Modeling the Dispersion of Light

Phillipe DONNAT
Commissariat de l'Energie Atomique, Centre d'Etude de Limeil Valenton
Villeneuve St. Georges 94195, FRANCE

Jeffrey RAUCH*
University of Michigan
Ann Arbor 48109 MI, U.S.A.

§1. Introduction.

This note discusses the modeling of dispersive phenomena in optics. The basic question is "Is there a reasonable differential equations model which explains the dispersive behavior of a prism?". Put another way the question is whether there is a differential equation model which includes the experimentally observed fact that the speed of light depends on its frequency. The index of refraction is equal $1/\text{speed}^2$ so this is equivalent to studying the dependence of the index on frequency.

The fundamental Maxwell equations have only one speed of propagation, the speed of light in vacuum. The origin of dispersive effects is the interaction with matter. Merely inserting electric and magnetic susceptibilities in time dependent models does not yield dispersion. Modeling matter via susceptibilites is not the answer.

One can model monochromatic problems by inserting an *ad hoc* dependence of the suspecptibilities on frequency. In a sense, what we do here is to place such an approach in a mathematically (and physically) reasonable setting. In particular, some such models arise from simple systems of partial differential equations.

A second motivation for (re)investigating this question is that current high power lasers are pushing the modeling of nonlinear effects to the limit. Here it is the dependence of the index on intensity that is at issue. Several distinct models are adequate for weak fields and yield different predictions for large fields (see [D]). As experimental data gets better, we should be able to choose among the options. All models must predict linear dispersion in the limit of weak fields, and, it is that case which is the subject of this note.

§2. Maxwell's equations in vacuum.

In gaussian (a.k.a. cgs) units, Maxwell's equations for the propagation of light in a vacuum are (see [P] pg. 263)

$$\frac{1}{c}\,E_t = \operatorname{curl} B\,, \qquad \frac{1}{c}\,B_t = -\operatorname{curl} E\,, \qquad c = 3\times 10^{10}\ \text{cm./sec.}$$

* Research partially supported by the U.S. National Science Foundation and the U.S. Office of Naval Research under grants number NSF-DMS-9203413 and OD-G-N0014-92-J-1245 respectively.

$$\operatorname{div} E = 0, \qquad \operatorname{div} B = 0\,. \tag{1}$$

The physical interpretation of the electric and magnetic fields is that a point charge is subject to a force which is the sum of two terms. One term is proportional to the product qE of the charge and the electric field, and the other to the cross product $qv \wedge B$ where v is the particle's velocity. The very large constant c is the speed of light.

To nondimensionalize the Maxwell system it is reasonable to leave the unit of length equal to centimeters as this is appropriate for optical devices like lenses and mirrors. In order to compensate for the large constant c, a the natural time unit, denoted $\mathcal{T}$ is the time it takes for light to travel one centimeter,

$$\mathcal{T} = \frac{1}{3 \times 10^{10}} = \frac{1}{3} \times 10^{-10}\, s\,, \qquad \text{one second} = 3 \times 10^{10}\, \mathcal{T}\,.$$

In these new cg$\mathcal{T}$ units, the dynamic Maxwell's equations are

$$E_t = \operatorname{curl} B\,, \qquad B_t = -\operatorname{curl} E\,. \tag{2}$$

Taking the divergence of (2) yields

$$\partial_t \operatorname{div} E = \partial_t \operatorname{div} B = 0\,. \tag{3}$$

Therefore, (1) holds as soon as it holds in the initial plane $\{t = 0\}$.

Equations (2) define a constant coefficient symmetric hyperbolic equation which determines the fields E, B from their values at $t = 0$. As with all such symmetric systems there is a simple L^2 estimate. Taking the scalar product of the system with E, B yields the identity

$$\frac{\partial}{\partial t} \int |E|^2 + |B|^2 \; dx = 0$$

expressing the conservation of energy. More generally, the change of energy in a nice subset of $\mathbf{R}^3$ is given by

$$\frac{\partial}{\partial t} \int_K |E|^2 + |B|^2 \, dx = \int_{\partial K} (E \wedge B).n \; d\sigma\,.$$

The flux of energy is described by the *Poynting vector* $S := E \wedge B$.

Since there are no lower order terms, the characteristic variety for this system is found by seeking plane wave solutions

$$\big(E(t,x), B(t,x)\big) = e^{i(\omega t + \xi . x)}(\mathbf{e}, \mathbf{b})\,, \tag{4}$$

where $\mathbf{e}$ and $\mathbf{b}$ are constant vectors. The Maxwell equation then take the form

$$\omega \mathbf{e} = \xi \wedge \mathbf{b}\,, \qquad \omega \mathbf{b} = -\xi \wedge \mathbf{e}\,. \tag{5}$$

There is a two dimensional family of solutions with $\omega = 0$ and both $\mathbf{e}$ and $\mathbf{b}$ parallel to ξ. These solutions do not satisfy (1) so are not relevant to the physics. They are steady states. Supperposition of these plane waves yields exactly the solutions of (2) which in addition satisfy

$$\operatorname{curl} E = \operatorname{curl} B = 0 \qquad \text{and therefore} \qquad E_t = B_t = 0\,.$$

The physically relevant solutions travel at the speed of light.

To find the other plane wave solutions, multiply the first equation in (5) by ω to get

$$\omega^2 \mathbf{e} = \xi \wedge \omega \mathbf{b} = \xi \wedge (-\xi \wedge\, \mathbf{e}) = \big(|\xi|^2 I - \;|\xi><\xi|\big)\,\mathbf{e}\,. \tag{6}$$

The 3×3 symmetric matrix $(\omega^2 - |\xi|^2)I + |\xi><\xi|$ has eigenvalue $\omega^2 - |\xi|^2$ with multiplicity two, the eigenspace being vectors orthogonal to ξ. It has eigenvalue ω^2 with eigenspace the vectors parallel to ξ. The conclusion is that the characteristic variety is equal to $\{\omega = 0\} \cup \{\omega^2 = |\xi|^2\}$.

The plane wave solutions with $\omega \neq 0$ satisfy $\omega^2 = |\xi|^2$, and are constructed by choosing $\mathbf{e} \perp \xi$, and defining, $\mathbf{b}$ by the second equation in (5). This is a two dimensional family of solutions for each ω, ξ. For ξ fixed there is a two dimensional space of solutions for each of the roots $\omega = \pm|\xi|$. The multiplicities show that the characteristic equation is $\omega^2(\omega^2 - |\xi|^2)^2 = 0$.

The phases of the physically relevant plane waves are given by

$$\omega t + \xi.x\,, \qquad \omega(\xi) = \pm|\xi|\,.$$

The phase velocity is computed from the diplacement of the hyperplanes of constant phase as time evolves. If the phase is equal to $\omega t + \xi.x$, then at time $t = 0$ the plane of phase equal to zero is the plane through the orgin which is orthogonal to ξ. At time $t = 1$ it is the plane with equation $\omega + \xi.x = 0$. This is the translate of the initial plane by $-\omega\xi/|\xi|^2$. Thus,

$$\text{phase velocity} := \frac{-\omega\,(\xi_1, \cdots, \xi_d)}{\xi_1^2 + \cdots + \xi_d^2}\,. \tag{7}$$

Actually, there are other answers. Any velocity perpendicular to ξ, and therefore parallel to the planes of constant phase could be added to (7). The result (7) is the smallest velocity that works and is the answer which comes to mind when faced with a family of parallel hyperplanes, namely motion perpendicular to the planes. For Maxwell's equation, $\omega = \pm|\xi|$ so

$$\text{phase velocity} = \frac{-\xi}{\pm|\xi|} = \mp 1\,.$$

The phase velocity is essentially constant, so by any definition, the equations not dispersive.

One reaches the same conclusion by taking the *curl* of each of the equations (2). Using (1) shows that each component of E and B is a solution of d'Alembert's equation $\square_{1+3} u = 0$. The singularities, high frequency oscillations, and support (by the strict Huyghen's principle) all proagate with speed equal to one.

§3. Dispersion.

When white light crosses obliquely an interface between air and glass, the beam is split into colored parts forming a rainbow effect. Different colors refract through different angles. The explanation of this effect was found very early. The refraction is explained by the geometric theory which asserts that light rays take paths of shortest time. This explains why rays bend on passing from air to glass. They want to spend more time in the faster medium, that is air. Furthermore the angle of refraction depends in a simple way on the ratio of the speeds. Thus the dispersion of white light is easily explained by the hypothesis that the different colors travel with different speeds. This dependence of speed on wavelength is also called *dispersion*. In glass, higher frequencies correspond to slower speeds and therefore large angles of refraction. This is called *normal dispersion*. The opposite monotonicity is observable but usually on only small bands of frequency often outside the visible (e.g. in the ultraviolet), and, is called *anomalous dispersion*.

The origin of dispersion is the interaction of light with the many charged particles which comprise matter. The standard model due to Lorenz ([L]) is based on the hypothesis that electrons are bound to the nuclei with restoring forces modeled by spring laws with perhaps a finite number of distinct natural frequencies. This model was invented before the advent of quantum mechanics, and it is a very pleasant surprise that quantum mechanical models yield indistinguishable predictions (see [B]). One of our goals is to show how Lorentz' model can be cast in a form which is very attractive from a mathematician's perspective. This presentation is also an excellent starting point for the modeling of nonlinear phenomena in optics.

The reason that light of different wavelengths travels at different speeds is that as the light travels through glass, the electric field pulls on the electrons and nuclei of the atoms of the glass. It pulls in opposite directions since the electrons have negative charge and the nucleus is positive. This results in a polarization of the atoms. The resulting dipoles oscillate at about the frequency of the light, and thereby radiate. The electric field that one sees is a superposition of the incoming field depleted a little, with the radiated fields of all the atoms. The reaction of the atoms to different frequencies is different so that one gets a different superposition for different frequencies and the net effect is the dispersion of light. A computation following exactly this point of view for plane waves of light passing through a planar slab of matter is presented in *The Feynmann Lectures on Physics I-§31, II-§32.* What Feynmann does not give us is a system of differential equations, which applies to more complicated geometric arrangements.

There is a key observation which is at the heart of the mathematical analysis. When the atoms are polarized, they tend to return to their unperturbed state. There is a restoring force. Taylor expansion shows that for small deformations, the restoring force is essentially proportional to the deformation. Thus the dipoles will oscillate as a forced harmonic oscillator. The natural frequency of the oscillator corresponds to spectral lines, which normally are just above the visible in the ultraviolet. This natural frequency is a large parameter which is larger than but of the same order of magnitude as the frequency of the light which is exciting the vibration. *There is a high frequency excitation of a high frequency oscillator. The small parameter which is the wavelength, appears as a coefficient in the model.* This is absolutely essential in the explanation of the dispersion of light.

To see that it is essential, note that a high frequency analysis of a standard hyperbolic equation, that is without large coefficients, results in rays with velocity which may depend on the direction of the oscillations but not on the frequency. The rays are the projections on space time of the *bicharacteristics*. In coordinates $x = x_0, x_1, \cdots, x_d$ with $x_0 := t$, $\xi = \xi_0, \xi_1, \cdots, \xi_d$ with $\xi_0 := \omega$, the bicharacteristics are integral curves of the Hamiltonian vector field

$$\partial_\xi p_m(x,\xi)\,\partial_x - \partial_x p_m(x,\xi)\,\partial_\xi$$

with Hamiltonian

$$p(x,\xi) := \det L(x,\xi)\,.$$

Here L is the principal symbol of the system and therefore p is a homogeneous polynomial in ξ. The spatial velocity of the rays is equal to

$$\frac{-(\partial_{\xi_1} p_m, \cdots, \partial_{\xi_d} p_m)}{\partial_{\xi_0} p_m}$$

which is homogeneous of degree zero in ξ, showing that the speed is independent of the magnitude of ξ. The same would be true of the propagation of ultra high frequency electromagnetic waves in glass. Then the geometric optics asymptotics would be appropriate since the wavelength would be small compared to all the other dimensions in the problem. However, in the case of visible light, this is not the case.

§4. Dispersive Maxwell equations.

The discussion of the last section reveals two crucial ingredients of the modeling of dispersion. First, it is the interaction of light with matter which underlies the phenomenon, and second, it is a high frequency phenomenon, for which the large parameter appears in the equations themselves.

The first step is to include the effect of charges on the fields. Maxwell's equations in cg$\mathcal{T}$ units and incorporating charges and currents are

$$E_t = \mathrm{curl}\, B - 4\pi \mathbf{j}\,, \qquad B_t = -\mathrm{curl}\, E\,. \tag{1}$$

$$\mathrm{div}\, E = 4\pi\rho, \qquad \mathrm{div}\, B = 0\,, \tag{2}$$

where ρ and $\mathbf{j}$ are the charge and current densities respectively. Taking the divergence of the first equation in (1) and the time derivative of the first equation in (2) shows that the *continuity equation*,

$$\partial_t \rho = -\mathrm{div}\, \mathbf{j}\,, \tag{3}$$

follows from the Maxwell system. Taking the divergence of (1) then yields

$$\partial_t\, \mathrm{div}\, B = \partial_t\, \mathrm{div}\,(E - 4\pi\rho) = 0\,. \tag{4}$$

The physically relevant solutions are those for which

$$\mathrm{div}\, B = \mathrm{div}\,(E - 4\pi\rho) = 0\,. \tag{5}$$

In principal one could compute the evolution of the fields in matter by using (1)-(2) taking into account all of the charges, that is all the electrons and protons. What one would find is an electric field with wild oscillations on an atomic scale. On one side of a nucleus and close in the field is enormous and points away from the nucleus. Thus on passing from one side to the other of a nucleus the field is huge and it reverses direction. In the interaction with macroscopic objects, these atomic scale variations play no role. One develops a *macroscopic Maxwell system* which describes the behavior of the field averaged over regions large compared to the atom and small compared to macroscopic objects. This is entirely analogous to the derivation of fluid mechanical equations for the macroscopic behavior of large collections of atoms.

On macroscopic scales, the field is not rapidly varying and the matter is neutral. Charges are not free to move. Electrons are relatively mobile but by and large are stuck with a fixed nucleus. Microscopic dipoles are formed which give the material a possibly nonzero dipole moment per unit volume, called the *polarization*.

A changing dipole is a current. To see this, consider a dipole consisting of a positive charge above an equal negative charge. Decreasing the moment means lowering the positive and raising the negative charges. Both yields downward pointing currents. The induced current is proportional to $-\partial_t P$ where P is the notation for the polarization per unit volume. The macroscopic Maxwell system has the form

$$E_t = \operatorname{curl} B - 4\pi\, P_t\,, \qquad B_t = -\operatorname{curl} E\,. \tag{6}$$

$$\operatorname{div}(E + 4\pi P) = 0, \qquad \operatorname{div} B = 0\,. \tag{7}$$

The essential ingredient in completing a model for the interaction is to prescribe a law relating P and the fields E and B.

The simplest model, and the one in your undergraduate physics course, is to suppose that the polarization is proportional to the electric field, $P = \chi E$. the constant is called the *dielectric constant*. The effect of the polarization term χE is to take the original Maxwell system and put a factor $1 + 4\pi\chi$ in front of the E in the first equation in (2.2). The computations leading to (3.6) show that the new characteristic equation is

$$\omega^2 \left((1 + 4\pi\chi)\omega^2 - |\xi|^2\right)^2 = 0\,.$$

This correctly predicts that in a dielectric, light travels at the slower speed $(1+4\pi\chi)^{-1/2}$. It explains why a knife looks bent at the surface when partially immersed in water. However, it does not explain why different colors go at different speeds in the same dielectric. The quantity $1 + 4\pi\chi$ is called the *dielectric constant* and is often denoted ϵ.

The standard treatment of dispersion depends crucially on linearity and time translation invariance of the equations. Fourier analysis in time yields

$$E(t,x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \hat E(\omega,x)\, e^{i\omega t}\, d\omega\,, \quad P(t,x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \hat P(\omega,x)\, e^{i\omega t}\, d\omega\,, \tag{8}$$

$$\hat E(\omega,x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} E(t,x) e^{-i\omega t}\, d\omega\,, \quad \hat P(\omega,x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} P(t,x) e^{-i\omega t}\, d\omega\,, \tag{9}$$

The electric field is a superposition of time harmonic fields, $\hat{E}(\omega,x)\,e^{i\omega t}$. Taking the Fourier transform of the Maxwell equations shows that each of these fields is a solution of the time harmonic Maxwell equations,

$$i\omega\hat{E} = \operatorname{curl} B - i\omega\hat{P}\,, \qquad i\omega\hat{B} = -\operatorname{curl} E\,, \tag{10}$$

$$\operatorname{div}(\hat{E} + 4\pi\hat{P}) = 0\,, \qquad \operatorname{div}\hat{B} = 0\,, \tag{11}$$

To explain dispersion, one supposes that

$$\hat{P}(\omega,x) = \chi(\omega)\hat{E}(\omega,x)\,. \tag{12}$$

with a $\chi = \chi(\omega)$ *which depends on* ω.

To see what this says in the original coordinates, insert the last expression into the formula for P in terms of its Fourier Transform to find

$$\begin{aligned} P =&(2\pi)^{-1/2}\int_{-\infty}^{\infty} \chi(\omega)\hat{E}(\omega,x)\,e^{i\omega t}\;d\omega \\ =&(2\pi)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi(\omega)\,E(s,x)\,e^{i\omega t}\;e^{-i\omega s}\;ds\,d\omega \end{aligned}\,.$$

Performing the ω integral first yields

$$P(t,x) = \int_{-\infty}^{\infty} a(t-s)\,E(s)\,ds\,, \tag{13}$$

where

$$a(t) := (2\pi)^{-1}\int_{-\infty}^{\infty} \chi(\omega)\,e^{i\omega t}\,d\omega\,. \tag{14}$$

Causality demands that the values of $P(t,x)$ be determined by the values of $E(s,x)$ with $s \le t$. The support of $a(t)$ must be contained in $t \ge 0$ which implies, by the Paley-Wiener Theorem, that χ is a holomorphic function in $\Im\omega > 0$. In particular, the real part of χ determines the imaginary part and vice versa. The identities, of Cauchy integral type, giving one in terms of the other are called the *Kramers-Krönig relations*. Other than causality, the function $\chi(\omega)$ is unspecified up to this point.

The Lorenz model predicts a specific form for the functions $a(t)$ and $\chi(\omega)$. The key idea from §3 is to suppose that the polarization arises as the result of harmonic oscillators distributed in space,

$$\partial_t^2\,P + \Omega^2 P = C E\,. \tag{15}$$

For a dilute gas of nonpolar molecules, this is a fine model. For dense gases and liquids, the equation must be modified because the field felt by an idividual molecule is actually stronger than the macroscopic field by an amount $4\pi\mathbf{P}/3$ (see [P] §9.13 or [FLS] II-11.4). Thus, the right hand side should be $C(E + 4\pi\mathbf{P}/3)$ which has the effect of replacing Ω^2 by $\Omega^2 - 4\pi C/3$. This leaves the form of (15) intact, and as we will see, is a small perturbation. Bearing in mind this alteration we keep equation (15).

Performing a Fourier transform in time in (15) yields

$$\chi(\omega) = \frac{C}{\Omega^2 - \omega^2}\,. \tag{16}$$

This is a useful expression so long as ω is not too near Ω. Otherwise, one is near resonance and fields oscillating at the resonant frequency will not lead to purely oscillatory steady states but to growing fields. Near resonance, the passage to the monchromatic case is more delicate and one is forced to consider the dynamics including (15). It is even wiser to refine the modeling of the interaction of the field and matter.

The singularity at $\omega = \Omega$ is removed if a small dissipation is added to the spring law yielding

$$\chi(\omega) = \frac{C}{i\mu\omega + \Omega^2 - \omega^2}\,. \tag{17}$$

One source of such damping is the loss of energy to heat.

It is normal to approximate the function χ as a linear combination of terms of form (17) with distinct resonant frequencies Ω_k and friction coefficients μ_k (e.g. [FLS] equation (31.20)). Exactly such expression are found in the quantum mechanical analysis (see [B]). The classical model yields such expressions when the polarization is written as a finite sum of terms each satisfying its own spring equation. One arrives at such an expression if several distinct types of atoms are present, or if one supposes that the dipoles have several distinct normal modes of oscillation. The latter point of view is closer to the quantum mechanical analysis. Finally note that linear combinations of expressions of the form (17) are dense in most any space of causal functions of ω which are small at $\pm\infty$. Thus, assuming such a form is not assuming much.

A crucial observation (see [D]) is that the positive coefficient Ω is large. Recall that the speed of light is $c = 3 \times 10^{10}$ cm./sec. A typical wavelength for visible light is $\lambda = 5000\mathring{A}$. One $\mathring{A} = 10^{-8}$cm. The frequency ω of such a wave is then computed from the relation

$$\omega = \frac{2\pi c}{\lambda} \approx \frac{2\pi \times 3 \times 10^{10}}{5 \times 10^{-5}} \approx 3 \times 10^{15}\, s^{-1}\,.$$

The resonant frequencies Ω which occur are typically in the ultraviolet so are larger than but of the same order of magnitude as the frequency of light. A knee jerk reaction is to say that these frequencies are very large indeed.

However, the Maxwell equations have been written in units so that the speed of light is equal to one. That choice was made to eliminate a large coefficient and it is possible that it neutralizes the large coefficient Ω^2 too. However, in our units,

$$\omega \approx 10^5\, \mathcal{T}^{-1}\,,$$

so is still large. The bottom line is that in units so that the speed of light is equal to one, visible light is still a high frequency phenomenon.

With

$$\epsilon := \frac{1}{\Omega} << 1\,, \tag{18}$$

the equation for the polarization becomes

$$(\epsilon\partial_t)^2\, P + P = \gamma E\,. \tag{19}$$

The physical constant $\gamma := C/\Omega^2$ is not very much smaller than 1. This follows from two facts. First, for light in glass, ω is smaller than but of the same order of magnitude as Ω, $\omega \approx 0.9\Omega$ gives the right idea. Second, the speed of light changes by abour fifty percent from one end to the other of the visible spectrum for which $\Delta\omega/\omega \approx 2$. Using the above expresssions for $\chi(\omega)$ shows that

$$\Delta c \approx \Delta\,(1 + 4\pi\chi)^{-1/2} \approx \frac{-1}{2}\,4\pi\frac{d\chi}{d\omega}\,\Delta\omega \approx \frac{-1}{2}\,\frac{4\pi\,(-2\omega)\,\Delta\omega\,C}{(\Omega^2-\omega^2)^2} \approx \frac{8\pi\,\omega^2 C}{(\Omega^2-\omega^2)^2}\,.$$

Thus,

$$\frac{C}{\Omega^2} \approx \frac{(\Omega^2-\omega^2)^2\Delta c}{8\pi\omega^2\Omega^2}$$

Thus γ is of the order .01.

Generalizing the derivation of the L^2 estimate for Maxwell's equations, multiply (6) by E, (7) by B, and (15) by $\partial_t P/\gamma$ to show that solutions of (6), (7), (15) which are smooth with values in $\mathcal{S}(\mathbf{R}^3)$ satisfy the energy conservation law

$$\partial_t \int |E|^2 + |B|^2 + \frac{1}{\gamma}\left(|\epsilon\partial_t P|^2 + \Omega|P|^2\right)dx = 0\,. \tag{20}$$

Note in particular the appearance of the operator $\epsilon\partial_t$ in this estimate.

Multiplying equations (6) and (7) by ϵ they become constant coefficient equations involving $\epsilon\partial$. Considering these two equations together with (15) as a system for

$$u := (E\,,\, B,\, \gamma^{-1/2}P,\, \gamma^{-1/2}\epsilon\partial_t P) \tag{21}$$

yields a symmetric hyperbolic system of the form

$$L(\epsilon\partial)u = 0\,, \tag{22}$$

where

$$L(\partial) := \sum_{j=0}^{d} A_j\frac{\partial}{\partial x_j} + L_0 = A_0\partial_0 + A_1\partial_1 + \cdots + A_d\partial_d + L_0\,. \tag{23}$$

The coefficients A_j are 12×12 hermitian symmetric matrices, L_0 is antihermitian, and, $A_0 = I$ so is strictly positive.

As indicated in the discussion of dispersion, we are interested in oscillatory solutions with period of order ϵ. As a first example seek plane wave solutions $u = \mathbf{u} e^{i\beta . x/\epsilon}$ where $\mathbf{u}$ is a constant vector. The expression is an exact solution if and only if

$$\big(L(\beta) + L_0\big)\, \mathbf{u} = 0\,. \tag{18}$$

Thus the natural $\epsilon-$characteristic variety is defined by the equation

$$\det\big(L(\beta) + L_0\big) = 0\,. \tag{24}$$

Write $\beta = (\beta_0, \beta_1, \cdots, \beta_d)$. Equation (24) asserts that β_0 is an eigenvalue of the hermitian symmetric matrix

$$-A_0^{-1/2}\,\Big(\sum_{j=1}^{d} A_j\Big)\, A_0^{-1/2}\,.$$

For $\beta_1, \cdots, \beta_d$ fixed there are only real roots β_0 and there are 12 of them counting multiplicity.

The computation of the plane waves can be finessed a little. The equations for $\mathbf{e}, \mathbf{b}, \mathbf{p}$ are

$$\omega\mathbf{e} = \xi \wedge \mathbf{b} - \omega 4\pi\mathbf{p}\,, \quad \omega\mathbf{b} = -\xi\wedge\mathbf{e}\,, \quad \big(-\omega^2 + 1\big)\mathbf{p} = \gamma\mathbf{e}\,. \tag{25}$$

For $\xi \in \mathbf{R}^3$ fixed, there is a two dimensional family of solutions with $\omega = 0$ obtained by choosing $\mathbf{e}$ and $\mathbf{b}$ to be parallel to ξ and setting $\mathbf{p} = -4\pi\mathbf{e}$. Superposition of these plane waves yields a subspace of stationary solutions of (6), (15) which satisfy in addition

$$\operatorname{curl} E = \operatorname{curl} B = E + 4\pi P = E_t = B_t = 0\,.$$

These stationary solutions are not physically relevant since they do not satisfy (5).

On the other hand, solutions with $\omega \neq 0$ automatically satisfy $\xi.(\mathbf{e} + 4\pi\mathbf{p}) = \xi.\mathbf{b} = 0$ so (5) is satisfied.

To find these solutions, multiply the first equation in (25) by ω. Eliminate $\omega\mathbf{b}$ using the second equation, and simplify the double cross product as in (3.6) to get

$$\omega^2(\mathbf{e} + 4\pi\mathbf{p}) = \xi \wedge (-\xi\wedge\mathbf{e}) = \big(\,|\xi|^2 I - |\xi><\xi|\,\big)\,\mathbf{e}\,. \tag{26}$$

Multiplying by γ and using the last of the equations in (25) to eliminate $\gamma\mathbf{e}$ yields

$$\Big(\omega^2(1 - \omega^2 + 4\pi\gamma)I - \big(|\xi|^2 I - |\xi><\xi|\big)\big(1-\omega^2\big)\Big)\mathbf{p} = 0\,. \tag{27}$$

For $\mathbf{p}$ parallel to ξ, the second term vanishes. Thus, in addition to the solutions with $\omega = 0$, there are two one dimensional spaces of solutions given by taking

$$\omega = \pm(1 + 4\pi\gamma)^{1/2}\,, \qquad \mathbf{p} \parallel \xi\,, \tag{28}$$

and, defining $\mathbf{e}$ and $\mathbf{b}$ by the last two equations in (25). The superposition in ξ of these plane waves gives solutions of (6), (7), (19) defined by

$$E + 4\pi P = 0\,, \qquad B = 0\,, \qquad \operatorname{curl} P = 0\,, \tag{29}$$

$$\Big((\epsilon\partial_t)^2 + (1 + 4\pi\gamma)\Big)P = 0\,. \tag{30}$$

These solutions are simply oscillatory in space time with frequency equal to $\sqrt{1+4\pi\gamma}$. This is an invariant subspace of the evolution defined by (6), (7), (19). This part of the solution oscillates in place indefinitely into the past and future. In particular, if it is not present in the initial conditions, the solution will remain orthogonal to these modes for all time. For this reason, these modes do not play a role in the dispersion of waves which move.

The interesting modes are those with $\mathbf{p}$ orthogonal to ξ. These satisfy $\operatorname{div} E = \operatorname{div} P = 0$. Equation (27) shows that ω must satisfy

$$0 = \omega^2(1 + 4\pi\gamma - \omega^2) - (1 - \omega^2)|\xi|^2 = (\omega^2 - |\xi|^2)(1 - \omega^2) + 4\pi\gamma\omega^2\,. \tag{31}$$

The roots are $\omega = \pm\sqrt{\rho_j}$ where $\rho_1(\xi)$ and $\rho_2(\xi)$ are the roots of the quadratic equation

$$0 = (\rho - |\xi|^2)(1 - \rho) + 4\pi\gamma\rho = -\rho^2 + (1 + 4\pi\gamma + |\xi|^2)\rho - |\xi|^2\,, \tag{32}$$

both of which are positive. For each root ω there is a two dimensional space of plane waves defined by taking $\mathbf{p}$ to lie in the two dimensional space of vectors perpendicular to ξ and then defining $\mathbf{e}$ and $\mathbf{b}$ by the last two equations of (25). That the first equation in (25) is satisfied is guaranteed by the choice of $\omega, \mathbf{p}$. This eight dimensional family of plane waves, together with two dimensions satisfying $\omega = 0$ and two dimensions from (28) is an exhaustive list.

Thus, the $\epsilon-$characteristic variety is given by

$$\{\omega = 0\} \,\cup\, \{\omega^2 = 1 + 4\pi\gamma\} \,\cup\, \{\omega^2(1 + 4\pi\gamma - \omega^2) = (1 - \omega^2)|\xi|^2\}\,. \tag{33}$$

Note that neither of the last two subvarieties is conic. The multiplicities computed above show that the $\epsilon-$characteristic equation (18) is

$$\omega^2\left(\omega^2 - 1 - 4\pi\gamma\right)\left(\omega^2(1 + 4\pi\gamma - \omega^2) - |\xi|^2(1 - \omega^2)\right)^2 = 0\,. \tag{34}$$

It is the branch given by the last of the sets in (33) which is most interesting.

Though the solution of (31) for ω as a function of ξ is not simple, the expression for ξ as a function of ω is simple,

$$|\xi|^2 = \omega^2\left(1 + \frac{4\pi\gamma}{1 - \omega^2}\right). \tag{35}$$

There a solution ξ if and only if the right hand side is nonnegative which occurs exactly when $1-\omega^2$ and $1+4\pi\gamma-\omega^2$ have the same sign. They are both positive when $0 < \omega^2 < 1$, and they are both negative when $\omega^2 > 1+4\pi\gamma$. For the intermediate range $1 < \omega^2 < 1+4\pi\gamma$ there are no solutions. For ξ fixed there are two solutions ω^2, one in each of the allowable ranges. The solution with $\omega^2 < 1$ corresponds to waves with frequency smaller than the resonant frequency of the oscillators, and it is these that are natural for visible light. The upper range is in the ultraviolet, and the intermediate range with no plane waves is a frequency band of opacity. Group velocities can be computed as the reciprocal of $d\xi/d\omega$. This is done in the 1-d case of scalar ξ in §4 of [FLS] I-48.

The transition from transparency to opacity to trasnparency is not opbserved in real materials. The reason is that real materials have many natural frequencies. In the above calculation, the natural frequency, Ω corresponds to $\omega = 1$. If the classical oscillator were repaced by a quantum oscillator, there would be natural frequencies at all the integers. For atoms is is more common that the natural frequencies cluster closer together as the frequency increases. The net effect is that for most materials, there is a transition from transparency to opacity, with no return to transparency.

§5. Asymptotic solutions of the $\epsilon\partial$ equations and group velocities.

As already remarked, standard geometric optics, that is wavelengths which are small compared to all other distances in the problem, yields speeds of transport which may depend on the direction of ξ but not on $|\xi|$. The goal of this section is to show that wavelength ϵ solutions of equations in $\epsilon\partial$ have speed of propagation which can depend on $|\xi|$.

For simplicity of exposition only linear phases are discussed below. Standard geometric optics solutions of wavelength ϵ are asymptotic solutions

$$L(\partial)u^\epsilon \sim 0\,, \qquad u^\epsilon \sim e^{i\beta.x/\epsilon} \sum_{j=0}^{\infty} \epsilon^j\, a_j(t,x)\,. \tag{1}$$

Here β belongs to the characteristic variety $\det L_1(\beta) = 0$ and the a_j are determined uniquely from the values of $\pi(\beta)\, a_j(0,x)$ where π denotes the orthogonal projection onto $\ker L_1$. The amplitudes are determined from transport equations along the rays of geometric optics.

For the dispersive case we seek solutions

$$L(\epsilon\partial)\, u^\epsilon \sim 0\,, \qquad u^\epsilon \sim e^{i\beta.x/\epsilon} \sum_{j=0}^{\infty} \epsilon^j\, a_j(t,x)\,. \tag{2}$$

The difference is the $\epsilon\partial$. Rather than work the general case or the relatively complicated Maxwell-Lorenz system, (4.6, 4.7, 4.19), we present the simplest case which reveals the essential outcome without the complications.

Seek asymptotic solutions as in the second equation in (2) to the scalar equation

$$(\epsilon^2\Box + 1)u = 0\,. \tag{3}$$

Plugging in yields

$$\left(\epsilon^2\Box + 1\right)\left(e^{i\beta.x/\epsilon}\sum_{j=0}^{\infty}\epsilon^j\,a_j(t,x)\right) = e^{i\beta.x/\epsilon}\sum_{j=0}^{\infty}\epsilon^j\,c_j(t,x)\,.$$

and the strategy is to choose the amplitudes a_j so that all the c_j vanish. The first term is given by

$$c_0 = (-\beta_0^2 + \beta_1^2 + \cdots + \beta_d^2 + 1)\,a_0\,.$$

For this to vanish with nonzero a_0, forces β to satisfy the $\epsilon - char$ equation

$$\beta_0^2 = \beta_1^2 + \cdots + \beta_d^2 + 1\,. \tag{4}$$

Note that this is not homogeneous in β. In fact if β is $\epsilon-$characteristic then $k\beta$ is not ϵ-characteristic as soon as $k \neq \pm 1$. Equation (4) guarantees that for a constant a, the function $ae^{i\beta.x/\epsilon}$ is exactly anihillated by $\epsilon^2\Box + 1$, so, represents an exact plane wave solution. The principal term in (2) can be viewed as plane wave solutions with slowly varying amplitude. The amplitude varies on length scale $O(1)$ while the phase varies much more rapidly with characteristic length scale $O(\epsilon) << 1$. The expansions are examples of the *slowly varying envelope approximation*, the amplitude $a_0(t,x)$ describing the envelope of the oscillations.

Assuming (4) is satisfied the next term in (3) is given by

$$c_1 = 2i\Big(\beta_0\,\partial_t - \sum_{j=1}^{d}\beta_j\,\partial_j\Big)\,a_0\,.$$

Setting this equal to zero yields the transport equation

$$\Big(\beta_0\,\partial_t - \sum_{j=1}^{d}\beta_j\,\partial_j\Big)\,a_0 = 0\,. \tag{5}$$

The velocity of this transport is given by

$$\text{transport velocity} = \frac{-(\beta_1,\cdots,\beta_d)}{\beta_0}\,. \tag{6}$$

For $j \geq 2$, one has

$$c_j = 2i\Big(\beta_0\,\partial_t - \sum_{j=1}^{d}\beta_j\,\partial_j\Big)\,a_{j-1} + \Box\,a_{j-2}\,.$$

The amplitudes are uniquely determined from their values at $t = 0$ by solving sucessively the transport equations

$$2i\Big(\beta_0\,\partial_t - \sum_{j=1}^{d}\beta_j\,\partial_j\Big)\,a_j + \Box\,a_{j-1} = 0\,, \tag{7}$$

where $a_{-1} := 0$.

Next compare the velocity given by (4) and (6) with the phase and group velocites of the plane wave solutions $e^{i\beta.x/\epsilon}$. As in (2.7), the phase velocity is given by

$$\text{phase velocity} := \frac{-\beta_0\,(\beta_1, \cdots, \beta_d)}{\beta_1^2 + \cdots + \beta_d^2}\,.$$

One computes the *group velocity* (see [W]) by setting the gradient with respect to β' of the phase equal to zero, that is

$$0 = \nabla_{\beta'}\Big(\beta_0(\beta')t + \beta'.x\Big)\,. \tag{8}$$

This yields

$$\text{group velocity} = \frac{x}{t} = -\nabla_{\beta'}\beta_0\,, \tag{9}$$

In the present case, implicit differentiation of equation (4) yields

$$2\beta_0\,\nabla_{\beta'}\beta_0 = 2\beta'\,, \quad \text{so} \quad \nabla_{\beta'}\beta_0 = \frac{\beta'}{\beta_0}\,. \tag{10}$$

plugging this into (9) shows that the group velocity is identical with velocity of the transport equation for the amplitudes as given in (6).

Asymptotic solutions of the $\epsilon\partial$-Klein-Gordon at wavelength ϵ gives the usual group velocity. Once this formal derivation is understood, it is easy to give rigorous results in the linear case. See [D], [DR], and [JMR] for the more difficult and more interesting nonlinear case where both the formal and rigorous theory pose challenges. In the nonlinear case, the rigorous analysis leans heavily on tools recently developed in the mathematical theory called *nonlinear geometric optics*.

A sample linear theorem is presented below. This clarifies two puzzles in the standard textbook derivations of group velocity. The first is that the velocity occurs in the large time asymptotics of a Fourier integral (see §6), while one observes these speeds at finite time. The second is that they are rigidly tied to Fourier integral expressions and thereby to constant coefficient operators. The group velocity appears in the applied literature in slowly varying envelope approximations for which it is often difficult to identify the natural approximation parameter(s). One consequence of the above remarks is that one should be on the lookout for operators in $\epsilon\partial$.

We turn next to the proof that the series constructed above gives an asymptotic expansion to all orders of exact solutions to (3). If smooth initial data supported in a fixed compact set are prescribed for the amplitudes $a_j(0,x)$, then smooth compactly supported a_j are determined for all time by solving the transport equations (7).

Borel's Theorem asserts that one can find a smooth function $a(\epsilon, t, x)$ whose support intersect $|t| \le T$ in a compact set for all $T > 0$ and so that

$$a(\epsilon, t, x) \sim \sum_{j=0}^{\infty} \epsilon^j\, a_j(t,x) \tag{11}$$

in $C^\infty(\mathbf{R}^{1+d})$.

Theorem. *Suppose that a smooth compactly supported amplitude $a(\epsilon, t, x)$ is constructed as in the last paragraph. Define*

$$u(\epsilon, t, x) := e^{i\beta.x/\epsilon}\, a(\epsilon, t, x)\,, \tag{12}$$

and, let $v(\epsilon, t, x)$ be the unique solution of the initial value problem

$$(\epsilon^2\,\Box + 1)\, v = 0\,, \qquad v(\epsilon, 0, x) = u(\epsilon, 0, x) \qquad v_t(\epsilon, 0, x) = u_t(\epsilon, 0, x)\,. \tag{13}$$

Then, for every $T > 0$, $\alpha \in \mathbf{N}^{1+d}$, and, $N \in \mathbf{N}$ there is a $C = C(\alpha, N, T)$ so that

$$\left|\, \partial^\alpha_{t,x}\big(\, u(\epsilon, t, x) - v(\epsilon, t, x)\,\big)\,\right| \le C\,\epsilon^N\,, \qquad \forall\ |t| \le T, \quad x \in \mathbf{R}^d\,. \tag{14}$$

Proof. We give the proof for real solutions. For the complex case, either work out a complex energy identity or apply the real case to the real and imaginary parts.
Use the energy identity

$$(\partial_t w)\,\big(\epsilon^2\Box + 1\big) w = \frac{1}{2}\,\frac{\partial}{\partial t}\,\big((\epsilon w_t)^2 + |\nabla_x w|^2\big) - \sum_{j=1}^{d} \frac{\partial}{\partial x_j}(\partial_t w\,\partial_j w)\,. \tag{15}$$

Fix $T > 0$ and N, and for $t \in]0, T]$ integrate over $]0, t[\times \mathbf{R}^d$ to find

$$\begin{aligned}\int_{\mathbf{R}^d} (\epsilon w_t(t,x))^2 + |\nabla_x w(t,x)|^2\ dx = \\ \int_{\mathbf{R}^d} (\epsilon w_t(0,x))^2 + |\nabla_x w(0,x)|^2\ dx + 2\int_0^t \int_{\mathbf{R}^d} \epsilon w_t\ \frac{(\epsilon^2\Box + 1)w}{\epsilon}\ dxdt\,.\end{aligned} \tag{16}$$

Apply this to $w := \partial^\mu_{t,x}\big(u(\epsilon,t,x) - v(\epsilon,t,x)\big)$. Then, the first term on the right is $O(\epsilon^N)$ as is the second factor in the last integral. Using the Schwarz intequality in the last integral and then Gronwall's method shows that for all β, and, $|t| \le T$,

$$\int_{\mathbf{R}^d} \left|\, \partial^\beta_{t,x}\big(u(\epsilon,t,x) - v(\epsilon,t,x)\big)\,\right|^2\ dx \le c(\beta, N, T)\,\epsilon^N\,.$$

Sobolev's inequality then implies (14). $\square$

It is important to note that hyperbolic equations in $\epsilon\partial$ can have solutions which grow exponentially like $e^{at/\epsilon}$. If that were the case for our problem, the fact that an approximate solution had residual $O(\epsilon^\infty)$ would not imply that it was a good approximation. Exactly this occurs if one considers the opposite sign $\epsilon^2\Box - 1$. Intriguingly, the formal asymptotic solutions in that case yield group velocities

$$-\frac{(\beta_1, \cdots, \beta_d)}{\sqrt{\beta_1^2 + \cdots + \beta_d^2 - 1}}\,.$$

which are larger than one for $|\beta| > 1$. Since signals cannot travel with such speed, there must be something wrong, and it is exactly the explosive growth which spoils the pudding.

§6. Long time asymptotics and group velocity.

The standard rigorous analysis which introduces the group velocity is the study or the long time evolution of amplitude modulated moderate frequency solutions. For example consider the solution of the initial value problem

$$\Box u + u = 0\,, \qquad u(0,x) = 0\,, \qquad u_t(0,x) = g \in \mathcal{S}(\mathbf{R}^3)\,.$$

The explicit solution is the sum of two terms

$$\frac{\pm 1}{2(2\pi)^{3/2}} \int e^{i(x.\xi \pm \langle\xi\rangle t)}\, \frac{\hat{g}(\xi)}{\langle\xi\rangle}\, d\xi\,, \qquad \langle\xi\rangle := (1 + |\xi|^2)^{1/2}\,. \tag{1}$$

Consider the term with the minus sign. The phase is stationary exactly when

$$x = t\nabla_\xi\langle\xi\rangle = \frac{t\xi}{\langle\xi\rangle}\,. \tag{2}$$

If $\hat{g}$ has support near $\underline{\xi}$ the solution propagates with speed $\underline{\xi}/\langle\underline{\xi}\rangle$. The solution is $O(t^{-\infty})$ along rays moving with velocity v such that $v \neq \xi/\langle\xi\rangle$ for all $\xi \in \operatorname{supp}\hat{g}$. This is easily proved by the prinicipal of nonstationary phase as in [R] pages 149-150. For an observer moving at one of the group velocities the solution decays as $t^{-3/2}$ the exact behavior given by the principal of stationary phase as in [W] §11.2.

The long time problem can be transformed to a special case of the of wavelength ϵ asymptotic solutions to equations in $\epsilon\partial$ as in (5.2). An observer moving with velocity v observes a wave which is spreading out linearly with time. This leads one to consider the behavior as $\lambda \to \infty$ of $u(\lambda, \lambda v + B_{\delta\lambda})$ where $B_{\delta\lambda}$ denotes the ball of radius $\delta\lambda$ with center at the origin. Introduce

$$u_\lambda := u(\lambda t, \lambda x)\,, \tag{3}$$

then the observer is interested in the behavior as $\lambda \to \infty$ of $u_\lambda(1, B_\delta)$. In addition, u_λ satisfies the $\epsilon\partial$ initial value problem

$$\Big(\epsilon^2\Box + 1\Big)u_\lambda = 0\,, \qquad \epsilon := \frac{1}{\lambda}\,, \tag{4}$$

$$u_\lambda(0,x) = 0\,, \qquad \partial_t u_\lambda(0,x) = \lambda g(x/\epsilon)\,. \tag{5}$$

Provided that the Fourier Transform of g has support disjoint from the origin, this is an initial value problem which can be analyzed by writing

$$g(x/\epsilon) = (2\pi)^{-d/2} \int_{\mathbf{R}^d} \hat{g}(\beta)\, e^{i\beta . x/\epsilon}\, d\beta\,.$$

The problem with initial data $g(\beta)\, e^{i\beta . x/\epsilon}$ is solved by the method of the last section, with initial amplitudes independent of x. Superposition shows that for times of order 1, u_λ is a sum of terms moving at the natural group velocity, recovering the result of the stationary phase argument by an argument which generalizes beyond the domain of applicability of the Fourier Transform.

References.

R. Boyd, *Nonlinear Optics*, Academic Press 1992.

P. Donnat, Quelque contributions mathématiques en optiques nonlinéaire, These doctoral de L'Ecole Poytechnique, Palaiseau, 1994.

P. Donnat and J. Rauch, Dispersive nonlinear geometric optics, preprint.

R. P. Feynmann, R. Leighton, and M. Sands, *The Feynmann Lectures on Physics Vols. I., II.*, Addison-Wesley Pub., 1964

J.-L. Joly, G. Metivier, and J. Rauch, in preparation.

H.A. Lorentz, *Theory of Electrons*, Teubner 1908, Reprinted by Dover 1952.

E. Purcell, *Electricity and Magnetism*, McGraw-Hill, 1965.

J. Rauch, *Partial Differential Equations*, Sringer-Verlag, 1991.

G. Whitham, *Linear and Nonlinear Waves*, John Wiley and Sons Inc. 1974.