

DIFFRACTIVE NONLINEAR GEOMETRIC OPTICS FOR SHORT PULSES*

DEBORAH ALTERMAN[†] AND JEFFREY RAUCH[‡]

Abstract. This paper considers the behavior of pulse-like solutions of length $\varepsilon \ll 1$ to semi-linear systems of hyperbolic partial differential equations on the time scale $t = O(1/\varepsilon)$ of diffractive geometric optics. The amplitude is chosen so that nonlinear effects influence the leading term in the asymptotics.

For pulses of larger amplitude so that the nonlinear effects are pertinent for times $t = O(1)$, accurate asymptotic solutions lead to transport equations similar to those valid in the case of wave trains (see [D. Alterman and J. Rauch, *J. Differential Equations*, 178 (2002), pp. 437–465]). The opposite is true here. The profile equation for pulses for $t = O(1/\varepsilon)$ is different from the corresponding equation for wave trains.

Formal asymptotics leads to equations for a leading term in the expansion and for correctors. The equations for the correctors are in general not solvable, being plagued by small divisor problems in the continuous spectrum. This makes the construction of accurate approximations subtle. We use low-frequency cutoffs depending on ε to avoid the small divisors.

Key words. pulses, diffraction, geometric optics, short wavelength asymptotics, hyperbolic partial differential equations

AMS subject classifications. 35B40, 35C20, 35L05, 35Q60

PII. S0036141002403584

1. Introduction. The simplest pulse-like solutions arise as plane wave solutions of constant coefficient homogeneous hyperbolic equations.

1.1. Linear plane waves.

Assumption 1.1 (symmetric hyperbolicity).

$$(1.1) \quad L(\partial_y) = \partial_t + \sum_{j=1}^d A_j \partial_{x_j},$$

where the coefficients A_j are constant $N \times N$ hermitian symmetric matrices.

The space-time variable is

$$y = (t, x) \in \mathbb{R}^{1+d} \quad \text{with dual variables} \quad (\tau, \xi).$$

If $f : \mathbb{R} \rightarrow \mathbb{C}^N$ is smooth and $\beta = (\tau, \xi)$, then the chain rule yields

$$L f(y.\beta) = L(\beta) f'(y.\beta).$$

Thus $L(f(y.\beta)) = 0$ when f' takes values in the nullspace of $L(\beta)$.

*Received by the editors March 5, 2002; accepted for publication (in revised form) November 18, 2002; published electronically DATE.

<http://www.siam.org/journals/sima/x-x/40358.html>

[†]Department of Applied Mathematics, University of Colorado, Boulder, CO 80309 (dalterman@earthlink.net). The research of this author was partially supported by the U.S. National Science Foundation under grant NSF-DMS-9810751.

[‡]Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 (rauch@umich.edu). The research of this author was partially supported by the U.S. National Science Foundation under grants NSF-DMS-9500823 and NSF-INT-9314095.

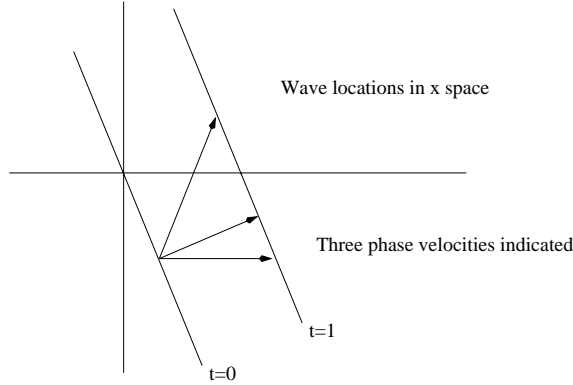


FIG. 1.1. Planar fronts and phase velocities.

Recall that the *characteristic variety*, $Char L$, is the set of $(\tau, \xi) \in \mathbb{R}^{1+d} \setminus 0$ satisfying the *dispersion relation* $\det L(\tau, \xi) = 0$. The characteristic variety is a real conic algebraic variety. For $\beta \in Char L$ one has the orthogonal decomposition

$$\mathbb{C}^N = \ker L(\beta) \oplus_{\perp} \text{range } L(\beta).$$

DEFINITION 1.2. For $\beta \in Char L$, $\pi = \pi(\beta)$ is the orthogonal projection of \mathbb{C}^N onto $\ker L(\beta)$. Define the partial inverse $Q(\beta)$ by

$$Q\pi = 0, \quad QL(\beta) = (I - \pi).$$

Then $u = f(y.\beta)$ is a *plane wave solution* of $Lu = 0$ when $\beta \in Char L$ and f satisfies the polarization $\pi(\beta)f = f$.

1.2. Plane pulses and group velocity. If, in addition,

$$f(s) \rightarrow 0 \quad \text{as} \quad s \rightarrow \pm\infty,$$

then the family of plane wave solutions

$$u^\varepsilon := f\left(\frac{y.\beta}{\varepsilon}\right)$$

describes pulses with planar wave fronts. If f has compact support, then the pulse u^ε is supported in an $O(\varepsilon)$ neighborhood of the hyperplane $y.\beta = 0$. The pulse cross section is given by the function $f^\varepsilon(s) := f(s/\varepsilon)$. The function f is called the *profile* of this pulse family. For profiles which tend to zero as $s \rightarrow \pm\infty$, the conditions $\pi(\beta)f = f$ and $\pi(\beta)f' = f'$ are equivalent.

At $t = 0$ (resp., $t = 1$) the pulse is supported near the planes $x.\xi = 0$ (resp., $x.\xi = -\tau$). This is indicated in Figure 1.1.

The phase is given by

$$y.\beta = t\tau + x.\xi = (x - \mathbf{v}t).\xi$$

for any velocity vector \mathbf{v} satisfying

$$\mathbf{v}.\xi = -\tau.$$

For any such \mathbf{v} , the pulse family is given by

$$h^\varepsilon(x - \mathbf{v}t), \quad \text{where} \quad h^\varepsilon(x) := f^\varepsilon(x, \xi) = f(x, \xi/\varepsilon).$$

The pulse family can be viewed as moving with any one of these *phase velocities*. Three such velocities are sketched in Figure 1.1. In dimension $d > 1$, the phase velocity is not uniquely determined.

In contrast Definition 1.4 below shows that the *group velocity* is well defined at smooth points β of the characteristic variety.

For $\xi \neq 0$ the points $(\tau, \xi) \in \text{Char } L$ which project to ξ are those points such that $-\tau$ is a real eigenvalue of $\sum \xi_j A_j$. Thanks to the hyperbolicity assumption this is a finite and nonempty set of points for each ξ , and so the variety has codimension 1. As such its singular points form a variety of codimension at least 2 so that most points of the variety are smooth in the sense of the next assumption.

Assumption 1.3 (smooth point of the characteristic variety). $\beta = (\tau_0, \xi_0)$ belongs to the characteristic variety, and there is a conic neighborhood of ξ_0 and a real analytic function $\tau(\xi)$ on that neighborhood so that on a conic neighborhood of β the variety is given by the equation $\tau = \tau(\xi)$.

DEFINITION 1.4. *For β a smooth point of the characteristic variety, the group velocity is defined by*

$$\text{group velocity} := \mathbf{v} := -\nabla_\xi \tau(\xi_0).$$

Since τ is homogeneous of degree 1 in ξ , the Euler homogeneity relation implies that

$$\xi \cdot \nabla_\xi \tau(\xi) = \tau(\xi).$$

This implies that the group velocity satisfies $\mathbf{v} \cdot \xi = -\tau$, the equation defining phase velocities. *The group velocity is the correct choice from among the possible phase velocities.*

1.3. Wave trains versus pulses. The geometric optics approximations which are most familiar concern the short wavelength limit of wave trains (see [24]). Wave trains and pulses are contrasted in Figure 1.2. Standard geometric optics yields equations for the envelope of wave trains. The methods go under the name of the slowly varying envelope approximation (SVEA) in science journals. A rule of thumb is that to use the SVEA the amplitude should not change more than 10% per wavelength. The wave train in Figure 1.2 is a borderline case for this rule. The rule of thumb suggests that one must have about ten to twenty wavelengths per pulse length before the SVEA is a reliable approximation.

For much shorter pulses like the one on the right in Figure 1.2 the SVEA is clearly inappropriate. Interest in short-pulse phenomena, which violate this slowly varying envelope assumption, has increased with the development of ultrafast lasers which produce few-cycle pulses. Rothenberg [25] clearly described the problems arising from treating short pulses as wave trains. Short-pulse solutions have been studied via full numerical simulation as in [28], [15], [16], [17], and [14]. A variety of asymptotic attacks are proposed and pursued in [13], [10], [21], [20], and [22]. In this paper the equation defining the leading order asymptotics is simple, and the approximation is proved to be accurate in the limit of small wavelength. Thus, from the above list only those which are consistent with our approximation can also be accurate. Only those whose equations are as simple can be competitive. It is our evaluation that with

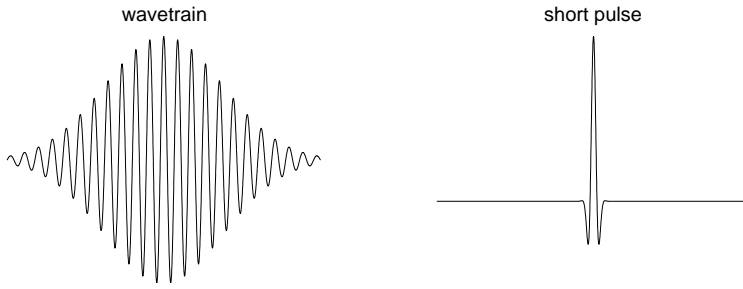


FIG. 1.2. *Example of a wave train and a short pulse.*

these two criteria in mind, a strong case can be made for our approach, but out of politeness we leave detailed comparison to the interested reader.

A difficulty in the study of pulses and wave trains is that the terms in the hyperbolic equation are of radically different magnitudes. Simply dropping the smaller ones and keeping the large ones usually leads to completely incorrect results. This is apparent in the simplest families of short pulses in \mathbb{R}^{1+1} ,

$$u^\varepsilon := e^{-t}f((x-t)/\varepsilon), \quad \partial_t u^\varepsilon + \partial_x u^\varepsilon + u^\varepsilon = 0.$$

The three terms of the equation are of sizes $O(1/\varepsilon)$, $O(1/\varepsilon)$, and $O(1)$. Dropping the relatively small $O(1)$ term yields the approximate solution $f((x-t)/\varepsilon)$, which is completely inaccurate.

A second difficulty in short-pulse asymptotics is that formally imitating the expansions of geometric optics generates equations for a leading term and correctors in an asymptotic expansion. Generically, the equations for the correctors cannot be solved. For the time scale $t = O(1)$, before the onset of diffractive effects, Yoshikawa [26], [27] showed that if one imposes physically unnatural assumptions guaranteeing that the corrector equations can be solved, then one does get an accurate description. In [5] we showed that the leading term is accurate without the unnatural assumption and extended the construction to the case of curved wavefronts. In a sequence of articles Carles and Rauch [7], [8], [9] studied the passage of spherical pulse solutions of semilinear wave equations across focal points.

For pulses on the scale of diffractive geometric optics, [1] includes some of the results of the present paper but notably does not prove a rate of convergence of the error as $\varepsilon \rightarrow 0$. There is also a study of pulses from Lannes' perspective of waves of broad spectrum (spectre large) in Barraillh and Lannes [6]. It is likely that the present analysis can be extended to nearly planar wavefronts as in the work of Dumas [12] for the diffractive wave train case. The present article contains the proofs of results described and used in [2] and [3].

Typical analytic expressions for the wave forms in Figure 1.2 are

$$\text{wave train :} \quad a(x) e^{ix_1/\varepsilon} \quad \text{with Fourier transform} \quad \hat{a}(\xi - (1/\varepsilon), 0);$$

$$\text{pulse :} \quad a(x_1/\varepsilon)b(x') \quad \text{with Fourier transform} \quad \varepsilon \hat{a}(\varepsilon \xi_1) \hat{b}(\xi').$$

The Fourier transform of the wave train is localized near $(1/\varepsilon, 0)$, which is called the carrier frequency in applications.

The Fourier transform of the pulse is spread over a box of dimensions $1/\varepsilon \times 1$ in (ξ_1, ξ') space. There is no carrier frequency. There is no exponential prefactor which

renders the quotient slowly varying. Some of the asymptotic approaches cited above are flawed because they insist on identifying a carrier frequency.

The approximations take the form

$$\begin{aligned} \text{wave train :} & \quad U(y, y.\beta/\varepsilon) & \text{with} & \quad U(y, \theta) \text{ periodic in } \theta; \\ \text{pulse :} & \quad U(y, y.\beta/\varepsilon) & \text{with} & \quad U(y, z) \rightarrow 0 \text{ as } z \rightarrow \infty. \end{aligned}$$

In both cases $\beta \in \text{Char } L$.

In the latter case the function $U(y, \cdot)$ represents the profile of the pulse. In the former case it gives the envelope of the wave train. The pulse approximation can be called the *slowly varying profile approximation* since the profiles vary on the scale $O(1)$, which is much longer than the pulse length $O(\varepsilon)$.

1.4. The basic problem. Consider the behavior for $t \sim 1/\varepsilon$ of solutions to a system of equations

$$(1.2) \quad L(\partial_y)u^\varepsilon + \Phi(u^\varepsilon) = 0, \quad u^\varepsilon(0, x) = \varepsilon^p f\left(x, \frac{x.\xi_0}{\varepsilon}\right),$$

where $\beta = (\tau_0, \xi_0)$ is a smooth point of the characteristic variety.

Assumption 1.5 (short pulse initial data). The function $f(x, z)$ satisfies

$$(1.3) \quad \forall N, \quad \langle \xi, \zeta \rangle^N \hat{f}(\xi, \zeta) \in L^\infty(\mathbb{R}^{d+1}).$$

This assumption is slightly stronger than the assumption $f(x, z) \in H^s(\mathbb{R}_{x,z}^{d+1})$ for all $s > 0$, used in [1] and [2], but is weaker than the Schwartz class. We will see that if one starts with f in the Schwartz class, then generically the pulse profile will not be Schwartz class for $t > 0$.

Assumption 1.6 (order J nonlinearity). The nonlinear function $\Phi(u)$ is of order $J \geq 2$ in the sense that for all $|\alpha| \leq J - 1$, $\partial^\alpha \Phi(0) = 0$. Denote by $\Phi_J(u)$ the homogeneous Taylor polynomial of degree J approximating $\Phi(u)$ near $u = 0$.

Assumption 1.7 (magnitude of the solution). The exponent p is chosen so that $p = 1/(J - 1)$. This insures that nonlinear effects become important on the time scale $t = O(1/\varepsilon)$.

To see that for waves of this amplitude it is reasonable that the nonlinear term is pertinent for times $t = O(1/\varepsilon)$ and not before, make the following back-of-an-envelope estimate. The nonlinear term is of size $\varepsilon^{pJ} = \varepsilon^{p+1}$. The accumulated effect of the nonlinear term for times $t = O(1/\varepsilon)$ is crudely estimated as

$$\frac{1}{\varepsilon} \varepsilon^{p+1} = \varepsilon^p.$$

Since ε^p is the size of our solution it is reasonable to expect the accumulated nonlinear effects to be important on these time scales.

Assumption 1.8 (polarization). The initial data f satisfy the polarization condition $\pi(\beta)f(x, z) = f(x, z)$.

DEFINITION 1.9. Define the scalar real second order homogeneous differential operator $P(\partial_x)$ by

$$(1.4) \quad P(\partial_x) := -\frac{1}{2} \sum_{l,m=1}^d \frac{\partial^2 \tau}{\partial \xi_l \partial \xi_m} \Big|_{\xi=\xi_0} \frac{\partial^2}{\partial x_l \partial x_m}.$$

With these assumptions and definitions, the approximate pulse-like solutions have the form

$$(1.5) \quad u_{\text{approx}}^\varepsilon = \varepsilon^p U_0 \left(\varepsilon t, t, x, \frac{\tau_0 t + \xi_0 \cdot x}{\varepsilon} \right), \quad \lim_{|z| \rightarrow \infty} U_0(T, t, x, z) = 0.$$

The slowly varying profile U_0 is polarized as usual, $\pi(\beta)U_0 = U_0$, and is determined from its initial data by the pair of evolution equations

$$(\partial_t + \mathbf{v} \cdot \partial_x) U_0 = 0, \quad \partial_{Tz} U_0 + P(\partial_x) U_0 + \pi(\beta) \partial_z \Phi_J(U_0) = 0.$$

The second equation, for which $T = 0$ is characteristic, is the pulse version of the nonlinear Schrödinger equation.

As these are the key equations that need to be solved in order to understand the behavior of solutions to (1.2), we pause briefly to discuss them. The first equation is handled by writing

$$(1.6) \quad U_0(T, t, x, z) = U_0(T, x - \mathbf{v}t, z).$$

The second equation is then equivalent to

$$(1.7) \quad \partial_{Tz} U_0 + P(\partial_x) U_0 + \pi(\beta) \partial_z \Phi_J(U_0) = 0.$$

On the face of it, this is a differential equation in the $d + 2$ variables T, x, z . In Proposition 4.1, it is shown that τ'' has rank $\leq d - 1$, so the differential operator has derivatives in at most $d + 1$ independent directions.

To see that (1.7) gives rise to a well-defined evolution, write it formally as

$$\partial_T U_0 + \partial_z^{-1} P(\partial_x) U_0 + \pi(\beta) \Phi_J(U_0) = 0.$$

The operator $\partial_z^{-1} P(\partial_x)$ is antisymmetric on the H^s , which for s large are invariant under Φ_J . Corollary 4.12 implies that for f as above, there is a $T_* \in]0, \infty]$ and a unique

$$U_0 \in C([0, T_*[; \cap_s H^s(\mathbb{R}_{x,z}^{d+1}))$$

satisfying (1.7) and the initial condition $U_0|_{T=0} = f$. If $T_* < \infty$, then for all $s > (d + 1)/2$,

$$\lim_{T \rightarrow T_*} \|U_0(T)\|_{H^s(\mathbb{R}_{x,z}^{d+1})} = \infty.$$

Having constructed U_0 , define an approximate solution by (1.6) and (1.5). Our main theorem asserts that the error in this approximation tends to zero as $\varepsilon \rightarrow 0$. To motivate a class of natural norms to measure this error, note that $u_{\text{approx}}^\varepsilon = O(\varepsilon^p)$. Differentiating u^ε costs a power of $1/\varepsilon$ but no worse, so one has

$$(1.8) \quad (\varepsilon \partial)^\alpha u_{\text{approx}}^\varepsilon = O(\varepsilon^p).$$

Denote by \mathbb{V} the $(d + 1)$ -dimensional space of constant coefficient vector fields. Choose a basis V_1, \dots, V_d of the d -dimensional subspace of fields which are tangent to the hyperplane $\{y \cdot \beta = 0\}$. Choose the basis so that V_1, \dots, V_{d-1} are tangent to $\{t = 0\}$. Then these $d - 1$ vectors are a basis for the constant fields on \mathbb{R}_x^d which are

tangent to $\{x.\xi_0 = 0\}$. Differentiating in the d directions V_1, \dots, V_d does not bring out a factor of $1/\varepsilon$, and one has

$$(1.9) \quad (V_1, \dots, V_d)^\alpha u_{\text{approx}}^\varepsilon = O(\varepsilon^p).$$

Choose a $(d+1)$ st field W , which completes the V_1, \dots, V_d to a basis of \mathbb{V} . Since $\xi \neq 0$, this vector field can be chosen tangent to $\{t = 0\}$. Define

$$(1.10) \quad V_{d+1} = \arctan(y.\beta) W.$$

This vector field vanishes on $\{y.\beta = 0\}$ and so is tangent to that hyperplane. Any smooth vector field tangent to this hyperplane is a linear combination of the V_j with smooth coefficients. Any smooth vector field on \mathbb{R}_x^d tangent to $\{x.\xi_0 = 0\}$ is a combination of $V_1, \dots, V_{d-1}, V_{d+1}|_{\{t=0\}}$ with smooth coefficients. One also has $V_{d+1} u_{\text{approx}}^\varepsilon = O(\varepsilon^p)$. Summarizing, one has for all $\alpha \in \mathbb{N}^{2(d+1)}$

$$(1.11) \quad (\varepsilon \partial_y, V_1, \dots, V_d, V_{d+1})^\alpha u_{\text{approx}}^\varepsilon = O(\varepsilon^p).$$

The next result is a straightforward consequence of our main result, Theorem 8.1. Note the technical point that the derivation V_d is not permitted in the error estimate.

THEOREM 1.10. *With the notation of the previous paragraphs, for any $\underline{T} < T_*$ there is an $\varepsilon_0 > 0$ so that for $0 < \varepsilon < \varepsilon_0$ problem (1.2) has a smooth solution $u^\varepsilon \in C^\infty(\{0 \leq T \leq \underline{T}/\varepsilon\} \times \mathbb{R}^d)$. The solution is well approximated by $u_{\text{approx}}^\varepsilon$ in the sense that for all $\alpha \in \mathbb{N}^{2(d+1)}$ there is a $C = C(\alpha)$ so that*

$$(1.12) \quad \|(\varepsilon \partial_y, V_1, \dots, V_{d-1}, V_{d+1})^\alpha (u^\varepsilon - u_{\text{approx}}^\varepsilon)\|_{L^\infty([0, \underline{T}/\varepsilon] \times \mathbb{R}^d)} \leq C \varepsilon^{p+\min(1/5, p)}.$$

Remark. Using the techniques of [5], one can show that there is a different family of exact solutions $u_{\text{ex}}^\varepsilon$ with error estimate including V_d , that is,

$$(1.13) \quad \|(\varepsilon \partial_y, V_1, \dots, V_{d-1}, V_d, V_{d+1})^\alpha (u_{\text{ex}}^\varepsilon - u_{\text{approx}}^\varepsilon)\|_{L^\infty([0, \underline{T}/\varepsilon] \times \mathbb{R}^d)} \leq C \varepsilon^{p+\min(1/5, p)}.$$

The initial data of the new family are small perturbations of the initial data for the family u^ε .

2. Formal asymptotics. Seek approximate solutions to the initial value problems with short-pulse initial data,

$$(2.1) \quad L u^\varepsilon + \Phi(u^\varepsilon) = 0, \quad u^\varepsilon(0, x) = \varepsilon^p f\left(x, \frac{x.\xi_0}{\varepsilon}\right), \quad p = \frac{1}{J-1}.$$

The initial function f is assumed to satisfy Assumptions 1.5 and 1.8 for a β satisfying Assumption 1.3. J is the order of the nonlinearity as in Assumption 1.6.

Motivated by its success in the analogous situation of wave train solutions for which f is periodic in z , a first attempt is to try to find a profile

$$(2.2) \quad U(\varepsilon, T, y, z) \sim \sum_{j=0}^{\infty} \varepsilon^j U_j(T, y, z) = \sum_{j=0}^{\infty} \varepsilon^j U_j(T, t, x, z),$$

where $U_j \rightarrow 0$ as $|z| \rightarrow \infty$, and

$$(2.3) \quad u^\varepsilon \sim \varepsilon^p U\left(\varepsilon, \varepsilon t, y, \frac{y.\beta}{\varepsilon}\right).$$

The chain rule implies that a sufficient condition guaranteeing that u^ε defined by $u^\varepsilon = U(\varepsilon, \varepsilon t, y, y \cdot \beta / \varepsilon)$ satisfies the differential equation $Lu^\varepsilon + \Phi(u^\varepsilon) = 0$ is that U satisfy

$$(2.4) \quad L\left((\varepsilon \partial_T, 0) + \partial_y + \frac{\beta}{\varepsilon} \partial_z\right) \varepsilon^p U(\varepsilon, T, y, z) + \Phi(\varepsilon^p U(\varepsilon, T, y, z)) = 0.$$

We pursue the less ambitious strategy, which is to satisfy

$$(2.5) \quad L\left((\varepsilon \partial_T, 0) + \partial_y + \frac{\beta}{\varepsilon} \partial_z\right) \varepsilon^p U(\varepsilon, T, y, z) + \Phi(\varepsilon^p U(\varepsilon, T, y, z)) \sim 0 \quad \text{as } \varepsilon \rightarrow 0,$$

in which case

$$(2.6) \quad Lu^\varepsilon + \Phi(u^\varepsilon) \sim 0 \quad \text{as } \varepsilon \rightarrow 0.$$

We take U to be a sum of only three terms, in which case the equivalence in (2.5) can be no smaller than $O(\varepsilon^{2p+1})$. Two crucial facts affect our implementation of this strategy:

1. A trio of equations, derived in section 3, determine U_0 from its initial data at $t = T = 0$, namely,

$$(2.7) \quad \begin{aligned} \pi(\beta) U_0 &= U_0, \\ \partial_t U_0 + \mathbf{v} \cdot \partial_x U_0 &= 0, \\ \partial_T \partial_z U_0 + P(\partial_x) U_0 + \pi(\beta) \partial_z \Phi_J(U_0) &= 0. \end{aligned}$$

The middle transport equation of (2.7) is solved by defining $U_0(T, x, z)$ as in (1.6) so that

$$(2.8) \quad U_0(T, x - \mathbf{v}t, z) = U_0(T, t, x, z).$$

2. The equations that one finds for the correctors U_1, U_2, \dots are *not in general solvable*. These equations involve the operator ∂_z^{-1} , which does not act well on a function whose Fourier transform with respect to z does not vanish at the origin, or equivalently, whose integral with respect to z is nonzero. For most choices of initial data, including those for which the transform vanishes on a neighborhood of zero at time $t = 0$, $\int U_0 dz$ does not vanish at later times, and hence the equations for the correctors are not solvable.

The second fact is the main difficulty this paper overcomes. In our study [5] of geometric optics before the onset of diffractive effects, a similar problem was encountered. In that case we were able to construct correctors which had a different form than the leading term. In the present case of diffractive geometric optics, we do not know how to find such modified correctors.

A crucial ingredient in the analysis is the representation of the exact solution u^ε in terms of an ‘‘exact profile’’ $\mathcal{V}(\varepsilon, t, x, \phi)$ as in [18] and [19] by setting

$$(2.9) \quad u^\varepsilon = \varepsilon^p \mathcal{V}\left(\varepsilon, t, x, \frac{x \cdot \xi_0}{\varepsilon}\right).$$

A key difference between \mathcal{V} and U is that in the phase variable slot $x \cdot \xi_0 / \varepsilon$ replaces $\beta \cdot y / \varepsilon$ for U . To maintain this distinction, the profile variable for which one inserts $x \cdot \xi_0 / \varepsilon$ is called ϕ and the profile variable associated with $\beta \cdot y / \varepsilon$ is z . The chain rule shows that u^ε defined by (2.9) satisfies $Lu^\varepsilon + \Phi(u^\varepsilon) = 0$ when

$$(2.10) \quad L\left(\partial_t, \partial_x + \frac{\xi_0}{\varepsilon} \partial_\phi\right) \varepsilon^p \mathcal{V} + \Phi(\varepsilon^p \mathcal{V}) = 0.$$

The exact profile $\mathcal{V}(\varepsilon, t, x, \phi)$ is then determined from (2.10) and the initial condition

$$(2.11) \quad \mathcal{V}(\varepsilon, 0, x, \phi) = f(x, \phi).$$

The existence and uniqueness of solutions to the initial value problem formed by (2.10) and (2.11) are examined in section 4. Our error estimates proceed by proving that

$$(2.12) \quad U_0\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) - \mathcal{V}(\varepsilon, t, x, \phi) \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0.$$

To establish (2.12), first note that the chain rule implies that if (2.5) holds, then as $\varepsilon \rightarrow 0$,

$$(2.13) \quad L\left(\partial_t, \partial_x + \frac{\xi_0}{\varepsilon} \partial_\phi\right) \varepsilon^p U\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) + \Phi\left(\varepsilon^p U\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right)\right) \sim 0.$$

Thus, if one has correctors U_1, U_2, \dots to the leading profile, then $U(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi)$ defines an accurate approximate solution of the equation for \mathcal{V} , and (2.12) would follow.

The difficulty is the absence of correctors—equivalently, the fact that we do not get (2.5). To circumvent this problem, we solve nearby problems with low-frequency cutoffs applied to the nonlinear term and the initial data. The cutoff problems propagate the property of having a Fourier transform with respect to z , which vanishes on a neighborhood of the origin.

Choose a cutoff function $\chi(\zeta) \in C^\infty(\mathbb{R})$ which vanishes for $|\zeta| < 1$ and is identically equal to 1 for $|\zeta| \geq 3/2$. Define

$$\chi^\delta := \chi^\delta(D_z) = \mathcal{F}_z^{-1} \chi(\zeta/\delta) \mathcal{F}_z,$$

where \mathcal{F}_z denotes the Fourier transform in z . Seek

$$(2.14) \quad U^\delta(\varepsilon, T, y, z) = U_0^\delta(T, y, z) + \varepsilon U_1^\delta(T, y, z) + \varepsilon^2 U_2^\delta(T, y, z)$$

as an approximate solution of the cutoff equation

$$(2.15) \quad \begin{aligned} L\left((\varepsilon \partial_T, 0) + \partial_y + \frac{\beta}{\varepsilon} \partial_z\right) \varepsilon^p U^\delta(\varepsilon, T, y, z) + \chi^\delta(D_z) \Phi(\varepsilon^p U^\delta(\varepsilon, T, y, z)) \\ = O(\varepsilon^{2p+1}) \end{aligned}$$

with initial data

$$(2.16) \quad U^\delta(0, 0, x, z) = \chi^\delta(D_z) f(x, z).$$

Then the main result is proved by showing that

$$(2.17) \quad U_0^\delta(T, t, x, z) - U_0(T, t, x, z) = O(\delta),$$

and that for $\delta = \varepsilon^{0.4}$,

$$(2.18) \quad U_0\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) - \mathcal{V}(\varepsilon, t, x, \phi) = O(\varepsilon^{\min\{p, 1/5\}}).$$

The proof has three main steps. First the approximate solution U^δ satisfying (2.15) and (2.16) is constructed in Proposition 4.3 and Corollary 4.12. Proposition 5.1 proves the convergence of U_0^δ to U_0 as in (2.17). Then Proposition 7.1 proves the error estimate (2.18). Propositions 5.1 and 7.1 combine to yield (2.12).

3. Derivation of the profile equations for U_j^δ . In this section we analyze (2.15). The construction of the correctors U_1^δ and U_2^δ works for $\delta > 0$. For $\delta = 0$ the analysis shows that the construction of such an approximate solution is in general not possible.

Define U^δ by (2.14). By convention set $U_{-1}^\delta = U_{-2}^\delta = U_3^\delta = U_4^\delta = 0$. Then computing the left-hand side of (2.15) yields

$$\varepsilon^p \left(\frac{1}{\varepsilon} L(\beta) \partial_z + L(\partial_y) + \varepsilon \partial_T \right) (U_0^\delta + \varepsilon U_1^\delta + \varepsilon^2 U_2^\delta) + \chi^\delta(D_z) \Phi(\varepsilon^p U^\delta).$$

Grouping by powers of ε yields

$$(3.1) \quad \sum_{j=-1}^{j=3} \varepsilon^{p+j} \{ \partial_T U_{j-1}^\delta + L(\partial_y) U_j^\delta + L(\beta) \partial_z U_{j+1}^\delta \} + \varepsilon^{p+1} \chi^\delta \Phi_J(U_0^\delta) + \chi^\delta \{ \Phi(\varepsilon^p U^\delta) - \Phi_J(\varepsilon^p U_0^\delta) \}.$$

We use formula (2.14) for times $t = O(1/\varepsilon)$. Thus if $U_1^\delta(T, t, x, z)$ grew linearly as $t \rightarrow \infty$, the term εU_1 would become as large as the U_0 term and would no longer be a small corrector. In order to represent a small correction it is necessary that U_1^δ grow sublinearly as $t \rightarrow \infty$. In fact, the correctors will be uniformly bounded in t, x . Sublinearity will play a crucial role in the derivation of the equations satisfied by the U_j .

3.1. Annihilating the ε^{p-1} term. This term is equal to

$$(3.2) \quad \varepsilon^{p-1} L(\beta) \partial_z U_0^\delta.$$

It is annihilated by imposing the polarization

$$(3.3) \quad U_0^\delta = \pi(\beta) U_0^\delta$$

from Definition 1.2. This is consistent with the polarization imposed on the initial data $f(x, z)$ in Assumption 1.8.

3.2. Annihilating the ε^p term. This term is equal to

$$(3.4) \quad \varepsilon^p \{ L(\partial_y) U_0^\delta + L(\beta) \partial_z U_1^\delta \}.$$

To annihilate (3.4), one annihilates in turn its image under $\pi(\beta)$ and then its image under $Q(\beta)$.

Multiplying (3.4) by $\pi(\beta)$ eliminates the U_1^δ term since $\pi(\beta)L(\beta) = 0$. Using the polarization of U_0^δ from (3.3), one finds

$$\pi(\beta)L(\partial_y)\pi(\beta)U_0^\delta = 0.$$

As shown in [11], whenever β is a smooth point of the characteristic variety,

$$\pi(\beta)L(\partial_y)\pi(\beta) = (\partial_t + \mathbf{v} \cdot \partial_x)\pi(\beta).$$

This identity yields the transport equation

$$(3.5) \quad \partial_t U_0^\delta + \mathbf{v} \cdot \partial_x U_0^\delta = 0.$$

Setting $Q(\beta)$ times the ε^p term equal to zero yields

$$(3.6) \quad (I - \pi(\beta)) \partial_z U_1^\delta = -Q(\beta) L(\partial_y) U_0^\delta.$$

This is the key troublesome equation for the correctors. In order for the equation to be solvable in $H^s(\mathbb{R}_{x,z}^{d+1})$, one needs

$$\frac{1}{\zeta} \mathcal{F}_z \left(Q(\beta) L(\partial_y) U_0^\delta \right)$$

to be square integrable near $\zeta = 0$. There is no reason to expect that this condition will be satisfied when $\delta = 0$.

However, (3.11) below defining U_0^δ for $\delta > 0$ implies that the Fourier transform of U_0^δ with respect to z vanishes on a neighborhood of $\zeta = 0$ as soon as it does so at $T = 0$. Thus we can solve (3.6) to find

$$(3.7) \quad (I - \pi(\beta)) U_1^\delta = -(\partial_z)^{-1} Q(\beta) L(\partial_y) U_0^\delta \quad \text{when } \delta > 0.$$

3.3. Annihilating the ε^{p+1} term. This term is equal to

$$(3.8) \quad \varepsilon^{p+1} \left(\partial_T U_0^\delta + L(\partial_y) U_1^\delta + L(\beta) \partial_z U_2^\delta + \chi^\delta \Phi_J(U_0^\delta) \right).$$

When $\delta > 0$, use (3.7) to write

$$U_1^\delta = \pi(\beta) U_1^\delta + (I - \pi(\beta)) U_1^\delta = \pi(\beta) U_1^\delta - (\partial_z)^{-1} Q(\beta) L(\partial_y) U_0^\delta.$$

Then setting $\pi(\beta)$ times (3.8) equal to zero yields

$$(3.9) \quad \begin{aligned} (\partial_t + \mathbf{v} \cdot \partial_x) \pi(\beta) U_1^\delta &= \pi L(\partial_y) \pi U_1^\delta \\ &= - \left(\partial_T U_0^\delta - \partial_z^{-1} \pi L(\partial_y) Q L(\partial_y) U_0^\delta + \pi \chi^\delta \Phi_J(U_0^\delta) \right). \end{aligned}$$

Thanks to (3.5) and (3.7), the right-hand side is constant along the integral curves of $\partial_t + \mathbf{v} \cdot \partial_x$. Therefore (3.9) implies that U_1^δ grows linearly along these straight lines unless the constant value is zero. As pointed out in the paragraph before section 3.1, such growth is unacceptable. Thus we must have

$$(3.10) \quad (\partial_t + \mathbf{v} \cdot \partial_x) \pi(\beta) U_1^\delta = 0$$

and

$$(3.11) \quad \partial_T \partial_z U_0^\delta - \pi L(\partial_y) Q L(\partial_y) U_0^\delta + \pi \chi^\delta \partial_z \Phi_J(U_0^\delta) = 0.$$

As shown in [11], at smooth points of the characteristic variety of $L(\partial_y)$,

$$(3.12) \quad \pi(\beta) L(\partial_y) Q(\beta) L(\partial_y) \pi(\beta) = -P(\partial_x) \pi(\beta),$$

where the second order differential operator $P(\partial_x)$ is defined in (1.4). Using (3.12) in (3.11), one gets the fundamental equation

$$(3.13) \quad \partial_T \partial_z U_0^\delta + P(\partial_x) U_0^\delta + \pi \chi^\delta \partial_z \Phi_J(U_0^\delta) = 0.$$

Note in passing that formulas (3.5), (3.7), and (3.10) imply that U_1^δ satisfies the transport equation

$$(3.14) \quad \partial_t U_1^\delta + \mathbf{v} \cdot \partial_x U_1^\delta = 0.$$

It remains to annihilate the product of $Q(\beta)$ with the ε^{p+1} term. This yields

$$(3.15) \quad (I - \pi(\beta))\partial_z U_2^\delta = -Q(\beta)\left(\partial_T U_0^\delta + L(\partial_y)U_1^\delta + \chi^\delta \Phi_J(U_0^\delta)\right).$$

Thanks to the polarization (3.3), $Q(\beta)\partial_T U_0^\delta = 0$. When $\delta > 0$, (3.15) is solvable and yields

$$(3.16) \quad (I - \pi(\beta))U_2^\delta = -Q(\beta)\partial_z^{-1}\left(L(\partial_y)U_1^\delta + \chi^\delta \Phi_J(U_0^\delta)\right).$$

The above calculations pose no constraints on $\pi(\beta)U_2^\delta$ and $\pi(\beta)U_1^\delta|_{t=0}$. For simplicity we set them equal to zero, and using (3.14) we find that

$$(3.17) \quad \pi(\beta)U_1^\delta = \pi(\beta)U_2^\delta = 0.$$

With these choices, (3.16) implies that

$$(3.18) \quad (\partial_t + \mathbf{v}\cdot\partial_x)U_j^\delta = 0 \quad \text{for } j = 0, 1, 2.$$

Corrector summary. Once U_0^δ is known with the Fourier transform vanishing on a neighborhood of $\zeta = 0$, the correctors are defined by

$$(3.19) \quad U_1^\delta = -\partial_z^{-1}Q(\beta)L(\partial_y)U_0^\delta,$$

$$(3.20) \quad U_2^\delta = -\partial_z^{-1}Q(\beta)\left(L(\partial_y)U_1^\delta + \chi^\delta \Phi_J(U_0^\delta)\right).$$

Residual summary. With the profiles defined in this way,

$$(3.21) \quad L\left((\varepsilon\partial_T, 0) + \partial_y + \frac{\beta}{\varepsilon}\partial_z\right)\varepsilon^p U^\delta + \chi^\delta(D_z)\Phi(\varepsilon^p U^\delta) \\ = \varepsilon^{p+2}\left(\varepsilon\partial_T U_2^\delta + L(\partial_y)U_2^\delta + \partial_T U_1^\delta\right) + \chi^\delta(D_z)\left[\Phi(\varepsilon^p U^\delta) - \Phi_J(\varepsilon^p U_0^\delta)\right].$$

4. Solvability of the profile equations for $U_0(T, x, z)$. U_0^δ must be constructed satisfying the equations

$$(4.1) \quad \begin{aligned} \pi(\beta)U_0^\delta &= U_0^\delta, \\ \partial_t U_0^\delta + \mathbf{v}\cdot\partial_x U_0^\delta &= 0, \\ \partial_T \partial_z U_0^\delta + P(\partial_x)U_0^\delta + \pi(\beta)\chi^\delta(D_z)\partial_z \Phi_J(U_0^\delta) &= 0. \end{aligned}$$

Equations (2.7) satisfied by U_0 are obtained by setting $\delta = 0$. Taking advantage of the middle equations of (2.7) and (4.1), define $U_0(T, x, z)$ and $U_j^\delta(T, x, z)$ by

$$U_0(T, t, x, z) := U_0(T, x - \mathbf{v}t, z), \quad U_j^\delta(T, t, x, z) := U_j^\delta(T, x - \mathbf{v}t, z).$$

The last equation in (2.7) is then equivalent to

$$(4.2) \quad \partial_T \partial_z U_0 + P(\partial_x)U_0 + \pi(\beta)\partial_z \Phi_J(U_0) = 0.$$

The *nonlinear diffractive pulse equation* (4.2) has $T = 0$ as a characteristic surface. In the wave train case one would have found an equation of nonlinear Schrödinger type at this stage. It too has $T = 0$ characteristic. In both cases the equation gives rise to a well-defined time evolution, at least locally in T .

Before proving this, we first prove that while the diffractive pulse equation appears to have $d + 2$ independent variables (t, x, z) , it actually involves one less direction of differentiation.

PROPOSITION 4.1. *The matrix $\partial^2\tau/\partial\xi_j\partial\xi_k$ has rank at most $d - 1$. In fact $\xi_0 \in \ker \partial^2\tau(\xi_0)/\partial\xi_j\partial\xi_k$.*

Proof. Since $\tau(\xi)$ is homogeneous of degree 1, it follows that for all j , $\partial\tau(\xi)/\partial\xi_j$ is homogeneous of degree zero. Therefore

$$\frac{d}{d\lambda} \frac{\partial\tau(\lambda\xi)}{\partial\xi_j} = 0.$$

Expanding the left-hand side using the chain rule yields

$$\sum_i \xi_i \frac{\partial^2\tau(\lambda\xi)}{\partial\xi_i\partial\xi_j} = 0.$$

Setting $\lambda = 1$ yields the desired result. \square

In coordinates so that $\xi_0 = (1, 0, \dots, 0)$ this implies that $\tau_{1,j} = \tau_{j,1} = 0$ so that

$$P(\partial_x) = -\frac{1}{2} \sum_{j,k=2}^d \frac{\partial^2\tau(1, 0, \dots, 0)}{\partial\xi_j\partial\xi_k} \frac{\partial^2}{\partial x_j\partial x_k}.$$

This decrease in the number of spatial dimensions decreases the complexity of the numerical implementation of the results of this paper.

The local solvability of the nonlinear diffractive pulse equation can be proved in the Sobolev spaces $H^s(\mathbb{R}^{1+d})$ with $s > (d + 1)/2$ by standard methods. The results would apply for general nonlinearities. A weakness is that for the Fourier transform of $U_0(T, x, z)$ with respect to the x, z variables one has only L^2 control locally.

The nonlinear term in the profile equation is always a polynomial. Because of this, we have the luxury of working in spaces related to the Wiener algebra which give us L^1 control of the Fourier transform. That in turn permits us to get L^∞ control of $\mathcal{F}_z U_0$ by estimates entirely on the Fourier side. These L^∞ estimates imply that $(I - \chi^\delta)U_0 = O(\delta)$ as $\delta \rightarrow 0$. The usual strategy to obtain sup norm estimates for Fourier transforms is to prove decay rates as $x, z \rightarrow \infty$. The argument completely on the Fourier side circumvents that avenue. We do not prove any decay rates beyond those implied by being in $\cap_s H^s(\mathbb{R}_{x,z}^{d+1})$.

DEFINITION 4.2. *The Wiener algebra $\mathbb{A}(\mathbb{R}^M)$ is the Banach space of tempered distributions on \mathbb{R}^M with the property that their Fourier transform belongs to $L^1(\mathbb{R}^M)$. The norm is the L^1 norm of the Fourier transform.*

Recall that for any $1 \leq p \leq \infty$ the map

$$L^1 \times L^p \ni f, g \rightarrow f * g \in L^p$$

is a continuous bilinear map from $L^1 \times L^p$ to L^p and

$$(4.3) \quad \|f * g\|_{L^p} \leq \|f\|_{L^1} \|g\|_{L^p}.$$

The inequality (4.3) with $p = 1$ shows that the map $U \rightarrow \Phi_J(U)$ maps bounded sets of \mathbb{A} to bounded sets of \mathbb{A} . To study the continuity of the map note that the difference $\Phi_J(U) - \Phi_J(V)$ can be expressed as

$$\Phi_J(U) - \Phi_J(V) = \sum_i^d P_i(U, V) (U_i - V_i)$$

with polynomials P_i of degree less than J . Then inequality (4.3) shows that the map $U \rightarrow \Phi_J(U)$ is uniformly Lipschitzian on bounded subsets of \mathbb{A} to \mathbb{A} .

The initial value problem for the profile equation (4.2) with initial value

$$U|_{T=0} = G(x, z)$$

is equivalent to the integral identities

$$(4.4) \quad \widehat{U}(T) = e^{-iP(\xi)T/\zeta} \widehat{G} + \int_0^T e^{-iP(\xi)(T-\sigma)/\zeta} \pi(\beta) \mathcal{F}(\Phi_J(U(\sigma))) d\sigma$$

for $0 \leq T \leq \underline{T}$.

The multipliers $e^{iP(\xi)t/\zeta}$ have modulus one so they define isometries on $\mathbb{A}(\mathbb{R}_{x,z}^{d+1})$. This, together with the uniform Lipschitzian property, is enough to make Picard's classical existence proof work, yielding the following result.

PROPOSITION 4.3. *For each $G \in \mathbb{A}$ there is a $T_* = T_*(G) \in]0, \infty]$ and a unique maximal solution $U \in C([0, T_*[; \mathbb{A})$ to the profile equation (4.2), which in addition satisfies the initial condition $U|_{T=0} = G$. The time T_* is uniformly strictly positive on bounded subsets of \mathbb{A} , and if $T_* < \infty$, then*

$$(4.5) \quad \lim_{T \rightarrow T_*} \|U(T)\|_{\mathbb{A}} = \infty.$$

The next result is a regularity theorem which asserts that if the initial data lies in a smaller Banach space \mathbb{B} , then the maximal solution is a continuous function with values in \mathbb{B} .

DEFINITION 4.4. *A Banach space $\mathbb{B} \subset \mathbb{A}$ is admissible if it has the following three properties:*

1. *The inclusion map $\mathbb{B} \rightarrow \mathbb{A}$ is continuous.*
2. *The map $U \rightarrow \Phi_J(U)$ maps \mathbb{B} to itself and is uniformly Lipschitzian on subsets of \mathbb{B} which are bounded in \mathbb{A} .*
3. *For $T \neq 0$, the Fourier multipliers $e^{iTP(\xi)/\zeta}$ are isometries from \mathbb{B} to itself.*

The following are examples of admissible Banach spaces.

Example 4.5. If $1 < p \leq \infty$, then $\mathbb{B} := \{U \in \mathbb{A} : \widehat{U} \in L^p\}$ is admissible.

Example 4.6. If $s > (d+1)/2$, then $H^s(\mathbb{R}^{d+1})$ is admissible.

Example 4.7. If \mathbb{B}_1 and \mathbb{B}_2 are admissible, then so is the intersection $\mathbb{B}_1 \cap \mathbb{B}_2$.

Proof for Example 4.5. Only property 2 in the definition is not immediate. One needs to prove that for every $R > 0$ there is a constant C so that if $\|U\|_{\mathbb{A}} \leq R$ and $\|V\|_{\mathbb{A}} \leq R$, then

$$(4.6) \quad \|\Phi_J(U) - \Phi_J(V)\|_{\mathbb{B}} \leq C \|U - V\|_{\mathbb{B}}.$$

Taylor's theorem implies that

$$\Phi_J(U) - \Phi_J(V) = \Psi(U, V)(U - V),$$

where Ψ is a matrix-valued homogeneous polynomial of degree $J - 1$.

To estimate the L^p norm of the Fourier transform use Young's inequality

$$\begin{aligned} \|\mathcal{F}(\Phi_J(U) - \Phi_J(V))\|_{L^p} &\leq \|\mathcal{F}(\Psi(U, V))\|_{L^1} \|\mathcal{F}(U - V)\|_{L^p} \\ &\leq C(R) \|U - V\|_{\mathbb{B}}. \quad \square \end{aligned}$$

PROPOSITION 4.8. *If \mathbb{B} is admissible and $G \in \mathbb{B}$, then the maximal solution found in Proposition 4.3 satisfies*

$$(4.7) \quad U \in C([0, T_*[; \mathbb{B})).$$

Proof. From the admissibility properties one easily demonstrates using Picard's method that the integral equation (4.4) has a maximal solution

$$U \in C([0, T^*(G)[; \mathbb{B})),$$

and if $T^* < \infty$, then

$$(4.8) \quad \lim_{T \rightarrow T^*} \|U(T)\|_{\mathbb{B}} = \infty.$$

Since this solution is continuous with values in \mathbb{A} it follows that $T^*(G) \leq T_*(G)$. The result of the proposition follows from establishing the inequality $T^*(G) \geq T_*(G)$.

The proof is indirect. We suppose on the contrary that $T^*(G) < T_*(G)$ and derive a contradiction.

If $T^*(G) < T_*(G)$, then U is continuous on $[0, T^*]$ with values in \mathbb{A} , and so there is an $R < \infty$ so that

$$\|U(T)\|_{\mathbb{A}} \leq R \quad \text{for } 0 \leq T \leq T^*.$$

Taking the \mathbb{B} norm of (4.4) and using the last two properties from the definition of admissibility yields

$$\|U(T)\|_{\mathbb{B}} \leq \|G\|_{\mathbb{B}} + \int_0^T C \|U(\sigma)\|_{\mathbb{B}} d\sigma \quad \text{for } 0 \leq T < T^*.$$

Gronwall's inequality implies that

$$\|U(T)\|_{\mathbb{B}} \leq \|G\|_{\mathbb{B}} e^{CT} \quad \text{for } 0 \leq T < T^*.$$

In particular, (4.8) is violated. This contradiction proves the proposition. \square

Our main existence result is a corollary of Propositions 4.9 and 4.10 below.

PROPOSITION 4.9. *Define \mathbb{B} to be the closed subspace of $\mathbb{A}(\mathbb{R}_{x,z}^{d+1})$ consisting of functions U such that (i) $\widehat{U} \in L^\infty(\mathbb{R}_{\xi,\zeta}^{d+1})$, and (ii) for all $\mu > 0$, \widehat{U} is uniformly continuous on $\{|\zeta| \geq \mu\}$. Then \mathbb{B} is admissible.*

Proof. \mathbb{B} is a closed subspace of Example 4.5 with $p = \infty$. Thus to prove that \mathbb{B} is admissible it suffices to show that Φ_J maps \mathbb{B} to itself. In fact, more is true. If $\widehat{U} \in L^1 \cap L^\infty$ (which is true if $U \in \mathbb{B}$), then $\mathcal{F}_{x,z}(\Phi_J(U))$ is bounded and uniformly continuous, which implies that $\Phi_J(U) \in \mathbb{B}$.

To prove the stronger assertion of the last sentence, write Φ_J as a sum of terms each of which is a product of a monomial of order $J-1$ and a monomial of order 1. The Fourier transform of the first factor belongs to L^1 and the Fourier transform of the second belongs to L^∞ . The desired result follows from the fact that the convolution of an element of L^1 with an element of L^∞ is uniformly continuous. \square

PROPOSITION 4.10. *If $0 \leq m \in \mathbb{Z}$ and $1 \leq p \leq \infty$, then the subspace*

$$\mathbb{B}^{m,p} := \left\{ U \in \mathbb{A} : \langle \eta \rangle^m \widehat{U}(\eta) \in L^p(\mathbb{R}_\eta^N) \right\}$$

is admissible.

Proof. This proof follows [23]. The only sticky point is estimate (4.6). Toward that end one must show that for K -fold products one has

$$(4.9) \quad \left\| \mathbf{U}_j \right\|_{\mathbb{A}} \leq R \quad \Rightarrow \quad \left\| \langle \eta \rangle^m \widehat{\mathbf{U}}_1 * \widehat{\mathbf{U}}_2 * \cdots * \widehat{\mathbf{U}}_K \right\|_{L^p} \leq C(K, R) \sum_j \left\| \langle \eta \rangle^m \widehat{\mathbf{U}}_j \right\|_{L^p}.$$

Written out, this becomes

$$(4.10) \quad \left\| \int \langle \eta \rangle^m \widehat{\mathbf{U}}_1(\eta - \eta_1) \widehat{\mathbf{U}}_2(\eta_1 - \eta_2) \cdots \widehat{\mathbf{U}}_K(\eta_K) \, d\eta_1 \, d\eta_2 \cdots d\eta_{K-1} \, d\eta_K \right\|_{L^p} \\ \leq C(K, R) \sum_j \left\| \langle \eta \rangle^m \widehat{\mathbf{U}}_j \right\|_{L^p}.$$

Note that

$$\eta = (\eta - \eta_1) + (\eta_1 - \eta_2) + \cdots + (\eta_{K-1} - \eta_K) + \eta_K,$$

so

$$|\eta| \leq |\eta - \eta_1| + |\eta_1 - \eta_2| + \cdots + |\eta_{K-1} - \eta_K| + |\eta_K|.$$

Define $\eta_0 := \eta, \eta_{K+1} := 0$, so the summands on the right are equal to $|\eta_{j-1} - \eta_j|$ for $1 \leq j \leq K+1$. The integral in (4.10) is split into $K+1$ pieces, the integrals over sets

$$E(j) := \left\{ (\eta, \eta_1, \dots, \eta_K) : |\eta_{j-1} - \eta_j| = \max_{1 \leq k \leq K} \{ |\eta_{k-1} - \eta_k| \} \right\}.$$

The sets E_j overlap in measure zero sets, so it suffices to show that

$$\left\| \int_{E(j)} \langle \eta \rangle^m \widehat{\mathbf{U}}_1(\eta - \eta_1) \widehat{\mathbf{U}}_2(\eta_1 - \eta_2) \cdots \widehat{\mathbf{U}}_K(\eta_K) \, d\eta_1 \, d\eta_2 \cdots d\eta_{K-1} \, d\eta_K \right\|_{L^p} \\ \leq C(K, R) \sum_j \left\| \langle \eta \rangle^m \widehat{\mathbf{U}}_j \right\|_{L^p}.$$

On $E(j)$, $|\eta| \leq K|\eta_{j-1} - \eta_j|$ so

$$\int_{E(j)} \langle \eta \rangle^m |\widehat{\mathbf{U}}_1(\eta - \eta_1)| \cdot |\widehat{\mathbf{U}}_2(\eta_1 - \eta_2)| \cdots |\widehat{\mathbf{U}}_K(\eta_K)| \, d\eta_1 \, d\eta_2 \cdots d\eta_{K-1} \, d\eta_K \\ \leq C(K) \int_{E(j)} \langle \eta_{j-1} - \eta_j \rangle^m |\widehat{\mathbf{U}}_1(\eta - \eta_1)| \cdot |\widehat{\mathbf{U}}_2(\eta_1 - \eta_2)| \\ \cdots |\widehat{\mathbf{U}}_K(\eta_K)| \, d\eta_1 \, d\eta_2 \cdots d\eta_{K-1} \, d\eta_K.$$

Young's inequality bounds the L^p norm of the integral on the right by

$$\left\| \langle \eta \rangle^m \widehat{\mathbf{U}}_j \right\|_{L^p} \prod_{k \neq j} \left\| \widehat{\mathbf{U}}_k \right\|_{L^1} = C(R) \left\| \langle \eta \rangle^m \widehat{\mathbf{U}}_j \right\|_{L^p}.$$

This completes the proof. \square

DEFINITION 4.11. \mathcal{B} is the Fréchet space of tempered distributions $\mathcal{V}(x, z)$ so that

1. $\widehat{\mathcal{V}}(\xi, \zeta) \in L^\infty(\mathbb{R}^{d+1})$ and for every $\mu > 0$ is uniformly continuous on the set $\{(\xi, \zeta) : |\zeta| \geq \mu\}$.

2. For every nonnegative integer m

$$\langle \xi, \zeta \rangle^m \widehat{V}(\xi, \zeta) \in L^\infty(\mathbb{R}^{d+1}).$$

Combining Propositions 4.9 and 4.10 shows that \mathcal{B} is the intersection of admissible spaces. By Example 4.7, this implies the following corollary.

COROLLARY 4.12. *Suppose that $U \in C([0, T_*[; \mathbb{A}))$ is a maximal solution of the profile equation (4.2) and that $U|_{T=0} \in \mathcal{B}$. Then*

$$(4.11) \quad U \in C([0, T_*[; \mathcal{B}).$$

This corollary gives us more than enough control on the leading profile to carry out our analysis. The most interesting aspect is the sup norm control near $\zeta = 0$ without continuity.

5. Construction of U_0^δ . A perturbation argument is the key to solving the $\delta > 0$ equations (4.1). The third equation in (4.1) is equivalent to

$$(5.1) \quad \partial_T \partial_z U_0^\delta + P(\partial_x) U_0^\delta + \pi(\beta) \chi^\delta(D_z) \partial_z \Phi_J(U_0^\delta) = 0.$$

Multiplying (4.2) by $\chi^\delta(D_z)$ yields

$$(5.2) \quad \partial_T \partial_z \chi^\delta(D_z) U_0 + P(\partial_x) \chi^\delta(D_z) U_0 + \pi(\beta) \chi^\delta(D_z) \partial_z \Phi_J(U_0) = 0.$$

This equation resembles (5.1), as demonstrated in the next proposition, which shows that the solution of (5.1) can be obtained as a small perturbation of $\chi^\delta(D_z) U_0$.

PROPOSITION 5.1. *Suppose that $U_0 = \pi(\beta) U_0 \in C([0, \underline{T}]; \cap_m \mathbb{B}^{m, \infty})$ satisfies (4.2). Then there is a $\delta_0 > 0$ so that for $0 < \delta < \delta_0$ the initial value problem defined by (5.2) with initial condition*

$$(5.3) \quad U_0^\delta|_{T=0} = \chi^\delta(D_z) U_0|_{T=0}$$

has a unique solution $U_0^\delta \in C([0, \underline{T}]; \cap_m \mathbb{B}^{m, \infty})$, and for all $1 \leq q < \infty$ and $0 \leq m < \infty$,

$$\sup_{0 \leq T \leq \underline{T}} \|U_0^\delta(T) - \chi^\delta U_0(T)\|_{\mathbb{B}^{m, q}(\mathbb{R}_{x, z}^{d+1})} = O(\delta^{1/q}).$$

Furthermore the Fourier transform $\mathcal{F}_z(U_0^\delta)$ vanishes identically on $|\zeta| \leq \delta$.

Proof of Proposition 5.1. Begin with the proof that U_0^δ has a Fourier transform with respect to z vanishing on $|\zeta| \leq \delta$. For any $\gamma(\zeta) \in C_0^\infty(\mathbb{R})$ supported on $|\zeta| \leq 1$ define $\gamma^\delta(\zeta) := \gamma(\zeta/\delta)$. It suffices to show that $\mathcal{F}_z \gamma^\delta(D_z) U_0^\delta(T) = 0$.

The choice of γ implies that $\gamma^\delta \chi^\delta = 0$. Thus multiplying (5.1) by $\gamma^\delta(D_z)$ annihilates the nonlinear term. This implies that $\gamma^\delta(D_z) U_0^\delta \in C([0, T_*[; \cap_s H^s)$ satisfies

$$(5.4) \quad \left(\partial_T \partial_z + P(\partial_x) \right) \gamma^\delta(D_z) U_0^\delta = 0.$$

In addition $\gamma^\delta(D_z) U_0^\delta$ vanishes when $T = 0$.

It follows from the basic H^s conservation law for the linear diffractive pulse equation [4] that for all s and all $T \in [0, T_*[$,

$$(5.5) \quad \|\gamma^\delta(D_z) U_0^\delta(T)\|_{H^s(\mathbb{R}_{x, z}^{d+1})} = \|\gamma^\delta(D_z) U_0^\delta(0)\|_{H^s(\mathbb{R}_{x, z}^{d+1})} = 0.$$

Identity (5.5) implies the second assertion of the proposition.

The strategy for proving the $O(\delta^{1/q})$ estimate in the proposition is to construct U_0^δ as a perturbation of $\chi^\delta(D_z)U_0$. Define the perturbation W^δ by

$$(5.6) \quad W^\delta := U_0^\delta - \chi^\delta(D_z)U_0.$$

Subtract (5.1) from (5.2) to show that U_0^δ is a solution if and only if W^δ satisfies the initial value problem

$$(5.7) \quad \begin{aligned} \partial_T \partial_z W^\delta + P(\partial_x) W^\delta + \pi(\beta) \chi^\delta(D_z) \partial_z (\Phi_J(U_0^\delta) - \Phi_J(U_0)) &= 0, \\ W^\delta|_{T=0} &= 0. \end{aligned}$$

Note that

$$\Phi_J(U_0^\delta) - \Phi_J(U_0) = \Phi_J(\chi^\delta U_0 + W^\delta) - \Phi_J(\chi^\delta U_0 + (I - \chi^\delta)U_0).$$

Since $\chi^\delta U_0(T)$ is bounded in the admissible subspace $\mathbb{B}^{m,\infty}$ uniformly for all $0 < \delta \leq 1$ and $0 \leq t \leq \underline{T}$, the second condition in the definition of admissibility implies that as long as $\|W^\delta\|_{\mathbb{B}^{m,\infty}} \leq 1$,

$$(5.8) \quad \begin{aligned} \|\Phi_J(U_0^\delta) - \Phi_J(U_0)\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} &\leq C(m,q) \left(\|W^\delta(T)\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} \right. \\ &\quad \left. + \|(I - \chi^\delta)U_0(T)\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} \right). \end{aligned}$$

Fix $m \geq 0$ and $1 \leq q \leq \infty$. If $\|W^\delta(T)\|_{\mathbb{B}^{m,q}} \leq 1$ for $0 \leq T \leq \underline{T}$, define $T_* = T_*(m, q, \delta) = \underline{T}$. Otherwise define

$$T_*(m, q, \delta) := \inf\{T \in [0, \underline{T}] : \|W^\delta(T)\|_{\mathbb{B}^{m,q}} = 1\}.$$

The homogeneous linear diffractive pulse equation generates a unitary group on each $\mathbb{B}^{m,q}$. The inhomogeneous version of this estimate implies that for $T \in [0, T_*]$,

$$(5.9) \quad \begin{aligned} \|W^\delta(T)\|_{\mathbb{B}^{m,q}} &\leq \int_0^T C(m) \left(\|W^\delta(\sigma)\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} \right. \\ &\quad \left. + \|(I - \chi^\delta)U_0(\sigma)\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} \right) d\sigma. \end{aligned}$$

Gronwall's inequality then shows that

$$(5.10) \quad \|W^\delta(T)\|_{\mathbb{B}^{m,q}} \leq C(m) \int_0^T \|(I - \chi^\delta)U_0(\sigma)\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} dt e^{C(m)T}.$$

To proceed we need the following lemma.

LEMMA 5.2. *For any $m \geq 0$, $1 \leq q < \infty$, and $M > q + (d+1)/q$ there is a constant $C = C(m, q, M)$ so that for all $\delta > 0$ and all $W \in \mathbb{B}^{M,\infty}$,*

$$\|(I - \chi^\delta)W\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})} \leq C \delta^{1/q} \|W\|_{\mathbb{B}^{M,\infty}}.$$

Remark. In contrast note that for $W \in \cap_s H^s$, $\|(I - \chi^\delta)W\|_{H^s} = o(1)$ as $\delta \rightarrow 0$, but there is no rate of convergence.

Proof of Lemma. By definition

$$\|(I - \chi^\delta)W\|_{\mathbb{B}^{m,q}(\mathbb{R}^{d+1})}^q = \int |1 - \chi(\zeta/\delta)|^q |\widehat{W}(\xi, \zeta)|^q \langle \xi, \zeta \rangle^{mq} d\xi d\zeta.$$

Use the estimate

$$|\widehat{W}(\xi, \zeta)| \leq \langle \xi, \zeta \rangle^{-M} \|\widehat{W}\|_{\mathbb{B}^{M, \infty}}$$

to find

$$\|(I - \chi^\delta)W\|_{\mathbb{B}^{m, q}}^q \leq \|\widehat{W}\|_{\mathbb{B}^{M, \infty}}^q \int |1 - \chi(\zeta/\delta)|^q \langle \xi, \zeta \rangle^{-Mq} \langle \xi, \zeta \rangle^{mq} d\xi d\zeta.$$

The integral on the right is over $|\zeta| \leq \delta$, so for $Mq > mq + d + 1$, the integral is $O(\delta)$, which proves the lemma. \square

This lemma implies that

$$\int_0^{\underline{T}} \|(I - \chi^\delta)U_0(T)\|_{\mathbb{B}^{m, q}(\mathbb{R}^{d+1})} dt = O(\delta^{1/q}).$$

Thus, one can choose $\delta_0(m, q)$ so that for $0 < \delta < \delta_0$,

$$C(m) \int_0^{\underline{T}} \|(I - \chi^\delta)U_0(T)\|_{\mathbb{B}^{m, q}(\mathbb{R}^{d+1})} dt e^{C(m)\underline{T}} < \frac{1}{2}.$$

Then (5.10) shows that if $T_* < \underline{T}$, then

$$\|W^\delta(T)\|_{\mathbb{B}^{m, q}} \leq \frac{1}{2}$$

for $0 < t < T_*$. Since this contradicts the definition of T_* we conclude that $T_* = \underline{T}$ and that (5.10) holds for $0 < T < \underline{T}$. In particular, the evolution equation for U_0^δ is solvable up to \underline{T} .

In addition, estimate (5.10) implies the $O(\delta^{1/q})$ convergence rate for the value m, q fixed at the start. Since m, q is arbitrary, the proof of Proposition 5.1 is complete. \square

Combining the convergence results of Proposition 5.1 and Lemma 5.2 yields

$$(5.11) \quad \sup_{0 \leq t \leq \underline{T}} \|U_0(T) - U_0^\delta(T)\|_{\mathbb{B}^{m, q}(\mathbb{R}_{x, z}^{d+1})} = O(\delta^{1/q}).$$

Note that the smallest upper bound occurs for the case $q = 1$ corresponding to the Wiener algebra.

6. Estimate for the residual. Suppose that \underline{T} is smaller than the maximal existence time for U_0 in the sense that a solution of (4.2) is known in the space $C([0, \underline{T}]; \cap_m \mathbb{B}^{m, \infty}(\mathbb{R}_{x, z}^{d+1}))$. Then Proposition 5.1 shows that for $0 < \delta < \delta_0$, $U_0^\delta(T, t, x, z) = U_0^\delta(T, x - \mathbf{v}t, z)$ exists on $[0, \underline{T}]$ and $\mathcal{F}_z U_0^\delta$ vanishes on a neighborhood of $\zeta = 0$. For $\delta > 0$, the equations for the correctors U_1^δ, U_2^δ are solvable, so

$$U^\delta(\varepsilon, T, t, x, z) := U_0^\delta + \varepsilon U_1^\delta + \varepsilon^2 U_2^\delta$$

is well defined for $0 \leq t \leq \underline{T}/\varepsilon$. The residual equation (3.21) is then

$$(6.1) \quad \begin{aligned} R^\delta(\varepsilon, T, t, x, z) &:= L\left((\varepsilon \partial_T, 0) + \partial_y + \frac{\beta}{\varepsilon} \partial_z\right) \varepsilon^p U^\delta + \chi^\delta(D_z) \Phi(\varepsilon^p U^\delta) \\ &= \varepsilon^{p+1} \left(\varepsilon^2 \partial_T U_2^\delta + \varepsilon L(\partial_y) U_2^\delta + \varepsilon \partial_T U_1^\delta \right) + \chi^\delta(D_z) \left[\Phi(\varepsilon^p U^\delta) - \Phi_J(\varepsilon^p U_0^\delta) \right]. \end{aligned}$$

This section is devoted to estimates for the right-hand side of (6.1). There are at least two subtle points. The first is that though the residual is formally $O(\varepsilon^{2p+1})$ it involves correctors which blow up as $\delta \rightarrow 0$. Care must be taken about small values of δ . The second point is that the residual involves the smooth function Φ , which need not be polynomial. Therefore, the Wiener algebra need not be invariant. The estimates are therefore done in the scale of Sobolev spaces.

PROPOSITION 6.1. *Suppose that U^δ and \underline{T} are as above, δ_0 is as in Proposition 5.1, and $s > (d+1)/2$. Then there is a constant $C = C(s)$ so that for all*

$$(6.2) \quad 0 \leq t < \infty \quad \cap \quad 0 \leq T \leq \underline{T} \quad \cap \quad 0 < \delta \leq \delta_0 \quad \cap \quad 0 \leq \varepsilon \leq \delta,$$

one has

$$(6.3) \quad \|R^\delta(\varepsilon, T, t, x, z)\|_{H^s(\mathbb{R}_{x,z}^{d+1})} \leq C \varepsilon^{p+1} \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p \right).$$

Proof. The first step is to estimate the size of the correctors U_1^δ, U_2^δ . The formulas for these functions involve U_0^δ and, most importantly, the operator ∂_z^{-1} . The formula for U_1^δ involves ∂_z^{-1} , while the formula for U_2^δ involves ∂_z^{-2} since it has a term with ∂_z^{-1} applied to U_1^δ . The application of the operator ∂_z^{-1} introduces a factor $1/\delta$ in estimates since the support of U_j^δ is in $|\zeta| > \delta$. The boundedness of the family U_0^δ in $\mathbb{B}^{m,\infty}$ yields the following estimates for the correctors for $0 < \delta \leq \delta_0$:

$$(6.4) \quad \forall m, \forall 0 \leq t < \infty, \quad \sup_{0 \leq T \leq \underline{T}} \|U_j^\delta(T, t, x, z)\|_{\mathbb{B}^{m,\infty}(\mathbb{R}_{x,z}^{d+1})} \leq \frac{C(m)}{\delta^j}.$$

Inserted in the definition of U^δ , this estimate proves that for $t \in \mathbb{R}$ and $0 \leq T \leq \underline{T}$,

$$\|U^\delta(T, t)\|_{\mathbb{B}^{m,\infty}(\mathbb{R}_{x,z}^{d+1})} \leq C(m) \left(1 + \frac{\varepsilon}{\delta} + \frac{\varepsilon^2}{\delta^2} \right).$$

Thus $U^\delta(T, t)$ is uniformly bounded in $\mathbb{B}^{m,\infty}$ for the parameter range (6.2). This is a key element in the estimate of the second term on the right-hand side of (6.1).

For the right-hand side of (6.1) we also need an estimate for $\varepsilon \partial_T U_1^\delta$ and $\varepsilon^2 \partial_T U_2^\delta$. Start with an estimate for the T derivative of U_0^δ . The evolution equation (5.1) yields

$$\partial_T U_0^\delta = -\partial_z^{-1} P(\partial_x) U_0^\delta - \pi(\beta) \chi^\delta(D_z) \Phi_J(U_0^\delta).$$

Together with the uniform boundedness of U_0^δ this yields

$$(6.5) \quad \|\partial_T U_0^\delta\|_{\mathbb{B}^{m,\infty}} \leq \frac{C(m)}{\delta}.$$

The factor $1/\delta$ comes from the norm of ∂_z^{-1} acting on functions with spectrum in $|\zeta| \geq \delta$. Differentiating (3.19) and (3.20) with respect to T yields

$$\partial_T U_1^\delta = -\partial_z^{-1} Q(\beta) L(\partial_y) \partial_T U_0^\delta$$

and

$$\partial_T U_2^\delta = -\partial_z^{-1} Q(\beta) \left(L(\partial_y) \partial_T U_1^\delta + \chi^\delta(D_z) \pi(\beta) \Phi'_J(U_0^\delta) \partial_T U_0^\delta \right).$$

Using estimate (6.5) in the equations above yields in turn

$$\|\partial_T U_1^\delta\|_{\mathbb{B}^{m,\infty}} \leq \frac{C(m)}{\delta^2} \quad \text{and} \quad \|\partial_T U_2^\delta\|_{\mathbb{B}^{m,\infty}} \leq \frac{C(m)}{\delta^3}.$$

Inserting these estimates into the first term on the right-hand side of (6.1) and using $\varepsilon \leq \delta$ yields

$$(6.6) \quad \left\| \varepsilon^2 \partial_T U_2^\delta + \varepsilon L(\partial_y) U_2^\delta + \varepsilon \partial_T U_1^\delta \right\|_{\mathbb{B}^{m,\infty}} \leq C(m) \left(\frac{\varepsilon^2}{\delta^3} + \frac{\varepsilon}{\delta^2} + \frac{\varepsilon}{\delta^2} \right) \leq \frac{C(m)\varepsilon}{\delta^2}.$$

Next turn to the second term on the right of (6.1). Taylor's theorem with remainder implies that there are smooth functions G_α so that

$$(6.7) \quad \Phi(u) = \sum_{|\alpha|=J} u^\alpha G_\alpha(u).$$

It follows that

$$(6.8) \quad \Phi(\lambda u) = \lambda^J \Psi(\lambda, u),$$

with smooth Ψ satisfying $\Psi(\lambda, 0) = 0$. An entirely analogous argument shows that there is a $\Xi(\lambda, u)$ vanishing when $u = 0$ so that

$$(6.9) \quad \Phi(\lambda u) - \Phi_J(\lambda u) = \lambda^{J+1} \Xi(\lambda, u).$$

Split the nonlinearity on the right-hand side of (6.1),

$$\Phi(\varepsilon^p U^\delta) - \Phi_J(\varepsilon^p U_0^\delta) = [\Phi(\varepsilon^p U^\delta) - \Phi(\varepsilon^p U_0^\delta)] + [\Phi(\varepsilon^p U_0^\delta) - \Phi_J(\varepsilon^p U_0^\delta)].$$

Using (6.8), (6.9), and the fact that $\varepsilon^{pJ} = \varepsilon^{p+1}$ yields

$$(6.10) \quad \Phi(\varepsilon^p U^\delta) - \Phi_J(\varepsilon^p U_0^\delta) = \varepsilon^{p+1} [\Psi(\varepsilon^p, U^\delta) - \Psi(\varepsilon^p, U_0^\delta)] + \varepsilon^{2p+1} \Xi(\varepsilon^p, U_0^\delta).$$

Since both U^δ and U_0^δ are H^s uniformly bounded, (6.10) and Schauder's lemma imply that for $s > (d+1)/2$,

$$(6.11) \quad \left\| \Phi(\varepsilon^p U^\delta) - \Phi_J(\varepsilon^p U_0^\delta) \right\|_{H^s(\mathbb{R}_{x,z}^{d+1})} \leq C(s) \varepsilon^{p+1} \left(\|U^\delta - U_0^\delta\|_{H^s(\mathbb{R}_{x,z}^{d+1})} + \varepsilon^p \right).$$

Using (6.4) and $\varepsilon \leq \delta$ yields

$$(6.12) \quad \|U^\delta - U_0^\delta\|_{H^s(\mathbb{R}_{x,z}^{d+1})} \leq C(s) \left(\frac{\varepsilon}{\delta} + \frac{\varepsilon^2}{\delta^2} \right) \leq \frac{C(s)\varepsilon}{\delta}.$$

Combining (6.6), (6.11), and (6.12) yields the estimate (6.3). \square

7. Proof of (2.12). In the next proposition we prove the third and last convergence result needed to establish our main result.

PROPOSITION 7.1. *Suppose that \underline{T} is smaller than the maximal existence time for U_0 and that U^δ, U_0^δ are as in the first paragraph of section 6, and hence in the space satisfying (1.3). Then there is a positive ε_1 so that for $0 < \varepsilon < \varepsilon_1$, the differential equation (2.10) has a unique solution $\mathcal{V}(\varepsilon, t, x, \phi) \in C([0, \underline{T}/\varepsilon]; \cap_s H^s(\mathbb{R}^{d+1}))$ satisfying the initial condition*

$$\mathcal{V}(\varepsilon, 0, x, \phi) = U_0(0, 0, x, \phi).$$

In addition, for all s , for $\delta = \varepsilon^{2/5}$, and for $0 < \varepsilon < \varepsilon_1$,

$$(7.1) \quad \sup_{0 \leq t \leq \underline{T}/\varepsilon} \left\| U^\delta \left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi \right) - \mathcal{V}(\varepsilon, t, x, \phi) \right\|_{H^s(\mathbb{R}_x^d \times \mathbb{R}_\phi)} \leq C(s) \varepsilon^{\max\{1/5, p\}}.$$

Proof. The construction of U^δ guarantees that $U^\delta(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi)$ satisfies

$$(7.2) \quad \begin{aligned} L\left(\partial_t, \partial_x + \frac{\xi_0}{\varepsilon} \partial_\phi\right) \varepsilon^p U^\delta + \Phi(\varepsilon^p U^\delta) \\ = R^\delta\left(\varepsilon, \varepsilon t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) + (I - \chi^\delta(D_z)) \Phi(\varepsilon^p U^\delta). \end{aligned}$$

The strategy is to construct \mathcal{V} as a perturbation, \mathcal{E}^δ , of U^δ . Define

$$(7.3) \quad \mathcal{E}^\delta(\varepsilon, t, x, \phi) := \mathcal{V}(\varepsilon, t, x, \phi) - U^\delta\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right).$$

The equation for \mathcal{V} is rewritten as an equation for \mathcal{E}^δ . The equation for \mathcal{E}^δ is then analyzed to show that the perturbation remains small for $0 \leq t \leq \underline{T}/\varepsilon$.

Subtracting (2.10) from (7.2) yields

$$(7.4) \quad \begin{aligned} L\left(\partial_t, \partial_x + \frac{\xi_0}{\varepsilon} \partial_\phi\right) \varepsilon^p \mathcal{E}^\delta = & -R^\delta\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) \\ & - (I - \chi^\delta(D_z)) \Phi\left(\varepsilon^p U^\delta\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right)\right) \\ & + \left[\Phi\left(\varepsilon^p U^\delta\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right)\right) - \Phi(\varepsilon^p \mathcal{V})\right]. \end{aligned}$$

The operator L is a symmetric hyperbolic operator with constant coefficients and coefficient of ∂_t equal to I . It follows that L generates a unitary evolution on the spaces $H^s(\mathbb{R}_{x,\phi}^{d+1})$. Thus for all s and for all t smaller than the maximal time of existence,

$$(7.5) \quad \begin{aligned} \|\varepsilon^p \mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} & \leq \|\varepsilon^p \mathcal{E}^\delta(0)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \\ & + \int_0^t \left\| L\left(\partial_t, \partial_x + \frac{\xi_0}{\varepsilon} \partial_\phi\right) \varepsilon^p \mathcal{E}^\delta(\sigma) \right\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} d\sigma. \end{aligned}$$

The key is to estimate the integral on the right-hand side of (7.5) using the expression (7.4).

7.1. Estimate for the R^δ term in (7.4). Estimate (6.3) implies that

$$(7.6) \quad \begin{aligned} \left\| R^\delta\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) \right\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} & = \left\| R^\delta(\varepsilon, \varepsilon t, x, z) \right\|_{H^s(\mathbb{R}_{x,z}^{d+1})} \\ & \leq C \varepsilon^{p+1} \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p \right). \end{aligned}$$

7.2. Estimate for the $(I - \chi^\delta(D_z)) \Phi(\varepsilon^p U^\delta(\varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi))$ term in (7.4). As in the derivation of (7.6), the translation invariance of the H^s norm yields

$$\begin{aligned} \left\| (I - \chi^\delta(D_z)) \Phi\left(\varepsilon^p U^\delta\left(\varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right)\right) \right\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \\ = \left\| (I - \chi^\delta(D_z)) \Phi(\varepsilon^p U^\delta(\varepsilon t, x, z)) \right\|_{H^s(\mathbb{R}_{x,z}^{d+1})}. \end{aligned}$$

Express $\Phi(\varepsilon^p U^\delta)$ as the sum of two terms,

$$(7.7) \quad \Phi(\varepsilon^p U^\delta) = \Phi_J(\varepsilon^p U^\delta) + (\Phi - \Phi_J)(\varepsilon^p U^\delta) = \varepsilon^{p+1} \Phi_J(U^\delta) + \varepsilon^{2p+1} \Xi(\varepsilon^p, U^\delta),$$

where we have used (6.9) to derive the second equality.

Since $\Xi(\varepsilon^p, U^\delta(\varepsilon t, x, z))$ is uniformly bounded in $H^s(\mathbb{R}_{x,z}^{d+1})$ it follows that

$$(7.8) \quad \|(I - \chi^\delta(D_z))\varepsilon^{2p+1}\Xi(\varepsilon^p, U^\delta)\|_{H^s(\mathbb{R}_{x,z}^{d+1})} \leq C \varepsilon^{2p+1}.$$

Since $\Phi_J(U^\delta(\varepsilon t, t, x, \frac{t\tau}{\varepsilon} + \phi))$ are uniformly bounded in $\mathbb{B}^{m,\infty}$, Lemma 5.2 implies that

$$(7.9) \quad \|(I - \chi^\delta(D_z))\varepsilon^{p+1}\Phi_J(U^\delta)\|_{H^s(\mathbb{R}_{x,z}^{d+1})} \leq C \sqrt{\delta} \varepsilon^{p+1}.$$

Adding (7.8) and (7.9) shows that

$$(7.10) \quad \|(I - \chi^\delta(D_z))\Phi\left(\varepsilon^p U^\delta\left(\varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right)\right)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq C \varepsilon^{p+1}(\sqrt{\delta} + \varepsilon^p).$$

7.3. Estimate for the $\Phi(\varepsilon^p U^\delta) - \Phi(\varepsilon^p \mathcal{V})$ term in (7.4). Using (6.8) yields

$$\Phi(\varepsilon^p U^\delta) - \Phi(\varepsilon^p \mathcal{V}) = \varepsilon^{p+1}[\Psi(\varepsilon^p, U^\delta) - \Psi(\varepsilon^p, \mathcal{V})] = \varepsilon^{p+1}[\Psi(\varepsilon^p, U^\delta) - \Psi(\varepsilon^p, U^\delta + \mathcal{E}^\delta)].$$

The proof of Proposition 5.1 shows that

$$\sup_{0 \leq t \leq T/\varepsilon, 0 < \varepsilon < \delta \leq \delta_0} \left\| U^\delta\left(\varepsilon, \varepsilon t, t, x, \frac{t\tau_0}{\varepsilon} + \phi\right) \right\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} < \infty.$$

Schauder's lemma then implies that *as long as* $\|\mathcal{E}^\delta\|_{H^s} \leq 1$,

$$(7.11) \quad \varepsilon^{p+1} \|\Psi(\varepsilon^p, U^\delta + \mathcal{E}^\delta) - \Psi(\varepsilon^p, U^\delta)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq C \varepsilon^{p+1} \|\mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})}.$$

Combining estimates (7.6), (7.10), and (7.11) yields the following estimate, valid *as long as* $\|\mathcal{E}^\delta\|_{H^s} \leq 1$:

$$(7.12) \quad \left\| L\left(\partial_t, \partial_x + \frac{\xi_0}{\varepsilon} \partial_\phi\right) \varepsilon^p \mathcal{E}^\delta \right\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq C \varepsilon^{p+1} \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} + \|\mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \right).$$

7.4. Estimate for $\mathcal{E}^\delta(0)$. The initial value of $\mathcal{E}^\delta(0)$ comes from the correctors in U^δ ,

$$\mathcal{E}^\delta(0) = f(x, \phi) - U^\delta(0, x, \phi) = \varepsilon U_1^\delta(0, x, \phi) + \varepsilon^2 U_2^\delta(0, x, \phi).$$

Using (6.4) and the fact that $\varepsilon \leq \delta$ yields

$$(7.13) \quad \|\mathcal{E}^\delta(0)\|_{H^s(\mathbb{R}^{d+1})} \leq C \left(\frac{\varepsilon}{\delta} + \frac{\varepsilon^2}{\delta^2} \right) \leq \frac{C\varepsilon}{\delta}.$$

8. End of proof. We now use the previous results to bound the error \mathcal{E}^δ between the exact solution \mathcal{V} and the approximate solution defined in terms of U^δ .

Fix $s > (d+1)/2$. Then as long as $\|\mathcal{E}^\delta\|_{H^s} \leq 1$, inserting (7.12) and (7.13) into (7.5) yields

$$\|\varepsilon^p \mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq \frac{C\varepsilon^{p+1}}{\delta} + C \int_0^t \varepsilon^{p+1} \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} + \|\mathcal{E}^\delta(\sigma)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \right) d\sigma.$$

Estimate the integral of the constant term using $t \leq \underline{T}/\varepsilon$ to find

$$\int_0^t \varepsilon^{p+1} \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} \right) d\sigma \leq \frac{\underline{T} \varepsilon^{p+1}}{\varepsilon} \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} \right) \leq C \varepsilon^p \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} \right).$$

Combine the last two estimates using $\frac{\varepsilon}{\delta^2} \geq \frac{\varepsilon}{\delta}$ and divide by ε^p to find

$$(8.1) \quad \|\mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq C \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} \right) + C \varepsilon \int_0^t \|\mathcal{E}^\delta(\sigma)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} d\sigma.$$

We need to show that \mathcal{E}^δ exists for $0 \leq T \leq \underline{T}/\varepsilon$ and that $\sup_{t \in [0, \underline{T}/\varepsilon]} \|\mathcal{E}^\delta(t)\|_{H^s}$ converges to zero as $\delta \rightarrow 0$. The integral inequality (8.1) leads to both of these goals by the ‘‘as long as’’ argument.

For any $0 < \varepsilon < \delta < \delta_0$, define $T_* = T_*(\varepsilon, s, \delta)$ by $T_* = \underline{T}/\varepsilon$ if $U^\delta(\varepsilon, T, t, x, z)$ exists for $0 \leq t \leq \underline{T}/\varepsilon$ and $\sup_{0 \leq t \leq \underline{T}/\varepsilon} \|\mathcal{E}^\delta(t)\|_{H^s} < 1$. Otherwise define

$$T_* := \inf \{t : \|\mathcal{E}^\delta(t)\|_{H^s} = 1\}.$$

Gronwall’s inequality applied to (8.1) implies that for $0 \leq t \leq T_*$,

$$(8.2) \quad \|\mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq C(s) \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} \right) e^{C(s)\varepsilon t} \leq C(s) \left(\frac{\varepsilon}{\delta^2} + \varepsilon^p + \sqrt{\delta} \right).$$

Now we can choose δ as a function of ε . Balancing the ε/δ^2 and $\sqrt{\delta}$ terms in (8.2) yields

$$\delta = \varepsilon^{0.4} \quad \text{and} \quad \frac{\varepsilon}{\delta^2} = \sqrt{\delta} = \varepsilon^{1/5}.$$

Then (8.2) yields for $\delta = \varepsilon^{0.4}$

$$(8.3) \quad \|\mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} \leq C(s) \varepsilon^{\min\{p, 1/5\}}.$$

Choose $\varepsilon(s) > 0$ so that

$$(8.4) \quad C(s) \varepsilon(s)^{\min\{p, 1/5\}} < \frac{1}{2}.$$

Combining (8.3) and (8.4) shows that for $0 \leq t \leq T_*$, $0 < \varepsilon < \varepsilon(s)$, and $\delta = \varepsilon^{0.4}$

$$(8.5) \quad \|\mathcal{E}^\delta(t)\|_{H^s(\mathbb{R}_{x,\phi}^{d+1})} < \frac{1}{2}.$$

If $T_* < \underline{T}$, setting $t = T_*$ violates the definition of T_* . It follows that for $0 < \varepsilon < \varepsilon(s)$, $\mathcal{E}^{\varepsilon^{0.4}}(t)$ has H^s norm less than $1/2$ for $0 \leq t \leq \underline{T}/\varepsilon$.

This proves the solvability for $0 \leq t \leq \underline{T}/\varepsilon$ of the initial value problem defining $\mathcal{V} \in C([0, T]; H^s)$. That the solution belongs to $C([0, \underline{T}/\varepsilon]; \cap_s H^s)$ is then a consequence of standard semilinear hyperbolic theory and the regularity of the initial data for \mathcal{V} .

In addition, since (8.5) holds, the ‘‘as long as’’ argument works, and it follows that inequality (8.3) is valid for $0 \leq t \leq \underline{T}/\varepsilon$, provided that $\varepsilon < \varepsilon(s)$. Since s is arbitrary this proves the convergence asserted in Proposition 7.1. \square

8.1. The main theorems. Combining Propositions 4.3, 5.1, and 7.1 and Corollary 4.12 proves the following result.

THEOREM 8.1 (main theorem). *Assume that the initial data in (2.1) satisfy Assumptions 1.5 and 1.8. Let $U_0 = \pi(\beta)U_0 \in C([0, T_*[; \mathbb{A}(\mathbb{R}^{1+d}))$ be the maximal solution of the principal profile equation (4.2) with initial value f . Let $\mathcal{V}(\varepsilon, t, x, \phi) \in \cap_s C([0, T_*[; H^s(\mathbb{R}^{d+1}))$ denote the maximal solution of the initial value problem (2.10), (2.11) defining the exact profile.*

Then, for any $\underline{T} < T_$ there is an $\varepsilon(\underline{T}) > 0$ so that for $0 < \varepsilon < \varepsilon(\underline{T})$ the solution u^ε of the initial value problem (2.1) exists for $0 \leq t \leq \underline{T}/\varepsilon$, $T_*' \geq \underline{T}/\varepsilon$, and u^ε is given by*

$$u^\varepsilon = \varepsilon^p \mathcal{V}\left(\varepsilon, t, x, \frac{x \cdot \xi_0}{\varepsilon}\right).$$

In addition the asymptotic behavior as $\varepsilon \rightarrow 0$ is given by

$$u^\varepsilon \sim \varepsilon^p U_0\left(\varepsilon, \varepsilon t, x - \mathbf{v}t, \frac{t\tau_0 + x \cdot \xi_0}{\varepsilon}\right)$$

in the sense that for all s , as $\varepsilon \rightarrow 0$

$$(8.6) \quad \sup_{0 \leq t \leq \underline{T}/\varepsilon} \left\| \mathcal{V}(\varepsilon, t, x, \phi) - U_0\left(\varepsilon t, x - \mathbf{v}t, \frac{t\tau_0}{\varepsilon} + \phi\right) \right\|_{H^s(\mathbb{R}_{x,\phi}^{1+d})} \leq C(s) \varepsilon^{\min\{p, 1/5\}}.$$

Proof of Theorem 1.10. Theorem 1.10 is an immediate consequence of this result. Simply write

$$u^\varepsilon - u_{\text{approx}}^\varepsilon = \varepsilon^p \mathcal{V}\left(\varepsilon, t, x, \frac{x \cdot \xi_0}{\varepsilon}\right) - \varepsilon^p U_0\left(\varepsilon t, x - \mathbf{v}t, \frac{t\tau_0 + x \cdot \xi_0}{\varepsilon}\right).$$

Then the estimate (1.12) follows from (8.6). \square

Note that the constant field V_d has a nonzero ∂_t component. It acts differently on the two terms in the expression for the error. That is why we have a reduced set of derivatives in the error estimate of Theorem 1.10.

REFERENCES

- [1] D. ALTERMAN, *Diffraction Nonlinear Geometric Optics for Short Pulses*, Ph.D. thesis, Department of Mathematics, University of Michigan, Ann Arbor, MI, 1999.
- [2] D. ALTERMAN AND J. RAUCH, *Diffraction short pulse asymptotics for nonlinear wave equations*, Phys. Lett. A., 264 (2000), pp. 390–395.
- [3] D. ALTERMAN AND J. RAUCH, *Correcting the failure of the slowly varying amplitude approximation for short pulses*, in Ultrafast Phenomena XII, T. Elsaesser, S. Mukamel, M. M. Murnane, and N. F. Scherer, eds., Springer-Verlag, New York, 2001, pp. 71–73.
- [4] D. ALTERMAN AND J. RAUCH, *The linear diffractive pulse equation*, Methods Appl. Anal., 7 (2000), pp. 263–274.
- [5] D. ALTERMAN AND J. RAUCH, *Nonlinear geometric optics for short pulses*, J. Differential Equations, 178 (2002), pp. 437–465.
- [6] K. BARRAILH AND D. LANNES, *A general framework for diffractive optics and its applications to lasers with large spectrums and short pulses*, SIAM J. Math. Anal., 34 (2003), pp. 636–674.
- [7] R. CARLES AND J. RAUCH, *Focusing of spherical nonlinear pulses in \mathbb{R}^{1+3}* , Proc. Amer. Math. Soc., 130 (2002), pp. 791–804.
- [8] R. CARLES AND J. RAUCH, *Absorption d'impulsions non-linéaires radiales focalisantes dans \mathbb{R}^{1+3}* , C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 985–990.
- [9] R. CARLES AND J. RAUCH, *Diffusion d'impulsions non-linéaires radiales focalisantes dans \mathbb{R}^{1+3}* , C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 1077–1082.

- [10] T. BRABEC AND F. KRAUSZ, *Nonlinear optical pulse propagation in the single-cycle regime*, Phys. Rev. Lett., 78 (1997), pp. 3282–3285.
- [11] P. DONNAT, J.-L. JOLY, G. MÉTIVIER, AND J. RAUCH, *Diffractive Nonlinear Geometric Optics*, in Séminaire sur les Equations aux Dérivées Partielles (1995–1996), Exp. 17, École Polytechnique, Paris, 1996.
- [12] E. DUMAS, *Periodic multiphase diffractive optics with curved phases*, Indiana Math. J., to appear.
- [13] E. ESAREY, P. SPRANGLE, M. PILLOFF, AND J. KRALL, *Theory and group velocity of ultrashort, tightly focused laser pulses*, J. Opt. Soc. Amer. B Opt. Phys., 12 (1995), pp. 1695–1703.
- [14] S. FENG, H. G. WINFUL, AND R. W. HELLWARTH, *Spatiotemporal evolution of focused single-cycle electromagnetic pulses*, Phys. Rev. E, 59 (1999), pp. 4630–4649.
- [15] R. W. HELLWARTH AND P. NOUCHI, *Focused one-cycle electromagnetic pulses*, Phys. Rev. E, 54 (1996), pp. 889–895.
- [16] C. HILE, *Comparisons between Maxwell’s equations and an extended nonlinear Schrödinger equation*, Wave Motion, 24 (1996), pp. 1–12.
- [17] C. V. HILE AND W. L. KATH, *Numerical solutions of Maxwell’s equations for nonlinear-optical pulse propagation*, J. Opt. Soc. Amer. B Opt. Phys., 13 (1996), pp. 1135–1145.
- [18] J.-L. JOLY, G. MÉTIVIER, AND J. RAUCH, *Coherent and focusing multidimensional nonlinear geometric optics*, Ann. Sci. École Norm. Sup. (4), 28 (1995), pp. 51–113.
- [19] J.-L. JOLY, G. MÉTIVIER, AND J. RAUCH, *Diffractive nonlinear geometric optics with rectification*, Indiana Univ. Math. J., 47 (1998), pp. 1167–1242.
- [20] A. E. KAPLAN, *Diffraction-induced transformation of near-cycle and subcycle pulses*, J. Opt. Soc. Amer. B Opt. Phys., 15 (1998), pp. 951–956.
- [21] M. A. PORRAS, *Ultrashort pulsed Gaussian light beams*, Phys. Rev. E, 58 (1998), pp. 1086–1093.
- [22] J. K. RANKA AND A. L. GAETA, *Breakdown of the slowly varying envelope approximation in the self-focusing of ultrashort pulses*, Opt. Lett., 23 (1998), pp. 534–536.
- [23] J. RAUCH, *An L^2 proof that H^s is invariant under nonlinear maps for $s > n/2$* , in Global Analysis: Analysis on Manifolds, T. M. Rassias, ed., Teubner, Leipzig, 1983, pp. 301–305.
- [24] J. RAUCH, *Lectures on geometric optics*, in Hyperbolic Equations and Frequency Interactions, L. Caffarelli and W. E. eds., IAS/Park City Math. Ser. 5, AMS, Providence, RI, 1999, pp. 383–466.
- [25] J. E. ROTHENBERG, *Space-time focusing: Breakdown of the slowly varying envelope approximation in the self-focusing of femtosecond pulses*, Opt. Lett., 17 (1992), pp. 1340–1342.
- [26] A. YOSHIKAWA, *Solutions containing a large parameter of a quasi-linear hyperbolic system of equations and their nonlinear geometric optics approximation*, Trans. Amer. Math. Soc., 340 (1993), pp. 103–126.
- [27] A. YOSHIKAWA, *Asymptotic expansions of the solutions to a class of quasilinear hyperbolic initial value problems*, J. Math. Soc. Japan, 47 (1995), pp. 227–252.
- [28] R. W. ZIOLKOWSKI, *Localized transmission of electromagnetic energy*, Phys. Rev. A, 39 (1989), pp. 2005–2033.