DIFFRACTIVE NONLINEAR GEOMETRIC OPTICS WITH RECTIFICATION *

Jean-Luc JOLY
MAB, Université de Bordeaux I
33405 Talence, FRANCE

Guy METIVIER
IRMAR, Université de Rennes I
35042 Rennes, FRANCE

Jeffrey RAUCH
Department of Mathematics, University of Michigan
Ann Arbor 48109 MI, USA

Outline.

§1. Introduction.

This paper continues the study, initiated in [DJMR], [D], of the behavior of high frequency solutions of nonlinear hyperbolic equations for time scales at which diffractive effects and nonlinear effects are both present in the leading term of approximate solutions. By diffractive effects we mean that the leading term in the asymptotic expansion has support which extends beyond the region reached by the rays of geometric optics. Our expansions are for problems where the time scale for nonlinear interaction is comparable to the time scale for the onset of diffractive effects. The key innovation is the analysis of rectification effects, that is the interaction of the nonoscillatory local mean field with the rapidly oscillating fields. On the long time scales associated with diffraction, these nonoscillatory fields tend to behave very differently from the oscillating fields. One of our main conclusions is that for oscillatory fields associated with wave vectors on curved parts of the characteristic variety, the interaction is negligible to leading order. For wave vector on flat parts of the variety, and in particular problems for problems in one dimensional space, the interaction cannot be ignored and is spelled out in detail. In all cases the leading term in an approximation

* Research partially supported by the U.S. National Science Foundation, U.S. Office of Naval Research, and the NSF-CNRS cooperation program under grants number NSF-DMS-9203413 and OD-G-N0014-92-J-1245 NSF-INT-9314095 respectively, and the CNRS through the Groupe de Recherche G1180 POAN.

is rigorously justified, the correctors are $o(1)$ as the wavelength ε tends to zero. The error is not $O(\varepsilon^\alpha)$ with $\alpha > 0$.

The point of departure are the papers [DJMR], [D] where infinitely accurate asymptotic expansions of the form

$$u^\varepsilon \;\sim\; \varepsilon^p \sum_{j\in\mathbb{N}p} \varepsilon^j\, a_j(\varepsilon x, x, x.\beta/\varepsilon)\,, \tag{1.1}$$

are justified. Here $x = (t, y) \in \mathbb{R}^{1+d}$, $\beta = (\tau, \eta) \in \mathbb{R}^{1+d}$ and the profiles $a_j(X, x, \theta)$ are periodic in θ. The exponent p is a critical exponent for which the the time scales for diffractive and nonlinear effects are equal. Note the three scales, a wavelength of order ε, variations on the scale ~ 1 coming from the dependence of the profiles on x and variations on the long and slow scales $\sim 1/\varepsilon$ from the dependence on εx. This three scale structure is typical of diffractive geometric optics. In [DJMR] the following key hypotheses guaranteed that nonlinear interaction did not create nonoscillatory terms,

- the nonlinear terms are odd functions of u, and
- the profiles $a_j(X, x, \theta)$ in the approximations (1.1) had spectrum contained in $\mathbb{Z}_{\rm odd}$.

The first hypothesis excludes quadratic nonlinearities and the second implies that the profiles have vanishing mean value with respect to θ. The main achievement of this paper is the analysis of solutions when these hypotheses are not made. Note that quadratic interaction of oscillatory terms tends to create nonoscillatory sources as the simple identities $e^{ik\theta}\, e^{-ik\theta} = 1$ and $\sin^2 k\theta + \cos^2 k\theta = 1$ illustrate. The creation of nonoscillatory waves from highly oscillatory sources is called *rectification*.

The spatial Fourier Transform of an expression as on the right of (1.1) is roughly localized at wave numbers $\zeta \in \mathbb{R}^d$ with $|\zeta - \eta/\varepsilon| = O(1)$. In the linear case solutions would be superpositions of plane waves moving with velocities with angular spread $O(\varepsilon)$. For times small compared with $1/\varepsilon$ this angular dispersion is not important. For times $O(1/\varepsilon)$ the angular dispersion becomes important since at those times the accumulated effects are $O(1)$. This scale $t \sim 1/\varepsilon$ is the time scale of diffractive geometric optics.

1. Example of creation of nonoscillating parts in the principal profile. Consider the semilinear initial value problem for for $u = (u_1, u_2)$

$$\frac{\partial u}{\partial t} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{\partial u}{\partial y} + \begin{pmatrix} u_1^2 \\ 0 \end{pmatrix} = 0\,, \qquad u^\varepsilon(0, y) = \varepsilon\,(f(y), g(y))\,\sin y/\varepsilon\,,$$

with $f, g \in C_0^\infty(\mathbb{R}; \mathbb{R})$. The equations for the u_j are decoupled. The linear part is a first order system which corresponds to the D'Alembert operator $\Box_{1+1}$ on $\mathbb{R}^{1+1}$. The characteristic variety is the union of the two lines $\tau = \pm\eta$. The exact solutions with $g = 0$ satisfies $u_2^\varepsilon = 0$ and

$$\begin{aligned} u_1^\varepsilon(t, y) =& \frac{\varepsilon\, f(t-y)\, \sin{(t-y)/\varepsilon}}{1 + \varepsilon t\, f(t-y)\, \sin{(t-y)/\varepsilon}} \\ =&\varepsilon\,\Big(\, f(t-y)\, \sin{(t-y)/\varepsilon} + \varepsilon t\, \big(f(t-y)\, \sin{(t-y)/\varepsilon}\big)^2\,\Big) + O(\varepsilon^3 t)\,. \end{aligned}$$

They have $p = 1$ and profile

$$a(T,t,y,\theta) = \Big(f(t-y)\,\sin\theta + T\,f(t-y)^2\,\sin^2\theta\,,\,0\Big)\,.$$

For $t \sim 1/\varepsilon$ the solution has nonoscillatory terms $\varepsilon^p \varepsilon t \pi f(t-y)^2$ of the same order as the oscillating part. The factor π is the mean value of $\sin^2\theta$. ∎

2. Example of the principal of nonoverlapping waves. The preceding example should be contrasted to the behavior of the system where the nonlinear term $(u_1^2, 0)$ is replaced by the nonlinear term $(u_2^2, 0)$. To see any effect one must take $g \neq 0$. The equation for u_1 then reads

$$\partial_t u_1 + \partial_y u_1 + \varepsilon^2\, g(t+y)^2 = 0\,.$$

The source term $\varepsilon^2 |g(t+y)|^2$ travels at speed minus one so crosses the integral curves of $\partial_t + \partial_y$ transversally. The u_1 wave and the u_2 wave overlap for only a finite period of time so the influence of the source term is $O(\varepsilon^2)$ therefore negligible compared to the leading $O(\varepsilon)$ term even for times $t \sim 1/\varepsilon$. If g were not integrable in y this argument would not be correct. ∎

In this example and for the remainder of the introduction the principal profiles will depend only on (T, t, y, θ), there is no dependence on Y.

3. Examples where rectification effects must be analysed. 1. The compressible Euler equations in fluid dynamics and all more complicated models containing variants of the Euler equations, have quadratic terms. **2.** From nonlinear optics the important phenomenon of second harmonic generation is a quadratic phenomenon. **3.** Even for odd nonlinearities initial profiles which contain even harmonics are not included in the previous analysis. ∎

When rectification effects are present, correctors to the leading approximation are required which do not have the structure of modulated high frequency wave trains. The correctors can look like typical nonoscillatory solutions of nonlinear or linear wave equations. In particular they have no reason to follow the rays of the oscillatory parts. They sometimes spread in all directions so are correspondingly smaller than the leading term in the approximation.

4. Example of the principal of spreading waves. Consider the two dimensional analogue of the equations in Examples 1 and 2,

$$L\,u + \Phi(u) := \frac{\partial u}{\partial t} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{\partial u}{\partial y_1} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{\partial u}{\partial y_2} + \Phi(u) = 0\,,$$

where Φ is quadratic and u^ε has oscillations of wavelength $\sim \varepsilon$. For this problem the critical exponent $p = 1$ so critical initial value problem is

$$u^\varepsilon(0,y) = \varepsilon\,\big(b(y, y_1/\varepsilon), 0\big)\,, \qquad b \in C_0^\infty(\mathbb{R}^2 \times \mathbb{T})\,.$$

Theorem 1.1 shows that the solution is given for $t \le C/\varepsilon$ by

$$u^\varepsilon(t,y) = \varepsilon\,\underline{a}(t,x) + \varepsilon\,\big(\,\mathrm{a}(T, t - y_1, y_2, \ldots, \theta)\,,\,0\big) + o(\varepsilon)\,.$$

where the oscillatory profile $\mathrm{a}(T,t,y,\theta)$ has mean zero with respect to θ and is determined by the initial value problem

$$\partial_T \mathrm{a} + \frac{1}{2}\,\Delta_{y_2,\ldots,y_d}\,\partial_\theta^{-1}\,\mathrm{a} + \Phi\big(\,(\mathrm{a},0)\,\big)^* = 0\,, \qquad a(0,y,\theta) = b(y,\theta)^*\,,$$

where * denontes oscillatory part, that is the function less its mean with respect to θ. The nonoscillatory part of the profile, $\underline{a}$, is determined by the linear hyperbolic dynamics

$$\frac{\partial \underline{a}}{\partial t} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{\partial \underline{a}}{\partial y_1} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{\partial \underline{a}}{\partial y_2} = 0\,, \qquad \underline{a}(0,x) = \frac{1}{2\pi}\int_0^{2\pi} b(0,y,\theta)\,d\theta\,.$$

The function $\underline{a}(t,x)$ decays to zero in $L^\infty(\mathbb{R}^2)$ as $t \to \infty$ but has $L^2(\mathbb{R}^2)$ norm independent of time. The conservation is maintained while the waves spread.. The L^∞ decay is responsible for the fact that the mean field does not influence the leading term of the oscillatory part. The oscillatory and nonosillatory part do not interact to leading order. If $\underline{a}|_{t=0} = 0$ as in Examples 1 and 2, it remains zero, in contrast to Example 2. However, even though the principal term in the expansion has mean zero in θ, to prove the accuracy of the leading term requires correctors which have nonoscillatory parts. □

The approximate solutions which we construct are not infinitely accurate and the proof of accuracy requires a fundamentally different analysis than our earlier papers on diffractive nonlinear geometric optics. Analogous difficulties arose when studying multiphase nonlinear geometric optics for example in [JMR 2]. For the longer time scales of diffractive geometric optics, the difficulty occurs for one phase problems. That one does not have a simple expansion like (1.1) is clear in the next example.

5. Example of a large corrector. Consider the coupled system

$$\Box_{1+2}\mathrm{u} = 0\,, \qquad \Box_{1+2}\mathrm{v} = |\nabla_{t,y}\mathrm{u}|^2\,.$$

With

$$L := \frac{\partial}{\partial t} + \begin{pmatrix} 0 & \partial/\partial y_1 & \partial/\partial y_2 \\ \partial/\partial y_1 & 0 & 0 \\ \partial/\partial y_2 & 0 & 0 \end{pmatrix},$$

this yields the triangular system of equations for $u := \nabla_{t,y}\mathrm{u}$ and $v := \nabla_{t,y}\mathrm{v}$,

$$L\,u = 0\,, \qquad L\,v = |u|^2\,.$$

Take Cauchy data for v^ε to be equal to zero. For $t \sim 1/\varepsilon$ one can take $u^\varepsilon = O(\varepsilon)$ to be a linearly diffractive solution and the computation in §7.5 shows that $v^\varepsilon = O(\varepsilon^{3/2})$ is a corrector but is large compared to the $O(\varepsilon^2)$ correctors in (1.1). □

Three scale expansions like those discussed in this paper arise in a variety of applied problems. Several from nonlinear optics are presented in [DR]. Others with origins in fluid dynamics including a variety of problems in shock diffraction where rigorous error estimates have yet to be proven can be found in [H]. The latter reference also contains an interesting historical discussion. We present examples in §7.

We now describe a typical result from our analysis. One must distinguish between hyperplanes in the characteristic variety and curved sheets since that determines the extent of the interaction between the mean field and the oscillatory parts.

Consider the constant coefficient semilinear equation

$$L(\partial_x)\, u + \Phi(u) = 0\,, \qquad L := \sum_{\mu=0}^{d} A_\mu \frac{\partial}{\partial x_\mu} \tag{1.2}$$

where u is a smooth $\mathbb{C}^N$ or $\mathbb{R}^N$ valued function and Φ is a polynomial of degree J with variables $(\Re u, \Im u)$. Quasilinear problems are treated in the body of the paper.

The system is assumed to be symmetric hyperbolic, that is

$$A_\mu = A_\mu^*, \qquad A_0 = I\,. \tag{1.3}$$

In the introduction we suppose that the characteristic variety consists of a finite number ℓ of smooth nonintersecting sheets $\tau = \tau_j(\eta)$, $j = 1, \dots, \ell$. This is the case for strictly hyperbolic equations, Maxwell's equations and the linearized Euler equations.

For J^{th} order semilinear problems

$$p := \frac{1}{J-1} \tag{1.4}$$

is the critical exponent for which the time of nonlinear interaction is $O(1/\varepsilon)$, which is the time for the onset of diffractive effects.

In this introduction we consider only profiles a_0 which are independent of Y. For $\beta \in Char\, L \subset \mathbb{R}^{1+d} \setminus \{0\}$ we construct approximate solutions

$$u^\varepsilon(x) = \varepsilon^p\, a_0(\varepsilon t, t, y, x.\beta/\varepsilon)\,, \tag{1.5}$$

where $a_0(T, t, y, \theta)$ is smooth, periodic in θ and has derivatives square integrable on $[0, T] \times \mathbb{R}^d_y \times \mathbb{T}$ uniformly for $t \in \mathbb{R}$.

Hyperplanes which are contained in the characteristic variety play a crucial role in describing diffractive nonlinear geometric optics. Decompose the index set

$$\{1, \dots, \ell\} = \mathcal{A}_f \cup \mathcal{A}_c \tag{1.6}$$

where $\mathcal{A}_c$ are the indices of curved sheets and $\mathcal{A}_f$ those of flat sheets. Then $\mathcal{A}_f$ is empty when there are no hyperplanes. Define a switch ι by $\iota = 1$ when β belongs to a hyperplane and $\iota = 0$ otherwise. When $\iota = 1$ relabel the sheets so that β belongs to the first hyperplane $\tau = \tau_1(\eta)$.

For $\xi \in Char\, L$ let $\pi(\xi)$ denote the orthogonal projection of $\mathbb{C}^N$ onto the kernel of $L(\xi)$. Define

$$E_\alpha(\eta) := \pi\big(\tau_\alpha(\eta), \eta)\big)\,, \qquad \alpha \in \mathcal{A}\,. \tag{1.7}$$

The Fourier multiplication operators $E_\alpha(D_y)$ are orthogonal projections on $H^s(\mathbb{R}^d)$ for all s.

One has $\beta = (\tau_{\underline{\alpha}}(\underline{\eta}), \underline{\eta})$ for a unique $1 \leq \underline{\alpha} \leq \ell$ and we associate two differential operators

$$V(\partial_x) := \partial_t + \mathbf{v}.\partial_y\,, \qquad \mathbf{v} := -\nabla_\eta \tau_{\underline{\alpha}}(\underline{\eta})\,, \tag{1.8}$$

and

$$R(\partial_y) := \frac{1}{2} \sum_{j,k=1}^{d} \frac{\partial^2 \tau_{\underline{\alpha}}(\underline{\eta})}{\partial \eta_j \partial \eta_k} \, \frac{\partial^2}{\partial y_j \partial y_k}\,. \tag{1.9}$$

For a periodic function a, the mean value in θ is denoted by either $\underline{a}$ or $\langle a \rangle$. The oscillating part is $a^* := a - \underline{a}$.

With the above definitions, we can write the equations for a_0. The first equation is

$$L(\partial_x)\, \underline{a}_0 = 0\,. \tag{1.10}$$

which allows us to decompose

$$\underline{a}_0 = \sum_{\alpha \in \mathcal{A}} \underline{a}_{0,\alpha}\,, \qquad \underline{a}_{0,\alpha} := E_\alpha(D_y)\, \underline{a}_0\,, \tag{1.11}$$

and the elements of this modal decomposition satisfy

$$\big(\, \partial_t + \tau_\alpha(D_y) \,\big)\, \underline{a}_{0,\alpha} = 0\,. \tag{1.12}$$

Define a switch $\kappa : \mathcal{A} \to \{0, 1\}$ by $\kappa(\alpha) = 1$ if and only if $\alpha \in \mathcal{A}_f$. The dynamics of the mean values is given by the equations

$$E_\alpha(D_y) \left(\partial_T\, \underline{a}_{0,\alpha} + \kappa(\alpha) \left\langle\, \Phi(\iota\, \delta_{1,\alpha}\, a_0^* + \underline{a}_{0,\alpha})\, \right\rangle \right) = 0\,, \tag{1.13}$$

where Kronecker's $\delta_{1,\alpha}$ has value 1 when $\alpha = 1$ and vanishes otherwise.

The oscillating part is polarized and has simple dynamics with respect to t,

$$\pi(\beta)\, a_0^* = a_0^*\,, \qquad V(\partial_x)\, a_0^* = 0\,. \tag{1.14}$$

These equations are equivalent to

$$a_0^*(T, t, y, \theta) := \mathrm{a}(T, y - \mathbf{v}t, \theta)\,, \qquad \mathrm{a} = \pi(\beta)\, \mathrm{a}^*\,. \tag{1.15}$$

The reduced profile, $\mathrm{a}(T, y, \theta)$, has dynamics given by

$$\partial_T\, a_0^* - R(\partial_y)\, \partial_\theta^{-1}\, a_0^* + \Big(\pi(\beta)\, \Phi(a_0^* + \iota \underline{a}_{0,1})\Big)^* = 0\,. \tag{1.16}$$

The * on the last term indicates that the nonoscillating term has been removed so the constraint $\mathrm{a} = \pi\, \mathrm{a}^*$ from (1.15) is satisfied as soon as it is satisfied in $\{T = 0\}$.

This paper is devoted to the derivation of more general versions of equations (1.10)-(1.16), the proof of their solvability, and the accuracy of the approximate solution in the limit $\varepsilon \to 0$.

The nature of equation (1.16) can be understood as follows. The linear part given by the first two terms acts as $\partial_T - iR(\partial_y)/n$ on the Fourier components $\hat{\mathrm{a}}(T, y, n)$. The operator R is symmetric so iR/n is antisymmetric so R generates unitary evolutions in H^s. The semilinear term is roughly of the form a^J.

The next results are a specialization of Theorems 5.1 and 6.1 to the semilinear case and profiles which do not depend on Y.

Theorem 1.1. *If $g(y,\theta) \in \cap_s H^s(\mathbb{R}^d \times \mathbb{T})$ satisfies $g^* = \pi(\beta) g^*$, then there is $T_* \in]0,\infty]$ and unique maximal solutions $(\underline{a}_0, \mathrm{a})$ to the profile equations (1.10-1.16) satisfying*

$$\forall\, \gamma\ \forall\, \underline{T} \in]0, T_*[\,, \quad \sup_{t\in\mathbb{R}} \Big\| \partial^\gamma_{T,y,\theta} \big(\underline{a}_0(T,t,y,\theta)\,,\, \mathrm{a}(T,t,y) \big) \Big\|_{L^2([0,T]\times\mathbb{R}^d_y\times\mathbb{T})} < \infty\,, \tag{1.17}$$

and the initial condition

$$\underline{a}_0(0,0,y) + \mathrm{a}(0,y,\theta) = g(y,\theta)\,.$$

Theorem 1.2. *With profiles $\{\underline{a}_0, \mathrm{a}\}$ from Theorem 1.1, define $a_0 = \underline{a}_0 + a_0^*$ with a_0^* given by (1.15) and a family of approximate solution u^ε by (1.5). Let $v^\varepsilon \in C^\infty([0, t_*(\varepsilon)[\times\mathbb{R}^d)$ be the maximal solutions of the initial value problems*

$$L(\partial_x)\, v^\varepsilon + \Phi(v^\varepsilon) = 0\,, \qquad v^\varepsilon|_{t=0} = u^\varepsilon|_{t=0}\,. \tag{1.18}$$

Then, for any $\underline{T} \in]0, T_*[$, there is an $\varepsilon_0 > 0$ so that if $\varepsilon < \varepsilon_0$, then $t_*(\varepsilon) > \underline{T}/\varepsilon$ and

$$\sup_{0\le t\le \underline{T}/\varepsilon,\ y\in\mathbb{R}^d} \big|\, v^\varepsilon(t) - u^\varepsilon \,\big| = o(\varepsilon^p)\,. \tag{1.19}$$

Remarks. 1. The discussion before the theorem suggests that $a_0 = O(1)$ for $0 \le t \le \underline{T}$ so that in (1.19) neither of the terms in the difference is $o(\varepsilon^p)$.

2. Typically for fixed t, the solutions $u^\varepsilon(t)$ and $v^\varepsilon(t)$ diverge to infinity in H^s when s is large. This prevents the application of elementary continuous dependence arguments. In addition, the residuals $L(\partial)\, u^\varepsilon + \Phi(u^\varepsilon)$, with $u^\varepsilon := \varepsilon^p a_0$, are $O(\varepsilon^{1+p})$ so the accumulated effect of the residual for times of order $1/\varepsilon$ is crudely estimated to be $O(\varepsilon^p)$ and therefore

nonnegligible. For the proof of Theorem 1.2 we construct a corrector of size $O(\varepsilon^{p+1})$ which reduces the residual to size $o(\varepsilon^{p+1})$.

3. The equations determining the profile a_0 are found so that such correctors exist.

4. To derive the equations for a_0 requires detailed information about the large time asymptotic behavior of solutions of the linear equation (1.10), and related inhomogeneous equations. For example, all solutions of a linear symmetric hyperbolic equation with C_0^∞ initial data tend to zero in sup norm if and only if the characteristic variety contains no hyperplane. The proof of such results in §3 is by a nonstationary phase arguments with attention paid to possible singular points in the characteristic variety.

5. Qualitative information can be gleaned from the profile equations. For example, if β does not lie on a hyperplane in the characteristic variety then the oscillatory part of the solution does not influence the nonoscillatory part to leading order. In particular, if $\underline{a}_0 = 0$ at $\{t = T = 0\}$, then it vanishes identically, and the profile equations simplify to the single equation (1.16). These reductions, applications, and extensions are discussed in §7.

Acknowledgement. *Thanks to J.-M. Delort, P. Donnat and J. Hunter for instructive discussions which were crucially important in carrying out this research.*

§2. Formulating the ansatz and the first profile equations.

§2.1. The ansatz and the profile equations.

In this section approximate solutions of the nonlinear symmetric hyperbolic equation

$$L(u, \partial_x)\, u + F(u) = 0\,. \tag{2.1}$$

are constructed. The principal part of (2.1) is the quasilinear operator

$$L(u, \partial_x)\, u := \sum_{\mu=0}^{d} A_\mu(u)\, \partial_\mu u \tag{2.2}$$

Symmetric hyperbolicity assumption. *The coefficients* A_μ *are smooth hermitian symmetric valued functions of* u *on a neighborhood of* $0 \in \mathbb{C}^N$ *, and, for each* u*,* $A_0(u)$ *is positive definite.*

Smoothness means that $A_\mu(u) = \mathcal{A}_\mu(\Re u, \Im u)$ with $\mathcal{A}_\mu \in C^\infty(\mathbb{R}^N \times \mathbb{R}^N\, ;\, Hom\,(\mathbb{C}^N))$.

Order of nonlinearity. *The quasilinear terms are of order* $2 \le K \in \mathbb{N}$ *in the sense that*

$$|\alpha| \le K - 2 \quad \Longrightarrow \quad \partial^\alpha_{\Re u, \Im u} \big(A_\mu(u) - A_\mu(0) \big)\big|_{u=0} = 0\,. \tag{2.3}$$

Then $A_\mu(u) - A_\mu(0) = O(|u|^{K-1})$ *and so the quasilinear term,* $(A_\mu(u) - A_\mu(0))\partial_\mu u$ *is of order* K*.*

The semilinear term F *is smooth on a neighborhood of* $0 \in \mathbb{C}^N$ *, and is of order* $J \ge 2$ *in the sense that*

$$|\gamma| \le J - 1 \quad \Longrightarrow \quad \partial^\gamma_{\Re u, \Im u} F(0) = 0\,. \tag{2.4}$$

Then $F(u) = O(|u|^J)$*.*

Making the change of dependent variable to $A_0(0)^{-1/2}\,u$ and multiplying the resulting equations by $A_0(0)^{-1/2}\,u$ preserves these hypotheses and reduces to the case $A_0(0) = I$.

Convention. $A_0(0) = I$.

If the solutions have amplitude of order ε^p with derivatives of order ε^{p-1}, then the size of the quasilinear terms is $\varepsilon^{p(K-1)}\,\varepsilon^{p-1}$. The accumulated effects after time T are crudely estimated to be of order $T\varepsilon^{pK-1}$. Setting this equal to the order of magnitude of the solutions, ε^p, yields the following estimate for the time of nonlinear interaction

$$T_{\text{quasilinear}} \sim \frac{1}{\varepsilon^{pK-p-1}}\,.$$

Similarly, the semilinear terms are of order ε^{pJ}, with accumulation $T\varepsilon^{pJ}$ and time of interaction estimated by setting this equal to ε^p,

$$T_{\text{semilinear}} \sim \frac{1}{\varepsilon^{pJ-p}}\,.$$

The goal is that the time of nonlinear interaction is of order ε^{-1} which is the time for the onset of diffractive effects in linear problems with linear phases and wavelengths $\sim \varepsilon$ (see [DJMR]). Thus one wants $pK - p - 1 \geq 1$ and $pJ - p \geq 1$ with equality in at least one of the two.

The standard normalization. *Consider waves of wavelength of order ε and amplitude of order ε^p satisfying*

$$p = \max\left\{\frac{2}{K-1}\,,\,\frac{1}{J-1}\right\}. \tag{2.5}$$

Examples. The classic case is quadratic nonlinearities. For quadratic quasilinear effects one has $K = 2$, and $p = 2$. A semilinear term with $J \geq 2$ will not affect the leading behavior.

For quadratic semilinear effects one has $J = 2$, $p = 1$, and $K \geq 3$.

For odd order quasilinear terms, the time of interaction for a semilinear term of order $J = (K+1)/2$ is of the same order as that of the quasilinear term. For even order quasilinear terms, there is never such agreement.

If $K > 3$ and $J > 2$, then p is a fraction with $0 < p < 1$. □

Definition of leading order nonlinear terms. *If $p = 2/(K-1)$, the degree $K-1$ Taylor polynomial of $A_\mu(u) - A_\mu(0)$ at $u = 0$ is denoted $\Lambda_\mu(u)$. If $p < 2/(K-1)$ set $\Lambda_\mu := 0$ for all μ.*

If $p = 1/(J-1)$, let $\Phi(u)$ denote the order J Taylor polynomial of F at $u = 0$. If $p < 1/(J-1)$ set $\Phi := 0$.

With this definition, the contribution of the quasilinear and semilinear terms to the dynamics of the leading term in the approximate solution is given in terms of Λ_μ and Φ. Note that Λ_μ and Φ are homogeneous polynomials of degree $2/p$ and $(p+1)/p$ respectively with the convention that if either degree is not an integer, the corresponding polynomial vanishes identically.

The basic *ansatz* has three scales

$$u^\varepsilon(x) = \varepsilon^p\, a(\varepsilon, \varepsilon x, x, x.\beta/\varepsilon)\,, \tag{2.6}$$

where

$$a(\varepsilon, X, x, \theta) = a_0(X, x, \theta) + \varepsilon\, a_1(X, x, \theta) + \varepsilon^2\, a_2(X, x, \theta)\,. \tag{2.7}$$

This expansion has terms which are chosen so that the leading terms in $L\,u^\varepsilon + F(u^\varepsilon)$ vanish.

Example of nonuniqueness of profiles. Different profiles can lead to the same approximate solution when multiple scale expansions are used. For example if $a(T,t) \in C_0^\infty(\mathbb{R}^2)$ define $g(T,t) \in C_0^\infty(\mathbb{R}^2)$ by

$$a(T,t) - a(0,t) = T g(T,t)\,.$$

Then the profiles $a(T,t)$ and $a(0,t) + \varepsilon t g(T,t)$ define the same function when $T = \varepsilon t$ is injected. This lack of uniqueness means the a_j are not determined so that in deriving equations we are obliged to make choices. We have tried to take a path where these choices are as natural as possible. ∎

Since the expansion (2.7) is used for times $t \sim 1/\varepsilon$ one must control the growth of the profiles in t. In order for the corrector a_1 in (2.7) to be smaller than the principal term for such times, it must satisfy

$$\lim_{\varepsilon\to 0}\ \sup_{[0,T/\varepsilon]\times\mathbb{R}^d\times\mathbb{T}} \big|\varepsilon\, a_1(\varepsilon x, x, \theta)\big| = \lim_{\varepsilon\to 0}\ \sup_{[0,T]\times\mathbb{R}^d\times\mathbb{T}} \big|\varepsilon\, a_1(X, X/\varepsilon, \theta)\big| = 0\,. \tag{2.8}$$

Our profiles satisfy the stronger condition of sublinear growth

$$\lim_{t\to\infty}\ \frac{1}{t}\ \sup_{X,y,\theta} \big|\, a_1(X,t,y,\theta)\,\big| \ =\ 0\,. \tag{2.9}$$

This condition plays a central role in the analysis to follow. The profiles $a_j(T,Y,x,\theta)$ are smooth on $[0,T_*[\times\mathbb{R}^d \times \mathbb{R}^{1+d} \times \mathbb{T}$.

§2.2. Equations for the profiles.

The chain rule implies that

$$L(u^\varepsilon, \partial_x)\, u^\varepsilon + F(u^\varepsilon) = \Big[\, L(\varepsilon^p a^\varepsilon, \varepsilon\partial_X + \partial_x + \frac{\beta}{\varepsilon}\,\partial_\theta\,)\, \varepsilon^p\, a + F(\varepsilon^p\, a)\Big]_{X=\varepsilon x,\ \theta = x.\beta/\varepsilon}\,. \tag{2.10}$$

The strategy is to expand the function of (X,x,θ) in brackets in powers of ε and to choose a so that the leading terms vanish.

Lemma 2.1. *Suppose that a_0, a_1, a_2 are smooth functions on $[0,T] \times \mathbb{R}^d \times \mathbb{R}^{1+d} \times \mathbb{T}$ and that a is given by (2.7). Then*

$$\begin{aligned} &L(\varepsilon^p a^\varepsilon, \varepsilon\partial_X + \partial_x + \frac{\beta}{\varepsilon}\,\partial_\theta\,)\,\varepsilon^p\, a + F(\varepsilon^p\, a) = \\ &\qquad \varepsilon^{p-1}\left\{\, r_0(X,x,\theta) + \varepsilon r_1(X,x,\theta) + \varepsilon^2 r_2(X,x,\theta) + \varepsilon^{2+\min\{1,p\}} h(\varepsilon,X,x,\theta)\,\right\}, \end{aligned} \tag{2.11}$$

where the $r_j(X,x,\theta)$ are given by

$$r_0 = L(0,\beta)\,\partial_\theta a_0\,, \qquad r_1 = L(0,\beta)\,\partial_\theta a_1 + L(0,\partial_x)\,a_0\,, \tag{2.12}$$

$$r_2 = L(0,\beta)\,\partial_\theta a_2 + L(0,\partial_x)\,a_1 + L(0,\partial_X)\,a_0 + \Phi(a_0) + \sum_{\mu=0}^{d} \beta_\mu\,\Lambda_\mu(a_0)\,\partial_\theta a_0\,. \tag{2.13}$$

The function $h(\varepsilon,X,x,\theta)$ depends on the a_j, is periodic in θ and each of its derivatives with respect to X,x,θ is continuous on $[-1,1] \times \mathbb{R}^{1+d} \times \mathbb{R}^{1+d} \times \mathbb{T}$.

Remark. For the a_j and h which we construct precise estimates are given in Theorem 5.1 and Proposition 5.4.

Proof. Compute

$$L(0,\varepsilon\partial_X + \partial_x + \frac{\beta}{\varepsilon}\,\partial_\theta)\,\varepsilon^p\, a = \varepsilon^{p-1}\Big(\,L(0,\beta)\,\partial_\theta a + \varepsilon\, L(0,\partial_x)\,a + \varepsilon^2\, L(0,\partial_X)\,a\,\Big). \tag{2.14}$$

The leading term in the Taylor expansion of $A_\mu(\varepsilon^p\, a) - A_\mu(0)$ is $\Lambda_\mu(\varepsilon^p\, a_0)$ and Λ_μ is homogeneous of degree $2/p$. Thus,

$$A_\mu(\varepsilon^p\, a) - A_\mu(0) = \varepsilon^2\,\Big(\,\Lambda_\mu(a_0(X,x,\theta)) + \varepsilon^{\min\{1,p\}} h_1(\varepsilon,X,x,\theta)\,\Big)\,. \tag{2.15}$$

Similarly, Φ is homogeneous of degree $(p+1)/p$ so,

$$F(\varepsilon^p\, a) = \varepsilon^{p+1}\,\Big(\,\Phi(a_0(X,x,\theta)) + \varepsilon^{\min\{1,p\}} h_2(\varepsilon,X,x,\theta)\,\Big)\,.$$

This together with equations (2.14), and (2.15) proves the Lemma. ∎

The profiles in (2.7) are chosen so that in (2.11) one has $r_j = 0$ for $j = 0,1,2$. Equations (2.12-2.13) show that r_0, r_1, r_2 vanish if and only if

$$L(0,\beta)\,\partial_\theta a_0 = 0\,, \tag{2.16}$$

$$L(0,\beta)\,\partial_\theta a_1 + L(0,\partial_x)\,a_0 = 0\,. \tag{2.17}$$

$$L(0,\beta)\,\partial_\theta a_2 + L(0,\partial_x)\,a_1 + L(0,\partial_X)\,a_0 + \Phi(a_0) + \sum_{\mu=0}^{d} \beta_\mu\,\Lambda_\mu(a_0)\,\partial_\theta a_0 = 0\,. \tag{2.18}$$

The analysis of these equations is as in [DJMR], with the important exception that the mean value or zero mode of the periodic functions must be treated. This is more subtle than meets the eye, and is the essential difficulty of the present paper.

Notation. *Introduce two notations for the mean value with respect to* θ

$$\underline{a} := \langle a \rangle := \frac{1}{2\pi} \int_0^{2\pi} a \, d\theta \, .$$

The oscillating part is denoted $a^* := a - \underline{a}\,.$

2.3. Analysis of equation 2.16.

Equation (2.16) holds if and only if a^* takes values in the kernel of $L(0,\beta)$. This kernel is nontrivial if and only if β is characteristic for the linear differential operator $L(0,\partial)$. The set of such β is denoted $Char\, L(0,\partial) := \{\beta \in \mathbb{R}^{d+1} \setminus 0 \, : \, \det\, L(0,\beta) = 0\}$. There are two naturally defined matrices which play a central role.

Definition. *For* $\beta \in \mathbb{R}^{1+d}$*,* $\pi(\beta)$ *denotes the linear projection on the kernel of* $L(0,\beta)$ *along the range of* $L(0,\beta)$*.* $Q(\beta)$ *is the partial inverse defined by*

$$Q(\beta)\,\pi(\beta) = 0\,, \qquad \text{and} \qquad Q(\beta)\,L(0,\beta) = I - \pi(\beta)\,. \tag{2.19}$$

Symmetric hyperbolicity implies that both $\pi(\beta)$ and $Q(\beta)$ are hermitian symmetric. In particular $\pi(\beta)$ is an orthogonal projector. With this notation, (2.16) is equivalent to

$$\pi(\beta)\, a_0^* = a_0^*\,. \tag{2.20}$$

This asserts that the oscillating part of the principal profile is polarized along the kernel of $L(0,\beta)$. Equation (2.16) imposes no constraints on the nonoscillating part $\underline{a}$.

§2.4. Analysis of equation 2.17.

Equation (2.17) involves both a_0 and a_1. Taking the mean value with respect to θ eliminates a_0 and yields a fundamental evolution equation for the nonoscillating part $\underline{a}_0(X,x)$,

$$L(0,\partial_x)\,\underline{a}_0 = 0\,. \tag{2.21}$$

Multiplying the oscillating part of (2.17) by $\pi(\beta)$ annihilates $L(0,\beta)$ so eliminates the a_1 term to give

$$\pi(\beta)\,L(\partial_x)\,\pi(\beta)\,a_0^* = 0\,. \tag{2.22}$$

A vector $w \in \mathbb{C}^N$ vanishes if and only if $\pi(\beta)w = 0$ and $Q(\beta)w = 0$. Thus the information in (2.17) complementary to (2.21-2.22) is obtained by multiplying the oscillatory part of (2.17) by $Q(\beta)$. This yields

$$(I - \pi(\beta))\,a_1^* = -Q(\beta)\,L_1(\partial_x)\,\partial_\theta^{-1} a_0^*\,. \tag{2.23}$$

Equations (2.20) and (2.22) are the fundamental equations of linear geometric optics (see [R]). They determine the dynamics of $a_0^* = \pi(\beta)a_0^*$ with respect to the time t. The following hypothesis guarantees that the linear geometric optics is simple. It excludes for example β along the optic axis of conical refraction.

Smooth characteristic variety hypothesis. *$\beta = (\underline{\tau}, \underline{\eta}) \in Char\, L(0, \partial)$ and there is a neighborhood ω of $\underline{\eta}$ in $\mathbb{R}^d$ (resp. $\mathcal{O}$ of β in $\mathbb{R}^{1+d}$) so that for each $\eta \in \omega$ there is exactly one point $(\tau(\eta), \eta) \in \mathcal{O} \cap Char\, L(0, \partial)$.*

The next proposition is virtually identical to Proposition 3.1 in [DJMR]. The proof proceeds by differentiating the identity $(\sum A_j \xi_j)\pi(\xi) = \tau(\xi)\pi(\xi)$ with respect to ξ_k. The formula (2.24) is simpler than in [DJMR] thanks to the convention that $A_0(0) = I$. The traditional presentation of this identity uses left and right eigenvectors as in §VI.3.11 of [C]. The eignevectors are not uniquely determined but the projectors $\pi(\beta)$ is.

Proposition 2.2. *If the smooth characteristic variety hypothesis is satisfied then the functions, $\tau(\eta)$, $\pi(\tau(\eta), \eta)$ and $Q(\tau(\eta), \eta)$ are real analytic on ω. If $A_0(0) = I$ then*

$$\pi(\beta)\, L(0, \partial_x)\, \pi(\beta) = \pi(\beta) \left(\frac{\partial}{\partial t} - \sum_{j=1}^{d} \frac{\partial \tau(\underline{\eta})}{\partial \eta_j} \frac{\partial}{\partial y_j} \right). \tag{2.24}$$

Definitions. *If $\underline{\tau}, \underline{\eta}$ belongs to the characteristic variety and satisfies the smoothness assumption, define the transport operator V and group velocity $\mathbf{v}$ by*

$$V(\underline{\tau}, \underline{\eta}\,; \partial_x) := \frac{\partial}{\partial t} - \sum_{j=1}^{d} \frac{\partial \tau(\underline{\eta})}{\partial \eta_j} \frac{\partial}{\partial x_j} := \partial_t + \mathbf{v}.\partial_y\,. \tag{2.25}$$

The $\underline{\tau}, \underline{\eta}$ dependence of π, Q, V, $\mathbf{v}$ will often be omitted when there is little risk of confusion. Equations (2.24) and (2.25) show that (2.22) is equivalent to

$$V(\partial_x)\, a_0^* = 0\,. \tag{2.26}$$

Corollary 2.3 *If the smooth characteristic variety hypothesis is satisfied at $\beta = (\underline{\tau}, \underline{\eta})$ then $Char\, V(\partial_x)$ is the tangent plane to $Char\, L(0, \partial_x)$ at $(\underline{\tau}, \underline{\eta})$.*

Proof. Near $(\underline{\tau}, \underline{\eta})$, the characteristic variety of $L(0, \partial_x)$ has equation $\tau = \tau(\eta)$. Therefore its tangent plane has equation $\tau - \underline{\tau} = \nabla_\eta \tau(\underline{\eta}) \cdot (\eta - \underline{\eta})$.

Formula (2.25) shows that the characteristic variety of $V(\partial_x)$ is the codimension 1 linear subspace with equation $\tau - \nabla\tau(\underline{\eta}) \cdot \eta = 0$. Thus $(\underline{\tau}, \underline{\eta}) + Char\, V$ is the tangent plane to $Char\, L$ at $(\underline{\tau}, \underline{\eta})$. To complete the proof it suffices to show that $(\underline{\tau}, \underline{\eta}) \in Char\, V$.

Since the function $\tau - \tau(\eta)$ is homogeneous of degree one in τ, η, Euler's identity implies that $\tau - \nabla\tau(\eta).\eta = \tau - \tau(\eta)$. At $(\underline{\tau}, \underline{\eta})$, the right hand side vanishes which proves that $(\underline{\tau}, \underline{\eta}) \in Char\, V$. ∎

§2.5. A first look at equation 2.18.

Equation (2.18) involves three profiles because there are three scales in the *ansatz.*

Taking the mean value with respect to θ eliminates the a_2 part and yields an equation for the nonoscillating part of a_1

$$L(0,\partial_x)\,\underline{a}_1 = -L(0,\partial_X)\,\underline{a}_0 - \Big\langle\, \Phi(a_0) + \sum \beta_\mu \Lambda_\mu(a_0)\partial_\theta a_0 \,\Big\rangle\,. \tag{2.27}$$

The last term on the right may be nonzero, even if $\underline{a}_0 = 0$*. This possibility of creating nonoscillating contributions from oscillating terms is called rectification.*

The analysis of equation (2.27) is the main difficulty in the derivation of the profile equations. The idea is that since a_0 must satisfy (2.21) and (2.26) its asymptotics as $t \to \infty$ are quite structured so the form of the source terms on the right of (2.27) can be described. Requiring that $\underline{a}_1$ satisfy the sublinear growth condition (2.9) places constraints on a_0. To understand these we must study the asymptotic behavior of solutions of the linear symmetric hyperbolic system (2.21) and also the asymptotics of the inhomogeneous equation $L(0,\partial_x)u = f$ where the source term f is constructed from solutions of the homogeneous equation as in the right hand side of (2.27). The next section is devoted to the study of these linear problems. In §4 we return to the study of (2.18) and in particular (2.27).

§3. Large time asymptotics for linear symmetric hyperbolic systems.

In this section we suppose that

$$L(\partial_x) = \sum_{\mu=0}^{d} A_\mu\, \frac{\partial}{\partial x_\mu}\,, \qquad A_\mu = A_\mu^*\,, \qquad A_0 = I \tag{3.1}$$

is a general constant coefficient symmetric hyperbolic operator with coefficient of ∂_t equal to the identity matrix. In our applications, L will be taken to be $L(0,\partial_x)$.

Proposition 3.2 decomposes solutions of $Lu = 0$ into modes which rigidly translate and modes which spread out in space. The latter decay in sup norm. The analysis is by stationary and nonstationary phase. Care is needed because the characteristic variety may have singular points. The basic stratification theorem of real algebraic geometry implies that these singularities form a lower dimensional variety and this implies that the contributions of the singular points can be treated as error terms.

Examples. The solutions of the equation

$$\frac{\partial v}{\partial t} + \frac{\partial v}{\partial y_1} = 0 \tag{3.2}$$

in dimension d and the equation

$$\frac{\partial v}{\partial t} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial v}{\partial y} = 0 \tag{3.3}$$

in dimension 1 have only purely translating modes.

In dimension 2, the first order analogue of D'Alembert's wave equation

$$\frac{\partial u}{\partial t} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial u}{\partial y_1} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial u}{\partial y_2} = 0\,. \tag{3.4}$$

has solutions which spread in time and whose L^∞ norm decays like $t^{-1/2}$ while the L^2 norm is conserved. □

The characteristic variety for (3.2) and (3.3) consist of hyperplanes while for (3.4) the variety is curved. This dichotomy is crucial as indicated by the following heuristic argument about the group velocities $-\nabla_\eta \tau(\eta)$. If the variety is flat with equation $\tau = -\mathbf{v}.\eta$ these velocities do not depend on η and wave packets will translate at this fixed velocity. If the variety is not flat, the variation of the speed spreads wavepackets leading to their decay in L^∞.

§3.1. Hyperplanes and singularities in the characteristic variety.

The characteristic variety is defined by

$$Char\, L := \{(\tau, \eta) \in \mathbb{R}^{1+d} \setminus 0 \,:\, \det\, L(\tau, \eta) = 0\,\}\,. \tag{3.5}$$

For each point (τ, η) in the variety and $\mathbf{a} \in \ker L(\tau, \eta)$, $u = e^{i(\tau t + \eta.y)}\mathbf{a}$ is a plane wave solution of $Lu = 0$.

Since $(\tau, 0)$ is noncharacteristic for L, any hyperplane $\{a\tau + b.\eta = 0\}$ contained in the characteristic variety must have $a \neq 0$ so the hyperplane is necessarily a graph $\{\tau = -\mathbf{v}.\eta\}$. Since over each $\eta \in \mathbb{R}^d$ there are at most N points in the characteristic variety, the variety contains at most N distinct hyperplanes $H_1, \ldots, H_M$,

$$H_j = \{\,(\tau, \eta) \,:\, \tau = -\mathbf{v}_j.\eta\,\}\,, \qquad j = 1, \ldots, M \leq N\,. \tag{3.6}$$

Examples. 1. When $d = 1$ the characteristic variety is a union of lines so consists only of hyperplanes. There are no curved sheets.

2. The operator from (3.4) has characteristic variety is given by $\tau^2 = |\eta|^2$ so the variety is the classical light cone, and there are no hyperplanes.

3. The characteristic varieties of Mawell's Equations and the the linearization at $u = 0$ of the compressible Euler equations are the union of a curved light cone and a single horizontal hyperplane $\tau = 0$. □

Definition. *A wave number $\omega \in \mathbb{R}^d \setminus \{0\}$ is* **good** *when there is a neighborhood Ω of ω and a finite number of smooth real valued functions $\lambda_1(\eta) < \lambda_2(\eta) < \cdots < \lambda_m(\eta)$ so that the spectrum of $\sum_{j=1}^d A_j \eta_j$ is $\{\lambda_1(\eta), \ldots, \lambda_m(\eta)\}$ for $\eta \in \Omega$. The complementary set consists of* **bad** *wave numbers. The set of bad wave numbers is denoted $\Gamma(L)$.*

Over a good η, the characteristic variety of L is m smooth nonintersecting sheets $\tau = -\lambda_j(\eta)$. The bad points are points where eigenvalues cross and therefore multiplicities change. The examples above have no bad points.

Examples. Consider the characteristic equation $(\tau^2 - |\eta|^2)(\tau - c\eta_1) = 0$ with $c \in \mathbb{R}$. If $|c| < 1$ then the variety is a cone and a hyperplane intersecting only at the origin and all points are good. If $|c| > 1$ the plane and cone intersect in a hyperbola whose projection on η space is the set of bad points

$$\Gamma = \left\{ \eta \,:\, (c^2 - 1)\eta_1^2 = \eta_2^2 + \ldots + \eta_d^2 \right\}.$$

When $|c| = 1$ the hyperbola degenerates to a line of tangency. □

Proposition 3.1. i. *$\Gamma(L)$ is a closed conic set of measure zero in $\mathbb{R}^d \setminus \{0\}$.*
ii. *The complementary set, $\mathbb{R}^d \setminus (\Gamma \cup \{0\})$, is the disjoint union of a finite family of conic connected open sets $\Omega_g \subset \mathbb{R}^d \setminus \{0\}$, $g \in \mathcal{G}$.*
iii. *The multiplicity of $\tau = -\mathbf{v}_j.\eta$ as a root of $\det L(\tau, \eta) = 0$ is independent of $\eta \in \mathbb{R}^d \setminus (\Gamma \cup \{0\})$.*
iv. *If $\lambda(\eta)$ is an eigenvalue of $\sum A_j \eta_j$ depending smoothly on $\eta \in \Omega_g$, then either there is $j \in \{1, \ldots, M\}$ such that $\lambda(\eta) = -\mathbf{v}_j \cdot \eta$ or $\nabla^2 \lambda \neq 0$ almost everywhere on Ω_g.*

Proof. i. Use the basic stratification theorem of real algebraic geometry (see [BR], [CR]). The characteristic variety is a conic real algebraic variety in $\mathbb{R}^{1+d}$. Since over each η it contains at least 1 and at most N points the dimension is d. The singular points are therefore a stratum of dimension at most $d - 1$. The bad frequencies are exactly the projection of this singular locus and so is a subvariety of $\mathbb{R}^d$ of dimension at most $d - 1$ and (i) follows.

ii. That there are at most a finite number of components in the complementary set is a classical theorem of Whitney (see [BR], [CR]).

iii. Denote by m the multiplicity on Ω_g and m' the multiplicity on $\Omega_{g'}$. By definition of multiplicity,

$$\eta \in \Omega_g \quad \Longrightarrow \quad \frac{\partial^{m-1} \det L(\tau, \eta)}{\partial \tau^{m-1}} \Bigg|_{\tau = -\mathbf{v}_j.\eta} = 0\,. \tag{3.7}$$

Then $\partial_\tau^{m-1} L(-\mathbf{v}_j.\eta, \eta)$ is a polynomial in η which vanishes on the nonempty open set Ω_g, so must vanish identically. Thus it vanishes on $\Omega_{g'}$ and it follows that $m' \geq m$. By symmetry one has $m \geq m'$.

iv. If λ is a linear function $\lambda = -\mathbf{v}.\eta$ on Ω_g, then $\det L(-\mathbf{v}.\eta, \eta) = 0$ for $\eta \in \Omega_g$ so by analytic continuation, must vanish for all η. It follows that the hyperplane $\tau = -\mathbf{v}.\eta$ lies in the characteristic variety and therefore that $\lambda = -\mathbf{v}_j.\eta$ for some j.

If λ is not a linear function, then the matrix $\nabla_\eta^2 \lambda$ is a real analytic function on Ω_g which is not identically zero and therefore vanishes at most on a set of measure zero in Ω_g. ∎

Definitions. *Enumerate the roots of* $\det L(\tau,\eta)=0$ *as follows. Let* $\mathcal{A}_f := \{1,\dots,M\}$ *denote the indices of the flat parts, and for* $\alpha\in\mathcal{A}_f$, $\tau_\alpha(\eta) := -\mathbf{v}_\alpha.\eta$. *For* $g\in\mathcal{G}$ *and* $\eta\in\Omega_g$, *number the roots other than the* $\{\tau_\alpha : \alpha\in\mathcal{A}_f\}$ *in order* $\tau_{g,1}(\eta)<\tau_{g,2}(\eta)<\cdots<\tau_{g,M(g)}$. *Multiple roots are not repeated in this list. Let* $\mathcal{A}_c$ *denote the indices of the curved sheets*

$$\mathcal{A}_c := \big\{\,(g,j)\,:\,g\in\mathcal{G}\ \text{ and }\ 1\le j\le M(g)\,\big\}\,. \tag{3.8}$$

Let $\mathcal{A} := \mathcal{A}_f\cup\mathcal{A}_c$. *For* $\alpha\in\mathcal{A}_f$ *and* $\eta\in\mathbb{R}^d$ *define* $E_\alpha(\eta) := \pi(-\mathbf{v}_j.\eta,\eta)$. *For* $\alpha\in\mathcal{A}_c$ *define*

$$E_\alpha(\eta) := \begin{cases} \pi(\tau_\alpha(\eta),\eta)) & \text{for}\quad \eta\in\Omega_g \\ \\ 0 & \text{for}\quad \eta\notin\Omega_g\,. \end{cases} \tag{3.9}$$

The next proposition decomposes an arbitrary solution of $Lu=0$ as a finite sum of simpler waves.

Proposition 3.2. **1.** *For each* $\alpha\in\mathcal{A}$, $E_\alpha(\eta)\in C^\infty(\mathbb{R}^d\setminus(\Gamma\cup\{0\}))$ *is an orthogonal projection valued function positive homogeneous of degree zero.*
2. *For each* $\eta\in\mathbb{R}^d\setminus(\Gamma\cup\{0\})$, $\mathbb{C}^N$ *is equal to the orthogonal direct sum*

$$\mathbb{C}^N = \oplus_{\alpha\in\mathcal{A}}\ \mathrm{Image}\,E_\alpha(\eta)\,. \tag{3.10}$$

3. *The operators* $E_\alpha(D_y) := \mathcal{F}^*E(\eta)\mathcal{F}$ *are orthogonal projectors on* $H^s(\mathbb{R}^d)$, *and for each* $s\in\mathbb{R}$, $H^s(\mathbb{R}^d)$ *is equal to the orthogonal direct sum,*

$$H^s(\mathbb{R}^d) = \oplus_{\alpha\in\mathcal{A}}\ \mathrm{Image}\,E_\alpha(D_y)\,. \tag{3.12}$$

4. *The solution of the initial value problem*

$$L(\partial_x)\,u=0\,,\qquad u|_{t=0}=f \tag{3.13}$$

is given by the formula

$$\hat u(t,\eta) = \sum_{\alpha\in\mathcal{A}} \hat u_\alpha(t,\eta) := \sum_{\alpha\in\mathcal{A}} e^{i\tau_\alpha(\eta)}\,E_\alpha(\eta)\,\hat f(\eta)\,. \tag{3.14}$$

Remarks. 1. The last decomposition is also written

$$u := \sum_{\alpha\in\mathcal{A}} u_\alpha := \sum_{\alpha\in\mathcal{A}} e^{i\tau_\alpha(D_y)}\,E_\alpha(D_y)f\,.$$

2. Since τ_α is real valued on the support of $E_\alpha(\eta)$ the operator $e^{i\tau_\alpha(D_y)}\,E_\alpha(D_y)$ is a contraction on $H^s(\mathbb{R}^d)$ for all s.
3. If $\alpha\in\mathcal{A}_f$ then $\tau_\alpha(D_y)=\mathbf{v}_\alpha.\partial_y$. For $\alpha=(g,j)\in\mathcal{A}_c$, since $|\tau_\alpha(\eta)|\le C|\eta|$ the operator $\tau_\alpha(D_y)f$ is continuous from H^s to H^{s-1}. The mode $u_\alpha = e^{i\tau_\alpha(D_y)}\,E_\alpha(D_y)f$ satisfies $\partial_t u = i\tau_\alpha(D_y)u$.

§**3.2. Large time asymptotics for the homogeneous equation.**

Theorem 3.3. $\mathbf{L}^\infty$ Asymptotics for symmetric systems. *Suppose that* $\hat f \in L^1(\mathbb{R}^d)$ *and* u *is the solution of the initial value problem* $L(\partial_x)u = 0$, $u|_{t=0} = f$. *Then with the notation introduced in the preceding definition,*

$$\lim_{t\to\infty} \Big\| u(t) - \sum_{\alpha\in\mathcal{A}_f} \left(E_\alpha(D_y)f\right)(y - \mathbf{v}_\alpha t) \Big\|_{L^\infty(\mathbb{R}^d)} = 0\,. \tag{3.15}$$

Remarks. 1. This result shows that a general solution of the Cauchy problem is the sum of M rigidly translating waves, one for each hyperplane in the characteristic variety, plus a term which tends to zero in sup norm. The last part decays because of the spreading of waves.

2. The Theorem does not extend to f whose Fourier Transform is a bounded measure. For example, $u := \big(e^{i(y_1 - t)}, 0\big)$ is a solution of (3.4) with $\hat f$ equal to a point mass. The characteristic variety contains no hyperplanes so (3.15) asserts that solutions with $\hat f \in L^1$ tend to zero in $L^\infty(\mathbb{R}^d)$ while $u(t)$ has sup norm equal to 1 for all t.

Proof. Step 1. Approximation-decomposition. Define a Banach space $\mathbb{A}$ and norm as the set of tempered distributions whose Fourier transform is in $L^1(\mathbb{R}^d)$ with

$$\|f\|_{\mathbb{A}} := (2\pi)^{-d/2} \int_{\mathbb{R}^d} |\hat f(\eta)|\; d\eta\,. \tag{3.16}$$

The Fourier Inversion Formula implies that

$$\|f\|_{L^\infty(\mathbb{R}^d)} \le \|f\|_{\mathbb{A}}\,. \tag{3.17}$$

Symmetric hyperbolicity implies that $\exp\big(it\sum A_j\eta_j\big)$ is unitary on $\mathbb{C}^N$ so the evolution operator $S(t) := \exp\big(-t\sum_j A_j\partial_j\big)$ is isometric on $\mathbb{A}$. Since the family of linear maps

$$f \quad\longmapsto\quad S(t)f - \sum_{\alpha\in\mathcal{A}_f} \left(E_\alpha(D_y)f\right)(y - \mathbf{v}_\alpha t)$$

is uniformly bounded from $\mathbb{A}$ to $L^\infty(\mathbb{R}^d)$, it suffices to prove (3.15) for a set of f dense in $\mathbb{A}$.

For $\alpha \in \mathcal{A}_c$, Propostion 3.1.iv shows that the matrix of second derivatives, $\nabla^2_\eta \tau_\alpha$ can vanish at most on a set of measure zero. The set of f we choose is those with

$$\hat f \;\in\; C_0^\infty\big(\mathbb{R}^d \setminus \big\{\, \Gamma \;\cup\; \{0\} \;\cup \bigcup_{\alpha\in\mathcal{A}_c} \{\eta \in \Omega_g \;:\; \nabla^2_\eta \tau_\alpha(\eta) = 0\,\}\;\big\}\,\big)\,,$$

which is dense since the set removed from $\mathbb{R}^d$ has measure zero.

To prove (3.15) for such f decompose

$$f = \sum_{\alpha\in\mathcal{A}} f_\alpha := \sum_{\alpha\in\mathcal{A}} E_\alpha(D_y)\, f\,, \qquad u(t) = S(t)f = \sum u_\alpha(t) := \sum S(t)\, f_\alpha\,. \tag{3.18}$$

For $\alpha \in \mathcal{A}_f$, $u_\alpha(t) = (E_\alpha(D_y)f)\,(y - \mathbf{v}_\alpha t)$ which recovers the summands in (3.15). To prove (3.18) it suffices to show that for $\alpha \in \mathcal{A}_c$

$$\lim_{t\to\infty} \|u_\alpha(t)\|_{L^\infty(\mathbb{R}^d)} = 0\,. \tag{3.19}$$

Step 2. Stationary and nonstationary phase. Proposition 3.2.4 shows that for $\alpha \in \mathcal{A}_c$,

$$u_\alpha(t,y) = \int_{\Omega_g} e^{i(\tau_\alpha(\eta)t + y.\eta)}\, \hat{f}_\alpha(\eta)\; d\eta\,, \qquad \hat{f}_\alpha \in C_0^\infty(\Omega_g)\,. \tag{3.20}$$

For each η in the support of f_α, there is a vector $\mathbf{r} \in \mathbb{R}^d$ so that $\langle \nabla_\eta^2 \tau(\eta)\,\mathbf{r}, \mathbf{r}\,\rangle \neq 0$ on a neighborhood of η. Using a partition of unity we can write $\hat{f}_\alpha$ as a finite sum of functions $\hat{h}_\mu \in C_0^\infty(\Omega_g)$ so that for each μ there is a $\mathbf{r}_\mu \in \mathbb{C}^N$ so that on an open ball containing the support of $\hat{h}_\mu$, $\langle \nabla_\eta^2 \tau(\eta)\,\mathbf{r}_\mu, \mathbf{r}_\mu\,\rangle \neq 0$. It suffices to show that for each μ

$$\lim_{t\to\infty} \Big\| \int e^{i(\tau_\alpha(\eta)t + y.\eta)}\, \hat{h}_\mu(\eta)\; d\eta \Big\|_{L^\infty(\mathbb{R}^d)} \;=\; 0\,. \tag{3.22}$$

To prove (3.22) first make a linear change of variables in η so that $\mathbf{r}_\mu = (1, 0, \ldots, 0)$ and therefore

$$\frac{\partial^2 \tau_\alpha}{\partial^2 \eta_1} \neq 0\,, \qquad \text{on} \quad \text{supp}\; \hat{h}_\mu\,. \tag{3.23}$$

For

$$K(t, y, \eta_2, \ldots, \eta_d) := \int e^{i(\tau_\alpha(\eta)t + y_1.\eta_1)}\, \hat{h}_\mu(\eta)\; d\eta_1$$

one has the simple estimate

$$|K(t, y_1, \eta_2, \ldots, \eta_d)| \;\leq\; \int |\hat{h}_\mu(\eta_1, \eta_2, \ldots, \eta_d)\,|\; d\eta_1 \;\in\; L^1(\mathbb{R}^{d-1})\,.$$

Lebegue's Dominated Convergence Theorem implies that to prove (3.22) it suffices to show that

$$\forall\, \eta_2, \ldots, \eta_d \in \mathbb{R}^{d-1}\,, \qquad \lim_{t\to\infty} \sup_{z\in\mathbb{R}} \Big| \int e^{it(\tau_\alpha(\eta) + z.\eta_1)}\, \hat{h}_\mu(\eta)\; d\eta_1 \Big| = 0\,. \tag{3.24}$$

The one dimensional oscillatory integral in (3.24) is analysed by the method of stationary phase. The derivative of the phase with respect to η_1 is equal to $\partial_1 \tau_\alpha + z$. Thus if

$$|z| \geq 1 + \sup_{\eta\in\text{supp}\,\mathcal{F}h_\mu} |\partial_1 \tau|$$

the principle of nonstationary phase implies that the integral in (3.24) is $o(t^{-\infty})$. Thus it suffices to prove (3.24) with z restricted to a compact set. For $\eta_2, \dots, \eta_d$ fixed, the family of smooth phases is compactly parametrized by z.

By construction $\partial_1^2 \tau \neq 0$ on a ball containing the support of $\hat{h}_\mu$ so for each $\eta_2, \dots, \eta_d$ there is at most one stationary point for the phase. The stationary phase theorem then implies that

$$\int e^{it(\tau_\alpha(\eta)+z.\eta)} \ \hat{h}_\mu(\eta) \ d\eta_1 = O(t^{-1/2})$$

uniformly for z belonging to the compact set. This proves (3.24) and therefore Theorem 3.3. ∎

Corollary 3.4. *If $L(\partial_x)$ is a constant coefficient symmetric hyperbolic operator, then the following are equivalent.*

1. The characteristic variety of L contains no hyperplanes.

2. For every smooth solution of $Lu = 0$ with $u\big|_{t=0} \in C_0^\infty(\mathbb{R}^d)$,

$$\lim_{t\to\infty} \|u(t)\|_{L^\infty(\mathbb{R}^d)} \ \to \ 0\,. \tag{3.25}$$

3. For every $f \in \mathbb{A}$ the solution of the Cauchy problem

$$Lu = 0\,, \qquad u|_{t=0} = f \tag{3.26}$$

satisfies (3.25).

4. If $\tau(\eta)$ is a smooth solution of $\det L(\tau,\eta) = 0$ defined on a open set of $\eta \in \mathbb{R}^d$ then for every $\mathbf{v} \in \mathbb{R}^d$, $\{\eta \in \mathbb{R}^d \ : \ \nabla_\eta \tau = -\mathbf{v}\}$ has measure zero.

Proof. That $\sim(1) \Longrightarrow \sim(4)$ is immediate. On the other hand if (4) is violated there is a smooth solution τ so that $\nabla_\eta \tau = -\mathbf{v}$ on a set of positive measure. It follows from the Fundamental Stratification Theorem (see [BR],[CR]) that $\nabla_\eta \tau = -\mathbf{v}$ on a conic open real algebraic set of dimension d in $\mathbb{R}^d \setminus 0$. Then $\tau = -\mathbf{v}.\eta$ on this set and we conclude that the polynomial $\det L(-\mathbf{v}.\eta, \eta)$ vanishes on this set and therefore everywhere. Thus the hyperplane $\{\tau = -\mathbf{v}.\eta\}$ is contained in the characteristic variety and (1) is violated. Thus (1) and (4) are equivalent.

That (3) implies (2) is immediate and the converse follows from the fact that $C_0^\infty(\mathbb{R}^d)$ is dense in $\mathbb{A}$. Thus (2) and (3) are equivalent.

Theorem 3.3 shows that (1) is equivalent to (3). ∎

Remark. Part four of this Corollary shows that for any given velocity $\mathbf{v}$ the group velocity $-\nabla_\eta \tau$ does not take the value $\mathbf{v}$ for a set of frequencies η of positive measure. □

§3.3 Asymptotics for the inhomogeneous equation.

Equation (2.27) is a linear inhomogeneous equation for $\underline{a}_1$ with source term defined in terms of a_0. Since $\underline{a}_0$ satisfies (2.21) It can be decomposed using Proposition 3.2 which expresses the first term on the right of (2.27) as a sum of terms of the form $e^{it\tau_\alpha(D_y)}f_\alpha$ with $\alpha \in \mathcal{A}$. For large time we will see that the second term in (2.27) is well approximated by a sum of rigidly translating waves. These can come from terms in $\underline{a}_0$ or from rectification of the oscillating parts a_0^* which thanks to (2.26) translates with velocity $-\nabla_\eta\tau(\underline{\eta})$. If this velocity is equal to one of the $\mathbf{v}_j$ then the characteristic variety of V is a hyperplane inside $Char\,L$ and the source terms are of the form $e^{it\tau_\alpha(D_y)}f_\alpha$. Otherwise, the characteristic variety of the transport operator V is a hyperplane which is not contained in $Char\,L$.

Main Lemma 3.5. *Consider the solution of the initial value problem*

$$Lu = f\,, \qquad u|_{t=0} = 0\,, \tag{3.27}$$

where $s \in \mathbb{R}$, and $f \in C([0,\infty[\,;\, H^s(\mathbb{R}^d))$.

1. *If the spatial Fourier Transform of f vanishes outside Ω_g, $\tau_\alpha(\eta)$ defines one of the smooth sheets of $Char\,L$ over Ω_g, and $\partial_t f = i\tau_\alpha(D_y)f$, then*

$$\lim_{t\to\infty}\ \frac{1}{t}\,\|\, u(t) - tE_\alpha(D_y)f(t)\,\|_{H^s(\mathbb{R}^d)} \ =\ 0\,. \tag{3.28}$$

2. *If $V(\partial_x) = \partial_t + \mathbf{v}.\partial_y$ is a constant coefficient vector field whose characteristic variety is not contained in the characteristic variety of L, and f satisfies $V(\partial_x)\,f = 0$, then*

$$\lim_{t\to\infty}\ \frac{1}{t}\,\|u(t)\|_{H^s(\mathbb{R}^d)} \ =\ 0\,. \tag{3.29}$$

Proof. 1. One has $f = \exp(i\tau_\alpha(D))\,f_0$ with $f_0 := f|_{t=0}$. Since the family of operators

$$f_0 \quad\longmapsto\quad \frac{1}{t}\Big(\ u(t) - tE_\alpha(D_y)f(t)\ \Big)\,, \qquad 1 \le t < \infty$$

is uniformly bounded from H^s to itself it suffices to prove (3.28) for f_0 belonging to the dense set $C_0^\infty(\Omega_g)$.

The Fourier Transform of u is supported in Ω_g so parts (2) and (3) of Proposition 3.2. imply that

$$u = \sum_{\gamma\in\mathcal{A}_f} E_\gamma(D_y)\,u + \sum_{j=1}^{M(g)} E_{(g,j)}(D_y)\,u\,.$$

The proof proceeds by computing the summands on the right. Multiplying equation (3.27) by $E_\mu(D_y)$ yields

$$L(\partial_x)\,E_\mu(D_y)\,u = E_\mu(D_y)\,f\,.$$

Simplifying both sides yields

$$(\partial_t - i\tau_\mu(D_y))E_\mu(D_y)\, u = e^{it\tau_\alpha(D_y)}\, E_\mu(D_y)\, f_0(y)\,.$$

The explicit solution is given by

$$E_\mu(\eta)\, \hat{u}(t,\eta) = \int_0^t e^{i(t-\sigma)\tau_\mu(\eta)}\, e^{i\sigma\tau_\alpha(\eta)}\, E_\mu(\eta)\, \hat{f}_0(\eta)\; d\sigma\,.$$

If $\mu = \alpha$ the integrand is independent of σ and the integral is equal to the Fourier transform of $tE_\alpha(D_y)f$ showing that

$$E_\alpha(D_y)\, u(t) = tE_\alpha(D_y)f(t)\,. \tag{3.30}$$

On the other hand, if $\mu \neq \alpha$, the integral is equal to

$$i\, e^{it\tau_\mu(\eta)}\; \frac{1 - e^{it(\tau_\alpha(\eta) - \tau_\mu(\eta))}}{\tau_\alpha(\eta) - \tau_\mu(\eta)}\; \hat{f}_0(\eta)\,.$$

Since the denominator is bounded away from zero on the support of $\hat{f}_0$ it follows that

$$\mu \neq \alpha \quad \Longrightarrow \quad \lim_{t\to\infty}\; \frac{1}{t}\; \big\| E_\mu(D_y)\, u(t)\, \big\|_{H^s(\mathbb{R}^d)} = 0\,. \tag{3.31}$$

The desired result, (3.28), follows from (3.30) and (3.31).

2. The basic energy estimates for V and L imply that

$$\|f(t)\|_{H^s(\mathbb{R}^d)} = \|f(0)\|_{H^s(\mathbb{R}^d)} \qquad \text{and} \qquad \|u(t)\|_{H^s(\mathbb{R}^d)} \le \int_0^t \|f(s)\|_{H^s(\mathbb{R}^d)}\, ds\,. \tag{3.32}$$

It follows that

$$\frac{1}{t}\, \|u(t)\|_{H^s(\mathbb{R}^d)} \le \|f(0)\|_{H^s(\mathbb{R}^d)}\,. \tag{3.33}$$

Thus it suffices to prove (3.29) for $f(0)$ belonging to a dense subset of $H^s(\mathbb{R}^d)$.

The real algebraic subvariety $Char\, L \cap Char\, V$ of the d dimensional hyperplane $Char\, V$ is defined by one polynomial equation. Thus (see [BR], [CR]) it is either all of $Char\, V$ or a subset of dimension $d-1$. By hypothesis the first alternative is ruled out. It follows that $Char\, L \cap Char\, V$ is a variety of dimension $d-1$ so its projection on $\mathbb{R}^d_\eta$,

$$\big\{\, \eta \in \mathbb{R}^d \setminus \Gamma(L)\; :\; (-\mathbf{v}.\eta, \eta) \in Char L\, \big\} \tag{3.34}$$

is of dimension no larger than $d-1$ and therefore is of measure zero.

The dense set of f is taken to be functions whose Fourier Transform is smooth and has compact support disjoint from the union of Γ and the measure zero set (3.34).

Decompose

$$u=\sum_{\alpha\in\mathcal{A}} u_\alpha := \sum_{\alpha\in\mathcal{A}} E_\alpha(D_y)\,u\,,\qquad f=\sum_{\alpha\in\mathcal{A}} f_\alpha := \sum_{\alpha\in\mathcal{A}} E_\alpha(D_y)\,f\,. \tag{3.35}$$

Then

$$\partial_t\, u_\alpha = i\tau_\alpha(D_y)\,u_\alpha + f_\alpha\,,\qquad V(\partial_x)\,f_\alpha = 0\,, \tag{3.36}$$

so

$$\hat f_\alpha(t,\eta) = e^{-it\mathbf{v}.\eta}\,\hat f_\alpha(0,\eta)\,,$$

and

$$\hat u_\alpha(t,\eta) = \int_0^t e^{i\tau_\alpha(\eta)(t-s)}\,\hat f_\alpha(s,\eta)\,ds\,.$$

Using these two formulas yields

$$\frac{1}{t}\,\hat u_\alpha(t,\eta) = e^{it\tau_m(\eta)}\,\hat f_\alpha(0,\eta)\,\frac{1}{t}\int_0^t e^{-is(\mathbf{v}.\eta+\tau_m(\eta))}\,ds\,. \tag{3.37}$$

By the choice of f,

$$\forall m\,,\quad \forall\,\eta\in \operatorname{supp}\hat f_\alpha\,,\quad -\mathbf{v}.\eta\neq\tau_\alpha(\eta)\,. \tag{3.38}$$

Equation (3.37) shows that $\hat f_\alpha$ is supported in a set where the factor multiplying s in the exponent of the integrand is bounded away from zero. Thus the mean value $\frac{1}{t}\int_0^t$ tends uniformly to zero on the support of $\hat f_\alpha(0)$. Multiplying the square of (3.37) by $\langle\xi\rangle^{2s}$ and integrating, (3.29) follows. ∎

Remark. Considering f whose Fourier Transform is localized where $Char\,L$ and $Char\,V$ are as close together as one likes, shows that there is no rate of convergence in (3.29). This shows that it is impossible to improve the $o(1)$ in (3.29) to $O(1/t^\sigma)$ or any other such rate. See also the example in §7.5. □

§4. Profile equations, continuation.

In this section we complete the derivation of the profile equations. In §2, the first equations were derived guaranteeing that the residuals r_0 and r_1 vanish. In this section we find equations guaranteeing that $r_2=0$.

The equations for a_0 are not obvious. One must control the evolution of the mean values and the interaction between the oscillations and the mean values. This interaction is negligible in the leading term unless β belongs to a hyperplane in $Char\,L(0,\partial)$. The equations encode this dichotomy.

An essential difficulty is to describe the mean values for times $t\sim 1/\varepsilon$. This is not obvious even in the absence of oscillations. The mean value satisfies $L(0,\partial_x)\,\underline{a}_0=0$ so that the dynamics of $\underline{a}_0$ is not simple except when L is given by uncoupled transport equations as in the one dimensional case. As soon as there are curved parts of the characteristic variety, there is a fundamental complication in the description. As a warmup we will discuss the dynamics for $t \sim 1/\varepsilon$ of nonoscillatory solutions of semilinear equations. This reveals two important principles.

§4.1. Nonoscillatory semilinear waves, a warmup problem.

Suppose that $\Phi(u)$ is homogeneous of degree J and $p = 1/(J-1)$ is the critical power yielding nonlinear interaction time $\sim 1/\varepsilon$. Consider the family of Cauchy problems

$$L(\partial_x)\, u^\varepsilon + \Phi(u^\varepsilon) = 0\,, \qquad u^\varepsilon|_{t=0} = \varepsilon^p\, g(y) \in C_0^\infty(\mathbb{R}^d)\,. \tag{4.1}$$

Rescale the equation by setting

$$v^\varepsilon := \varepsilon^{-p}\, u^\varepsilon, \tag{4.2}$$

to find

$$L(\partial_x)\, v^\varepsilon + \varepsilon\, \Phi(v^\varepsilon) = 0\,, \qquad v^\varepsilon|_{t=0} = g(y)\,. \tag{4.3}$$

A first approximation is the solution v^0 of the linear problem obtained by setting $\varepsilon = 0$. Since L is conservative, $\|v^0(t)\|_{H^s(\mathbb{R}^d)}$ is independent of time.

Computing $\partial v/\partial\varepsilon|_{\varepsilon=0}$, or performing one Picard iteration yields the next approximation

$$v^\varepsilon \approx v^0 - \varepsilon\, L^{-1}(\Phi(v^0)) := v^0 - \varepsilon\, w$$

where w is the unique solution of

$$L(w) = \Phi(v^0)\,, \qquad w\big|_{t=0} = 0\,. \tag{4.4}$$

Since H^s norm of $v^0(t)$ is independent of t, the H^s norm of $\Phi(v^0(t)$ is bounded provided $s > d/2$. It follows that $\|w(t)\|_{H^s} \le c_s t$. The corrector, $-\varepsilon t w$, is small compared to the principal term, v^0 for times $t = o(1/\varepsilon)$ which verifies again that the nonlinear effects small for $t = o(1/\varepsilon)$.

We find approximate solutions of (4.4) by decomposing v^0 into modes and observing that the nonlinear interaction is simplified using two fundamental principles.

Example illustrating the principle of nonoverlapping waves. Consider the system of two equations in space of dimension one,

$$\frac{\partial v}{\partial t} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial v}{\partial y} + \varepsilon\, \Phi(v) = 0\,. \tag{4.5}$$

The linear approximation

$$v^0 = \big(\, v_1^0(y-t)\,,\, v_2^0(y+t)\,\big)\,. \tag{4.6}$$

contains two waves, a v_1^0 component which moves to the right with speed one and a v_2^0 component moving to the left.

In the source term in (4.4), products of the waves moving leftward with waves moving rightward are essentially zero except for a time interval of order 1 so make a contribution $O(\varepsilon)$ to w and are therefore negligible.

In this way one shows that the corrector w is given by $w = (w_1, w_2) + O(\varepsilon)$ where w is determined by the uncoupled equations

$$(\partial_t + \partial_y)\, w_1 + \Phi_1(v_1^0(y-t), 0) = 0\,, \qquad (\partial_t - \partial_y)\, w_2 + \Phi_2(0, v_1^0(y+t)) = 0\,. \tag{4.7}$$

with initial condition $w(0,y) = 0$. Then $w = O(t)$ so εw is an appreciable corrector for times $t \sim 1/\varepsilon$. *In the nonlinear term only products of terms propagating in the same way through space time need be retained.* □

Example illustrating the principle of ignoring spreading waves. Consider in two dimensions the first order analogue of a semilinear wave equation

$$\frac{\partial u}{\partial t} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial u}{\partial y_1} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial u}{\partial y_2} + \Phi(u) = 0\,, \qquad u(0,y) = \varepsilon^p\, g(y) \in C_0^\infty(\mathbb{R}^2)\,. \tag{4.8}$$

Rescaling yields

$$\frac{\partial v}{\partial t} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{\partial v}{\partial y_1} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial v}{\partial y_2} + \varepsilon\, \Phi(v) = 0\,. \tag{4.9}$$

This case is strikingly different from the preceding example. The reason is that the solution v^0 of the linear equation is $O(t^{-1/2})$ in $L^\infty(\mathbb{R}^2)$ and has $H^s(\mathbb{R}^2)$ norm independent of time. Thus the L^2 norm of $\Phi(v^0)$ is at most $O(t^{-(J-1)/2})$. More generally the H^s norm is $o(1)$ as $t \to \infty$. It follows that the corrector $w = o(t)$ so εw makes a small correction even for times $\sim 1/\varepsilon$. *For modes which decay because they spread over larger and larger regions of space time, the nonlinear term can be neglected.* □

The next result is a quantitative result incorporating the two principles.

Proposition 4.1. *Suppose that $\Psi(u)$ is a smooth function which vanishes along with its partial derivatives of order 1 when $u = 0$. If $f_k \in \cap_s H^s(\mathbb{R}^d)$ is a finite family of functions, $\mathbf{v}_k \in \mathbb{R}^d$ a family of pairwise distinct velocities, $u = \sum_k f_k(y - \mathbf{v}_k t) + \rho(t,y)$, and $\forall\, s, \alpha$*

$$\sup_{t \in [0,\infty[} \|\, \rho(t,y)\, \|_{H^s(\mathbb{R}^d_y)} < \infty\,, \qquad \lim_{t\to\infty} \|\, \partial_y^\alpha \rho(t,y)\, \|_{L^\infty(\mathbb{R}^d_y)} = 0\,, \tag{4.10}$$

then for all s

$$\lim_{t\to\infty} \Big\|\, \Psi(u) - \sum_k \Psi\big(f_k(y - \mathbf{v}_k t)\big)\, \Big\|_{H^s(\mathbb{R}^d_y)} = 0\,. \tag{4.11}$$

Proof. Let $u_1 := \sum_k f_k(y - \mathbf{v}_k t)$. Since the first derivatives of Ψ vanish at the origin, Taylor's Theorem expresses

$$\Psi(u) - \Psi(u_1) = \sum_\nu \psi_\nu(u_1, \rho)\ \ell_\nu(\rho) \tag{4.12}$$

where the ψ_ν are smooth function which vanish when $u_1 = \rho = 0$ and the ℓ_ν are real linear functions.

Leibniz' rule expresses derivatives of $\Psi(u) - \Psi(u_1)$ as a finite sum of terms $\ell_\nu(\partial^\mu \rho)\, \partial^\gamma(\psi_\nu)$. Since ψ_ν vanishes at the origin and u_1 and ρ are bounded in H^s as t tends to infinity, the L^2 norm of the second factor is bounded as $t \to \infty$. At the same time, the L^∞ norm of the first factor tends to zero, so the product tends to zero in L^2. Thus $\|\Psi(u) - \Psi(u_1)\|_{H^s(\mathbb{R}^d)} \to 0$.

For $s > d/2$, the mapping $w \mapsto \Psi(w)$ is uniformly lipshitzean from bounded sets in H^s to H^s. Given a challenge number ε, one can choose $h_k \in C_0^\infty(\mathbb{R}^d)$ so that $f_k - h_k$ is as small as one likes in H^s, and therefore for all t

$$\Big\|\, \Psi\big(\sum h_k(y - \mathbf{v}_k t) \big) - \Psi\big(\sum f_k(y - \mathbf{v}_k t)\big) \,\Big\|_{H^s} < \varepsilon\,,$$

and

$$\Big\|\, \Psi(u_1) - \Psi\Big(\sum h_k(y - \mathbf{v}_k t) \Big) \,\Big\|_{H^s} < \varepsilon\,.$$

Since the h_k have compact support and the speeds $\mathbf{v}_k$ are distinct it follows that for t large,

$$\Psi(\sum h_k(y - \mathbf{v}_k t)) = \sum_k \Psi(h_k(y - \mathbf{v}_k t))\,.$$

The triangle inequality yields

$$\limsup_{t \to \infty}\ \Big\|\Psi(u_1) - \sum_k \Psi(f_k(y - \mathbf{v}_k t))\Big\|_{H^s} \le 2\,\varepsilon$$

and the result follows. ∎

§4.2. Profile equations, endgame.

In this section we complete the description of the equations determining the profiles a_j which guarantee that r_0, r_1, r_2 vanish. It remains to analyse r_2. The equation $r_2 = 0$ is split into three pieces, $\underline{r}_2 = 0$, $\pi(\beta)\, r_2^* = 0$, and $Q(\beta)\, r_2^* = 0$.

§4.2.1. The equation $\underline{r}_2 = 0$.

Use (2.21) and Theorem 3.3 to decompose $\underline{a}_0$ into modes of $L(0, \partial_x)$,

$$a_0 = a_0^* + \underline{a}_0 = a_0^* + \sum_{\alpha \in \mathcal{A}} \underline{a}_{0,\alpha}\,, \qquad \underline{a}_{0,\alpha} = E_\alpha(D_y)\, \underline{a}_0 \tag{4.13}$$

with $\|\underline{a}_{0,\alpha}(t)\|_{L^\infty} \to 0$ as $t \to \infty$ if $\alpha \in \mathcal{A}_c$ and $(\partial_t + \mathbf{v}_\alpha.\partial_x)\underline{a}_{0,\alpha} = 0$ if $\alpha \in \mathcal{A}_f$. This decomposition is injected into equation (2.27). The constraint (2.9) that $a_1(T, Y, t, y)$ must grow sublinearly in t imposes conditions on a_0.

The key step is to simplify the nonlinear term

$$\Big\langle\ \Phi(a_0) + \sum_{\mu=0}^{d} \beta_\mu\, \Lambda_\mu(a_0)\, \partial_\theta a_0\ \Big\rangle$$

using the principles of the last sections. The principle of ignoring spreading terms suggests that the terms $\underline{a}_{0,\alpha}$ with $\alpha \in \mathcal{A}_c$ can be dropped. The other terms involve a_0^* and $\underline{a}_{0,\alpha}$ with $\alpha \in \mathcal{A}_f$ each of which satisfies a transport equation. The principle of nonoverlapping waves suggests that we ignore the interaction of summands which move at different speeds. For this it is important to know whether the speed of V is equal to one of the $\mathbf{v}_j$. The speed is one of the $\mathbf{v}_j$ if and only if the hyperplane $Char\, V(\partial_x)$ is contained in $Char\, L(0, \partial_x)$. Since $Char\, V(\partial_x)$ is the tangent plane to the variety at β by Corollary 2.3, $Char\, V(\partial_x)$ is contained in $Char\, L(0, \partial_x)$ if and only if β lies on a hyperplane in the variety.

This dichotomy strongly affects the nature of the profile equation since if β belongs to a hyperplane in the variety, the oscillations travel at exactly the same speed as one of the nonspreading modes for L. This encourages interaction between the mean field and the oscillations. For β belonging to a curved sheet in the variety, the interaction is weaker.

Definitions. *Suppose that $\beta = (\underline{\tau}, \underline{\eta})$ satisfies the simple characteristic variety hypothesis. Define a switch ι by*

$$\iota = \begin{cases} 1 & \text{if} \quad Char\, V(\partial_x) \subset Char\, L(0, \partial_x) \\[2ex] 0 & \text{if} \quad Char\, V(\partial_x) \not\subset Char\, L(0, \partial_x)\,. \end{cases} \tag{4.14}$$

When $Char\, V \subset Char\, L$, enumerate the projections E_α, $\alpha \in \{1, \ldots, M\} = \mathcal{A}_f$ from the definition before Proposition 3.2 so that V corresponds to $\alpha = 1$, that is $V = \partial_t + \mathbf{v}_1 . \partial_y$.

When $Char\, V \not\subset Char\, L$ the term a_0^* translates at a speed different than the $\mathbf{v}_j$. Then the principle of ignoring spreading waves and the interaction of nonoverlapping waves suggest the replacement

$$\Phi(a_0) \quad \longmapsto \quad \Phi(a_0^*) + \sum_{j=1}^{M} \Phi(\underline{a}_{0,j})\,.$$

When $Char\, V(\partial_x) \subset Char\, L$, the term a_0^* translates at the same speed as the term $\underline{a}_{0,1}$ and the principles suggest the replacement

$$\Phi(a_0) \quad \longmapsto \quad \Phi(a_0^* + \underline{a}_{0,1}) + \sum_{j=2}^{M} \Phi(\underline{a}_{0,j})\,.$$

This dichotomy is summarized by

$$\Phi(a_0) \quad \longmapsto \quad (1-\iota)\Phi(a_0^*) + \Phi(\iota\, a_0^* + \underline{a}_{0,1}) + \sum_{j=2}^{M} \Phi(\underline{a}_{0,j})\,. \tag{4.15}$$

For the quasilinear term, reasoning as above suggests the replacement

$$\Lambda_\mu(a_0)\,\partial_\theta a_0 \quad\longmapsto\quad \Big((1-\iota)\Lambda_\mu(a_0^*) + \Lambda_\mu(\iota a_0^* + \underline{a}_{0,1}) + \sum_{j=2}^{M} \Lambda_\mu(\underline{a}_{0,j})\Big)\,\partial_\theta a_0\,.$$

The factor $\partial_\theta a_0 = \partial_\theta a^*$ moves with the speed of the vector field V, and each of the summands in the other factor also translates at constant velocity. The principle of nonoverlapping waves suggests that only the summands annihilated by V need be retained. When $\iota = 0$ this is $\Lambda_\mu(a_0^*)\partial_\theta a_0$. while when $\iota = 1$ it is $\Lambda_\mu(\iota a_0^* + \underline{a}_{0,1})$. These choices are summarized by the simplification

$$\Lambda_\mu(a_0)\,\partial_\theta a_0 \;\longmapsto\; \Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\,\partial_\theta a_0^*\,,$$

which generates the quasilinear term

$$\sum_{\mu=0}^{d} \beta_\mu\,\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\,\partial_\theta a_0^*\,, \tag{4.16}$$

Adding (4.15) and (4.16) yields

$$\mathcal{N}(a_0,\partial_\theta a_0) := (1-\iota)\Phi(a_0^*) + \Phi(\iota\, a_0^* + \underline{a}_{0,1}) + \sum_{j=2}^{M}\Phi(\underline{a}_{0,j}) + \sum_{\mu=0}^{d}\beta_\mu\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\partial_\theta a_0. \tag{4.17}$$

With these definitions one has for a_0 satisfying (2.21)

$$\Phi(a_0) + \sum \beta_\mu\Lambda_\mu(a_0)\partial_\theta a_0 = \mathcal{N}(a_0,\partial_\theta a_0) + o(1) \qquad \text{as}\;\; t\to\infty\,. \tag{4.18}$$

Here and in succeeding estimates the terms $o(1)$ tend to zero in $L^\infty(\mathbb{R}^d)$ as $t\to\infty$. This follow from Proposition 4.1 and Main Lemma 3.5. Estimate (5.26) is a more precise version. Inserting (4.18) in (2.27) yields

$$L(0,\partial_x)\,\underline{a}_1 = \Big\langle -L(0,\partial_X)\,\underline{a}_0 - \mathcal{N}(a_0,\partial_\theta a_0)\Big\rangle + o(1)\,. \tag{4.19}$$

Solving with initial data $\underline{a}_1\big|_{t=0} = 0$ yields

$$\underline{a}_1 = L(0,\partial_x)^{-1}\Big(-\sum_{\alpha\in\mathcal{A}} L(0,\partial_X)\,\underline{a}_{0,\alpha} - \Big\langle \mathcal{N}(a_0,\partial_\theta a_0)\Big\rangle\Big) + o(t)\,. \tag{4.20}$$

Each source term on the right hand side of (4.20) satisfies a differential equation which allows us to use Main Lemma 3.5. Introduce the switch

$$\kappa(\alpha) := \begin{cases} 1 & \text{if}\;\; \alpha\in\mathcal{A}_f \\ 0 & \text{if}\;\; \alpha\in\mathcal{A}_c \end{cases} \tag{4.21}$$

Decompose the source term

$$S := \sum_{\alpha\in\mathcal{A}} L(0,\partial_X)\,\underline{a}_{0,\alpha} + \Big\langle \mathcal{N}(a_0,\partial_\theta a_0)\Big\rangle = S_* + S_1 + \sum_{\alpha\in\mathcal{A}\setminus\{1\}} S_\alpha\,, \tag{4.22}$$

where S_α for $\alpha = 1$ is defined by

$$S_1 := L(0,\partial_X)\underline{a}_{0,1} + \Big\langle\, \kappa(1)\,\Phi(\iota a_0^* + \underline{a}_{0,1})\,\Big\rangle\,, \tag{4.23}$$

while

$$S_\alpha := L(0,\partial_X)\underline{a}_{0,\alpha} + \kappa(\alpha)\,\Phi(\underline{a}_{0,\alpha}) \qquad \text{for} \quad \alpha \neq 1\,, \tag{4.24}$$

$$S_* := \Big\langle\, (1-\iota)\Phi(a_0^*) + \sum_{\mu=0}^{d} \beta_\mu\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\partial_\theta a_0\,\Big\rangle\,, \qquad (4.25)\,.$$

They satisfy

$$V(\partial_x)\,S_* = 0\,, \qquad \big(\,\partial_t + \tau_\alpha(D_y)\,\big)\,S_\alpha = 0\,.$$

The second part of Main Lemma 3.5 together with Proposition 4.1 yield

$$\begin{aligned} L(0,\partial_x)^{-1}\,S_\alpha &= t\,E_\alpha(D_y)\,S_\alpha + o(t)\,,\\ L(0,\partial_x)^{-1}\,S_* &= \iota\,t\,E_1(D_y)\,S_* + o(t)\,. \end{aligned} \tag{4.26}$$

Note that since $\iota(1-\iota) = 0$ the $(1-\iota)\Phi(a_0^*)$ term in S_* does not contribute to the right hand side of the second equation in (4.26).
Equations (4.20), (4.22), and (4.26) show that $\underline{a}_1$ is sublinear in time if and only if $0 = E_\alpha(D_y)\,S_\alpha = E_1(D_y)\,S_*$ that is

$$E_1(D_y)\bigg(\,L(0,\partial_X)\,\underline{a}_{0,1} + \Big\langle\,\kappa(1)\,\Phi(\iota\,a_0^* + \underline{a}_{0,1}) + \iota\sum_\mu \beta_\mu\Lambda_\mu(a_0^* + \underline{a}_{0,1})\,\partial_\theta a_0^*\,\Big\rangle\bigg) = 0\,, \tag{4.27}$$

$$E_\alpha(D_y)\,\Big(\,L(0,\partial_X)\,\underline{a}_{0,\alpha} + \kappa(\alpha)\,\Phi(\underline{a}_{0,\alpha})\Big) = 0\,, \qquad \text{for} \quad \alpha\in\mathcal{A}\setminus\{1\}\,. \tag{4.28}$$

Equations (4.27), (4.28) determine the dynamics in T of the mean values.

§4.2.2. The equation $\pi(\beta)\,\mathbf{r}_2^* = \mathbf{0}$

Multiplying the expression for r_2 on the left of (2.18) by $\pi(\beta)$ eliminates the $L(0,\beta)\partial_\theta$ term. Taking the oscillatory part yields

$$\pi\,L(0,\partial_X)\,\pi\,a_0^* + \pi\,L(0,\partial_x)a_1^* + \pi\,\Big(\Phi(a_0) + \sum \beta_\mu\Lambda_\mu(a_0)\,\partial_\theta a_0\Big)^* = 0\,. \tag{4.29}$$

Using (2.23) write

$$a_1^* = \pi(\beta)\,a_1^* + (1-\pi(\beta))\,a_1^* = \pi(\beta)\,a_1^* - Q(\beta)\,L(0,\partial_x)\,\partial_\theta^{-1}a_0^*\,.$$

Inserting this into (4.29) yields

$$\pi\, L(0,\partial_X)\,\pi\, a_0^* - \pi\, L(0,\partial_x)\, Q(\beta)\, L(0,\partial_x)\,\pi a_0^* + \pi\,\Big(\Phi(a_0) + \sum \beta_\mu \Lambda_\mu(a_0)\,\partial_\theta a_0\Big)^* = -\pi\, L(0,\partial_x)\,\pi\, a_1^*\,.$$

Use Proposition 2.2 two times to find

$$V(\partial_X)\,\pi\, a_0^* - \pi\, L(0,\partial_x)\, Q(\beta)\, L(0,\partial_x)\,\pi a_0^* + \pi\,\Big(\Phi(a_0) + \sum \beta_\mu \Lambda_\mu(a_0)\,\partial_\theta a_0\Big)^* = -V(\partial_x)\,\pi\, a_1^*\,. \tag{4.30}$$

The constraint that $a_1^*(T,Y,t,y,\theta)$ grow sublinearly in t places restrictions on the left hand side of (4.30) which determines our next profile equation. Equation (2.26) shows that the first two summands on the left are annihilated by $V(\partial_x)$. Thus,

$$V(\partial_x)^{-1}\Big(V(\partial_X)\,\pi\, a_0^* - \pi\, L(0,\partial_x)\, Q(\beta)\, L(0,\partial_x)\,\pi a_0^*\Big) = t\,\Big(V(\partial_X)\,\pi\, a_0^* - \pi\, L(0,\partial_x)\, Q(\beta)\, L(0,\partial_x)\,\pi a_0^*\Big)\,. \tag{4.31}$$

Equation (4.18) implies that

$$V(\partial_x)^{-1}\,\pi\,\Big(\Phi(a_0) + \sum \beta_\mu \Lambda_\mu(a_0)\,\partial_\theta a_0\Big)^* = V(\partial_x)^{-1}\,\pi\,\mathcal{N}(a_0,\partial_\theta a_0)^* + o(t)\,. \tag{4.32}$$

We next consider the terms from the right hand side of (4.32).

Since $\Phi(a_0^*)$ and $\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\partial_\theta a_0$ are annihilated by $V(\partial_x)$ one has

$$V(\partial_x)^{-1}\Big((1-\iota)\,\pi\,\Phi(a_0^*) + \sum_{\mu=0}^{1} \beta_\mu\,\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\partial_\theta a_0\Big)^* = t\,\Big((1-\iota)\,\pi\,\Phi(a_0^*) + \sum_{\mu=0}^{d} \beta_\mu\,\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\partial_\theta a_0\Big)^*\,. \tag{4.33}$$

Similarly if $\iota = 1$ one has

$$V(\partial_x)^{-1}\,\Phi(\iota\, a_1^* + \underline{a}_{0,1})^* = t\,\Phi(\iota\, a_1^* + \underline{a}_{0,1})^*\,. \tag{4.34}$$

If $\iota = 0$, $\Phi(\iota\, a_1^* + \underline{a}_{0,1})$ is nonoscillatory so the contribution vanishes.

The terms $\Phi(\underline{a}_{0,j})$ in $\mathcal{N}$ are nonoscillating so make no contribution to (4.32). Combined with (4.30-34) this shows that πa_1^* grows sublinearly in t if and only if

$$V(\partial_X)\,\pi\, a_0^* - \pi\, L(0,\partial_x)\, Q(\beta)\, L(0,\partial_x)\,\pi\partial_\theta^{-1} a_0^* + (1-\iota)\,\pi\,\Phi(a_0^*)^* + \iota\,\pi\,\Phi(\iota\, a_1^* + \underline{a}_{0,1})^* + \Big(\pi\sum_{\mu=0}^{d} \beta_\mu\,\Lambda_\mu(a_0^* + \iota\underline{a}_{0,1})\partial_\theta a_0\Big)^* = 0\,. \tag{4.35}$$

When the smooth characteristic variety hypothesis is satisfied, the operator $\pi\, L\, Q\, L\,\pi$ appearing in (4.35) is essentially scalar. That is the content of the next proposition whose proof proceeds by differentiating twice the identity $(\sum A_j\xi_j)\pi(\xi) = \tau(\xi)\pi(\xi)$. The details can be found in [DJMR]. The expression here is simpler than in that reference thanks to the convention that $A_0(0) = I$.

Proposition 4.2. *If the smooth characteristic variety hypothesis is satisfied at* $\beta = (\tau(\eta), \eta)$ *and* $A_0(0) = I$, *then*

$$\pi(\beta)\, L(0, \partial_x)\, Q(\beta)\, L_1(\partial_x)\, \pi(\beta) \;=\; -\frac{1}{2}\, \pi(\beta) \left(\sum_{j,k} \frac{\partial^2 \tau(\eta)}{\partial \eta_j \partial \eta_k} \frac{\partial^2}{\partial y_j \partial y_k} \right). \tag{4.36}$$

Define the scalar differential operator

$$R(\partial_y) := \frac{1}{2} \sum_{j,k=1}^{d} \frac{\partial^2 \tau(\underline{\eta})}{\partial \eta_j \partial \eta_k} \frac{\partial^2}{\partial y_j \partial y_k}. \tag{4.37}$$

The dynamics (4.35) of $a_0^* = \pi\, a_0^*$ with respect to T simplifies to

$$\begin{aligned} &V(\partial_X)\, a_0^* - R(\partial_y)\, \partial_\theta^{-1} a_0^* + (1-\iota)\, \pi(\beta)\, \Phi(a_0^*)^* \\ &\qquad + \iota\, \pi(\beta)\, \Phi(a_0^* + \underline{a}_{0,1})^* + \Big(\pi(\beta) \sum_{\mu=0}^{d} \beta_\mu\, \Lambda_\mu(a_0^* + \iota \underline{a}_{0,1}) \partial_\theta a_0\Big)^* = 0\,. \end{aligned} \tag{4.38}$$

Remarks. 1. In all cases $V(\partial_x)\, a_0^* = 0$ and the operator $V(\partial_X)$ in (4.38) is converted to ∂_T by writing $a_0^* = \mathrm{a}(T, Y - \mathbf{v}T, t, y - \mathbf{v}t, \theta)$. Then (4.38) is equivalent to

$$\partial_T\, \mathrm{a} - R(\partial_y)\, \partial_\theta^{-1}\, \mathrm{a} + (I-\iota)\, \pi\, \Phi(\mathrm{a})^* + \iota\, \pi\, \Phi(\mathrm{a} + \underline{a}_{0,1})^* + \Big(\pi \sum_{\mu=0}^{d} \beta_\mu\, \Lambda_\mu(\mathrm{a} + \iota \underline{a}_{0,1}) \partial_\theta \mathrm{a}\Big)^* = 0\,.$$

2. In the semilinear case, the Λ terms are absent, yielding (1.16) of the introduction. □

§4.2.3. The equation $\mathbf{Q}(\beta)\, \mathbf{r}_2^* = \mathbf{0}$.

The equation in the title is satisfied if and only if

$$\begin{aligned} &(I - \pi)\, a_2^* = \\ &\quad -Q(\beta)\, \partial_\theta^{-1} \Big[L(0, \partial_x) a_1^* + L(0, \partial_X)\, a_0^* + \Big(\Phi(a_0) + \sum \beta_\mu \Lambda_\mu(a_0)\, \partial_\theta a_0\Big)^* \Big]. \end{aligned} \tag{4.39}$$

Summary. *The principal profile* a_0 *is determined from its initial data at* $t = T = 0$ *from equations (2.20), (2.21), (2.26), (4.27), (4.28) and (4.38). The parts* $\underline{a}_1$ *and* $\pi(\beta)\, a_1^*$ *of the second profile are determined by solving (2.27) and (4.30) with initial data vanishing at* $t = 0$ *Finally,* $(I - \pi(\beta)) a_1^*$ *and* $a_2 = (I - \pi) a_2^*$ *are given by (2.23) and (4.39) respectively.*

In the next section we verify that the initial value problems referred to in this summary are in fact well posed locally in time, and the terms expected to be sublinear are in fact sublinear. That the solutions provide accurate approximate solutions is proved in §6.

§5. Solvability of the profile equations.

§5.1. Solvability of the equations for the leading profile $\mathbf{a_0}$.

Theorem 5.1. *Suppose that $g(Y,y,\theta)$ belongs to $\cap_s H^s(\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{T})$ and satisfies the polarization condition $\pi(\beta)\,g^*=g^*$. Then there is a $T_*>0$ and a unique $a_0(T,Y,t,y,\theta)$ so that for all $0<\underline{T}<T_*$ and all $\alpha\in\mathbb{N}^{2d+3}$*

$$\sup_{0\le T\le \underline{T}}\ \sup_{t\in\mathbb{R}}\ \|\,\partial^\alpha_{T,Y,t,y,\theta}\,a_0(T,.,t,.,.)\,\|_{L^2(\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T}_\theta)}\ <\ \infty\,, \tag{5.1}$$

a_0 satisfies (2.20), (2.21), (2.26), (4.27), (4.28) and (4.38) together with the initial condition

$$a_0(0,Y,0,y,\theta)\ =\ g(Y,y,\theta)\,. \tag{5.2}$$

Proof. The strategy is the following. Equations (4.27), (4.28), and (4.38) involve $\partial_{T,Y,y,\theta}$ but not ∂_t. In the plane $\{t=0\}$ one solves the quasilinear initial value problem defined by (4.27), (4.28), (4.38) with time variable T. Then the Cauchy problem defined by equations (2.21), (2.26) with initial values determined by $a_0|_{t=0}$ extends the solution to all t. The validity of (4.27), (4.28), (4.38) for all t is proved using a commutation argument together with uniqueness for hyperbolic initial value problems. The next lemma performs the first step.

Lemma 5.2. Solvability in $\{\mathbf{t=0}\}$. *There is a $\underline{T}>0$ and unique solution $w(T,Y,y,\theta)\in\cap_s C^s([0,\underline{T}]\ ;\ H^s(\mathbb{R}^{2d}_{Y,y}\times\mathbb{T}))$ to equations (4.27), (4.28), (4.38) and satisfying the initial condition $w(0,Y,y,\theta)=g$. In addition, $\pi(\beta)\,w^*=w^*$.*

Proof. Step 1. Existence follows from a differential inequality.

We construct solutions as limits as $h\to 0^+$ of the nonlinear ordinary differential equations obtained by replacing $\partial_{Y,y,\theta}$ in (4.27), (4.28), and (4.38) by the associated symmetric finite difference operators, $\delta^h_{Y,y,\theta}$. The replacement is not made in the expressions $E_\alpha(D_y)$ and ∂_θ^{-1} which are bounded operators as is. With $w_\alpha:=E_\alpha(D_y)\,w$, the resulting equations are

$$E_1(D_y)\bigg(L(0,\partial_T,\delta^h_Y)\underline{w}^h_1+ \\ +\Big\langle\kappa(1)\,\Phi(\underline{w}^h_1+\iota\,w^{h,*})+\iota\sum_\mu\beta_\mu\Lambda_\mu(w^{h,*}+\underline{w}_1)\,\delta^h_\theta w^{h,*}\Big\rangle\bigg)=0\,, \tag{5.3}$$

$$E_\alpha(D_y)\Big(L(0,\partial_T,\delta^h_Y)\underline{w}_\alpha+\kappa(\alpha)\,\Phi(\underline{w}_\alpha)\Big)=0\,,\qquad \text{for}\quad \alpha\in\mathcal{A}\setminus\{1\}\,, \tag{5.4}$$

$$V(\partial_T,\delta^h_Y)\, w^{h,*} - R(\delta^h_y)\, \partial_\theta^{-1}\, w^{h,*} + (1-\iota)\, \pi(\beta)\, \Phi(w^{h,*})^* \\ + \iota\, \pi(\beta)\, \Phi(w^{h,*} + \underline{w}_1)^* + \Big(\pi(\beta) \sum_{\mu=0}^{d} \beta_\mu\, \Lambda_\mu(w^{h,*} + \iota\, \underline{w}_1)\delta^h_\theta w\Big)^* = 0\,. \tag{5.5}$$

These three equations together express

$$\partial_T w^h \;=\; E_1(D_y)\, \partial_T\, \underline{w}^h_1 + \sum_{\alpha\neq 1} E_\alpha(D_y)\, \partial_T\, \underline{w}^h_\alpha + \partial_T w^{h,*}$$

as a function $G_h(w^h)$.

If $s > (2d+1)/2$ and h are fixed, Schauder's Lemma implies that the nonlinear map G_h is uniformly lipshitzean from bounded subsets of $H^s(\mathbb{R}^{2d}\times\mathbb{T})$ to $H^s(\mathbb{R}^{2d}\times\mathbb{T})$. Picard's Fundamental Existence Theorem for ordinary differential equations implies that there is a $T_*(s,h) > 0$ and a unique local solution

$$w^h \in C^1\big(\, [0, T_*(s,h)[\ ;\ H^s(\mathbb{R}^{2d}\times\mathbb{T})\,\big) \tag{5.6}$$

and that if $T_*(s,h) < \infty$ then

$$\lim_{t\nearrow T_*(s,h)} \ \|\, w^h(T,\cdot,\cdot,\cdot)\,\|_{H^s(\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{T}))} \;=\; \infty\,. \tag{5.7}$$

Equations (5.3)-(5.5) imply that $\partial_T\,\big\{(I-\pi(\beta))w^{h,*}\,\big\} = 0$, so the polarization $\pi(\beta)w^{h,*} = w^{h,*}$ holds as soon as it holds at $T=0$.

Lemma 5.2 follows from the following *a priori* estimate. Fix $\underline{s} > (2d+3)/2$. For each $s \geq \underline{s}$, there is a continuous function $K_s : \mathbb{R}_+ \to \mathbb{R}_+$ independent of h so that the local solutions satisfy for $T < T_*(s,h)$

$$\frac{d\,\|w^h(T)\|^2_{H^s(\mathbb{R}^{2d}\times\mathbb{T})}}{dT} \leq K_s\big(\,\|\, w^h(T)\,\|_{H^{\underline{s}}(\mathbb{R}^{2d}\times\mathbb{T})}\,\big)\ \|w^h(T)\|^2_{H^s(\mathbb{R}^{2d}\times\mathbb{T})}\,. \tag{5.8}$$

To see that this suffices first take $s = \underline{s}$. Let $\zeta(t) \in C([0,T_1[\,;\,\mathbb{R})$ be the maximal solution of

$$2\,\frac{d\zeta}{dT} = K_{\underline{s}}(\zeta(T))\,\zeta(T)\,, \qquad \zeta(0) = ||g||_{H^s(\mathbb{R}^{2d}\times\mathbb{T})}\,.$$

Then inequality (5.8) implies that for $T < T_1$

$$\|w^h(T)\|_{H^{\underline{s}}(\mathbb{R}^{2d}\times\mathbb{T})} \leq \zeta(T)\,. \tag{5.9}$$

This proves that $T_*(\underline{s},h) \geq T_1$ which gives an h independent domain of existence for w^h.

For any $\underline{T} \in [0,T_1[$ and $s \geq \underline{s}$ let

$$C_1 = C_1(\underline{T}) := \sup_{0\leq T\leq \underline{T}} |\zeta(T)|\,, \qquad C_2 = C_2(s,\underline{T}) := \sup_{|\rho|\leq C_1(\underline{T})} |K_s(\rho)|\,.$$

Inequality (5.8) implies that for $T \leq \underline{T}$,

$$\|w^h(T)\|^2_{H^s(\mathbb{R}^{2d}\times\mathbb{T})} \leq e^{C_2 T}\,\|g\|^2_{H^s(\mathbb{R}^{2d}\times\mathbb{T})}\,. \tag{5.10}$$

Using the differential equation to express time derivatives in terms of space derivatives implies corresponding estimates

$$\|\partial_T^k w^h(T)\|_{H^s(\mathbb{R}^{2d}\times\mathbb{T})} \leq C(s,k,\underline{T})\,\|g\|_{H^{s+2k}(\mathbb{R}^{2d}\times\mathbb{T})}\,, \qquad 0\leq T\leq \underline{T}\,. \tag{5.11}$$

With these bounds uniform in h it is routine to pass to the limit constructing a solution in $\cap_s C^s([0,\underline{T}]\,;\,H^s(\mathbb{R}^{2d}\times\mathbb{T})$. Passing to the limit also yields the polarization identity.

Step 2. Proof of the differential inequality (5.8).

On $\mathbb{T}$ use the measure $d\theta/2\pi$ of total mass equal to one. To prove (5.8) add the results of taking the real part of the $L^2(\mathbb{R}^{2d}\times\mathbb{T})$ scalar product of

$$\begin{aligned}
&(5.3)\quad \text{with}\quad (1-\Delta_{Y,y})^s\underline{w}_1^h \\
&(5.4)\quad \text{with}\quad (1-\Delta_{Y,y})^s\underline{w}_\alpha^h \\
&(5.5)\quad \text{with}\quad (1-\Delta_{Y,y,\theta})^s w^{h,*}\,.
\end{aligned}\tag{5.12}$$

The symmetry of $A(0)$ and the antisymmetry of δ_Y^h show that from (5.3) and (5.4) one has for all $1\le j\le d$ and $\alpha\in\mathcal{A}$

$$\Re\Big(\,(1-\Delta_{Y,y})^s\underline{w}_\alpha^h\;,\;E_\alpha(D_y)\,A_j(0)\,\delta_{Y_j}^h\,\underline{w}_{1\alpha}^h\,\Big)_{L^2(\mathbb{R}^{2d})}\;=\;0\,.$$

Similarly from (5.5)

$$\Re\Big(\,(1-\Delta_{Y,y,\theta})^s w^{h,*}\;,\;\mathbf{v}.\delta_Y^h\,w^{h,*}-R(\delta_y^h)\,\partial_\theta^{-1}\,w^{h,*}\,\Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\;=\;0\,.$$

Thus summing the real parts from (5.12) yields

$$\frac{1}{2}\,\frac{d\,\|w^h(T)\|^2_{H^s(\mathbb{R}^{2d}\times\mathbb{T})}}{dT}\;=\;\Re I_1+\iota\sum_{\mu=0}^{d}\beta_\mu\,\Re I_{2,\mu}\,,\tag{5.13}$$

where I_1 contains the Φ terms

$$\begin{aligned}
I_1 :=\;&\Big(\,(1-\Delta_{Y,y})^s\underline{w}_1^h\,,\,\big\langle\,\kappa(1)\,\Phi(\underline{w}_1^h+\iota\,w^{h,*})\big\rangle\,\Big)_{L^2(\mathbb{R}^{2d})}\\
&+\sum_{j=2}^{M}\Big(\,(1-\Delta_{Y,y})^s\underline{w}_j^h\,,\,\Phi(\underline{w}_j)\,\Big)_{L^2(\mathbb{R}^{2d})}\\
&+\big(\,(1-\Delta_{Y,y,\theta})^s w^{h,*}\;,\;(1-\iota)\,\pi\,\Phi(w^{h,*})^*+\iota\,\pi\,\Phi(\iota\,w^{h,*}+\underline{w}_1)\,\Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,.
\end{aligned}\tag{5.14}$$

The more troublesome terms are the quasilinear terms $I_{2,\mu}$

$$\begin{aligned}
I_{2,\mu} :=\;&\Big(\,(1-\Delta_{Y,y})^s\,\underline{w}_1^h\,,\,\iota\,\big\langle\,\Lambda_\mu(w_0^*+\underline{w}_1)\,\delta_\theta^h\,w^{h,*}\,\big\rangle\,\Big)_{L^2(\mathbb{R}^{2d})}\\
&+\Big(\,(1-\Delta_{Y,y,\theta})^s w^{h,*}\,,\,\big(\,\Lambda_\mu(\iota\,\underline{w}_1^h+w^{h,*})\,\delta_\theta^h w^{h,*}\,\big)^*\,\Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,.
\end{aligned}\tag{5.15}$$

Moser's inequality shows that

$$|I_1|\;\le\;K_s(\,\{||w_\alpha^h(T)||_{L^\infty}\;:\;\alpha\in\mathcal{A}\}\,)\,||w^h||^2_{H^s}\;\le\;\text{r.h.s. of (5.8)}\,.\tag{5.16}$$

The crucial point for estimating the real part of $I_{2,\mu}$ is the identity

$$I_{2,\mu} = \Big((1-\Delta_{Y,y,\theta})^s (\underline{w}_1^h + \iota\, w^{h,*})\,,\, \Lambda_\mu(\underline{w}_1^h + \iota\, w^{h*})\, \delta_\theta^h (\underline{w}_1^h + \iota\, w^{h,*}) \Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,, \tag{5.17}$$

which is proved in the next lines.

First consider the case $\iota = 1$. In the scalar product in the first summand on the right in (5.15), the second factor is a mean value so is orthogonal to purely oscillating terms. Thus in this summand $\underline{w}_1$ can be replaced by $\underline{w}_1 + w^{h,*}$. Similarly in the second summand on the right of (5.15) the second factor has mean zero so that in the first factor $w^{h,*}$ can be replaced by $\underline{w}_1 + w^{h,*}$. Finally use that $\delta_\theta^h w^{h,*} = \delta_\theta^h(\underline{w}_1 + w^{h,*})$ to find that when $\iota = 1$,

$$\begin{aligned} I_{2,\mu} := &\Big((1-\Delta_{Y,y,\theta})^s\, (\underline{w}_1^h + w^{h,*})\,,\, \big\langle\, \Lambda_\mu(w_0^* + \underline{w}_1)\, \delta_\theta^h\, (\underline{w}_1^h + w^{h,*})\,\big\rangle\, \Big)_{L^2(\mathbb{R}^{2d})} \\ &+ \Big((1-\Delta_{Y,y,\theta})^s (\underline{w}_1^h + w^{h,*})\,,\, \big(\Lambda_\mu(\underline{w}_1^h + w^{h,*})\, \delta_\theta^h (\underline{w}_1^h + w^{h,*})\big)^*\, \Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,. \end{aligned}$$

The first factors in the two scalar products are equal so that the sum simplifies and one has

$$I_{2,\mu} = \Big((1-\Delta_{Y,y,\theta})^s (\underline{w}_1^h + w^{h,*})\,,\, \Lambda_\mu(\underline{w}_1^h + w^{h*})\, \delta_\theta^h (\underline{w}_1^h + w^{h,*}) \Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,. \tag{5.18}$$

This proves (5.17) when $\iota = 1$.

When $\iota = 0$, use the fact that $w^{h,*}$ has mean zero to find that

$$\iota = 0 \quad \Longrightarrow \quad I_{2,\mu} = \Big((1-\Delta_{Y,y,\theta})^s\, w^{h,*}\,,\, \Lambda_\mu(w^{h*})\, \delta_\theta^h (w^{h,*}) \Big)_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,.$$

This proves (5.17) when $\iota = 0$.

The expression (5.17) is like those encountered in the energy method for quasilinear symmetric hyperbolic systems. The crucial fact is that the smooth matrix valued functions Λ_μ have hermitian symmetric values. An integration by parts shows that the leading term in the real part of $I_{2,\mu}$ vanishes. Estimating the commutator $[\Lambda_\mu, \delta^h]$, the Lipshitz norm of Λ_μ is needed which explains the lower bound $(2d+3)/2$ on s. In total, the Gagliardo-Nirenberg inequalities are used to prove $\Re\, I_{2,\mu} \le \text{r.h.s. } (5.8)$. This together with (5.16) completes the proof of (5.8) and therefore the existence part of Lemma 5.2.

Step 3. Sketch of uniqueness proof.

It suffices to show that the $L^2(\mathbb{R}^{2d}\times\mathbb{T})$ norm of the difference of two solutions satisfies a differential inequality

$$\frac{d\, \|w_1 - w_2\|^2_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}}{dt} \le K(\|\,(w_1, w_2)(T)\,\|_{H^{\underline{s}}})\, \|w_1 - w_2\|^2_{L^2(\mathbb{R}^{2d}\times\mathbb{T})}\,.$$

This is done by subtracting the two equations and using multipliers analogous to (5.12) but without the operator $1-\Delta$. The details are left to the reader. ∎

End of proof of Theorem 5.1.

Once w is constructed on $[0,T_*[\,\times\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T}$, a_0 is uniquely determined in $[0,T_*[\,\times\mathbb{R}^d_Y\times\mathbb{R}_t\times\mathbb{R}^d_y\times\mathbb{T}$ by solving the linear initial value problems determined from (2.21) and (2.26), namely

$$V(\partial_x)\,a_0^*=0\,,\qquad a_0^*|_{t=0}=w^*\,,\tag{5.19}$$

$$L(0,\partial_x)\,\underline{a}_0=0\,,\qquad \underline{a}_0|_{t=0}=\underline{w}\,,\tag{5.20}$$

The resulting function a_0 satisfies (5.1) and (5.2) as well as the constraints $\pi\,a_0^*=a_0^*$. This proves uniqueness of the solution a_0, since the values of w were uniquely determined. To complete the proof of Theorem 5.1, it suffices to show that equations (4.27), (4.28), and (4.38) are satisfied when $t\neq 0$.

To verify these three equations it is sufficient to show that

$$L(0,\partial_x)\Big(\text{l.h.s. of (4.27) and (4.28)}\Big)=0\,,\qquad V(\partial_x)\Big(\text{l.h.s. of (4.38)}\Big)=0\,.\tag{5.21}$$

To see that this suffices note that for each $\underline{T}\in[0,T_*[$ the restriction of the left hand side of (4.27) (resp (4.28), and (4.38)) to $\{T=\underline{T}\}$ is is annihilated by a first order symmetric hyperbolic operator with t timelike, and has vanishing initial data at $t=0$. It follows that they vanishe at all points of $\{T=\underline{T}\}$.

The verifications of the three equations in (5.21) rely on the identities

$$L(0,\partial_x)\,E_\alpha(D_y)=E_\alpha(D_y)\,L(0,\partial_x)=E_\alpha(D_y)\,\big(\,\partial_t+\tau_\alpha(D_y)\,\big)\,.\tag{5.22}$$

For (4.27) compute

$$\begin{aligned}L(0,\partial_x)\,(\text{l.h.s. of (4.27)})=&L(0,\partial_x)\,E_1(D_y)\,\Bigg(\,L(0,\partial_X)\underline{a}_{0,1}\\&+\Big\langle\,\kappa(1)\Phi(\underline{a}_{0,1}+\iota\,a_0^*)+\iota\sum_\mu\beta_\mu\Lambda_\mu(a_0^*+\underline{a}_{0,1})\,\partial_\theta a_0^*\,\Big\rangle\Bigg)\,.\end{aligned}\tag{5.23}$$

The right hand side of (5.23) is equal to

$$E_1(D_y)\,\big(\,\partial_t+\tau_1(D_y)\,\big)\,\Bigg(\,L(0,\partial_X)\,\underline{a}_{0,1}+\Big\langle\,\kappa(1)\Phi(\underline{a}_{0,1}+\iota\,a_0^*)+\iota\sum_\mu\beta_\mu\Lambda_\mu(a_0^*+\underline{a}_{0,1})\,\partial_\theta a_0^*\,\Big\rangle\Bigg)\,.$$

When $\iota=1$, $\partial_y+\tau_1(D_y)=\partial_t+\mathbf{v}_1.\partial_y=V(\partial_x)$ and so $(\partial_t+\mathbf{v}_1.\partial_y)\,a_0^*=(\partial_t+\mathbf{v}_1.\partial_y)\,a_{0,1}=0$. It follows that the right hand side of (5.23) vanishes. When $\iota=0$, the right hand side simplifies to

$$L(0,\partial_x)\,E_1(D_y)\,\Big(\,L(0,\partial_X)\underline{a}_{0,1}+\kappa(1)\langle\Phi(\underline{a}_{0,1})\rangle\,\Big)\,.\tag{5.24$_1$}$$

For (4.28) one has $\alpha\neq 1$ and

$$L(0,\partial_x)\,\big(\text{l.h.s. of (4.28)}\big)=L(0,\partial_x)\,E_1(D_y)\,\Big(\,L(0,\partial_X)\underline{a}_{0,\alpha}+\kappa(\alpha)\Phi(\underline{a}_{0,\alpha})\,\Big)\,.\tag{5.24$_2$}$$

The first term on the right of (5.24) is equal to

$$E_1(D_y)\,(\partial_t+\tau_\alpha(D_y))\,L(0,\partial_X)\,\underline{a}_{0,\alpha} = E_1(D_y)\,L(0,\partial_X)\,(\partial_t+\tau_\alpha(D_y))\,\underline{a}_{0,\alpha} \;=\; 0\,.$$

The second term can be nonzero only if $\alpha\in\mathcal{A}_f$, in which case $\partial_t+\tau_\alpha(D_y)=\partial_t+\mathbf{v}_\alpha.\partial_y$, and therefore $(\partial_t+\tau_\alpha(D_y))\underline{a}_{0,\alpha}=0$ implies that $(\partial_t+\tau_\alpha(D_y))\Phi(\underline{a}_\alpha)=0$ proving that (5.24) vanishes.

Reasoning as above shows that $V(\partial_x)$ annihilates every term on the left hand side of (4.38). This verifies the final assertion in (5.21) and completes the proof of Theorem 5.1. ∎

§5.2. The correctors a_1 and a_2.

Recall from the summary at the end of §4, that $(I-\pi)a_1^*$ is given by (2.23) while $\underline{a}_1$ and $\pi\,a_1^*$ are defined by solving the hyperbolic initial value problems (2.27) and (4.30) with time like variable t and initial values equal to zero at $t=0$. The corrector $a_2=(I-\pi)\,a_2^*$ is given by formula (2.28).

Proposition 5.3. *The first corrector a_1 grows sublinearly in the sense that for all $\underline{T}\in\,]0,T_*[$ and $\gamma\in\mathbb{N}^{2d+3}$,*

$$\lim_{t\to\infty}\quad\frac{1}{t}\,\|\,\partial^\gamma_{X,x,\theta}\,a_1(T,Y,t,y,\theta)\,\|_{L^2([0,\underline{T}]\times\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})}=0\,. \tag{5.25}$$

Proof. We must prove estimate (5.25) for each of the three parts, $\underline{a}_1$, $\pi(\beta)\,a_1^*$, and $(I-\pi(\beta))a_1^*$. Formula (2.23) together with (5.1) immediately yield the estimate for $(I-\pi(\beta))a_1^*$. It suffices to treat the other two.

Step 1. Estimate (5.25) for $\underline{\mathbf{a}}_1$. Define $\underline{a}_{0,\alpha}:=\pi_\alpha(D_y)\,\underline{a}_0$. Then define ρ by

$$\underline{a}_0 \;=\; \sum_{\alpha\in\mathcal{A}_f}\underline{a}_{0,\alpha}(x-\mathbf{v}_\alpha t)\;+\;\rho(T,Y,t,y)\,.$$

Estimate (5.1) for a_0 shows that $\underline{a}_{0,\alpha}$ and ρ satisfy estimates analogous to (5.1). Theorem 3.3 implies that

$$\lim_{t\to\infty}\;\|\,\rho(T,Y,t,y)\,\|_{L^\infty([0,\underline{T}]\times\mathbb{R}^d_Y\times\mathbb{R}^d_y)}\;=\;0\,.$$

Using (2.26) for a_0^* one can apply the proof of Proposition 4.1 to find the following quantitative version of (4.18). For all $\underline{T}<T_*$, and $\gamma\in\mathbb{N}^{2d+3}$

$$\lim_{t\to\infty}\;\Big\|\,\partial^\alpha_{X,x,\theta}\Big(\Phi(a_0)+\sum\beta_\mu\Lambda_\mu(a_0)\partial_\theta a_0-\mathcal{N}(a_0,\partial_\theta a_0)\Big)\Big\|_{L^2([0,\underline{T}]\times\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})}=0\,. \tag{5.26}$$

The quantity estimated to be small here, serves as an error term below so denote

$$Err:=\Phi(a_0)+\sum\beta_\mu\Lambda_\mu(a_0)\partial_\theta a_0-\mathcal{N}(a_0,\partial_\theta a_0)\,.$$

Equation (4.20) reads

$$\underline{a}_1 = L(0,\partial_x)^{-1}\Big(- \sum_{\alpha\in\mathcal{A}} L(0,\partial_X)\,\underline{a}_{0,\alpha} - \Big\langle \mathcal{N}(a_0,\partial_\theta a_0)\Big\rangle\Big) + L(0,\partial_x)^{-1}\,(Err)\,.$$

With (5.26) this yields the quantitative version of (4.20)

$$\lim_{t\to\infty}\frac{1}{t}\bigg\|\,\partial^\alpha_{X,x,\theta}\Big(\underline{a}_1 \tag{5.27}$$

$$+ L(0,\partial_x)^{-1}\Big(\sum_{\alpha\in\mathcal{A}} L(0,\partial_X)\,\underline{a}_{0,\alpha} + \langle\mathcal{N}(a_0,\partial_\theta a_0)\,\rangle\Big)\Big)\bigg\|_{L^2([0,\underline{T}]\times\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})} = 0\,.$$

Using the notation (4.22-4.25), decompose

$$L^{-1}S = L^{-1}S_* + \sum_\alpha L^{-1}S_\alpha := v_*(T,Y,y,\theta) + \sum_\alpha v_\alpha(T,Y,t,y)\,.$$

For each T fixed, S_*, S_α are continuous functions of t with values in $H^s(\mathbb{R}^d_Y\times\mathbb{R}^d_y)$ and satisfy (4.26). The profile equations guarantee that $E_\alpha(D_y)S_\alpha = 0$, and $\iota E_1(D_y)S_* = 0$. Viewing $L(0,\partial_x)$ and $\partial_t - i\tau_\alpha(D_y)$ as differential operators in t, Y, y, and applying Main Lemma 3.5 yields a quantitative version of (4.27)

$$\lim_{t\to\infty}\ \frac{1}{t}\,\|\, v_*(T,Y,t,y,\theta)\,,\ v_\alpha(T,Y,t,y)\,\|_{L^2(\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})}\ =\ 0\,.$$

Nearly the same reasoning applies to $L^{-1}(\partial^\gamma_{X,x}\{S_*,S_\alpha\})$ showing that

$$\lim_{t\to\infty}\ \frac{1}{t}\,\|\, L^{-1}(\partial^\gamma_{X,x}\{S_*,S_\alpha\})\,\|_{L^2(\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})}\ =\ 0\,. \tag{5.28}$$

In this derivation, the Main Lemma is applied with source terms f whose initial values at $t=0$ are

$$f_0 := \{\partial^\gamma_{X,x}v_*(T,Y,0,y)\,,\ \partial^\gamma_{X,x}v_\alpha(T,Y,0,y)\}\,. \tag{5.29}$$

These are continuous functions of T with values in $L^2(\mathbb{R}^{2d})$ so their values lie in a compact subset of L^2 for $0\le T\le \underline{T}$. It follows that the convergence in (5.28) is uniform in T. Thus,

$$\lim_{t\to\infty}\ \frac{1}{t}\,\|\, L^{-1}(\partial^\gamma_{X,x}\{S_*,S_\alpha\})\,\|_{L^2([0,\underline{T}[\times\mathbb{R}^{2d}_{Y,y}\times\mathbb{T})}\ =\ 0 \tag{5.30}$$

The difference $w := \partial^\gamma_{X,x}\{v_*,v_\alpha\} - L^{-1}(\partial^\gamma_{X,x}\{S_*,S_\alpha\})$ is the solution of $Lw=0$ whose initial data agrees with those of $\partial^\gamma_{X,x}\{v_*,v_\alpha\}$. Thus the $L^2([0,\underline{T}]\times\mathbb{R}^{2d})$ norm of w is bounded independent of t. Together with (5.30) this shows that

$$\lim_{t\to\infty}\ \frac{1}{t}\,\Big\|\,\partial^\gamma_{X,x}\{v_*(T,Y,t,y)\,,\ v_\alpha(T,Y,t,y)\}\Big\|_{L^2([0,\underline{T}]\times\mathbb{R}^{2d}_{Y,y}\times\mathbb{T})}\ =\ 0\,. \tag{5.31}$$

This together with (5.27) yields (5.25) for $\underline{a}_1$.

Step 2. Estimate (5.25) for $\pi(\beta)\,\mathbf{a}_1^*$. The profile equation (4.38) for a_0 is derived exactly so that

$$\pi(\beta)\,a_1^* = V(\partial_x)^{-1}\left(\pi(\beta)\Big(\Phi(a_0) + \sum \beta_\mu \Lambda_\mu(a_0)\,\partial_\theta a_0 - \pi(\beta)\,\mathcal{N}(a_0, \partial_\theta a_0)\Big)^*\right).$$

That this is sublinear follows from the quantitative version of (4.18) as in the proof of (5.31). ∎

§5.3. Estimate for the residual.

Once profiles are defined, the approximate solution is given by (2.6) and (2.7). The residual is computed in (2.10)-(2.11). In the present case, $r_0 = r_1 = r_2 = 0$. Our estimates for the profiles a_0, a_1, a_2 yield estimates for the the residual.

Proposition 5.4. *With the profiles constructed in this section the residual*

$$k(\varepsilon, X, x, \theta) := L\big(\varepsilon^p a\,,\, \varepsilon\partial_X + \partial_x + \frac{\beta}{\varepsilon}\partial_\theta\big)\,\varepsilon^p a + F(\varepsilon^p a) \tag{5.32}$$

satisfies the following estimates. For all $\gamma \in \mathbb{N}^{2d+3}$ and $\underline{T} < T_*$,

$$\sup_{0\le t<\underline{T}/\varepsilon} \big\|\,\partial^\gamma_{X,x,\theta}\,k(\varepsilon, X, x, \theta)\big\|_{L^2([0,\underline{T}]\times\mathbb{R}^{2d}_{Y,y}\times\mathbb{T})} = o(\varepsilon^{p+1}) \tag{5.33}$$

as $\varepsilon \to 0$.

Proof. Since $r_0 = r_1 = r_2 = 0$, the calculations of Lemma 2.1 show that

$$\begin{aligned} k(\varepsilon, X, x, \theta) :=& L(\varepsilon^p a, \varepsilon\partial_X + \partial_x)\,\varepsilon^{2+p} a_2 + L(\varepsilon^p a, \varepsilon\partial_X)\,\varepsilon^{1+p} a_1 \\ &+ \Big(\,F(\varepsilon^p\,a) - \Phi\big(\varepsilon^p\,a_0(X, x, \theta)\big)\Big) \\ &+ \sum_\mu \Big(\,A_\mu(\varepsilon^p\,a) - \sum_\mu A_\mu(0) - \Lambda_\mu\big(\varepsilon^p\,a_0(X, x, \theta)\big)\Big)\,\frac{\beta_\mu}{\varepsilon}\,\frac{\partial\,\varepsilon^p a_0}{\partial\theta}\,. \end{aligned} \tag{5.34}$$

Estimates (5.1) and (5.25) imply that

$$\sup_{0\le t\le\underline{T}/\varepsilon} \Big\|\,\partial^\gamma_{X,x,\theta}\,\varepsilon^j a_j\,\Big\|_{L^2([0,\underline{T}]\times\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})} \le \begin{cases} C & \text{for} \quad j = 0 \\ o(1) & \text{for} \quad j = 1, 2 \end{cases} \tag{5.35}$$

as $\varepsilon \to 0$. This suffices to shows that the first two terms in (5.34) contribute terms $o(\varepsilon^{p+1})$ to the left hand side of (5.33).

To treat the last two expressions on the right of (5.34) use Taylor's Theorem together with the definition of p in (2.5) There are smooth functions $H_\nu, M_{\kappa,\mu}$ and B_ρ such that

$$
\begin{aligned}
F(u) &= \Phi(u) + \sum_{|\nu|=J+1} (\Re u, \Im u)^\nu \, H_\nu(u) \\
F(u) - F(v) &= \sum_{|\rho|=J-1} B_\rho(u,v)(\Re u, \Im u, \Re v, \Im v)^\rho \, (u-v)
\end{aligned}
\tag{5.36}
$$

and

$$
A_\mu(u) - \sum_\mu A_\mu(0) - \Lambda_\mu(u) = \sum_{|\kappa|=K+1} (\Re u, \Im u)^\kappa M_{\kappa,\mu}(u)\,. \tag{5.37}
$$

Write

$$
F(\varepsilon^p\, a) - \Phi\big(\varepsilon^p\, a_0(X,x,\theta)\big) = \Big(F(\varepsilon^p\, a) - F(\varepsilon^p\, a_0)\Big) + \Big(F(\varepsilon^p\, a_0) - \Phi\big(\varepsilon^p\, a_0\big)\Big)\,. \tag{5.38}
$$

The first summand on the right of (5.38) is a finite sum of terms

$$
\varepsilon^{p+p|\rho|} B(\varepsilon^p a, \varepsilon^p a_0)\, (\Re a, \Im a, \Re a_0 \Im a_0)^\rho\, (\varepsilon a_1 + \varepsilon^2 a_2)\,. \tag{5.39}
$$

Formula (2.5) shows that $p|\rho| = p(J-1) \geq 1$. Denote by $|||\ \ |||_s$ the norm

$$
|||\ \cdot\ |||_s := \sum_{j\leq s}\ \sup_{0\leq t\leq \underline{T}/\varepsilon}\ \Big\|\, \partial_t^j\ \cdot\ \Big\|_{H^s([0,\underline{T}]\times\mathbb{R}^d_Y\times\mathbb{R}^d_y\times\mathbb{T})}\,. \tag{5.40}
$$

A variant of Schauder's Lemma shows that for $s > (2d+1)/2$ and smooth G with $G(0) = 0$,

$$
|||\, G(w)\, |||_s \leq C(|||w|||_s)\, |||w|||_s\,. \tag{5.41}
$$

Estimate (5.35) shows that the last factor in (5.39) is $o(1)$ in $|||\ \ |||_s$. This together with the Schauder lemma shows that the summands (5.39) make a contribution $o(\varepsilon^{p+1})$ to the left hand side of (5.33)

The second summand on the right of (5.38) is a finite sum of terms

$$
\varepsilon^{p|\nu|}\, (\Re a_0, \Im a_0)^\nu\, H(\varepsilon^p a_0)\,. \tag{5.42}
$$

Formula (2.5) shows that $p|\nu| = p(J+1) = p(J-1) + 2p \geq 1 + 2p$ and (5.35) shows that this too is $o(\varepsilon^{p+1})$.

Finally using (5.37), the quasilinear term in (5.34) is a finite sum of terms

$$
\varepsilon^{p|\kappa|}\, M_{\kappa,\mu}(\varepsilon^p a)\, (\Re a, \Im a)^\kappa\, \frac{\beta_\mu}{\varepsilon}\, \frac{\partial\, \varepsilon^p a_0}{\partial\theta}\,. \tag{5.44}
$$

Formula (2.5) shows that $p|\kappa| = p(K+1) = p(K-1) + 2p \geq 2 + 2p$. One power of ε is sacrificed to the β/ε factor but the remaining ε^{1+2p} suffices to show that the last term contributes $o(\varepsilon^{1+p})$ to the left hand side of (5.34) and the proof is complete. ∎

§6. Convergence and stability.

In the last section we constructed a family u^ε of approximate solutions to (2.1) with wavelengths of order ε and amplitude of order ε^p where p given in (2.5) is determined from the nonlinear terms in such a way that the nonlinearities play an essential role for times $t \sim 1/\varepsilon$. The approximate solution is constructed on $[0, T_*/\varepsilon[\times \mathbb{R}^d$.

In this section we prove that as $\varepsilon \to 0$, exact solutions of (2.1) with initial values converging to those of the approximate solution exist for times T/ε for any $T < T_*$ and the relative error in the approximate solutions tends to zero as ε tends to zero.

Theorem 6.1. *Suppose that $\beta = (\tau, \eta)$ and $u^\varepsilon = \varepsilon^p\, a(\varepsilon, \varepsilon x, x, \beta.x/\varepsilon)$ is a family of approximate solutions with profile $a = a_0 + a_1 + a_2$ constructed in the last section. Let v^ε be the exact solution of the initial value problem*

$$L(v^\varepsilon, \partial_x)\, v^\varepsilon + F(v^\varepsilon) = \varepsilon^{p+1}\, \ell^\varepsilon(x, y.\eta/\varepsilon)\,, \qquad v^\varepsilon|_{t=0} = u^\varepsilon|_{t=0} + \varepsilon^p\, g^\varepsilon(y, y.\eta/\varepsilon)\,, \tag{6.1}$$

where $g^\varepsilon \to 0$ and $\ell^\varepsilon \to 0$ in the sense that for all $s \in \mathbb{R}$ and $j \in \mathbb{N}$,

$$\|\, g^\varepsilon \,\|_{H^s(\mathbb{R}^d\times\mathbb{T})} \to 0\,, \qquad \sup_{0\le t\le T_*/\varepsilon} \|\, (\varepsilon\partial_t)^j \ell^\varepsilon(t) \,\|_{H^s(\mathbb{R}^d\times\mathbb{T})} \to 0\,.$$

i. *For each $T \in]0, T_*[$ there is an $\varepsilon_0 > 0$ so that for all $\varepsilon \in]0, \varepsilon_0]$, v^ε exists and is smooth on $[0, T/\varepsilon] \times \mathbb{R}^d$.*

ii. *There are functions $\mathcal{U}^\varepsilon, \mathcal{V}^\varepsilon \in \cap_s\, C^\infty\big([0, T/\varepsilon]\ ;\ H^s(\mathbb{R}^d \times \mathbb{T})\big)$ with*

$$v^\varepsilon = \varepsilon^p\, \mathcal{V}^\varepsilon(t, y, y.\eta/\varepsilon)\,, \qquad u^\varepsilon = \varepsilon^p\, \mathcal{U}^\varepsilon(t, y, y.\eta/\varepsilon)\,,$$

and for all s

$$\lim_{\varepsilon\to 0}\ \sup_{0\le t\le T/\varepsilon}\ \|\, \mathcal{U}^\varepsilon(t) - \mathcal{V}^\varepsilon(t) \,\|_{H^s(\mathbb{R}^d\times\mathbb{T})} = 0\,.$$

In particular, for all α

$$\sup_{0\le t\le T/\varepsilon}\ \|\, (\varepsilon\partial_x)^\alpha (v^\varepsilon(t) - u^\varepsilon(t)) \,\|_{L^2(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)} = o(\varepsilon^p)\,.$$

Theorem 6.1 shows that the approximate solution is stable. In particular, replacing the small residual by zero and perturbing the initial data by suitable small terms yields exact solutions which are close to the approximate solution.

Example. Taking initial values $v^\varepsilon(0, y) = \varepsilon^p\, a_0(0, \varepsilon y, y, y.\eta/\varepsilon)$ from the leading term only, corresponds to taking $g^\varepsilon(y, \theta) = \varepsilon\, a_1(0, \varepsilon y, y, \theta) + \varepsilon^2\, a_2(0, \varepsilon y, \theta)$ which satisfies the hypotheses of the theorem. Thus knowing only a_0 one has an exact solution v^ε so that $v^\varepsilon - \varepsilon^p a_0(\varepsilon X, x, x.\beta/\varepsilon) = o(\varepsilon^p)$. This is the case described in Theorem 1.2. ∎

The proof starts by first transforming the problem. The approximate solution has initial data equal to

$$\varepsilon^p\Big(\,a_0(0,Y,0,y,y.\eta/\varepsilon)+\varepsilon\,a_1(0,Y,0,y,y.\eta/\varepsilon)+\varepsilon^2\,a_2(0,Y,0,y,y.\eta/\varepsilon)\Big)\,.$$

The strategy, as in [JMR2], is to find the exact solution in the form

$$v^\varepsilon=\varepsilon^p\,\mathcal{V}^\varepsilon(t,y,y.\eta/\varepsilon)\,, \tag{6.2}$$

where for all ε,t,

$$\mathcal{V}^\varepsilon(t,y,\theta)\ \in\ \cap_s\,H^s(\mathbb{R}^d_y\times\mathbb{T})\,, \tag{6.3}$$

The approximate solution has this form,

$$u^\varepsilon=\varepsilon^p\,\mathcal{U}^\varepsilon(t,y,y.\eta/\varepsilon)\,, \tag{6.4}$$

where with a from (2.6) and (2.7)

$$\begin{aligned}\mathcal{U}^\varepsilon(t,y,\theta)&:=a(\varepsilon,\varepsilon x,x,\frac{\tau t}{\varepsilon}+\theta)\\&=a_0(\varepsilon t,\varepsilon y,t,y,\frac{\tau t}{\varepsilon}+\theta)+\varepsilon\,a_1(\varepsilon t,\varepsilon y,t,y,\frac{\tau t}{\varepsilon}+\theta)+\varepsilon^2\,a_2(\varepsilon t,\varepsilon y,t,y,\frac{\tau t}{\varepsilon}+\theta)\,.\end{aligned} \tag{6.5}$$

Note that $\mathcal{U}^\varepsilon$ is a continuous function of t with values in $H^s(\mathbb{R}^d\times\mathbb{T})$ which is uniformly bounded for times $t\sim 1/\varepsilon$ but its derivatives with respect to time explode like $1/\varepsilon$ even for small times. The solution $\mathcal{V}^\varepsilon$ is constructed continuous in time with values in H^s with the expectation that its time derivatives are also $\sim 1/\varepsilon$.
In order for v^ε from (6.2) to satisfy (6.1) it is sufficient that

$$\begin{aligned}&L(\,\varepsilon^p\,\mathcal{V}^\varepsilon,\,\partial_t,\,\partial_y+\frac{\eta}{\varepsilon}\,\partial_\theta\,)\,\varepsilon^p\mathcal{V}^\varepsilon+F(\varepsilon^p\,\mathcal{V}^\varepsilon)=\varepsilon^{p+1}\,\ell^\varepsilon(t,y,\theta)\,,\qquad\text{for all}\quad t,y,\theta\,,\\&\mathcal{V}^\varepsilon(0,y,\theta)=\mathcal{U}^\varepsilon(0,y,\theta)+g^\varepsilon(y)\,.\end{aligned} \tag{6.6}$$

The construction of the profiles shows that $\mathcal{U}^\varepsilon$ defined in (6.5) is an approximate solution of (6.6). Equations (2.10) and (2.11) show that

$$L(\,\varepsilon^p\mathcal{U}^\varepsilon,\,\partial_t,\,\partial_y+\frac{\eta}{\varepsilon}\,\partial_\theta\,)\,\varepsilon^p\mathcal{U}^\varepsilon+F(\varepsilon^p\mathcal{U}^\varepsilon)=k(\varepsilon,\varepsilon x,x,t\tau/\varepsilon+\theta)$$

with k defined in (5.33). Thus

$$L(\,\varepsilon^p\mathcal{U}^\varepsilon,\,\partial_t,\partial_y+\frac{\eta}{\varepsilon}\,\partial_\theta\,)\,\varepsilon^p\mathcal{U}^\varepsilon+F(\varepsilon^p\,\mathcal{U}^\varepsilon)=\varepsilon^{p+1}\,H^\varepsilon(t,y,\theta)\,, \tag{6.7}$$

with

$$H^\varepsilon(t,y,\theta):=\varepsilon^{-p-1}\,k(\varepsilon,\varepsilon t,\varepsilon y,t,y,\frac{\tau t}{\varepsilon}+\theta)\,. \tag{6.8}$$

The strategy is to construct the solution $\mathcal{V}^\varepsilon$ as a perturbation of $\mathcal{U}^\varepsilon$,

$$\mathcal{V}^\varepsilon = \mathcal{U}^\varepsilon + \mathcal{W}^\varepsilon . \tag{6.9}$$

It suffices to show that $\mathcal{W}^\varepsilon$ exists for $0 \le t \le T/\varepsilon$ and converges to zero uniformly on that interval.

The first step is to show that a combination of the powers of ε^p and the hypothesis on the orders of the nonlinearities, simplifies (6.6) and (6.7) and shows why one expects such an approximation result for times $t \sim 1/\varepsilon$.

The order of nonlinearity hypothesis as in the proof of Lemma 2.1, implies that

$$A_\mu(\varepsilon^p\,\mathcal{V}) - A_\mu(0) = \varepsilon^2\, B_\mu(\varepsilon, \mathcal{V})\,, \qquad F(\varepsilon^p\,\mathcal{V}) = \varepsilon^{p+1}\, G(\varepsilon, \mathcal{V})$$

where B_μ and G are smooth functions of $\mathcal{V} \in \mathbb{C}^d$, G taking values in $\mathbb{C}^d$ and B_μ taking values in the hermitian symmetric $N \times N$ matrices. Both depend smoothly on ε^p and therefore continuously on $\varepsilon \in [0, 1]$. Define a family of quasilinear symmetric hyperbolic operators by

$$B(\varepsilon, \mathcal{V}, \partial)\,\mathcal{V} := \sum_{\mu=0}^{d} B_\mu(\varepsilon, \mathcal{V})\,\partial_\mu \mathcal{V}\,.$$

After division by ε^p, the equation for $\mathcal{V}^\varepsilon$ takes the form

$$L(0, \partial_t,\, \partial_y + \frac{\eta}{\varepsilon}\,\partial_\theta)\,\mathcal{V}^\varepsilon + \varepsilon\, B(\varepsilon, \mathcal{V}^\varepsilon,\, \varepsilon\partial_t, \varepsilon\partial_y + \eta\,\partial_\theta\,)\,\mathcal{V}^\varepsilon + \varepsilon\, G(\varepsilon, \mathcal{V}^\varepsilon) = \varepsilon\,\ell^\varepsilon(t, y, \theta)\,. \tag{6.10}$$

In the first term of (6.10), $L(0, \partial_t, \partial_y + \frac{\eta}{\varepsilon}\,\partial_\theta)$ is a singular constant coefficient symmetric hyperbolic operator, and the second and third terms have small coefficients. The approximate solution $\mathcal{U}^\varepsilon$ satisfies

$$L(0, \partial_t, \partial_y + \frac{\eta}{\varepsilon}\,\partial_\theta)\,\mathcal{U}^\varepsilon + \varepsilon\, B(\varepsilon, \mathcal{U}^\varepsilon,\, \varepsilon\partial_t, \varepsilon\partial_y + \eta\,\partial_\theta\,)\,\mathcal{U}^\varepsilon + \varepsilon\, G(\varepsilon, \mathcal{U}^\varepsilon) = \varepsilon\, H^\varepsilon(t, y, \theta)\,. \tag{6.11}$$

We show that $\mathcal{U}^\varepsilon$ and $\mathcal{V}^\varepsilon$ stay close for times $t \sim 1/\varepsilon$. There are two singularities in the problem. First there are the operators in $\varepsilon^{-1}\partial_\theta$, and second we are interested in times $t \sim 1/\varepsilon$. The fact that there are two distinct singularities is because we have a three scale problem rather than a two scale problem as in standard geometric optics. Note that the factors ε^p have been removed and the solutions $\mathcal{U}^\varepsilon$ and $\mathcal{V}^\varepsilon$ are $O(1)$.

At this stage the argument is made clearer by passing to a more general framework. In addition results with several angle variables θ are needed to treat the multiphase case. The key structure in equation (6.10) is that the zero order term term G and all the quasilinear terms have a factor ε in front of them and the quasilinear term in ∂_t has an ε^2 in front. In fact all the derivatives in x have such a factor but it is only for the ∂_t term that it is needed because the derivatives in time are typically larger by a factor $1/\varepsilon$.

The spatial variables and angular variables play identical roles in the analysis to follow even though they arise in different ways. With this in mind, introduce the variable

$$z := (z_1, z_2, \ldots, z_{d+m}) := (y, \theta) \in \mathbb{R}^d \times \mathbb{T}^m\,.$$

The spacetime variable is still denoted

$$x := (t, z) = (x_0, x_1, \ldots, x_{d+m})\,.$$

The index μ runs from 0 to $d + m$, and the index j runs from 1 to $d + m$.

Hypotheses. *Suppose that*

$$\mathcal{M}^\varepsilon(\partial_x) = \mathcal{M}^\varepsilon(\partial_{t,z}) := A_0\,\frac{\partial}{\partial t} + \sum_{j=1}^{d+m} \Big(A_j + \frac{1}{\epsilon}\,\tilde{A}_j \Big) \frac{\partial}{\partial z_j} + \mathcal{L}_0 \tag{6.12}$$

is a singular family of constant coefficient dissipative linear symmetric hyperbolic operators on $\mathbb{R}^{1+d} \times \mathbb{T}^m$ with time variable $t = x_0$. Here symmetry means that the coefficients A_μ and $\tilde{A}_\mu$ are hermitian symmetric with A_0 strictly positive definite. Dissipativity means that the constant matrix $\mathcal{L}_0$ satisfies

$$\mathcal{L}_0 + \mathcal{L}_0^* \geq 0\,. \tag{6.13}$$

Suppose that for $0 \le \mu \le d+m$ and $\varepsilon \in [0,1]$ $\mathcal{B}_\mu(\varepsilon, V)$ is a smooth hermitian symmetric matrix valued function on $\mathbb{C}^N$ and $G(\varepsilon, V)$ is smooth with values in $\mathbb{C}^N$. The partial derivatives of these functions with respect to $\Re V, \Im V$ are assumed to be continuous on $[0,1] \times \mathbb{C}^N$. Introduce the singular family of quasilinear symmetric hyperbolic operators

$$\mathcal{L}^\varepsilon(V)\,V := \mathcal{M}^\varepsilon(\partial_{t,z})\,V + \varepsilon^2 \mathcal{B}_0(\varepsilon, V)\,\frac{\partial V}{\partial t} + \varepsilon \sum_{j=1}^{d+m} \mathcal{B}_j(\varepsilon, V) \frac{\partial V}{\partial z_j} + \varepsilon G(\varepsilon, V)\,. \tag{6.14}$$

With $T > 0$ a family of solutions

$$U^\varepsilon \in C^\infty([0, T/\varepsilon] \times \mathbb{R}^d \times \mathbb{T}^m)) \tag{6.15}$$

with residuals

$$\mathcal{L}^\varepsilon(U^\varepsilon)\,U^\varepsilon := R^\varepsilon(t, y, \theta) \tag{6.16}$$

is given. Suppose that the family U^ε is bounded for times of order $1/\varepsilon$ in the sense that for all s,

$$\sup_{0\le\varepsilon\le\varepsilon_0} \ \sup_{t\in[0,T/\varepsilon]} \ \big\|\, \varepsilon\partial_t U^\varepsilon(t)\,,\, U^\varepsilon(t)\,\big\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \ < \ \infty \tag{6.17}$$

The next theorem shows that U^ε is stable for times of order $1/\varepsilon$ under perturbations of R^ε and initial data. Note that there are strong hypotheses on the size of U^ε but other than that no hypotheses on the size of R^ε. The hypotheses are only on the size of U^ε and and the size of the perturbations.

Stability Theorem 6.2. *Suppose that $\mathcal{L}^\varepsilon$ is as above and that a family of solutions U^ε satisfies (6.15)-(6.17). Suppose that $T > 0$ and that for $0 < \varepsilon \le 1$, $g^\varepsilon \in \cap_s H^s(\mathbb{R}^d \times \mathbb{T}^m)$ and $K^\varepsilon \in \cap_s L^1\big([0, T/\varepsilon]\ ;\ H^s(\mathbb{R}^d \times \mathbb{T}^m)\big)$ are $o(1)$ in the sense that for all s*

$$\lim_{\varepsilon\to 0} \|\, g^\varepsilon(y,\theta)\,\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} = 0\,, \qquad \text{and} \qquad \lim_{\varepsilon\to 0} \int_0^{T/\varepsilon} \|\, K^\varepsilon(\sigma)\,\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, d\sigma \ = 0\,. \tag{6.18}$$

Define $V^\varepsilon \in \cap_s C([0,T_*(\varepsilon)[\ ; H^s(\mathbb{R}^d\times\mathbb{T}^m))$ to be the unique maximal solutions of the initial value problems

$$\mathcal{L}^\varepsilon(V^\varepsilon)\,V^\varepsilon = R^\varepsilon + K^\varepsilon\,, \tag{6.19}$$

$$V^\varepsilon\big|_{t=0} \;=\; U^\varepsilon\big|_{t=0} + g^\varepsilon\,. \tag{6.20}$$

Then, there is a $0<\varepsilon_1\le 1$ so that for $0<\varepsilon\le\varepsilon_1$, $T_*(\varepsilon)\ge T/\varepsilon$, and for all s,k

$$\lim_{\varepsilon\to 0}\quad \sup_{0\le t\le T/\varepsilon} \big\|\,(\varepsilon\partial_t)^k\big(V^\varepsilon(t)-U^\varepsilon(t)\big)\,\big\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \;=\; 0\,. \tag{6.21}$$

Proof of Theorem 6.1 assuming Theorem 6.2. Apply Stability Theorem 6.2 to (6.10) and (6.11) with $K^\varepsilon=\varepsilon\,H^\varepsilon$, $R^\varepsilon(t,Y,\theta)=-K^\varepsilon(t,Y,\theta)+\varepsilon\,\ell^\varepsilon(t,Y,\theta)$ and $g^\varepsilon=g^\varepsilon(y,\theta)$. That $K^\varepsilon=\varepsilon\,H^\varepsilon$ satisfies (6.18) follows from formula (6.18) and estimate (5.33). ∎

Proof of Theorem 6.2. Define $W^\varepsilon=V^\varepsilon-U^\varepsilon$. To derive an equation for the perturbation, W^ε, subtract (6.16) from (6.19). Write

$$\mathcal{B}_\mu(\varepsilon,V^\varepsilon)\,\partial_\mu V^\varepsilon-\mathcal{B}_\mu(\varepsilon,U^\varepsilon)\,\partial_\mu U^\varepsilon=\mathcal{B}_\mu(\varepsilon,V^\varepsilon)\,\partial_\mu W^\varepsilon+\big(\mathcal{B}_\mu(\varepsilon,V^\varepsilon)-\mathcal{B}_\mu(\varepsilon,U^\varepsilon)\big)\,\partial_\mu U^\varepsilon\,.$$

Use Taylor's Theorem in the last term to find an equation of the form

$$\mathcal{L}^\varepsilon(U^\varepsilon+W^\varepsilon)\,W^\varepsilon+\varepsilon\,G\big(\varepsilon,U^\varepsilon,\varepsilon\partial_t U^\varepsilon\,\partial_z U^\varepsilon,W^\varepsilon\big)\,W^\varepsilon=K^\varepsilon \tag{6.22}$$

where G is continuous in ε and smooth in its other arguments. The initial value is given by

$$W^\varepsilon\big|_{t=0}=g^\varepsilon\,. \tag{6.23}$$

Fix an integer $s>1+(d+m)/2$. The standard local existence theory for quasilinear symmetric hyperbolic systems shows that there is a unique maximal local solution

$$W^\varepsilon\in\cap_k\,C^k([0,T_*(\varepsilon)[\,;\,H^k(\mathbb{R}^d\times\mathbb{T}))$$

and if the time of existence $T_*(\varepsilon)<\infty$ then

$$\lim_{t\nearrow T_*(\varepsilon)}\ \|\,W^\varepsilon(t)\,\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}=\infty\,. \tag{6.24}$$

To prove that $T_*\ge T/\varepsilon$ it suffices to bound the H^s norm of $W^\varepsilon(t)$ for $0\le t\le T/\varepsilon$. The next Lemma is the key element in the stability proof. The important fact is that the amplification factor C is uniform for times $t\sim 1/\varepsilon$.

Lemma 6.3. $\mathbf{H^s}(\mathbb{R}^{\mathbf{d}}\times\mathbb{T}^{\mathbf{m}})$ estimate. *For $s-1>(d+m)/2$, there is a constant C so if $\varepsilon\in]0,1]$, $\underline{t}\in[0,T/\varepsilon]$, and $W\in C^1\big([0,\underline{t}]\,;\,H^{1+s}(\mathbb{R}^d\times\mathbb{T}^m)\big)$ satisfies*

$$\sup_{0\le t\le \underline{t}/\varepsilon}\ \big\|\,W(t),\partial_z W(t),\varepsilon\partial_t W(t)\,\big\|_{L^\infty(\mathbb{R}^d\times\mathbb{T}^m)}\le 1\,, \tag{6.26}$$

then for $0\le t\le\underline{t}$

$$\|W(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\le C\left(\,\|W(0)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}+\int_0^t\|\mathcal{L}^\varepsilon(U^\varepsilon+W)\,W(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\,dt\right). \tag{6.27}$$

Proof of Theorem 6.2 assuming Lemma 6.3. Equations (6.18), (6.22), and (6.23), imply that there is a $0 < \varepsilon_2 \le 1$ so that for $0 < \varepsilon < \varepsilon_2$,

$$\big\| \, W^\varepsilon(0),\, \partial_z W^\varepsilon(0),\, \varepsilon\partial_t W^\varepsilon(0) \, \big\|_{L^\infty(\mathbb{R}^d \times \mathbb{T}^m)} \le 1/2 \,,$$

For $0 < \varepsilon < \varepsilon_2$ define $0 < \underline{t}_*(\varepsilon)$ by

$$\begin{aligned} \underline{t}_*(\varepsilon) = \sup \Big\{ \, \underline{t} < \min\{T_*(\varepsilon), T/\varepsilon\} \ : \\ \sup_{0 \le t \le \underline{t}} \big\| \, W^\varepsilon(t),\, \partial_z W^\varepsilon(t),\, \varepsilon\partial_t W^\varepsilon(t) \, \big\|_{L^\infty(\mathbb{R}^d \times \mathbb{T}^m)} \le 1 \Big\} . \end{aligned} \tag{6.28}$$

Fix $\mathbb{N} \ni s \ge 1 + (d+m)/2$. Inequality (6.27) then implies that for $0 \le t \le \underline{t}_*(\varepsilon)$,

$$\begin{aligned} &\|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \le \\ &C\bigg(\|W^\varepsilon(0)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} + \int_0^t \|K^\varepsilon(t) - \varepsilon\, G(\varepsilon, U^\varepsilon, \varepsilon\partial_t U^\varepsilon, \partial_z U^\varepsilon, W^\varepsilon)\, W^\varepsilon\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, dt \bigg). \end{aligned} \tag{6.29}$$

Hypothesis (6.17) bounds U^ε and $\varepsilon\partial_t U^\varepsilon$ in $H^s(\mathbb{R}^d \times \mathbb{T}^m)$. Then Moser's inequality shows that for times no larger than $\underline{t}_*(\varepsilon)$

$$\|G(\varepsilon, U^\varepsilon, \varepsilon\partial_t U^\varepsilon, \partial_z U^\varepsilon, W^\varepsilon)\, W^\varepsilon\, \|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \ \le\ C\, \|\, W^\varepsilon\, \|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \,.$$

Then (6.29) yields

$$\begin{aligned} \|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \ \le\ C\, \Big(\, &\|W^\varepsilon(0)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} + \\ &\varepsilon \int_0^t \|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, dt + \int_0^t \|K^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, dt \Big) . \end{aligned} \tag{6.30}$$

Hypothesis (6.18) shows that as $\varepsilon \to 0$,

$$\begin{aligned} &\|W^\varepsilon(0)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} + \int_0^t \|K^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, dt \ \le \\ &\qquad \|g^\varepsilon\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} + \int_0^{T/\varepsilon} \|\, K^\varepsilon(t)\, \|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, dt \ = o(1) \,. \end{aligned}$$

Inserting this in (6.30) yields

$$\|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \ \le\ C\, \Big(\, o(1) + \varepsilon \int_0^t \|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)}\, dt \Big) . \tag{6.31}$$

Gronwall's lemma yields $\|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \ \le\ e^{C\varepsilon t}\, o(1)$. Since $\varepsilon t \le T$, this implies

$$\sup_{0 \le t \le t_*(\varepsilon)} \ \|W^\varepsilon(t)\|_{H^s(\mathbb{R}^d\times\mathbb{T}^m)} \ =\ o(1) \,. \tag{6.32}$$

Then Sobolev's Lemma shows that

$$\|W^\varepsilon(t), \partial_{z,\theta} W^\varepsilon(t)\|_{L^\infty(\mathbb{R}^d\times\mathbb{T}^m)} = o(1)\,.$$

Using equation (6.19) to express the time derivative in terms of z derivatives and using hypothesis (6.15) then shows that $\|\varepsilon\partial_t W(t)\|_{L^\infty(\mathbb{R}^d\times\mathbb{T}^m)} = o(1)$. Thus, one can choose $0<\varepsilon_3<\varepsilon_2$ so that for $0<\varepsilon<\varepsilon_3$ and $0<t\le \underline{t}_*(\varepsilon)$,

$$\sup_{0\le t\le \underline{t}_*(\varepsilon)} \big\| W^\varepsilon(t),\, \partial_z W^\varepsilon(t),\, \varepsilon\partial_t W^\varepsilon(t)\big\|_{L^\infty(\mathbb{R}^d\times\mathbb{T}^m)} \le 1/3\,.$$

It follows from the definition (6.28) that for these ε, $\underline{t}_* = T/\varepsilon$, and therefore that (6.32) holds for all $0\le t\le T/\varepsilon$. In particular $T_*(\varepsilon) > t_*(\varepsilon) = T/\varepsilon$ and

$$\sup_{0\le t\le T/\varepsilon} \| W^\varepsilon(t), \partial_{z,\theta} W^\varepsilon(t)\|_{L^\infty(\mathbb{R}^d\times\mathbb{T})} \le 1\,.$$

Next consider larger integer values s. Given the sup norm bound for $W^\varepsilon, \partial_{z,\theta}W^\varepsilon$ one can repeat the argument leading to (6.32) to shows that (6.32) holds for the new s and $0\le t\le T/\varepsilon$ which proves the desired conclusion (6.21) ∎

Proof of Lemma 6.3. Use the energy method for symmetric hyperbolic systems. There are three key ingredients in the computation. First is that one estimates $z=y,\theta$ derivatives and not t derivatives. Second, the singular terms are dissipative and have constant coefficients, so commute with ∂_z. Third the scales in ε must be carefully followed. They work but with no margin for error because p is a critical exponent.

Write

$$\mathcal{L}(U^\varepsilon + W) = \sum_{\mu=0}^{d+m} \mathrm{A}^\varepsilon_\mu(x)\,\partial_\mu + \mathcal{L}_0\,,$$

where

$$\mathrm{A}^\varepsilon_0(x) := A_0 + \varepsilon^2\, B_0\big(\varepsilon,\, U^\varepsilon(x) + W(x)\,\big)\,,$$

and for $j=1,\dots,d+m$

$$\mathrm{A}^\varepsilon_j(x) := A_j + \frac{1}{\varepsilon}\,\tilde{A}_j + \varepsilon\, B_j\big(\varepsilon,\, U^\varepsilon(x) + W(x)\,\big)\,.$$

Step 1. $\mathbf{L^2(\mathbb{R}^d\times\mathbb{T}^m)}$ energy estimate. The standard energy identity for the linear operator $\mathcal{L}(U^\varepsilon+W)$ applied to $Z\in C^1([0,\underline{t}]\,;\,H^1(\mathbb{R}^d\times\mathbb{T}^m)\,)$ reads

$$\partial_t\,(\mathrm{A}^\varepsilon_0 Z, Z)_{L^2(\mathbb{R}^d)} + \int_{\mathbb{R}^d\times\mathbb{T}^m} \Big(\big(\mathcal{L}_0 + \mathcal{L}_0^* + 2\sum_{\mu=0}^{d+m} \partial_\mu \mathrm{A}^\varepsilon_\mu\,\big)\, Z\,,\, Z\Big)\, dz = 2\,\mathrm{Re} \int_{\mathbb{R}^d\times\mathbb{T}^m} \Big(Z\,,\, \mathcal{L}(U^\varepsilon + W)\, Z\Big)\, dz\,. \tag{6.33}$$

The term on the right is estimated using the Schwartz inequality to give

$$\Big| \int_{\mathbb{R}^d\times\mathbb{T}^m} \Big(Z\,,\, \mathcal{L}(U^\varepsilon + W)\, Z \Big)\, dz \Big| \le \ \|\, Z(t)\,\|_{L^2}\, \|\, \mathcal{L}\, Z(t)\,\|_{L^2}\,. \tag{6.34}$$

Hypothesis (6.13) implies that

$$\int_{\mathbb{R}^d\times\mathbb{T}^m} \Big((\mathcal{L}_0 + \mathcal{L}_0^*)W\,,\, W \Big)\, dz \ge 0\,. \tag{6.35}$$

The explicit form of the coefficients A_0 shows that

$$\|\, \partial_0 \mathrm{A}_0^\varepsilon(t,z)\,\|_{L^\infty(z)} \ \le\ \varepsilon\, C(\|U^\varepsilon + W\|_{L^\infty})\ \big\|\, \varepsilon\partial_0(U^\varepsilon(t) + W)\,\big\|_{L^\infty} \ \le\ C\varepsilon\,, \tag{6.36}$$

thanks to (6.17) and (6.26). Similarly

$$\|\, \partial_j \mathrm{A}_j^\varepsilon(t,z)\,\|_{L^\infty(z)} \ \le\ \varepsilon\, C(\|U^\varepsilon + W\|_{L^\infty})\ \big\|\, \partial_j(U^\varepsilon + W)(t)\,\big\|_{L^\infty} \ \ \le\ C\varepsilon\,. \tag{6.37}$$

It is important that in (6.36) one has $\varepsilon\partial_0$ derivatives on the right hand side. The extra power of ε compared to (6.37) comes from the ε^2 in the $\mathcal{B}_0$ term in (6.14) as compared to the ε in the $\mathcal{B}_j$ terms.

Combining the last five equations yields a differential inequality

$$\partial_t\, (\mathrm{A}_0^\varepsilon Z, Z)_{L^2(\mathbb{R}^d)} \le\ C\, \Big(\varepsilon \|\, Z(t)\,\|_{L^2}^2 + \|\, Z(t)\,\|_{L^2} \|\, \mathcal{L}(U^\varepsilon + W) Z(t)\,\|_{L^2} \Big)\,. \tag{6.38}$$

Define

$$\Phi(t) := \big(\mathrm{A}_0^\varepsilon Z(t)\,,\, Z(t)\,\big)^{1/2}_{L^2(\mathbb{R}^d\times\mathbb{T}^m)}\,. \tag{6.39}$$

The symmetric hyperbolicity assumption guarantees that so long as $U^\varepsilon + W$ is pointwise bounded, Φ is equivalent to the L^2 norm in the sense that

$$\frac{1}{C} \|\, Z(t)\,\|_{L^2(\mathbb{R}^d\times\mathbb{T}^m)} \le \Phi(t) \le C \|\, Z(t)\,\|_{L^2(\mathbb{R}^d\times\mathbb{T}^m)}\,. \tag{6.40}$$

Using this in (6.38) yields the differential inequality

$$\frac{d}{dt}\, \Phi^2(t) \le \varepsilon\, C\, \Phi(t)^2 + C\, \Phi(t)\, \|\, \mathcal{L}(U^\varepsilon + W)\, Z(t)\,\|_{L^2}\,. \tag{6.41}$$

As usual this yields

$$\Phi(t) \le\ e^{\varepsilon C t}\, \Phi(0) + \int_0^t\ e^{\varepsilon C(t-\sigma)} \|\, \mathcal{L}(U^\varepsilon + W)\, Z(\sigma)\,\|_{L^2}\ d\sigma\,. \tag{6.42}$$

Since $\varepsilon t \le T$ in our domain this proves the crucial L^2 estimate

$$\|Z(t)\|_{L^2} \le C \|Z(0)\|_{L^2} + C \int_0^t \|\mathcal{L}(U^\varepsilon + W) Z(\sigma)\|_{L^2}\ d\sigma\,. \tag{6.43}$$

The constants in this equation depend on the sup norms of $U^\varepsilon + W$, $\partial_j(U^\varepsilon + W)$, and $\varepsilon\partial_t(U^\varepsilon + W)$.

Step 2. A commutator estimate. The strategy is to apply estimate (6.43) to $Z(t) := \partial_z^\alpha W$ with $|\alpha| \le s$. Toward that end write

$$\mathcal{L}(U^\varepsilon + W)\, \partial_z^\alpha W = \partial_z^\alpha \, \big(\mathcal{L}(U^\varepsilon + W) W\, \big) + \big[\mathcal{L}(U^\varepsilon + W)\,,\, \partial_z^\alpha\, \big]\, W\,. \tag{6.44}$$

The next step is to estimate the commutator.

Commutator estimate. *There is a constant $C(s, U^\varepsilon, \mathcal{L}^\varepsilon)$ so that for all $\underline{t} \le T$, $|\alpha| \le s$ and W satisfying (6.26) one has for $0 \le t \le \underline{t}/\varepsilon$,*

$$\begin{aligned} &\| \left[\mathcal{L}(U^\varepsilon + W)\,,\, \partial_z^\alpha \right] W(t) \|_{L^2(\mathbb{R}^d \times \mathbb{T}^m)} \ \le \\ &\qquad \varepsilon\, C \Big(\| W(t) \|_{H^s(\mathbb{R}^d \times \mathbb{T}^m)} + \varepsilon \| \mathcal{L}^\varepsilon \big(U^\varepsilon(t) + W(t) \big) W(t) \|_{H^s(\mathbb{R}^d \times \mathbb{T}^m)} \Big) . \end{aligned} \tag{6.45}$$

Proof of the commutator estimate. The commutator on the left of (6.45) is a sum of

$$\varepsilon^2 \left[B_0(U^\varepsilon + W)\, \partial_0\,,\, \partial_z^\alpha \right] \tag{6.46}$$

and the terms

$$\varepsilon \left[B_j(U^\varepsilon + W)\, \partial_j\,,\, \partial_z^\alpha \right], \qquad j = 1, 2, \ldots, d+m\,. \tag{6.47}$$

First consider (6.47). Leibniz' rule for differentiating a product shows that

$$\partial_z^\alpha \Big(B_j(U^\varepsilon + W) \partial_j W \Big) = B_j(U^\varepsilon + W) \partial_j \partial_z^\alpha W + \sum_{0 \ne \nu = \alpha - \gamma} c_{\nu,\gamma}\, (\partial_z^\nu B_j)\, \partial_z^\gamma\, \partial_j W\,. \tag{6.48}$$

Thus (6.47) is a linear combination of terms $\varepsilon\, (\partial_z^\nu B_j)\, \partial_z^\gamma \partial_j W$. Leibniz' rule expresses $\partial_z^\nu B_j(U^\varepsilon + W)$ as a sum of terms of the form

$$\varepsilon\, G(U^\varepsilon + W) \prod \partial_z^{\nu_k} (U^\varepsilon + W)\,, \qquad \sum \nu_k = \nu\,.$$

The G factor and all of $\partial_z^{\nu_k} U^\varepsilon$ factors are bounded thanks to (6.17) and (6.26). It remains to estimate the $L^2(\mathbb{R}^d \times \mathbb{T}^m)$ norm of

$$\varepsilon \Big(\prod \partial_z^{\nu_k} W \Big)\, \partial_j \partial_z^\gamma W\,, \qquad |\nu_k| \ge 1\,, \ \ 1 \le |\gamma| \le s-1\,, \ \ |\gamma| + \sum_k |\nu_k| \le s\,. \tag{6.49}$$

Note the total order $1 + |\gamma| + \sum_k |\nu_k|$ could be as large as $s+1$. We use the Gagliardo-Nirenberg inequalities together with Hölder's inequality.

If $|\nu_k| = 1$ then (6.26) bounds the corresponding factor in L^∞ so those factors can be eliminated. Thus it suffices to consider expressions (6.49) with all $|\nu_k| \ge 2$.

For such expressions, let

$$r := |\gamma| + \sum_k (|\nu_k| - 1) \le s - 1, \qquad \text{so} \qquad \frac{|\gamma|}{2r} + \sum_k \frac{|\nu_k| - 1}{2r} = \frac{1}{2}\,. \tag{6.50}$$

Interpolating between $\partial_j W \in L^\infty$ and $|\partial_z|^r \partial_j W \in L^2$ one has

$$\| \partial_z^\gamma \partial_j W \|_{L^{2r/|\gamma|}} \ \le \ C\, \| \partial_j W \|_{L^\infty}^{1 - |\gamma|/r}\ \| |\partial_z|^r \partial_j W \|_{L^2}^{|\gamma|/r}\,. \tag{6.51}$$

Similarly writing $\partial_z^{\nu_k}$ as the product of a first order partial and a partial of order $|\nu_k| - 1$ yields

$$\| \partial_z^{\nu_k} W \|_{L^{2r/(|\nu_k|-1)}} \ \le\ C \| \nabla_z W \|_{L^\infty}^{1-(|\nu_k|-1)/r} \ \ \| |\partial_z|^{r+1} W \|_{L^2}^{(|\nu_k|-1)|/r} . \tag{6.52}$$

Applying Hölder's inequality yields

$$\varepsilon \Big\| \Big(\prod \partial_z^{\nu_k} W \Big) \partial_j \partial_z^\gamma W \Big\|_{L^2(\mathbb{R}^d \times \mathbb{T}^m)} \le \varepsilon\, F(\|W\, , \, \nabla_z W\|_{L^\infty}) \, \| |\partial_z|^{r+1} W \|_{L^2(\mathbb{R}^d \times \mathbb{T}^m)} . \tag{6.53}$$

This suffices to estimate the terms (6.47) by the the first summand on the right hand side of (6.45).

The analysis of the terms (6.46) follows the same lines with ∂_j replaced by $\varepsilon\partial_0$ leading to

$$\Big\| \Big(\varepsilon \prod \partial_z^{\nu_k} W \Big) \varepsilon\partial_0 \partial_z^\gamma W \Big\|_{L^2(\mathbb{R}^d \times \mathbb{T}^m)} \le \varepsilon\, F(\|W\, , \, \varepsilon\partial_0 W\, , \, \nabla_z W\|_{L^\infty}) \, \| \varepsilon\partial_0 |\partial_z|^r W \|_{L^2} . \tag{6.54}$$

Expressing the time derivative

$$\varepsilon \partial_0 \, W \ = \ A_0^{-1} \Big(\varepsilon\, \mathcal{L}^\varepsilon (U^\varepsilon + W) W - \sum_{j=1}^{d+m} \Big(\varepsilon \mathcal{A}_j + \tilde{\mathcal{A}}_j \Big) \frac{\partial W}{\partial z_j} + \varepsilon \mathcal{L}_0 \, W \Big) .$$

yields

$$\| \varepsilon\partial_0 W(t) \|_{H^{s-1}(\mathbb{R}^d \times \mathbb{T}^m)} \ \le\ C \Big(\| W(t) \|_{H^s(\mathbb{R}^d \times \mathbb{T}^m)} + \varepsilon \| \mathcal{L}^\varepsilon (U^\varepsilon + W) W \|_{H^s(\mathbb{R}^d \times \mathbb{T}^m)} \Big) . \tag{6.55}$$

Estimates (6.54) and (6.55) suffice to estimate the terms (6.46) by the right hand side of (6.45). This completes the proof of the commutator estimate.

Step 3. Endgame. Applying (6.43) to $Z := \partial_z^\alpha W$ with $|\alpha| \le s$ and using the commutator estimate (6.45) yields with norms on $\mathbb{R}^d \times \mathbb{T}$,

$$\|Z(t)\|_{L^2} \le C \Big(\|Z(0)\|_{L^2} + \int_0^t \|\partial_z^\alpha \big(\mathcal{L}(U^\varepsilon + W) W \big)(\sigma)\|_{L^2} \, d\sigma + \varepsilon \int_0^t \|W(\sigma)\|_{H^s} \, d\sigma \Big) .$$

Summing over all $|\alpha| \le s$ yields

$$\|W(t)\|_{H^s} \le C \Big(\|W(0)\|_{H^s} + \int_0^t \| \big(\mathcal{L}(U^\varepsilon + W) W \big)(\sigma)\|_{H^s} \, d\sigma + \varepsilon \int_0^t \|W(\sigma)\|_{H^s} \, d\sigma \Big) .$$

Gronwall's inequality implies that

$$\|W(t)\|_{H^s} \le C \Big(e^{\varepsilon C t} \|W(0)\|_{H^s} + \int_0^t e^{\varepsilon C (t-\sigma)} \, \| \big(\mathcal{L}(U^\varepsilon + W) W \big)(\sigma)\|_{H^s} \, d\sigma \Big) .$$

Since $0 \le \sigma \le \underline{t}/\varepsilon \le T/\varepsilon$ on the domain in consideration, this proves (6.27) and completes the proof of Lemma 6.3, and therefore Stability Theorem 6.2. ∎

§7. Applications, examples, and extensions.

Once the construction of the approximate solution and the proof of its validity is complete, it remains to study the structure of the approximate solutions. In this section we present a selection of qualitative results concerning the dynamics defined by the profile equations. Several examples show the forms these equations can take. In addition, we describe several extensions of the results whose proof does not require essentially new ideas. In particular we discuss nonoscillatory solutions and profiles periodic in Y.

§7.1. Nonoscillatory solutions.

The dynamics of the oscillatory part a_0^* is determined by equations (2.26) and (4.38). These equations always have the solution $a_0^* = 0$. The solutions with $a_0^* = 0$ describe the behavior for times $t \sim 1/\varepsilon$ of hyperbolic systems with nonoscillatory initial data $u(0,y) = \varepsilon^p \underline{a}(0, \varepsilon y, 0, y)$. We record some features of the result as a Proposition.

Proposition 7.1. *If the oscillatory part $a_0^*(T,Y,y,\theta)$ vanishes for $T = t = 0$ then it vanishes for all time, and the nonoscillatory part is determined by the uncoupled system of equations for $\underline{a}_{0,\alpha} = E_\alpha(D_y)\, \underline{a}_{0,\alpha}$, $\alpha \in \mathcal{A}$*

$$\big(\, \partial_t + \tau_\alpha(D_y) \,\big)\, \underline{a}_{0,\alpha}(T,Y,t,y) = 0\,. \tag{7.1}$$

$$E_\alpha(D_y) \Big(\, L(0,\partial_X)\underline{a}_{0,\alpha} + \kappa(\alpha)\, \Phi(\underline{a}_{0,\alpha}) \,\Big) = 0\,. \tag{7.2}$$

Proof. Equations, (2.21), (2.26), (4.27), (4.28) and (4.38) with initial data $a_0(0,Y,y,\theta) = \underline{a}_0(0,Y,0,y)$ are solved if one takes $a_0^* = 0$ and $\underline{a}_{0,\alpha}$ to be the unique solution of (7.1), (7.2) with initial value $\underline{a}_{0,\alpha}(0,Y,0,y) = E_\alpha(D_y)\, \underline{a}_0(0,Y,0,y)$. By uniqueness this yields *the* solution. ∎

Remarks. 1. The modes are uncoupled.

2. The nonlinear parts of the quasilinear terms do not contribute to the principal term.

3. For curved parts of the characteristic variety, $\kappa(\alpha) = 0$ and the equation for $\underline{a}_\alpha$ is linear. If all sheets of the variety are curved the dynamics is linear.

4. In the one dimensional case, one can diagonalize $A_1(0)$

$$A_1(0) = \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_N)\,.$$

In this basis we have components

$$a_0 = \underline{a}_0 = (\mathrm{a}_1, \ldots, \mathrm{a}_N)\,, \qquad \Phi = (\Phi_1, \ldots, \Phi_N)\,.$$

The profile equations are uncoupled nonlinear transport equations

$$\big(\, \partial_t + \lambda_n\, \partial_y \,\big)\, \mathrm{a}_n = 0\,, \qquad \big(\, \partial_T + \lambda_n\, \partial_Y \,\big)\, \mathrm{a}_n + \Phi_n(\mathrm{a}_n) = 0\,, \qquad n = 1, \ldots, N\,.$$

Writing $\mathrm{a}_n = b_n(T, Y - \lambda_n T, y - \lambda_n t, \theta)$ with $b_n = b_n(T, Y, y, \theta)$ these equations reduce to

$$\partial_T b_n + \Phi_n(b_n) = 0\,, \qquad n = 1, \ldots, N\,.$$

§**7.2. Profiles independent of or periodic in Y**.

The equations for the profile $a_0(X, x, \theta)$ are translation invariant in the sense that for any $\underline{X}, \underline{x}, \underline{\theta} \in \mathbb{R}^{2d+3}$, $a_0(X + \underline{X}, x + \underline{x}, \theta + \underline{\theta})$ is a solution if and only if a_0 is a solution. This suggests the following argument. If the initial data $a_0(0, Y, t, y, \theta)$ is independent of Y, then the solution a_0 is also independent of Y since $a_0(T, Y + \underline{Y}, t, y, \theta)$ is a solution with the same initial data, so must be equal to a_0. The error in the argument is that we have only studied the initial value problem for data belonging to $\cap_s H^s(\mathbb{R}^d_Y \times \mathbb{R}^d_y \times \mathbb{T})$ which excludes functions independent of Y. However, the arguments which work for $Y \in \mathbb{R}^d$ work exactly the same way for $Y \in \mathbb{T}^d$. In this way a theory exactly analogous to that we have developed works for modulations on the slow scale which are periodic rather than square integrable in Y. One finds exactly the same profile equations and perfectly analogous stability results where $H^s(\mathbb{R}^d_Y)$ is systematically replaced by $H^s(\mathbb{T}^d_Y)$. Once this is remarked the preceding uniqueness argument is correct and we have the following important case of profiles independent of Y for which the profile equations simplify.

Proposition 7.2. *If the initial data $g(y, \theta) \in \cap_s H^s(\mathbb{R}^d \times \mathbb{T})$ with $\pi(\beta)\, g^* = g^*$ is independent of Y then there is a unique profile $a_0(T, t, y, \theta)$ independent of Y determined as the solution of (2.21), (2.26) and the T dynamic equations*

$$E_\alpha(D_y)\Big(\partial_T \underline{a}_{0,\alpha} + \kappa(\alpha)\, \Phi(\underline{a}_{0,\alpha})\Big) = 0\,, \qquad \text{for} \quad \alpha \in \mathcal{A} \setminus \{1\}\,. \tag{7.3}$$

$$E_1(D_y)\left(\partial_T\, \underline{a}_{0,1} + \Big\langle \kappa(1)\, \Phi(\iota\, a_0^* + \underline{a}_{0,1}) + \iota \sum_\mu \beta_\mu \Lambda_\mu(a_0^* + \underline{a}_{0,1})\, \partial_\theta a_0^* \Big\rangle \right) = 0\,. \tag{7.4}$$

$$\begin{aligned} &\partial_T\, a_0^* - R(\partial_y)\, \partial_\theta^{-1} a_0^* + (1-\iota)\, \pi\, \Phi(a_0^*)^* \\ &\qquad + \iota\, \pi\, \Phi(a_0^* + \underline{a}_{0,1})^* + \Big(\pi \sum_{\mu=0}^d \beta_\mu\, \Lambda_\mu(a_0^* + \iota \underline{a}_{0,1}) \partial_\theta a_0\Big)^* = 0\,. \end{aligned} \tag{7.5}$$

which are formally identical to equations (4.28), (4.27) and (4.38) for Y independent functions.

Combining the simplifications of §7.1 and §7.2 yields the situation of the warmup problem in §4.1.

Profiles independent of Y arise in another natural way. Whenever one has a principal profile which is rapidly decreasing in y, the asymptotic description can be achieved with a principal profile which is independent of Y. In particular this is the case of the profiles constructed in the published (as opposed to internet) version of [DJMR]. The argument is the following. Write

$$a_0(T, Y, t, y, \theta) = a_0(T, 0, y, \theta) + Y.b(T, Y, t, y, \theta)\,, \qquad b := \int_0^1 \nabla_Y a(T, \sigma Y, t, y, \theta)\, d\sigma\,.$$

Let $c(T, Y, t, y, \theta) := y.b(T, Y, t, y, \theta)$. When profiles are used one substitutes $Y = \varepsilon y$ so that the profile

$$a_0(T, 0, t, y, \theta) + \varepsilon c(T, Y, t, y, \theta)$$

describes the same asymptotic behavior as a_0 and has principal profile $a_0(T, 0, t, y, \theta)$ which is independent of Y. In order that c be suitably square integrable in y requires that $y \nabla_Y a_0$ be square integrable in y which is a decay condition on $\nabla_Y a_0$.

If the initial values $a_0(0, Y, 0, y, \theta)$ are rapidly decaying in the sense that

$$y^\gamma a_0(0, Y, 0, y, \theta) \in \cap_s H^s(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{T})$$

then the process can be repeated yielding an expansion

$$a_0(0, Y, 0, y, \theta) \sim a_0(0, 0, 0, y, \theta) + \sum_{j=1}^{\infty} \varepsilon^j \, c_j(y, \theta)$$

all of whose profiles are independent of y. Solving the profile equations with the initial data for $\underline{a}_1, \pi(\beta)\, a_1^*$ changed from zero to

$$\underline{a}_1\big|_{t=T=0} = \underline{c}_1(y)\,, \qquad \pi(\beta)\, a_1^*\big|_{t=T=0} = \pi(\beta)\, c_1^*$$

yields an asymptotic description with profiles independent of Y.

Summary. *If a_0 is rapidly decaying in y, the above recipe constructs an asymptotic description of the solution with profiles independent of Y.*

§7.3. Curved sheet implies that a_0^* does not influence $\underline{a}_0$.

Proposition 7.3. *If β satisfies the smooth variety hypothesis and belongs to a sheet of the characteristic variety which is not a hyperplane, then the evolution of the nonoscillating part $\underline{a}_0$ is not influenced by the oscillating part a_0^*. For such β if $\underline{a}_0$ vanishes at $T = t = 0$ then it vanishes identically.*

For solutions with $\underline{a}_0 = 0$, the profile equations simplify appreciably which makes this Proposition particularly interesting.

Proof. The condition on β is equivalent to $\iota = 0$. In that case the profile equations for $\underline{a}_0$ are

$$\big(\,\partial_t + \tau_\alpha(D_y)\,\big)\, \underline{a}_\alpha = 0\,, \tag{7.6}$$

$$E_\alpha(D_y)\,\big(\,L(0, \partial_X)\, \underline{a}_\alpha + \kappa(\alpha)\Phi(\underline{a}_{0,j})\,\big) = 0\,, \tag{7.7}$$

These equations do not involve a_0^* proving the first assertion of the Proposition.

If the initial data vanishes at $T = t = 0$ then (7.6) implies that $\underline{a}_\alpha$ vanishes in $\{T = 0\}$. Then the second equation implies that it vanishes for all T. ∎

This result shows that creation of nonoscillating contributions from oscillating terms is restricted to the case of β belonging to hyperplanes in the characteristic variety.

Examples of hyperplanes in the variety. 1. When $d = 1$ all sheets of the characteristic variety are hyperplanes. For this reason the one dimensional case is not a good guide to the general case.

2. The characteristic varieties of both the compressible Euler equations and Maxwell's Equations are the union of a light cone and the horizontal plane $\tau = 0$. For the Euler equation the hyperplane corresponds to entropy waves. For the Maxwell equations, the hyperplane in the characteristic variety correspond to unphysical solutions which are eliminated by the constraints $\operatorname{div} E = \operatorname{div} B = 0$.

3. It is possible that in a conic neighborhood of β the characteristic variety contains a curved sheet and its tangent plane. An example is the characteristic equation

$$(\tau^2 - |\eta|^2)(\tau + \eta_1) = 0 \qquad \text{with} \qquad \beta = (1, -1, 0, \dots, 0)\,.$$

In such cases the simple characteristic variety hypothesis is violated and our construction of approximate solutions does not apply. ∎

§7.4. Two examples with $\underline{\mathrm{a}}_0 = 0$.

We compute the form taken by the profile equations for two examples satisfying the hypotheses of Proposition 7.3 and for which the mean value $\underline{a}_0$ vanishes identically.

Example 7.1. Consider the semilinear system

$$\frac{\partial u}{\partial t} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{\partial u}{\partial y_1} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{\partial u}{\partial y_2} + \Phi(u) = 0\,, \tag{7.8}$$

where $\Phi(u_1, u_2) = (\Phi_1, \Phi_2)$ is homogeneous of degree J. The standard normalization is then $p = 1/(J-1)$. For quadratic (resp. cubic) nonlinearity one has $p = 1$ (resp. $p = 1/2$). The characteristic variety is given by $\tau^2 = |\eta|^2$ so the simple characteristic variety and curved characteristic variety hypotheses are satisfied at all β. Take $\beta = (1, -1, 0, 0)$ then

$$L(\beta) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}, \quad \text{and,} \quad \pi(\beta) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The transport operator is $V(\partial_x) = \partial_t + \partial_1$. The principal profile satisfies $\pi(\beta) a_0^* = a_0$ so has vanishing second component. Risking confusion write $(\mathrm{a}(T, Y_1 - T, Y_2, y_1 - t, y_2, \theta), 0)$ for the leading profile. Then (4.38) shows that the scalar valued function $\mathrm{a}(T, Y, y, \theta)$ is determined from its values at $T = 0$ by the equation

$$\frac{\partial\, \mathrm{a}}{\partial T} + \frac{1}{2}\, \Delta_y\, \partial_\theta^{-1}\, \mathrm{a} + \Phi_1(\mathrm{a}, 0)^* = 0\,. \tag{7.9}$$

If p is an odd integer and $\Phi_1(\mathrm{a}, 0) = \phi(|\mathrm{a}|^{p-1})\mathrm{a}$, there are monochromatic solutions $\mathrm{a}(T, Y, y, \theta) = e^{ik\theta}\, v(T, Y, y)$ whose profile v is determined as a solution of the classical Nonlinear Schrödinger Equation

$$\frac{\partial\, v}{\partial T} + \frac{1}{2ik}\, \Delta_y\, v + \phi(|v|^{p-1})\, v = 0\,. \qquad ∎$$

Example 7.2. The inviscid compressible Euler equations.

The isentropic inviscid $2d$ compressible Euler equations describe flows with negligible heat conduction.* The unknowns are the velocity $v = (v_1(t,y), v_2(t,y))$ and the density $\rho(t,y)$. The dynamics is governed by the system of equations

$$\Big(\frac{\partial}{\partial t} + v_1\frac{\partial}{\partial y_1} + v_2\frac{\partial}{\partial y_2}\Big)\, v_1 + \frac{p'(\rho)}{\rho}\,\frac{\partial}{\partial y_1}\,\rho = 0\,. \tag{7.10}$$

$$\Big(\frac{\partial}{\partial t} + v_1\frac{\partial}{\partial y_1} + v_2\frac{\partial}{\partial y_2}\Big)\, v_2 + \frac{p'(\rho)}{\rho}\,\frac{\partial}{\partial y_2}\,\rho = 0\,. \tag{7.11}$$

$$\Big(\frac{\partial}{\partial t} + v_1\frac{\partial}{\partial y_1} + v_2\frac{\partial}{\partial y_2}\Big)\, \rho + \rho\,\Big(\frac{\partial v_1}{\partial y_1} + \frac{\partial v_2}{\partial y_2}\Big) = 0\,, \tag{7.12}$$

where $p = p(\rho)$ gives the pressure as a function of the density. This system is of the form $\sum B_\mu(v,\rho)\partial_\mu(v,\rho) = 0$ with coefficients

$$B_0 = I\,, \qquad B_1 = \begin{pmatrix} v_1 & 0 & p'(\rho)/\rho \\ 0 & v_1 & 0 \\ \rho & 0 & v_1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} v_2 & 0 & p'(\rho)/\rho \\ 0 & v_2 & 0 \\ \rho & 0 & v_2 \end{pmatrix}. \tag{7.13}$$

The system is symmetrized by multiplying by $\tilde{B}_0$ yielding coefficients

$$\tilde{B}_0 := \begin{pmatrix} \rho^2/p'(\rho) & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \tilde{B}_j = \begin{pmatrix} \rho^2 v_j/p'(\rho) & 0 & \rho \\ 0 & v_j & 0 \\ \rho & 0 & v_j \end{pmatrix}. \tag{7.14}$$

The background state is $v = 0,\ \rho = \underline{\rho} > 0$ with constant density. Introduce

$$u := \tilde{B}_0^{1/2}(0,\underline{\rho})\,(v, \rho - \underline{\rho}) = \Big(\frac{\underline{\rho} v_1}{c}, v_2, \rho - \underline{\rho}\Big)\,, \qquad c := \sqrt{p'(\underline{\rho})}$$

to find an equivalent system with coefficients

$$A_\mu(u) = \tilde{B}_0^{-1/2}(0,\underline{\rho})\ \tilde{B}_\mu\Big(\frac{c}{\underline{\rho}}\,u_1, u_2, u_3 + \underline{\rho}\Big)\,\tilde{B}_0^{-1/2}(0,\underline{\rho})\,.$$

satisfying the conventions that the background state is $u = 0$ and the coefficient of ∂_t is equal to I when $u = 0$. This is usually a simple zero so the nonlinearity is of order $K = 2$ and the critical exponent for diffractive nonlinear geometric optics is $p = 2/(K-1) = 2$.

The system is strictly hyperbolic with characteristic equation

$$\tau\,\big(\,\tau^2 - c^2|\eta|^2\,\big) = 0\,. \tag{7.15}$$

* The nonisentropic, and $3d$ equations can be analysed in the same manner.

For $\beta = (c, -1, 0) = (\tau, \eta)$ the basic operators are

$$V(\partial_x) = \frac{\partial}{\partial t} + c\,\frac{\partial}{\partial y_1}\,, \qquad R(\partial_y) = \Delta_y\,, \qquad \pi(\beta) = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 0 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}\,. \tag{7.16}$$

The last is orthogonal projection on $\mathbb{R}(1, 0, 1)$.

Since the variety is curved, Proposition 7.3 yields profiles with vanishing nonoscillatory part. In that case, the principal profiles

$$a_0 = \mathrm{a}^*(T, Y_1 - cT, Y_2, y_1 - ct, y_2, \theta)\,\big(1/\sqrt{2}, 0, 1/\sqrt{2}\big)$$

with scalar valued amplitude $\mathrm{a}(T, Y, y, \theta)$ are determined as solutions of the quasilinear evolution equations

$$\partial_T\,\mathrm{a} + \Delta_y \partial_\theta^{-1}\mathrm{a} + \sigma\,\mathrm{a}\,\partial_\theta \mathrm{a} \;=\; 0\,, \tag{7.17}$$

with $\sigma \in \mathbb{R}$ determined from the identity

$$\pi(\beta)\,\Big(\sum \beta_\mu\, A'_\mu(0)\Big)\,\pi(\beta) = \pi(\beta)\,\Big(c\,A'_0(0) - A'_1(0)\Big)\,\pi(\beta) = \sigma\pi(\beta)\,. \qquad \Box$$

§7.5. An example of a large corrector.

We present the computations for Example 5 of the introduction. In the notation of that example, either by analysing the exact solution with oscillatory initial data as in [DJMR], or by applying Theorem 1.1 in the absence of nonlinear terms one finds solutions u^ε with

$$u^\varepsilon \;\sim\; \varepsilon\, e^{i(y_1 - t)/\varepsilon} \sum_{j=0}^{\infty} \varepsilon^j\, a_j(\varepsilon t, y_1 - t, y_2)\,.$$

with $a_0(T, y)$ determined as the solution of the Schrödinger equation

$$2i\,\partial_T a_0 + \;\Delta_y\, a_0 = 0\,.$$

Consider gaussian initial data, $a_0(0, y) = e^{-y^2}$. Then $|a(t, y|$ is a spreading gaussian with maximum at the origin. Thus for $c > 0$ sufficiently small, and $0 \le T \le 1$,

$$-(\operatorname{sgn} y_1)\frac{\partial}{\partial y_1}\,|a_0(T, y)|^2 \ge c\,e^{-y^2/c} := g(y)\,.$$

Take Cauchy data for v^ε to be equal to zero. The for $t \sim 1/\varepsilon$

$$\mathrm{v}^\varepsilon \;\sim\; \varepsilon^2\,\Box^{-1}|a_0(\varepsilon t, y_1 - t, y_2)|^2\,, \qquad v^\varepsilon = (v_0^\varepsilon, v_1^\varepsilon, v_2^\varepsilon) = \nabla_{t,y}\mathrm{v}^\varepsilon\,.$$

Then in $\{y_1 \ge t\}$,

$$-v_1^\varepsilon = -\frac{\partial \mathrm{v}^\varepsilon}{\partial y_1} = \varepsilon^2\,\Box^{-1}\frac{\partial |a_0(\varepsilon t, y_1 - t, y_2)|^2}{\partial y_1} \ge \varepsilon^2\,w\,.$$

where w is the solution of the intial value problem

$$\Box_{1+2} w = g(y_1 - t, y_2) \,, \qquad w^\varepsilon|_{t=0} = w^\varepsilon_t|_{t=0} = 0 \,,$$

whose source term is a pulse moving at speed 1. Define $G(t,y)$ to be the solution of

$$\Box_{1+2}\, G = g(t,y) \,, \qquad G|_{t=0} = G_t|_{t=0} = 0 \,.$$

Then

$$w(t,y) = \int_0^t G(t-s, y_1 - s, y_2)\, ds \,,$$

so

$$w(t,t,0) = \int_0^t G(t,t,0)\, ds \,,$$

For $t > 0$,

$$G(t,t,0) = \frac{1}{2\pi} \int_{D(t;(t,0))} \frac{g(y)}{\sqrt{t^2 - (y_1 - t)^2 + y_2^2}}\ dy$$

where $D(r;z)$ denotes the disc of center z and radius r.
Let $d^2 := (y_1 - t)^2 + y_2^2$ so in the intersection of the disc and $\operatorname{supp} g$,

$$\begin{aligned} t^2 - (y_1 - t)^2 - y_2^2 &= t^2 - d^2 = (t+d)(t-d) = 2t(t-d) - (t-d)^2 \\ &= 2t \operatorname{dist}(y, \partial D(t;(t,0)) - \operatorname{dist}(y, \partial D(t;(t,0))^2 = 2t \operatorname{dist}(y, \partial D(t;(t,0)) + O(1) \,, \end{aligned}$$

as $t \to \infty$. For t large, $D(t;(t,0)) \cap \operatorname{supp} g$ approaches the intersection $\operatorname{supp} g \cap \{y_1 \ge 0\}$. In particular

$$\operatorname{dist}(y, \partial D(t;(t,0)) = y_1 + O(1/t) \,,$$

on $D(t;(t,0)) \cap \operatorname{supp} g$, so

$$G(t,t,0) = \frac{1}{2\pi\sqrt{2t}} \int_{y_1 > 0} \frac{g(y)}{\sqrt{y_1}}\ dy + O(t^{-1})$$

and therefore

$$w(t,t,0) = \frac{\sqrt{t}}{\pi\sqrt{2}} \int_{y_1 > 0} \frac{g(y)}{\sqrt{y_1}}\ dy + O(\log t) \,.$$

In the same way

$$\frac{\partial w(t,t,0)}{\partial y_1} = \frac{\sqrt{t}}{\pi\sqrt{2}} \int_{y_1 > 0} \frac{\partial g(y)/\partial y_1}{\sqrt{y_1}}\ dy + O(\log t) \,.$$

Thus, for $t \sim 1/\varepsilon$, $v_1^\varepsilon(t,t,0) = \varepsilon^2 \partial w(t,t,0)/\partial y_1 \sim \varepsilon^{3/2}$. Thus v^ε is small compared to the $O(\varepsilon)$ principal term which is u^ε, but v^ε large compared the $O(\varepsilon^2)$ correctors in (1.1).

§7.6. A scalar higher order equation.

An analysis like that for first order systems applies to equations of higher order. One has the option of reducing to a first order system or to start from scratch deriving profile equations. The equations for the profiles for scalar equations are a little easier to find because there are no polarizations and projectors π. There are two changes which should be born in mind for higher order equations.

For an equation of order m, nonlinearities can be present in derivatives of all orders between 0 and m. Each nonlinear term must have a time of interaction no smaller than $O(1/\varepsilon)$

For a linear m^{th} order equation with source term oscillating with wavelength ε the response will be of size ε^m if the phase has noncharacteristic differential. The response is $O(\varepsilon^{m-1})$ in the characteristic case, that is when the phase satisfies the eikonal equation. Our phases fall in the latter category. In the applied literature phases of this sort are sometimes called phase matched.

Example 7.3. Consider the equation

$$\Box_{t,y}\, u + \Phi(\partial_{t,y} u) = 0\,, \tag{7.18}$$

where Φ is a homogeneous polynomial of degree J. The characteristic variety is $\tau^2 = |\eta|^2$ so is everywhere simple and curved.

The equation can be converted to the semilinear symmetric hyperbolic system

$$\partial_t \mathbf{w}_j - \partial_j \mathbf{w}_0 = 0\,, \qquad j = 1, \ldots, d\,,$$

$$\partial_t \mathbf{w}_0 - \sum_{j=1}^{d} \partial_j \mathbf{w}_j - \Phi(\mathbf{w}) = 0\,,$$

for $\mathbf{w} = \nabla_{t,y} u$. Since Φ is of order J the critical exponent is $q = 1/(J-1)$ and have solutions

$$\mathbf{w}^\varepsilon \;=\; \varepsilon^q\, \mathbf{w}_0(\varepsilon x, x, \beta.x/\varepsilon) + \text{h.o.t.}$$

The characteristic variety is given by the equation $\tau^{d-1}(\tau^2 - |\eta|^2) = 0$ with the extraneous hyperplane $\{\tau = 0\}$ of multiplicity $d-1$. These roots are easily understood since solutions of the system satisfy

$$\partial_t\, d\big(\mathbf{w}_1 dx_1 + \ldots + \mathbf{w}_d dx_d\big) = 0\,,$$

and the solutions of interest are those in the invariant subspace defined by $d(\mathbf{w}_1 dx_1 + \ldots + \mathbf{w}_d dx_d) = 0$. The $L^2(\mathbb{R}^d)$-orthogonal complement of these solutions are stationary solutions satisfying $\operatorname{div}(\mathbf{w}_1, \ldots, \mathbf{w}_d) = 0$. These correspond to the extraneous roots.

If one chooses $\beta = (1, -1, 0, \ldots, 0)$, then

$$\pi(\beta) = \frac{|(1,1,0\ldots,0)\rangle\langle(1,1,0\ldots,0)|}{2}\,, \qquad V(\partial) = \partial_0 + \partial_1\,, \qquad R(\partial_y) = \frac{1}{2}\Delta_y\,,$$

and for profiles whose mean value is initially zero and independent of Y it remains so for positive time by Propositions 7.2 and 7.3. Profiles with vanishing mean and independent of Y are given

$$\mathbf{a}_0(T,t,y,\theta) \;=\; \mathbf{a}_0^*(T,t,y,\theta) \;=\; \mathrm{a}(T, t-y_1, y_2, \ldots, y_d, \theta)\Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\ldots, 0\Big)$$

with scalar valued a satisfying

$$\partial_T \mathrm{a} + \frac{1}{2}\Delta_y \partial_\theta^{-1}\mathrm{a} + \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \ldots, 0\Big) \cdot \Phi\Big(\mathrm{a}\big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \ldots, 0\big)\Big) \;= 0\,.$$

An alternate approach is to reason directly on the scalar equation seeking approximate solution of the form (2.6-2.7). To compute the critical exponent p one must estimate the accumulated effect of the nonlinear term $\Phi(\partial u)$. For u of wavelength $\sim \varepsilon$ and amplitude $\sim \varepsilon^p$ as in (2.6), $\partial u \sim \varepsilon^{p-1}$ so $\Phi(\partial u) \sim \varepsilon^{(p-1)J}$. The phase will satisfy the eikonal equation, and then linear geometric optics shows that the effect of the oscillatory part is $\Box^{-1}\left(\Phi(\partial u)^*\right) \sim t\varepsilon\, \varepsilon^{(p-1)J}$. Setting this equal to ε^p for times $t \sim 1/\varepsilon$ yields $p(J-1) = p$ so $p = J/(J-1)$. Note that $p+1 = q$ is the critical power for a semilinear first order system with nonlinearity of degree J.

Inject the ansatz (2.6-2.7) into the differential equation and insist that the nonlinear term contribute at the same power in ε that yields linear diffractive effects. A first computation yields

$$\Box u^\varepsilon = \varepsilon^{p-2}\left(\beta_0^2 - \beta_1^2 - \ldots - \beta_d^2\right)\frac{\partial^2 a_0}{\partial^2\theta} + O(\varepsilon^{p-1})\,.$$

This forces the eikonal equation $\beta_0^2 = \beta_1^2 + \ldots + \beta_d^2$. Then one has

$$\Box u^\varepsilon = 2\varepsilon^{p-1}\left(\beta_0\frac{\partial}{\partial t} - \sum_{j=1}^d \beta_j\frac{\partial}{\partial y_j}\right)\frac{\partial a_0}{\partial\theta} + O(\varepsilon^p)\,,$$

which yields

$$\beta_0\frac{\partial a_0^*}{\partial t} - \sum_j \beta_j\frac{\partial a_0^*}{\partial y_j} = 0\,. \tag{7.19}$$

The nonlinear term should appear in the next term which gives the laws of diffractive geometric optics. The next term is of order ε^p and the nonlinear term is of order $\varepsilon^{(p-1)J}$. Equating these orders gives again the critical exponent $p = J/(J-1)$.

Choosing $\beta = (1, -1, 0, \ldots, 0)$, yields the vector field $V(\partial_x) = \partial_t + \partial_1$. With the choice of p above one finds

$$\Box u^\varepsilon + \Phi(\nabla_{t,y}u^\varepsilon) = \varepsilon^p\left(2V(\partial_X)\frac{\partial a_0}{\partial\theta} + \Box_x a_0 + \Phi\Big(\beta\frac{\partial a_0}{\partial\theta}\Big) + 2V(\partial_x)\frac{\partial a_1}{\partial\theta}\right) + O(\varepsilon^{p+1})\,. \tag{7.20}$$

This suggests imposing the equation

$$2V(\partial_X)\frac{\partial a_0}{\partial\theta} + \Box_x a_0 + \Phi\Big(\beta\frac{\partial a_0}{\partial\theta}\Big) + 2V(\partial_x)\frac{\partial a_1}{\partial\theta} = 0\,. \tag{7.21}$$

Thanks to (7.19), applying $V(\partial_x)$ to the oscillating part of (7.21) yields $V(\partial_x)^2 a_1^* = 0$. In order that a_1 grow sublinearly in x one must have $V(\partial_x)a_1^* = 0$ and one finds the profile equation

$$2V(\partial_X)\frac{\partial a_0}{\partial\theta} + \Box_x a_0^* + \Phi\Big(\beta\frac{\partial a_0}{\partial\theta}\Big)^* \;=\; 0\,. \tag{7.22}$$

Equations (7.19) and (7.22) suffice to determine the oscillating part $\underline{a}_0^*$ from its initial values at $\{t = T = 0\}$. Note that the mean values of $\underline{a}_0$ do not enter in this determination, and that the profile equation just derived is equivalent to the equation found from reducing to a first order system with the identification $\mathbf{w}_0 = \partial \mathbf{a}_0/\partial\theta = \beta\partial \mathrm{a}_0/\partial\theta$. Differentiating one has

$$\nabla u^\varepsilon \;=\; \varepsilon^{p-1}\,\frac{\partial \mathbf{a}_0(\varepsilon x, x, \beta.x/\varepsilon)}{\partial\theta} + O(\varepsilon^p)\,,$$

and one recovers the answer from the first order system computation. ∎

References

[BR] R. Benedetti and J.-J. Risler, *Real Algebraic and Semi-algebraic Sets*, Actualités Mathématiques, Hermann (1990).

[CR] M. Coste and M.F. Roy, *Géometrie Algébrique Réelle*, Ergebnisse der Mathematik, Springer-Verlag, 1987

[C] R. Courant, *Methods of Mathematical Physics Vol II*, Interscience, 1962.

[D] P. Donnat, Quelques contributions mathématiques en optique non linéaire, Thèse Doctoral, École Polytechnique, 1994.

[DJMR] P. Donnat, J.-L. Joly, G. Métivier, and J. Rauch, Diffractive nonlinear geometric optics I, Seminaire Equations aux Dérivees Partielles, École Polytéchnique, Palaiseau, 1995-1996, XVII 1-XVII 23. Revised version available at `www.math.lsa.umich.edu/~rauch`.

[DR 1] P. Donnat and J. Rauch, Modeling the dispersion of light, in *Ocillations and Singularities* eds J. Rauch and M. Taylor, IMA vol Math. and Appls., to appear

[DR 2] P. Donnat and J. Rauch, Dispersive nonlinear geometric optics, Journ. Math, Phys. to appear

[DR 3] P. Donnat and J. Rauch, Diffractive nonlinear geometric optics for dispersive hyperbolic systems, in preparation.

[F] R. P. Feynmann, R. Leighton, and M. Sands, *The Feynmann Lectures on Physics Vols. I,II,* Addison-Wesley (1964)

[G 1] O. Guès, Développements asymptotiques de solutions exactes de systèmes hyperboliques quasilinéaires, Asymptotic Anal., 6(1993)241-269.

[G 2] O. Guès, Ondes multidimensionnelles ε-stratifiées et oscillations, Duke Math. J., 68(1992)401-446.

[H] J. Hunter, Asymptotic equations for nonlinear hyperbolic waves, in Surveys in Applied Mathematics Vol 2, eds. J.B. Keller, D.W. McLauglin, and G.C. Papanicolaou, Plenum Press, 1995.

[HMR] J. Hunter, A. Majda, and R. Rosales, Resonantly interacting weakly nonlinear hyperbolic waves II: several space variables, Stud. Appl. Math. 75(1986)187-226.

[JMR 1] J.-L. Joly, G. Métivier, and J. Rauch, Resonant one dimensional nonlinear geometric optics J. of Funct. Anal., 114(1993)106-231.

[JMR 2] J.-L. Joly, G. Métivier, and J. Rauch, Coherent and focusing multidimensional nonlinear geometric optics, Ann. de L'École Norm. Sup. de Paris, 28(1995)51-113.

[JMR 3] J.-L. Joly, G. Métivier, and J. Rauch, Generic rigorous asymptotic expansions for weakly nonlinear multidimensional oscillatory waves, Duke Math. J., 70(1993)373-404.

[JMR 4] J.-L. Joly, G. Métivier, and J. Rauch, Coherent nonlinear waves and the Wiener algebra, Ann. Inst. Fourier, 44(1994)167-196.

[JMR 5] J.-L. Joly, G. Métivier, and J. Rauch, Nonlinear oscillations beyond caustics, Comm. Pure and Appl. Math., 48(May 1996)443-529.

[JR] J.-L. Joly and J. Rauch, Justification of multidimensional single phase semilinear geometric optics, Trans. Amer. Math. Soc. 330(1992)599-625.

[Lax] P.D. Lax, Asymptotic solutions of oscillatory initial value problems, Duke Math. J., 24(1957)
627-645.

[Lo] H. A. Lorentz, *The Theory of Electrons*, Teubner 1908, reprinted by Dover 1952.

[MRS] A. Majda, R. Rosales and M. Schonbeck, A canonical system of integrodifferential equations arising in resonant nonlinear acoustics, Stu. Appl. Math., 79(1988)205-262.

[O] A. Owyoung, *The Origin of the Nonlinear Refractive Indices of Liquids and Glasses* , Ph.D dissertation, Caltech, 1971.

[R] J. Rauch, Lectures on geometric optics, IAS-Park City Summer School on Nonlinear Waves, July, 1995. Also available at `www.math.lsa.umich.edu/~rauch`

[Se] D. Serre, Oscillations non linéaires des systèmes hyperboliques : méthodes et résultats qualitatifs, Ann. Inst. Henri Poincaré, 8(1991)351-417.

[Sch] S. Schochet, Fast singular limits of hyperbolic partial differential equations, J.Diff.Eq. 114(1994)474-512.

[T] L. Tartar, Solutions oscillantes des équations de Carleman, Séminaire Goulaouic-Schwartz, École Polytechnique, Paris, année 1980-1981.

[Wa] T. Wagenmaker, Analytic solutions and resonant solutions of hyperbolic partial differential equations, PhD Thesis, University of Michigan, 1993.

[Wi] M. Williams, Resonant reflection of multidimensional semilinear oscillations, Comm. in Partial Diff. Eq. 18(1993)1901-1959.