsilon}\,, \qquad w^\epsilon = 0 \ \text{ for } \ t \le -2a/3\,. \tag{67}$$

To analyse this equation start by using Taylor's Theorem to show that

$$F(y, u + \epsilon w) = F(y, u) + \epsilon F_u(y, u)\, w + \epsilon^2\, J(\epsilon, u, w)\,,$$

where J is a smooth function of its arguments with $J(\epsilon, u, 0) = 0$. Define

$$B^\epsilon(y) := F_u(y, u^\epsilon(y))\,, \qquad g^\epsilon(y) := \frac{(1 - \chi(t)) r^\epsilon}{\epsilon}\,, \qquad L^\epsilon := L + B^\epsilon\,.$$

Then for all α,

$$\sup_{0<\epsilon<\epsilon_0} \| \big(Z_1, \dots, Z_K\big)^\alpha \{B^\epsilon, g^\epsilon\} \|_{L^\infty(\Omega_{a,T})} < \infty .$$

Injecting the definitions into (67) yields the equivalent problem

$$L^\epsilon w^\epsilon = g^\epsilon - \epsilon J(\epsilon, u^\epsilon, w^\epsilon), \qquad w^\epsilon = 0 \ \text{ for } \ t \leq -2a/3 . \tag{68}$$

Define $G^\epsilon(y)$ to be the solution of the linear problem

$$L^\epsilon \, G^\epsilon = g^\epsilon , \qquad G^\epsilon = 0 \ \text{ for } \ t \leq -2a/3 . \tag{69}$$

The estimates for g^ϵ together with Theorem 4.1 imply that

$$\sup_{0<\epsilon<\epsilon_0} \| \big(Z_1, \dots, Z_K\big)^\alpha G^\epsilon \|_{L^\infty(\Omega_{a,T})} < \infty . \tag{70}$$

With the goal of showing that w^ϵ is a small perturbation of G^ϵ, define $z^\epsilon(y)$ and $\tilde{J}$ by

$$w^\epsilon := G^\epsilon + z^\epsilon , \qquad \tilde{J}(\epsilon, u, G, z) = -J(\epsilon, u, G + z) .$$

Then, the initial value problem for w^ϵ is equivalent to

$$L^\epsilon \, z^\epsilon = \epsilon \tilde{J}(\epsilon, u^\epsilon, G^\epsilon, z^\epsilon), \qquad z^\epsilon = 0 \ \text{ for } \ t \leq -2a/3 . \tag{71}$$

This nonlinear problem for z^ϵ can be solved by a contraction mapping argument. Fix $(3+d)/2 < s \in \mathbb{N}$ and let

$$\mathbf{S} := \Big\{ z \in L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}} : z\big|_{\{t<-2a/3\}} = 0, \text{ and } \|z\|_{L^\infty(\Omega_{a,T})} + \|z\|_{H^s_{\mathcal{Z}}} \leq 1 \Big\} .$$

The bounds for u^ϵ, G^ϵ and Gagliardo-Nirenberg estimates in the conormal category (see for example [MR]) show that the family of maps

$$z \quad \mapsto \quad \tilde{J}(\epsilon, u^\epsilon, G^\epsilon, z)$$

is uniformly Lipshitzean from $\mathbf{S} \subset L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}$ to $L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}$. At the same time the estimate from Theorem 4.9 shows that the family of linear maps $(L^\epsilon)^{-1}$ are uniformly bounded from $\mathbf{S} \subset L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}$ to $L^\infty(\Omega_{a,T}) \cap H^s_{\mathcal{Z}}$. It follows that there is a $\epsilon_1 \in]0, \epsilon_0[$ so that the family of maps

$$z \quad \mapsto \quad \epsilon \, (L^\epsilon)^{-1} \, \tilde{J}(\epsilon, , u^\epsilon, G^\epsilon, z) , \qquad 0 < \epsilon \leq \epsilon_1$$

is uniformly contractive from $\mathbf{S}$ to itself.

Thus, for these values of ϵ the equation (71) has a solution $z^\epsilon \in \mathbf{S}$. This proves that the exact solution $v^\epsilon = u^\epsilon + \epsilon G^\epsilon + \epsilon z^\epsilon$ exists througout $\Omega_{a,T}$ and that $\|v^\epsilon - u^\epsilon\|_{L^\infty \cap H^s_{\mathcal{Z}}} = O(\epsilon)$.

To complete the proof it remains to show that the tangential derivatives of $v^\epsilon - u^\epsilon$ are $O(\epsilon)$ in sup norm. For this, it suffices to show that the tangential derivative of w^ϵ are $O(1)$ in sup norm. To do that use (44) applied to (68) to find

$$\sum_{|\alpha|\le k} \|(Z_1,\dots,Z_K)^\alpha w^\epsilon\|_{L^\infty\cap H^s_{\mathcal{Z}}} \le$$
$$C_1\Big(1+\sum_{|\alpha|\le k} \epsilon\,\|(Z_1,\dots,Z_K)^\alpha \tilde J(\epsilon,u^\epsilon,w^\epsilon)\|_{L^\infty\cap H^s_{\mathcal{Z}}}\Big)\,.$$

Given our bounds for u^ϵ and the fact that $\tilde J(\epsilon,u,0)=0$, the Gagliargo-Nirenberg estimates show that there is a $C_2 = C_2(k,s,u^\epsilon,\tilde J)$ so that

$$\sum_{|\alpha|\le k} \|(Z_1,\dots,Z_K)^\alpha \tilde J(\epsilon,u^\epsilon,w^\epsilon)\|_{L^\infty\cap H^s_{\mathcal{Z}}} \le C_2 \sum_{|\alpha|\le k} \|(Z_1,\dots,Z_K)^\alpha w^\epsilon\|_{L^\infty\cap H^s_{\mathcal{Z}}}\,.$$

Thus when $\epsilon C_1 C_2 < 1/2$ one has

$$\sum_{|\alpha|\le k} \|(Z_1,\dots,Z_K)^\alpha w^\epsilon\|_{L^\infty(\Omega_{a,T})\cap H^s_{\mathcal{Z}}} \le 2C_1\,.$$

Since $0\le k\in\mathbb{N}$ is arbitrary this shows that for all α, $\|(Z_1,\dots,Z_K)^\alpha w^\epsilon\|_{L^\infty} = O(1)$ which completes the proof. $\Box$

6 References.

[A] D. Alterman, Diffractive nonlinear geometric optics for short pulses, Ph.D. Thesis, University of Michigan, May 1999.

[Be] M. Beals, *Propagation and Interaction of Singularities in Nonlinear Hyperbolic Problems*, Birkhäuser, Boston, 1989.

[Bo1] J.-M. Bony, Interaction des singularités pour les équations aux dérivées partielles non-linéaires, Sem. Gaoulaouic-Meyer-Schwartz, Ecole-Polytéchnique, Paris, 1979-1980, exposé 22.

[Bo2] J.-M. Bony, Interaction de singularités pour les équations aux dérivées partielles non-linéaires, 1981-82, exposé no. 2.

[DJMR] P. Donnat, J.-L. Joy, G. Métivier, and J. Rauch, Diffractive nonlinear geometric optics, Séminaire Equations aux Dérivées Partielles, Ecole Polytechnique, Paris, 1995-1996.

[Ma] A. Majda, Nonlinear geometric optics for hyperbolic systems of conservation laws, *Oscillation theory, computation, methods of compensated compactness*, IMA Vol. Math. Appl. 2, Springer, New York, 1986, pp. 115-165.

[MaR] A. Majda and R. Rosales, Resonantly interacting weakly nonlinear hyperbolic waves I, Stud. appl. Math., 71(1984), 149-179.

[MeR] R. Melrose and N. Ritter, The interaction of nonlinear progressing waves, Annals of Math (121) 1985, 187-213.

[Me] G., Métivier, The Cauchy problem for semilinear hyperbolic systems, Duke Math. J. 53(1986), 983-1011.

[R] J. E. Rothenberg, Space-time focusing: breakdown of the slowly varying envelope approximation in the self-focusing of fentosecond pulses, Optics Letters, 17(1992), 1340-1342.
[RR] J. Rauch and M. Reed, Bounded stratified. and striated solutions of hyperbolic systems, in *Nonlinear Parital Differential Equations and Their Applications Vol. IX,* H. Brezis and J.L. Lions eds, Pitman Research Notes in Math. 181, 1989.
[Y1] A. Yoshikawa, Solutions containing a large parameter of a quasi-linear hyperbolic system of equations and their nonlinear geometric optics approximation, Trans. A.M.S., 340(1993), 103-126.
[Y2] A. Yoshikawa, Asymptotic expansions of the solutions to a class of quasilinear hyperbolic initial value problems, J. Math. Soc. Japan, (47)1995, 227-252.