Direct and Inverse Scattering of Electromagnetic Waves in Nonlinear Media

Sjöberg, Daniel

1999

Direct and Inverse Scattering of Electromagnetic Waves in Nonlinear Media

by

Daniel Sjöberg

Thesis for the degree of Licentiate in Engineering

CODEN:LUTEDX/(TEAT-1012)/1-25/(1999)

Department of Electromagnetic Theory
Lund Institute of Technology

Abstract

Models for materials with nonlinear electromagnetic response are examined. Three simple physical causes for nonlinear behavior are presented: electronic polarization, molecular direction and electrostriction. Some introductory results and approximations for nonlinear, dispersive media are given, such as the phase-matching criterion, Miller's rule, the Born approximation and the slowly varying amplitude approximation. However, the emphasis is on instantaneously reacting media, *i.e.*, materials with no or negligible memory. For these models, the inverse scattering problem of determining the field dependent permittivity and permeability from the scattered field is solved.

List of papers

This thesis consists of an introduction and the following two papers:

- Paper I

  Reconstruction of nonlinear material properties for homogeneous, isotropic slabs using electromagnetic waves

  D. Sjöberg

  Inverse Problems, 15(2), 431–444, 1999 (in press)

- Paper II

  Simple wave solutions for the Maxwell equations in bianisotropic, nonlinear media, with application to oblique incidence

  D. Sjöberg

  Technical Report LUTEDX/(TEAT-7078)/1–17/(1999)

Contents

Linear

Nonlinear

Figure 1: To the discussion on nonlinearity in the supermarket.

1 Introduction

How can we explain the term *nonlinear*? As the name suggests, it is a *non*-property, a nonlinear material is characterized by not having a property called "linear". This is a rather odd way of classifying things; instead of saying what they *are*, we try to say what they are *not*. However, the concept of linearity is deeply rooted in our everyday experiences. When we go to the supermarket and weigh the fruit we want to buy, we expect two oranges to weigh twice as much as one orange. This is the essence of linearity; when we get one output for a certain input, we expect to get twice as much when we double the input.

Strangely enough, the concept of nonlinearity is upset in the same supermarket, when we come to the candy department. This is the residence of the signs saying "Buy three, pay for two!" Thus, we pay say 5 crowns for one chocolate bar, but when we buy three, we only pay 10 crowns, not 15! This is an example of *non*-linearity.

Another situation where we see that a linear model cannot apply, is the speed of cars on a highway. When there are few cars on the road, each driver is relatively free to choose his/her own speed. Though, when the number of cars increases, there is less space to maneuver in. This results in an overall reduction of speed, which ultimately may turn into a total standstill, as all commuters probably are aware of. This is an example of *saturation*, which is a typically nonlinear phenomenon.

The latter example of nonlinearity, *i.e.*, cars influencing each others speed on the highway, actually has many similarities with electromagnetic waves propagating in nonlinear materials. The nonlinearity of the cars is that the speed depends on how many cars there are. In the nonlinear materials we study in this thesis, the speed of the electromagnetic waves depends on the electromagnetic energy in the media, *i.e.*, the density of the electromagnetic fields. This causes waves with high amplitudes to travel slower than waves with low amplitudes, just as cars on an empty highway can travel much faster than the cars on a jammed one.

In this thesis, we also discuss the so called inverse scattering problem. This is a term that we use for a kind of scattering problem where we can measure the effects of something, *e.g.*, the reflected field from a surface, and wish to extract some information on the cause of this effect, *e.g.*, what material the surface was made of. We have all experienced this problem, for instance when being X-rayed at the dentists in order to determine whether our teeth have cavities or not. The inverse scattering problem treated in this thesis, is to determine the nonlinear properties of

Nonlinear process	Response time [s]
Electronic polarization	10^{-15}
Molecular direction	10^{-12}
Electrostriction	10^{-9}

Table 1: Approximate response times for different nonlinear processes. From [6, p. 163]

a slab or half space based on measurements of the reflected and transmitted fields. A few general references to inverse problems are [3, 12, 22, 35, 36], and some inverse problems regarding nonlinear materials are treated in [23, 38, 46].

The topic of nonlinear wave propagation has been the subject of considerable investigation during the years. Much analytical work was made during the 60s and 70s, *e.g.*, [7, 8, 26–28, 34, 49]. Many results from this period are collected in [4], as well as suggestions for further investigations. Nonlinear optics is a well established field with vast literature, of which only some is given here [2, 6, 51]. Methods based on the theory of partial differential equations as well as variational principles, suited for wave propagation in nonlinear materials, are presented in [14]. Some basic theory on nonlinear partial differential equations is found in [42], some more advanced in [29]. Recent results regarding nonlinear hyperbolic differential equations are reported in [24]. A suitable numerical method for nonlinear wave propagation is finite differences in the time domain, FDTD [30, 54, 61].

In this introductory part of the thesis, we explain in Section 2 some physical processes which lead to nonlinear behavior, and present a way to mathematically model these in Section 3. In Section 4 we give some more detailed theory regarding the enclosed papers, and the inverse scattering problem of determining the material parameters from oblique incidence of a plane wave is solved in Section 5.

2 Physical causes for nonlinear phenomena

In this section we describe three simple processes, which contribute to a nonlinear electromagnetic response from a material. These examples are electronic polarization, molecular direction and electrostriction. In Table 1 we list the approximate response times of the different processes.

2.1 Electronic polarization

A very simple model of the electromagnetic properties of a material is a collection of individual atoms, where the different atoms do not have a significant influence on each other. This is the same assumption as for an ideal gas, and if the wave functions for the individual atoms are known, it is possible to explicitly calculate the constitutive relations from quantum mechanics [6, Chap. 3]. In this thesis, we are content with a simpler, phenomenological, model of the atoms.

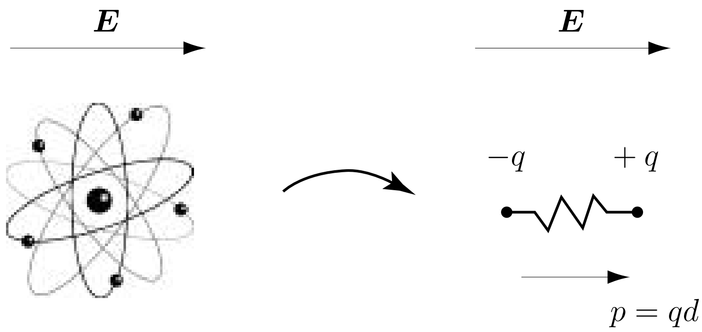

Figure 2: The atom is subjected to an electric field $\boldsymbol{E}$. Its response can be modeled with two separated charges $\pm q$, creating an electric dipole moment p proportional to the separation d. A restoring force acts on the charges, modeled with a spring connecting the charges in the figure.

Negatively charged electrons are orbiting a nucleus with positive charge. Since the electrons are orbiting the nucleus very fast, the mean positions of the electrons coincide with the position of the nucleus, and the atom appears electrically neutral without any electric dipole moment. However, when an external electric field is applied, the charges are separated and an electric dipole moment is created, see Figure 2.

The strength of this electric dipole moment depends on the restoring force between the electron and the nucleus. For small displacements, this force is proportional to the separation between the charges, *i.e.*, it behaves as a linear spring. However, when we apply a very strong external field, the situation becomes so asymmetric that we must add some additional terms to the restoring force in order to properly model the material response. For isotropic materials, it can be shown that only terms with odd exponents appear in a power series of the restoring force, see [38], [6, p. 29] and [50, p. 740],

$$F = a_1 r + a_3 r^3 + a_5 r^5 + \dots \tag{2.1}$$

where F denotes the restoring force and r denotes the separation between the charges. The term $a_1 r$ correspond to the linear response, and is in general much larger than the other terms.

2.2 Molecular direction

Many molecules are not symmetric, which results in an intrinsic electric dipole moment. A typical example of such a molecule is water, H_2O, where the center of the negative charges is close to the oxygen atom and the center of the positive charges is close to the two hydrogen atoms. Normally, the thermal excitation causes the molecules to vibrate so vividly, that the electric dipole moments are randomly oriented and make no contribution to the total electric dipole moment. However, when a strong external field is applied, the dipoles tend to be aligned along the electric field. The effect of this can be modeled with a field dependent refractive

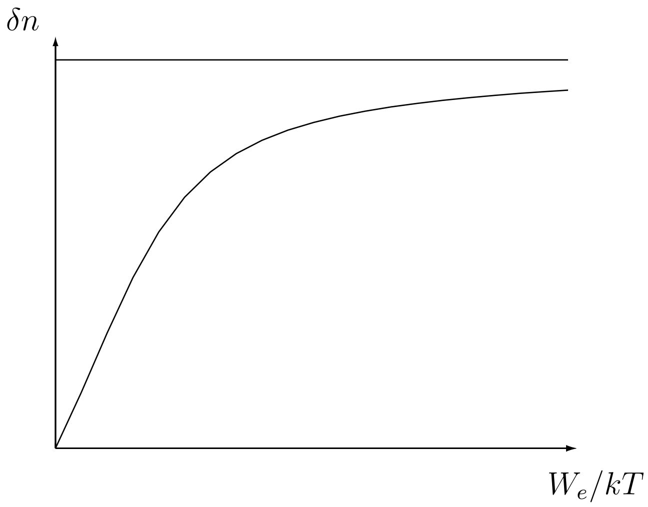

Figure 3: Dependence of the nonlinear refractive index on the electromagnetic energy $W_e \sim E^2$ and the thermal energy kT, where k is the Boltzmann constant and T is the absolute temperature. The horizontal line indicates refractive index when the molecules are completely aligned, *i.e.*, when the material is saturated. Based on [6, p. 183].

index,

$$n(E) = n_0 + \delta n(E),$$

where n_0 correspond to the refractive index at low field strengths, *i.e.*, the linear response. The nonlinear response δn depends on how strong the electromagnetic energy is compared to the thermal energy, as is seen in Figure 3.

Since the asymmetric molecules become aligned along the electric field, the material has different properties in different directions. The nonlinear effect of molecular direction is thus an *anisotropic* effect. From Figure 3 we also see that the material can be expected to saturate for large fields, *i.e.*, when the field is large enough, the molecules are totally aligned and cannot contribute more to the materials electromagnetic response. A more extensive treatment of molecular direction can be found in [5], [6, pp. 178–185] and [10, p. 203].

2.3 Electrostriction

The third nonlinear effect presented here, is the interaction between an applied electric field and mechanic forces called *electrostriction*. This is a macroscopic effect, where the dipoles in the material induced by an external field attract each other so strongly, that the material is compressed. Since the refractive index depends on the density of the material, this implies a refractive index which grows with the electric field. Some general remarks on electrostriction are found in [40, pp. 55–56] and [53, pp. 149–151], and a thorough treatment of electromechanical coupling is given in [16, 45].

Since electrostriction implies a coupling between electric fields and mechanic forces, this effect may generate acoustic waves in the material. This is the origin of Brillouin scattering in optical fibers, where electromagnetic waves are scattered by acoustic waves with much slower propagation speed [1, p. 371].

3 Constitutive relations and some common approximations

In this section we study a nonlinear, anisotropic, homogeneous, dispersive dielectric material with local constitutive relations. The analysis is easily generalized to bianisotropic materials, but we only wish to point out some general aspects. We do not treat wave propagation in nonlinear dispersive materials in this thesis, but some early results can be found in [43, 59, 60]. Numerous examples of nonlinear constitutive relations can be found in [10, 16, 17].

3.1 Expansion of the polarization functional

We assume a constitutive relation between polarization and electric field as

$$\boldsymbol{P}(\boldsymbol{r},t) = \{\mathcal{P}\boldsymbol{E}\}(\boldsymbol{r},t),$$

where $\mathcal{P}$ is an operator, for which we assume that the following calculations are valid. For weak electric fields, the operator may be expanded in a Volterra series, see [9], [50, p. 783] and [57, p. 21],

$$\begin{aligned}\{\mathcal{P}\boldsymbol{E}\}_i(\boldsymbol{r},t) =&\epsilon_0\Big\{\int dt_1 \chi^{(1)}_{ij}(t_1)E_j(\boldsymbol{r},t-t_1)\\ &+\int dt_1\int dt_2 \chi^{(2)}_{ijk}(t_1,t_2)E_j(\boldsymbol{r},t-t_1)E_k(\boldsymbol{r},t-t_2)\\ &+\int dt_1\int dt_2\int dt_3 \chi^{(3)}_{ijkl}(t_1,t_2,t_3)E_j(\boldsymbol{r},t-t_1)E_k(\boldsymbol{r},t-t_2)E_l(\boldsymbol{r},t-t_3)\\ &+\dots\Big\},\end{aligned} \tag{3.1}$$

where we have assumed the Einstein convention that summation occurs over multiple indices, and alphabetical indices refer to the spatial coordinates x, y and z. The integrations are over all real numbers, which implies that the kernels $\chi^{(n)}$ must contain step functions to ensure causality, *i.e.*, $\chi^{(1)}_{ij}(t) = H(t)\chi^{(1)}_{ij}(t)$, see *e.g.*, [25, p. 332] and [32, 56]. The Volterra series is a generalization of the Taylor series for functionals, and functionals which can always be expanded in such a manner are called analytic functionals [57, p. 21].

From the Volterra series we see an immediate effect of the nonlinear constitutive relations: the generation of harmonics. It suffices to study what happens with a field that varies as a cosine in time, *i.e.*, $E(t) \sim \cos(\omega_0 t)$. The second term in the above series then contains terms proportional to

$$\begin{aligned}
&\cos(\omega_0(t-t_1))\cos(\omega_0(t-t_2)) \\
&\qquad = \frac{1}{2}[\cos(\omega_0(t-t_1)+\omega_0(t-t_2))+\cos(\omega_0(t-t_1)-\omega_0(t-t_2))] \\
&\qquad = \frac{1}{2}[\cos(2\omega_0 t-\omega_0(t_1+t_2))+\cos(\omega_0(t_2-t_1))],
\end{aligned}$$

which implies terms with frequencies $2\omega_0$ and zero. In the next section we treat these processes more systematically by applying a Fourier transform in time and space to the constitutive relation.

3.2 Fourier transformation of the polarization functional

We define the four-dimensional Fourier transform of a scalar function f of time and space by

$$\begin{cases} f(\boldsymbol{r},t) = \dfrac{1}{(2\pi)^4}\displaystyle\int d^3\boldsymbol{k}\int d\omega f(\boldsymbol{k},\omega)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \\ f(\boldsymbol{k},\omega) = \displaystyle\int d^3\boldsymbol{r}\int dt f(\boldsymbol{r},t)e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}, \end{cases}$$

where we distinguish the transform from the function by their arguments. We apply this transform to the n:th term in the Volterra series, which we denote by $P_i^{(n)}(\boldsymbol{r},t)$, where

$$P_i^{(n)}(\boldsymbol{r},t) = \epsilon_0\int dt_1\cdots\int dt_n \chi^{(n)}_{ip_1\ldots p_n}(t_1,\ldots,t_n)E_{p_1}(\boldsymbol{r},t-t_1)\ldots E_{p_n}(\boldsymbol{r},t-t_n).$$

We emphasize again that the indices $p_1,\ldots,p_n$ are summed over, allowing the material to be anisotropic. Following the approach in [9], we apply the four-dimensional Fourier transform to this expression, use the inverse Fourier transform to express the space-time dependence of the electric fields, and change the order of integration:

$$
\begin{aligned}
P_i^{(n)}(\boldsymbol{k},\omega) &= \epsilon_0 \int d^3\boldsymbol{r} \int dt \int dt_1 \cdots \int dt_n \chi^{(n)}_{ip_1\ldots p_n}(t_1,\ldots,t_n) \\
&\quad \cdot E_{p_1}(\boldsymbol{r}, t-t_1)\ldots E_{p_n}(\boldsymbol{r}, t-t_n) e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \\
&= \epsilon_0 \int d^3\boldsymbol{r} \int dt \int dt_1 \cdots \int dt_n \chi^{(n)}_{ip_1\ldots p_n}(t_1,\ldots,t_n) e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \\
&\quad \cdot \frac{1}{(2\pi)^{4n}} \int d^3\boldsymbol{k}_1 \cdots \int d^3\boldsymbol{k}_n \int d\omega_1 \cdots \int d\omega_n \\
&\quad \cdot E_{p_1}(\boldsymbol{k}_1,\omega_1)\ldots E_{p_n}(\boldsymbol{k}_n,\omega_n) e^{i\sum_{j=1}^n [\boldsymbol{k}_j\cdot\boldsymbol{r}-\omega_j(t-t_j)]} \\
&= \epsilon_0 \frac{1}{(2\pi)^{4n}} \int d^3\boldsymbol{k}_1 \cdots \int d^3\boldsymbol{k}_n \int d\omega_1 \cdots \int d\omega_n \\
&\quad \cdot \int dt_1 \cdots \int dt_n \chi^{(n)}_{ip_1\ldots p_n}(t_1,\ldots,t_n) e^{i\sum_{j=1}^n \omega_j t_j} \\
&\quad \cdot E_{p_1}(\boldsymbol{k}_1,\omega_1)\ldots E_{p_n}(\boldsymbol{k}_n,\omega_n) \\
&\quad \cdot \int d^3\boldsymbol{r} \int dt e^{-i[(\boldsymbol{k}-\sum_{j=1}^n \boldsymbol{k}_j)\cdot\boldsymbol{r}-(\omega-\sum_{j=1}^n \omega_j)t]} \\
&= \epsilon_0 \frac{1}{(2\pi)^{4n-4}} \int d^3\boldsymbol{k}_1 \cdots \int d^3\boldsymbol{k}_n \int d\omega_1 \cdots \int d\omega_n \\
&\quad \cdot \chi^{(n)}_{ip_1\ldots p_n}(\omega_1,\ldots,\omega_n) E_{p_1}(\boldsymbol{k}_1,\omega_1)\ldots E_{p_n}(\boldsymbol{k}_n,\omega_n) \\
&\quad \cdot \delta(\boldsymbol{k}-\sum_{j=1}^n \boldsymbol{k}_j)\delta(\omega-\sum_{j=1}^n \omega_j),
\end{aligned}
$$

where we have used the well-known representation of the δ-distribution

$$\frac{1}{(2\pi)^4}\int d^3\boldsymbol{r}\int dt e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \delta(\boldsymbol{k})\delta(\omega).$$

Since the δ-distribution contributes to the integrals only when its argument is zero, we interpret the final expression as a sum of all terms $\epsilon_0 \frac{1}{(2\pi)^{4n-4}} \chi^{(n)}_{ip_1\ldots p_n}(\omega_1,\ldots,\omega_n) \cdot E_{p_1}(\boldsymbol{k}_1,\omega_1)\ldots E_{p_n}(\boldsymbol{k}_n,\omega_n)$ which satisfy

$$\sum_{j=1}^n \boldsymbol{k}_j = \boldsymbol{k} \quad \text{and} \quad \sum_{j=1}^n \omega_j = \omega.$$

These are the only ones contributing to the n:th order field with wave vector $\boldsymbol{k}$ and frequency ω. The result for the wave vectors is called *phase-matching*, and is very important in nonlinear optics. To create a wave with frequency equal to the sum of two frequencies, we must orientate the beams containing the constituent frequencies so that the phase-matching criterion is satisfied. Since many nonlinear materials are anisotropic, this involves a careful choice of polarization. Observe that since the integrations are performed over both positive and negative frequencies, the new frequency can also be the difference between two frequencies.

In [6] the phase-matching criterion is derived using the Maxwell equations, but we have shown it to be a consequence of the constitutive relation only. This has earlier been shown in [9].

3.3 Miller's rule in the time domain

It is possible to express the nonlinear kernels $\chi^{(n)}$ from the expansion (3.1) in the linear kernel $\chi^{(1)}$, see *e.g.*, [6, p. 27] and [50, p. 786]. This is called *Miller's rule* in nonlinear optics, and was found empirically in the 60s. We show how it can be derived in our time domain analysis.

In this section we temporarily neglect the vector character of the fields, and study linearly polarized fields, *e.g.*, $\boldsymbol{P} = P\hat{x}$. Homogeneous dispersive materials can be modeled by a differential equation, see [19] and [50, p. 784],

$$\mathcal{L}P + f(P) = E, \tag{3.2}$$

where $\mathcal{L}$ is a linear differential operator in time and $f(P)$ is a nonlinear term of order P^2 for small P. For a Lorentz model with one resonance frequency, *i.e.*, where the atoms are modeled with charges subjected to a restoring force as in Section 2.1, we have $\mathcal{L} = k(\frac{\partial^2}{\partial t^2} + \nu\frac{\partial}{\partial t} + \omega_0^2)$, where ω_0 denotes the resonant frequency and ν the collision frequency, *i.e.*, the linewidth. $f(P)$ then represents the nonlinear correction to the restoring force. Assuming small nonlinearities, we expand the function f as

$$f(P) = a_2P^2 + a_3P^3 + \dots$$

and make the perturbation Ansatz

$$P = \lambda P^{(1)} + \lambda^2 P^{(2)} + \lambda^3 P^{(3)} + \dots,$$

with the right hand side of (3.2) equal to λE. The scalar parameter λ varies continuously from 0 to 1, where 1 corresponds to the physical situation at hand. Collecting terms of corresponding orders in λ, we have for λ and λ^2

$$\begin{cases} P^{(1)} = \mathcal{L}^{-1}E \\ P^{(2)} = -a_2\mathcal{L}^{-1}[P^{(1)}]^2 = -a_2\mathcal{L}^{-1}[\mathcal{L}^{-1}E]^2. \end{cases}$$

We assume that the linear operator $\mathcal{L}^{-1}$ exists and is bounded and invariant under time translations. The operator can then be represented by a convolution [32], and by comparing with (3.1) we identify

$$\mathcal{L}^{-1}E = \epsilon_0 \int \chi^{(1)}(t_1)E(t-t_1)dt_1.$$

The equation for $P^{(2)}$ is now

$$\begin{aligned} P^{(2)} &= -a_2\epsilon_0^3 \int \chi^{(1)}(t_1)\left[\int \chi^{(1)}(t_2)E(t-t_1-t_2)dt_2\right]^2 dt_1 \\ &= -a_2\epsilon_0^3 \iiint \chi^{(1)}(t_1)\chi^{(1)}(t_2)\chi^{(1)}(t_3)E(t-t_1-t_2)E(t-t_1-t_3)dt_1dt_2dt_3 \\ &= -a_2\epsilon_0^3 \iint \left[\int \chi^{(1)}(t_1)\chi^{(1)}(t_2-t_1)\chi^{(1)}(t_3-t_1)dt_1\right] E(t-t_2)E(t-t_3)dt_2dt_3, \end{aligned}$$

and once again by comparing with (3.1) we conclude that

$$\chi^{(2)}(t_1,t_2) = -a_2\epsilon_0^2 \int \chi^{(1)}(t')\chi^{(1)}(t_1-t')\chi^{(1)}(t_2-t')dt'. \tag{3.3}$$

Thus we can express the nonlinear susceptibility $\chi^{(2)}$ in the linear kernel $\chi^{(1)}$. For isotropic materials the coefficient a_2 is zero, and the lowest order nonlinear susceptibility $\chi^{(3)}$ is

$$\chi^{(3)}(t_1,t_2,t_3) = -a_3\epsilon_0^3 \int \chi^{(1)}(t')\chi^{(1)}(t_1-t')\chi^{(1)}(t_2-t')\chi^{(1)}(t_3-t')dt'. \tag{3.4}$$

Since the linear susceptibility is rather thoroughly investigated, this result may simplify the analysis of dispersive nonlinear materials, and provide reasonable models for small nonlinear effects. We see that *all* susceptibilities $\chi^{(n)}$ are products of $\chi^{(1)}$ in some sense, but the expressions are more complicated for the higher order susceptibilities.

3.4 The Born approximation

Having discussed properties of the constitutive relations alone in the previous sections, we now turn to the propagation of electromagnetic waves in nonlinear dielectric media. In a dielectric material the electric flux density $\boldsymbol{D}$ depends on electric field strength $\boldsymbol{E}$ only, *i.e.*, $\boldsymbol{D} = \epsilon_0\boldsymbol{E}+\boldsymbol{P}$, and the magnetic flux density $\boldsymbol{B}$ is proportional to the magnetic field strength $\boldsymbol{H}$, *i.e.*, $\boldsymbol{B} = \mu_0\boldsymbol{H}$. With these prerequisites, the source free Maxwell equations

$$\begin{cases} \nabla\times\boldsymbol{E} + \dfrac{\partial}{\partial t}\boldsymbol{B} = 0 \\ \nabla\times\boldsymbol{H} - \dfrac{\partial}{\partial t}\boldsymbol{D} = 0 \end{cases} \tag{3.5}$$

can be written as a second order differential equation in $\boldsymbol{E}$

$$\nabla^2\boldsymbol{E} - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}\boldsymbol{E} = \mu_0\frac{\partial^2}{\partial t^2}\boldsymbol{P} = \mu_0\frac{\partial^2}{\partial t^2}\mathcal{P}\boldsymbol{E}.$$

A common method to analyze this equation, is to split the polarization functional in a linear and nonlinear part, *i.e.*, $\mathcal{P} = \mathcal{P}^{\mathrm{L}} + \mathcal{P}^{\mathrm{NL}}$, where the linear part is

$$\{\mathcal{P}^{\mathrm{L}}\boldsymbol{E}\}_i(\boldsymbol{r},t) = \epsilon_0 \int dt_1\chi_{ij}^{(1)}(t_1)E_j(\boldsymbol{r},t-t_1).$$

We write the above differential equation as

$$\nabla^2\boldsymbol{E} - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}(1+\frac{1}{\epsilon_0}\mathcal{P}^{\mathrm{L}})\boldsymbol{E} = \mu_0\frac{\partial^2}{\partial t^2}\mathcal{P}^{\mathrm{NL}}\boldsymbol{E}.$$

We denote the left hand side of this equation by $\mathcal{L}\boldsymbol{E}$ and the right hand by $\mathcal{S}\boldsymbol{E}$, where $\mathcal{L}$ is a linear differential operator, for which there are standard methods to calculate a solution for many classes of materials [31, 33, 54, 61], especially for one-dimensional wave propagation. Arguing that the nonlinear contribution $\boldsymbol{\mathcal{S}E}$ is small, we make the Ansatz

$$\boldsymbol{E} = \boldsymbol{E}_0 + \boldsymbol{E}_1 + \boldsymbol{E}_2 + \dots,$$

where $\boldsymbol{\mathcal{L}E}_0 = 0$, $\boldsymbol{\mathcal{L}E}_1 = \boldsymbol{\mathcal{S}E}_0$ and

$$\boldsymbol{\mathcal{L}E}_n = \mathcal{S}\sum_{j=0}^{n-1}\boldsymbol{E}_j - \mathcal{L}\sum_{j=0}^{n-1}\boldsymbol{E}_j = \mathcal{S}\sum_{j=0}^{n-1}\boldsymbol{E}_j - \mathcal{S}\sum_{j=0}^{n-2}\boldsymbol{E}_j,$$

for $n \geq 2$. Observe that the last expression does *not* reduce to $\boldsymbol{\mathcal{S}E}_{n-1}$, since the operator $\mathcal{S}$ is nonlinear. When we have solved the starting equation $\boldsymbol{\mathcal{L}E}_0 = 0$, we calculate $\boldsymbol{\mathcal{S}E}_0$, solve $\boldsymbol{\mathcal{L}E}_1 = \boldsymbol{\mathcal{S}E}_0$ for $\boldsymbol{E}_1$ and then keep iterating for higher order terms. This is the Born iterative procedure, which is likely to converge for small nonlinearities. In this thesis, we are not concerned with convergence of this series, since we use a different approach to the nonlinear wave propagation, which is further developed in Section 4 and the enclosed papers.

3.5 Slowly varying amplitude and solitons

Before turning to the analysis used in the enclosed papers, we briefly mention a common approximation for fixed frequency. We have previously mentioned that a prominent feature of nonlinear media is the generation of new frequencies. For materials with a cubic term in the polarization functional, there is a nonlinear contribution at the basic frequency. This has inspired the *slowly varying amplitude approximation*, where the (linearly polarized) electric field propagating in the z-direction is assumed to have a carrier frequency ω_0 with an envelope A,

$$\boldsymbol{E}(\boldsymbol{r}, t) = \hat{x}\,\mathrm{Re}\,A(z,t)e^{i(kz-\omega_0 t)},$$

where $A(z,t)$ is supposed to vary slowly in space and time compared to the exponential factor. This approximation leads to the nonlinear Schrödinger equation for the envelope [1, 6, 21],

$$i\frac{\partial A}{\partial z} - \frac{1}{2}k_2\frac{\partial^2 A}{\partial \tau^2} + \gamma|A|^2 A = 0,$$

where $\tau = t - kz/\omega_0$ is a time variable relative to the wave front. The factor $k_2 = \frac{d^2k}{d\omega^2}|_{\omega=\omega_0}$ is a measure of the dispersion, and γ is a measure of the nonlinearity. This is the basic equation modeling the propagation of solitons in lossless media. The existence of soliton solutions was experimentally verified by Mollenauer *et al* [44], and a theoretical treatment is found in [20, 21].

4 Instantaneous constitutive relations

This thesis is concerned with electromagnetic wave propagation in nonlinear media which responds instantaneously to excitation. This means that the Volterra series in the previous section is reduced to a power series, *i.e.*, in the isotropic case

$$P(\boldsymbol{r},t) = \epsilon_0(\chi^{(1)}E(\boldsymbol{r},t) + \chi^{(2)}E(\boldsymbol{r},t)^2 + \chi^{(3)}E(\boldsymbol{r},t)^3 + \dots),$$

where we have assumed a linear polarization.[1] We generalize this to an instantaneous constitutive relation between the electric and magnetic fluxes $\boldsymbol{D}$ and $\boldsymbol{B}$ and the electric and magnetic field strengths $\boldsymbol{E}$ and $\boldsymbol{H}$, *i.e.*,

$$\begin{cases} \boldsymbol{D}(\boldsymbol{r},t) = \boldsymbol{D}(\boldsymbol{E}(\boldsymbol{r},t),\boldsymbol{H}(\boldsymbol{r},t)) \\ \boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{B}(\boldsymbol{E}(\boldsymbol{r},t),\boldsymbol{H}(\boldsymbol{r},t)). \end{cases}$$

A note of caution is appropriate here: the above model permits an immediate coupling between the $\boldsymbol{D}$ and $\boldsymbol{H}$ fields, and between the $\boldsymbol{B}$ and $\boldsymbol{E}$ fields. This implies a *nonreciprocal* medium, see *e.g.*, [37, p. 403], [32] and [47, Chap. 8]. There is an ongoing debate regarding the existence of such materials, see *e.g.*, [39, 48, 52, 55], but no definite results can be presented at this stage. The possible existence of nonreciprocal media is easily handled in the framework presented in this thesis, and is therefore not excluded.

Following the approach in Paper II, we introduce the six-vectors [41]

$$\boldsymbol{e} = \begin{pmatrix} \sqrt{\epsilon_0}\boldsymbol{E} \\ \sqrt{\mu_0}\boldsymbol{H} \end{pmatrix} \quad \text{and} \quad \boldsymbol{d} = \begin{pmatrix} \frac{1}{\sqrt{\epsilon_0}}\boldsymbol{D} \\ \frac{1}{\sqrt{\mu_0}}\boldsymbol{B} \end{pmatrix},$$

where ϵ_0 and μ_0 are the permittivity and permeability in vacuum, respectively. All components of $\boldsymbol{e}$ and $\boldsymbol{d}$ now have the same dimension, *i.e.*, $\sqrt{\text{energy/volume}}$, and the instantaneous constitutive relations are written

$$\boldsymbol{d}(\boldsymbol{r},t) = \boldsymbol{d}(\boldsymbol{e}(\boldsymbol{r},t)).$$

Using the operator

$$\mathbf{J} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}$$

the source free Maxwell equations (3.5) are now written

$$[\nabla \times \mathbf{J}] \cdot \boldsymbol{e} + \frac{1}{c_0}\boldsymbol{\varepsilon}(\boldsymbol{e}) \cdot \partial_t \boldsymbol{e} = 0, \tag{4.1}$$

where $1/c_0 = \sqrt{\epsilon_0\mu_0}$ is the speed of light in vacuum, and the dimensionless $\boldsymbol{\varepsilon}(\boldsymbol{e})$ denotes the field gradient of the constitutive relation $\nabla_e \boldsymbol{d}$, as in Paper II. The components of $\nabla_e \boldsymbol{d}$ are $[\nabla_e \boldsymbol{d}]_{nm} = \frac{\partial}{\partial e_m} d_n(\boldsymbol{e})$. We see that (4.1) is a quasilinear partial differential equation, *i.e.*, it is linear in the derivatives of $\boldsymbol{e}$. The quasilinearity implies that it is possible to use the superposition principle for the derivatives, which permits us to use linear wave-splitting techniques [13, 15, 58]. However, the nonlinearity causes the fields to couple through $\boldsymbol{\varepsilon}(\boldsymbol{e})$, which makes the splitting local

[1] For isotropic materials $\chi^{(2)}$ is always zero, since P must be an odd function of E, see (2.1).

	Known	Unknown
Direct	$\boldsymbol{e}^i$, $\boldsymbol{\varepsilon}(\boldsymbol{e})$	$\boldsymbol{e}^s$
Inverse	$\boldsymbol{e}^i$, $\boldsymbol{e}^s$	$\boldsymbol{\varepsilon}(\boldsymbol{e})$

Table 2: The known and unknown quantities for the direct and the inverse scattering problem, respectively. $\boldsymbol{e}^i$ denotes the incident field, $\boldsymbol{e}^s$ denotes the scattered field (which may be the reflected and/or the transmitted field, depending on the situation), and $\boldsymbol{\varepsilon}(\boldsymbol{e})$ denotes the material parameters.

in both space and time, which complicates the propagation. This is clearly seen in Paper I, where a problem of one-dimensional nonlinear wave propagation is treated, and waves traveling in different directions couple through a field dependent wave speed.

It can be shown via thermodynamic considerations that $\boldsymbol{\varepsilon}(\boldsymbol{e})$ must be a symmetric, positive definite operator for a passive material, *i.e.*, a material which does not produce energy [11]. For such $\boldsymbol{\varepsilon}(\boldsymbol{e})$, the Maxwell equations are hyperbolic and the waves propagate with finite speed [18].

Further properties of the instantaneous constitutive relations are shown in the enclosed papers.

5 The inverse scattering problem

We solve the inverse scattering problem of finding the nonlinear permittivity and permeability of an isotropic, homogeneous half space, using oblique incidence of plane electromagnetic waves from vacuum. The pertinent constitutive relations are

$$\boldsymbol{\varepsilon}(\boldsymbol{e}) = \begin{pmatrix} \epsilon(\sqrt{\epsilon_0}E)\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mu(\sqrt{\mu_0}H)\mathbf{I} \end{pmatrix}.$$

The scattering situation is as in Paper II, Figure 1: an incident plane wave $\boldsymbol{e}^i(\boldsymbol{r},t) = \boldsymbol{e}^i(\hat{k}^i \cdot \boldsymbol{r} - c_0 t)$ is impinging from vacuum on a half space of nonlinear material. The unit vector $\hat{k}^i$ denotes the propagation direction of the incident field. The nonlinear material is contained in $z > 0$, and the plane of incidence is in the y-z plane. This implies that there is no propagation in the x direction. The incident plane wave causes a reflected field $\boldsymbol{e}^r$ and a transmitted field $\boldsymbol{e}^t$, with unit propagation directions $\hat{k}^r$ and $\hat{k}^t$, respectively. The difference between the direct and the inverse scattering problem is summarized in Table 2.

5.1 Analysis of the boundary conditions

Using the results in Paper II, we have the following condition on the tangential field strengths:

$$\dot{\boldsymbol{e}}^i_{\parallel} + \dot{\boldsymbol{e}}^r_{\parallel} = \dot{\boldsymbol{e}}^t_{\parallel}, \tag{5.1}$$

where the dots denote time differentiation and the subscript $\|$ denotes components parallel to the interface between the materials, *i.e.*, x and y components. The fields satisfy the equations

$$\begin{cases} \dot{\boldsymbol{e}}^i = [\hat{k}^i \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^i \\ \dot{\boldsymbol{e}}^r = [\hat{k}^r \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^r \\ \boldsymbol{\varepsilon} \cdot \dot{\boldsymbol{e}}^t = \dfrac{c_0}{c}[\hat{k}^t \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^t, \end{cases} \tag{5.2}$$

where $c_0/c = \sqrt{\epsilon\mu}$ is the propagation speed in the nonlinear material relative the speed of light in vacuum. Snell's law [25, p. 303] implies

$$k_y^i = k_y^r = \frac{c_0}{c} k_y^t, \tag{5.3}$$

where $k_y^{i,r,t} = \hat{y} \cdot \hat{k}^{i,r,t}$, *i.e.*, the sine of the angles of incidence, reflection and transmission, respectively.

Observing that $k_z^i = -k_z^r$, we use Snell's law to write (5.2) as

$$\begin{cases} \dfrac{1}{k_z^i}[\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^i = [\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^i \\ -\dfrac{1}{k_z^i}[\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^r = [\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^r \\ \dfrac{c}{c_0}\dfrac{1}{k_z^t}[\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^t = [\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^t. \end{cases} \tag{5.4}$$

Since $[\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^{i,r,t} = [\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_{\|}^{i,r,t}$, we multiply (5.1) with $[\hat{z} \times \mathbf{J}]$ and use (5.4) to obtain

$$\frac{1}{k_z^i}[\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}] \cdot \left\{\dot{\boldsymbol{e}}^i - \dot{\boldsymbol{e}}^r\right\} = \frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^t. \tag{5.5}$$

This equation involves the normal component of the fields, which can be eliminated by using (5.2),

$$\begin{aligned} \dot{\boldsymbol{e}}_z^t = \hat{z}\hat{z} \cdot \dot{\boldsymbol{e}}^t &= \hat{z}\hat{z} \cdot \frac{c_0}{c}\boldsymbol{\varepsilon}^{-1} \cdot [\hat{k}^t \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^t \\ &= \frac{c_0}{c}\boldsymbol{\varepsilon}^{-1} \cdot \hat{z}\hat{z} \cdot [\hat{k}^t \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}^t \\ &= \frac{c_0}{c}\boldsymbol{\varepsilon}^{-1} \cdot \hat{z}\hat{z} \cdot [k_y^t \hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_{\|}^t \\ &= k_y^i \boldsymbol{\varepsilon}^{-1} \cdot \hat{z}\hat{z} \cdot [\hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_{\|}^t, \end{aligned}$$

where we have used the isotropy of the material to conclude that $\boldsymbol{\varepsilon}^{-1}$ commutes with $\hat{z}\hat{z}$, and Snell's law (5.3). We use this result for the right hand side of (5.5),

$$\begin{aligned}
\frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon}-k_y^i\hat{y}\times\mathbf{J}]\cdot\dot{\boldsymbol{e}}^t &= \frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon}-k_y^i\hat{y}\times\mathbf{J}]\cdot\left\{\dot{\boldsymbol{e}}_\parallel^t+\dot{\boldsymbol{e}}_z^t\right\}\\
&= \frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon}-k_y^i\hat{y}\times\mathbf{J}]\cdot\{\mathbf{I}+k_y^i\boldsymbol{\varepsilon}^{-1}\cdot\hat{z}\hat{z}\cdot[\hat{y}\times\mathbf{J}]\}\cdot\dot{\boldsymbol{e}}_\parallel^t\\
&= \frac{1}{k_z^t}\frac{c}{c_0}\big[\boldsymbol{\varepsilon}-[\mathbf{I}-\hat{z}\hat{z}]\cdot k_y^i[\hat{y}\times\mathbf{J}]-(k_y^i)^2[\hat{y}\times\mathbf{J}]\cdot\boldsymbol{\varepsilon}^{-1}\cdot\hat{z}\hat{z}\cdot[\hat{y}\times\mathbf{J}]\big]\cdot\dot{\boldsymbol{e}}_\parallel^t\\
&= \frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon}+(k_y^i)^2\mathbf{J}\cdot\boldsymbol{\varepsilon}^{-1}\cdot\mathbf{J}\cdot\hat{x}\hat{x}]\cdot\dot{\boldsymbol{e}}_\parallel^t,
\end{aligned}$$

where the last line follows from explicitly expanding the cross product operations and using the isotropy properties of $\boldsymbol{\varepsilon}$. The product $\mathbf{J}\cdot\boldsymbol{\varepsilon}^{-1}\cdot\mathbf{J}$ is

$$\begin{pmatrix}\mathbf{0} & -\mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{pmatrix}\cdot\begin{pmatrix}\frac{1}{\epsilon}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \frac{1}{\mu}\mathbf{I}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{0} & -\mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{pmatrix}=\begin{pmatrix}-\frac{1}{\mu}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & -\frac{1}{\epsilon}\mathbf{I}\end{pmatrix}=-\frac{1}{\epsilon\mu}\begin{pmatrix}\epsilon\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mu\mathbf{I}\end{pmatrix},$$

and we conclude that $(k_y^i)^2\mathbf{J}\cdot\boldsymbol{\varepsilon}^{-1}\cdot\mathbf{J}=-(k_y^i)^2(c/c_0)^2\boldsymbol{\varepsilon}=-(k_y^t)^2\boldsymbol{\varepsilon}$, since $1/\epsilon\mu=(c/c_0)^2$ and $k_y^i c/c_0=k_y^t$. Thus the right hand side of (5.5) is

$$\begin{aligned}
\frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon}-k_y^i\hat{y}\times\mathbf{J}]\cdot\dot{\boldsymbol{e}}^t &= \frac{1}{k_z^t}\frac{c}{c_0}[\boldsymbol{\varepsilon}-(k_y^t)^2\boldsymbol{\varepsilon}\cdot\hat{x}\hat{x}]\cdot\dot{\boldsymbol{e}}_\parallel^t\\
&= \frac{1}{k_z^t}\frac{c}{c_0}\boldsymbol{\varepsilon}\cdot[\mathbf{I}-(k_y^t)^2\hat{x}\hat{x}]\cdot\dot{\boldsymbol{e}}_\parallel^t\\
&= \frac{1}{k_z^t}\frac{c}{c_0}\boldsymbol{\varepsilon}\cdot[(k_z^t)^2\hat{x}\hat{x}+\hat{y}\hat{y}]\cdot\dot{\boldsymbol{e}}_\parallel^t.
\end{aligned}$$

The same calculations for the left hand side of (5.5) implies

$$\frac{1}{k_z^i}[(k_z^i)^2\hat{x}\hat{x}+\hat{y}\hat{y}]\cdot\left\{\dot{\boldsymbol{e}}_\parallel^i-\dot{\boldsymbol{e}}_\parallel^r\right\}=\frac{1}{k_z^t}\frac{c}{c_0}\boldsymbol{\varepsilon}\cdot[(k_z^t)^2\hat{x}\hat{x}+\hat{y}\hat{y}]\cdot\dot{\boldsymbol{e}}_\parallel^t.$$

From (5.1) we have $\dot{\boldsymbol{e}}_\parallel^t=\dot{\boldsymbol{e}}_\parallel^i+\dot{\boldsymbol{e}}_\parallel^r$, and we write the above equation for both x and y components:

$$\begin{cases} k_z^i\left\{\dot{\boldsymbol{e}}_x^i-\dot{\boldsymbol{e}}_x^r\right\}=k_z^t\dfrac{c}{c_0}\boldsymbol{\varepsilon}\cdot\left\{\dot{\boldsymbol{e}}_x^i+\dot{\boldsymbol{e}}_x^r\right\}\\ \dfrac{1}{k_z^i}\left\{\dot{\boldsymbol{e}}_y^i-\dot{\boldsymbol{e}}_y^r\right\}=\dfrac{1}{k_z^t}\dfrac{c}{c_0}\boldsymbol{\varepsilon}\cdot\left\{\dot{\boldsymbol{e}}_y^i+\dot{\boldsymbol{e}}_y^r\right\}.\end{cases}\tag{5.6}$$

Since we can measure the incident and reflected fields and the angle of incidence, the only unknowns in the above equations are k_z^t and $\frac{c}{c_0}\boldsymbol{\varepsilon}$. The latter is

$$\frac{c}{c_0}\boldsymbol{\varepsilon}=\frac{1}{\sqrt{\epsilon\mu}}\begin{pmatrix}\epsilon\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mu\mathbf{I}\end{pmatrix}=\begin{pmatrix}\sqrt{\frac{\epsilon}{\mu}}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \sqrt{\frac{\mu}{\epsilon}}\mathbf{I}\end{pmatrix}=\begin{pmatrix}\frac{1}{\eta}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \eta\mathbf{I}\end{pmatrix}$$

where we have introduced the dimensionless wave impedance $\eta = \sqrt{\mu/\epsilon}$. The explicit representation of (5.6) is now written, where we have moved k_z^i to the right hand side of the equations,

$$\begin{cases} \dot{E}_x^i - \dot{E}_x^r = \dfrac{k_z^t}{k_z^i}\dfrac{1}{\eta}(\dot{E}_x^i + \dot{E}_x^r) \\ \dot{H}_x^i - \dot{H}_x^r = \dfrac{k_z^t}{k_z^i}\eta(\dot{H}_x^i + \dot{H}_x^r) \\ \dot{E}_y^i - \dot{E}_y^r = \dfrac{k_z^i}{k_z^t}\dfrac{1}{\eta}(\dot{E}_y^i + \dot{E}_y^r) \\ \dot{H}_y^i - \dot{H}_y^r = \dfrac{k_z^i}{k_z^t}\eta(\dot{H}_y^i + \dot{H}_y^r). \end{cases} \tag{5.7}$$

We see that if we use TE or TM polarization of the incident field, either the second and the third, or the first and the fourth, equations will be zero on both sides. The equations left can then determine only one of the quantities $\frac{k_z^t}{k_z^i}\eta$ and $\frac{k_z^i}{k_z^t}\eta$. However, if we use an incident field that is a mixture of TE and TM polarizations, we have two linearly independent equations and can obtain both quantities. The alternative is to make two measurements, with TE and TM polarizations of the incident field, respectively.

When we have determined both k_z^t and η, we can determine the nonlinear permittivity ϵ and permeability μ from the relations

$$\begin{cases} \dfrac{1}{\epsilon\mu} = \left(\dfrac{c}{c_0}\right)^2 = \left(\dfrac{k_y^t}{k_y^i}\right)^2 = \dfrac{1-(k_z^t)^2}{(k_y^i)^2} \\ \dfrac{\mu}{\epsilon} = \eta^2. \end{cases}$$

At this stage, we have determined the material parameters ϵ and μ as functions of the tangential fields only. By considering the continuity of the flux components normal to the surface, we are able to conclude the total field strengths in the material, and thus determine ϵ and μ as functions of the total field strengths E and H. This procedure is not shown explicitly.

5.2 Numerical results

A program has been written in MATLAB based on the methods in Paper II to calculate the reflected field, and we use the formulae in (5.7) to find the permittivity and permeability. The constitutive functions used in the calculations are

$$\begin{aligned} \epsilon(E) &= 2 + \frac{E^2}{1+E^2} \\ \mu(H) &= 1 + \frac{1}{2}\frac{H^2}{1+H^2}, \end{aligned}$$

where we have scaled the fields to be non-dimensional (and therefore excluded the factors $\sqrt{\epsilon_0}$ and $\sqrt{\mu_0}$), see Paper I for a discussion on the scaling. The incident field

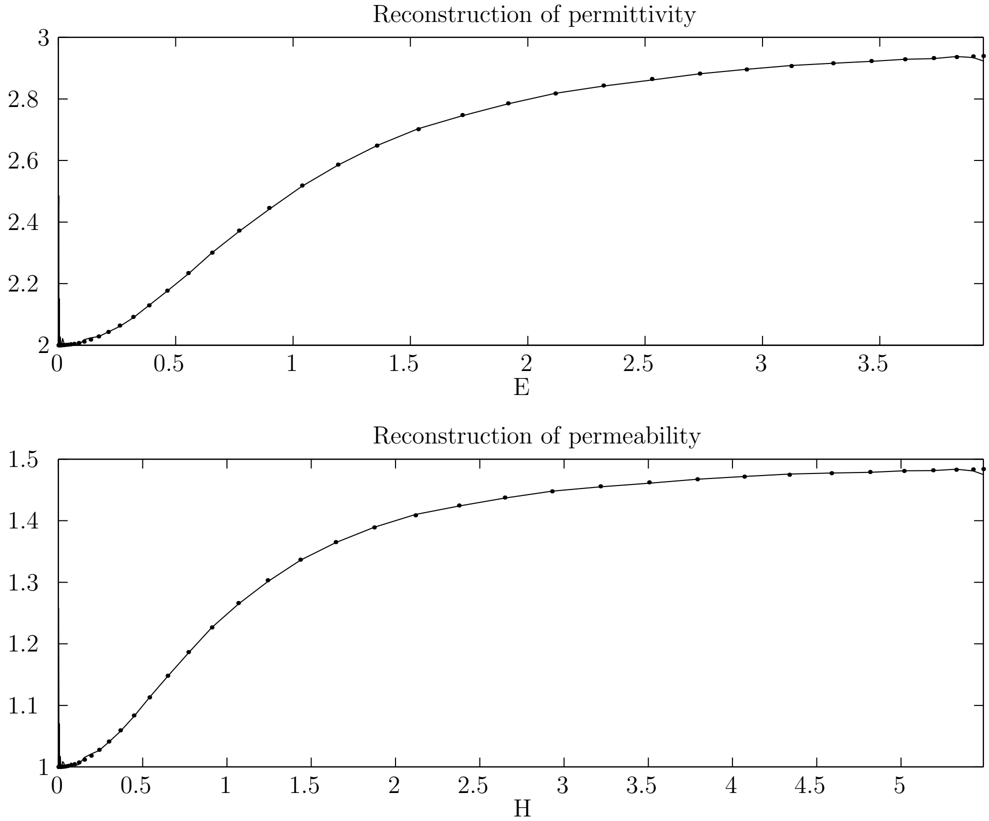

Figure 4: Reconstruction of the field dependent permittivity and permeability with noise of amplitude $\pm 10^{-5}$. Solid line is reconstruction and the dots are exact values.

is a Gaussian pulse in time of amplitude 5, where both the E and the H field have components in both the tangential directions, *i.e.*, we use both TE and TM polarization at the same time. The angle of incidence is 30°, and Figure 4 demonstrate the field dependence of the permittivity and permeability of the nonlinear medium. Note that at this stage we commit the inverse crime [12, p. 121], *i.e.*, using the same algorithm for both direct and inverse problem. Work is at progress to remedy this deficiency. To demonstrate at least the continuity of our algorithm, we perturb the data from the direct problem with uniformly distributed random numbers in the range $\pm 10^{-5}$. Using the range $\pm 10^{-3}$ causes the reconstructions to seem useless, but one must remember that the material parameters are to be integrated in order to get the fluxes, see Paper II Section 5. Since integration is a smoothening procedure, we expect the flux reconstruction to remain good, which is verified in Figure 5.

6 Acknowledgments

The author thanks his supervisor Prof. Gerhard Kristensson, for excellent guidance through the scientific education, constant encouragement and many valuable suggestions to the work presented in this thesis. This achievement would have been a

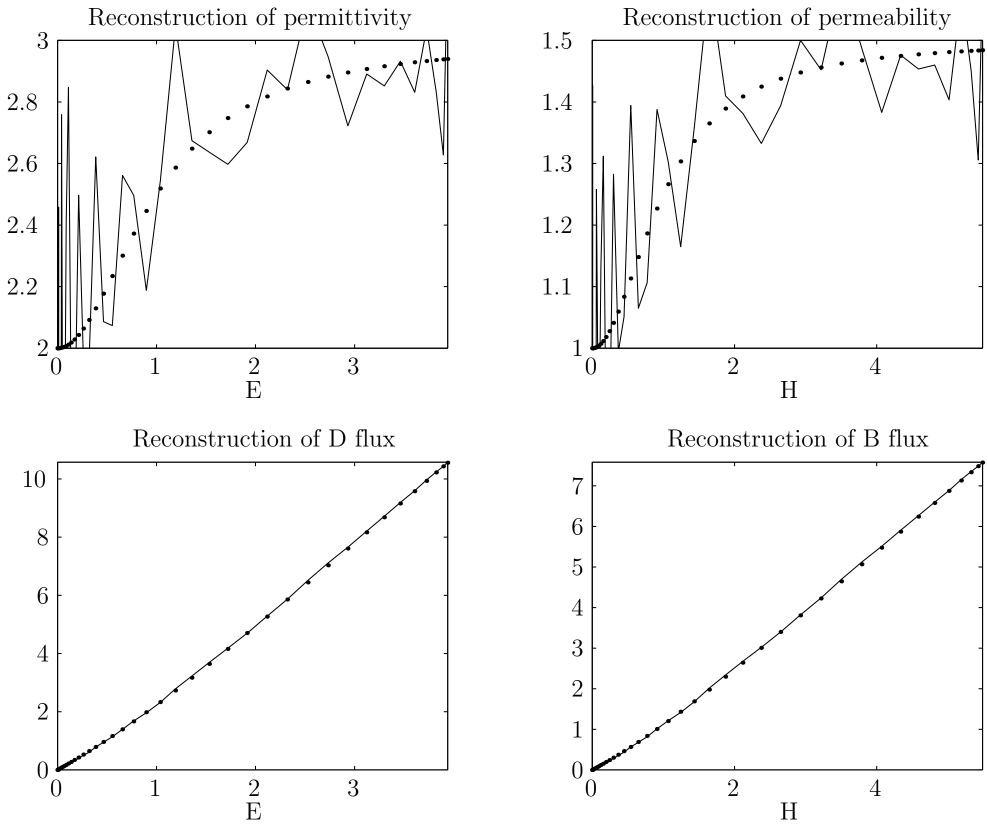

Figure 5: Upper row: Reconstruction of the permittivity and permebility with noise of amplitude $\pm 10^{-3}$. Lower row: Reconstruction of the constitutive functions $D(E)$ and $B(H)$.

lot harder without you!

The staff at the Department of Electromagnetic Theory in Lund is warmly acknowledged for an inspiring and understanding atmosphere. The author would like to especially mention Mats Gustafsson for many interesting discussions on electromagnetic theory and numerical methods.

This work would not have been possible were it not for the author's family, whose never-dying patience, belief and trust have been very valuable. Thank you ever so much!

The author also thanks his friends from Hultsfred, undergraduate studies, the student union at the Lund Institute of Technology and all the others. Without you life would be a great deal less interesting!

This work is partially supported by a grant from the Swedish Research Council for Engineering Sciences and their support is gratefully acknowledged.

References

[1] G. P. Agrawal. *Nonlinear fiber optics.* Academic Press, San Diego, 1995.

[2] N. Bloembergen. *Nonlinear Optics.* W. A. Benjamin, Inc., Reading, Massachusetts, 1977.

[3] H. Blok and M. Oristaglio. Wavefield imaging and inversion in electromagnetics and acoustics. Technical report, Centre for Technical Geoscience, Delft University of Technology, 1995.

[4] F. Bloom. *Mathematical problems of classical nonlinear electromagnetic theory.* Longman Scientific & Technical, Burnt Mill, Harlow, England, 1993.

[5] C. F. Bohren and D. R. Huffman. *Absorption and Scattering of Light by Small Particles.* John Wiley & Sons, New York, 1983.

[6] R. W. Boyd. *Nonlinear Optics.* Academic Press, San Diego, 1992.

[7] L. J. F. Broer. Wave propagation in non-linear media. *Z. angew. Math. Phys.*, **16**, 18–26, 1965.

[8] L. J. F. Broer and P. H. A. Sarluy. On simple waves in non-linear dielectric media. *Physica*, **30**, 1421–1432, 1964.

[9] D. Censor and Y. Ben-Shimol. Wave propagation in weakly nonlinear bianisotropic and bi-isotropic media. *J. Electro. Waves Applic.*, **11**, 1763–1779, 1997.

[10] A. Chelkowski. *Dielectric Physics.* Elsevier Science Publishers, Amsterdam, 1980.

[11] B. D. Coleman and E. H. Dill. Thermodynamic restrictions on the constitutive equations of electromagnetic theory. *Z. angew. Math. Phys.*, **22**, 691–702, 1971.

[12] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer-Verlag, Berlin, 1992.

[13] J. P. Corones, M. E. Davison, and R. J. Krueger. Direct and inverse scattering in the time domain via invariant imbedding equations. *J. Acoust. Soc. Am.*, **74**(5), 1535–1541, 1983.

[14] J. L. Davis. *Wave Propagation in Electromagnetic Media.* Springer-Verlag, Berlin, 1990.

[15] M. E. Davison and R. C. Winther. A general approach to splitting and invariant imbedding for linear wave equations. *J. Math. Anal. Appl.*, **188**(1), 158–181, 1994.

[16] A. C. Eringen and G. A. Maugin. *Electrodynamics of Continua I.* Springer-Verlag, New York, 1990.

[17] A. C. Eringen and G. A. Maugin. *Electrodynamics of Continua II.* Springer-Verlag, New York, 1990.

[18] M. Gustafsson. Time domain theory of the macroscopic maxwell equations. Technical Report LUTEDX/(TEAT-7062)/1–24/(1997), Lund Institute of Technology, Department of Electromagnetic Theory, P.O. Box 118, S-211 00 Lund, Sweden, 1997.

[19] M. Gustafsson. *Inverse Electromagnetic Scattering Problems – A Time-Domain Optimization Approach.* Licentiate thesis, Lund Institute of Technology, Department of Electromagnetic Theory, P.O. Box 118, S-211 00 Lund, Sweden, 1998.

[20] A. Hasegawa. *Optical Solitons in Fibers.* Springer-Verlag, Berlin Heidelberg, 1990.

[21] A. Hasegawa and Y. Kodama. *Solitons in Optical Communications.* Clarendon Press, Oxford, 1995.

[22] S. He, S. Ström, and V. H. Weston. *Time Domain Wave-splittings and Inverse Problems.* Oxford University Press, Oxford, 1998.

[23] S. He. Newton-Kantorovich approach to an inverse problem for a quasi-linear acoustic wave equation. *J. Inverse and Ill-posed Problems*, **4**(3), 223–232, 1996.

[24] L. Hörmander. *Lectures on Nonlinear Hyperbolic Differential Equations.* Number 26 in Mathémathiques & Applications. Springer-Verlag, Berlin, 1997.

[25] J. D. Jackson. *Classical Electrodynamics.* John Wiley & Sons, New York, third edition, 1999.

[26] A. Jeffrey. Wave propagation and electromagnetic shock wave formation in transmission lines. *J. Math. and Mech.*, **15**(1), 1–13, 1966.

[27] A. Jeffrey. Non-dispersive wave propagation in nonlinear dielectrics. *Z. angew. Math. Phys.*, **19**, 741–745, 1968.

[28] A. Jeffrey and V. P. Korobeinikov. Formation and decay of electromagnetic shock waves. *Z. angew. Math. Phys.*, **20**, 440–447, 1969.

[29] F. John. *Partial Differential Equations.* Springer-Verlag, New York, fourth edition, 1982.

[30] R. M. Joseph and A. Taflove. FDTD Maxwell's equations models for nonlinear electrodynamics and optics. *IEEE Trans. Antennas Propagat.*, **45**, 364–374, 1997.

[31] A. Karlsson. Wave propagators for transient waves in one-dimensional media. *Wave Motion*, **24**(1), 85–99, 1996.

[32] A. Karlsson and G. Kristensson. Constitutive relations, dissipation and reciprocity for the Maxwell equations in the time domain. *J. Electro. Waves Applic.*, **6**(5/6), 537–551, 1992.

[33] A. Karlsson and R. Stewart. Wave propagators for transient waves in periodic media. *J. Opt. Soc. Am. A*, **12**(7), 1513–1521, 1995.

[34] J. Y. Kazakia and R. Venkataraman. Propagation of electromagnetic waves in a nonlinear dielectric slab. *Z. angew. Math. Phys.*, **26**, 61–76, 1975.

[35] J. B. Keller. Inverse problems. *Am. Math. Mon.*, **83**, 107–118, 1976.

[36] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Springer-Verlag, New York, 1996.

[37] J. A. Kong. *Electromagnetic Wave Theory.* John Wiley & Sons, New York, 1986.

[38] G. Kristensson and D. J. N. Wall. Direct and inverse scattering for transient electromagnetic waves in nonlinear media. *Inverse Problems*, **14**, 113–137, 1998.

[39] A. Lakhtakia and W. S. Weiglhofer. Are linear, nonreciprocal, bi-isotropic media forbidden? *IEEE Trans. Microwave Theory Tech.*, **42**(9), 1715–1716, 1994.

[40] L. D. Landau and E. M. Lifshitz. *Electrodynamics of Continuous Media.* Pergamon, Oxford, first edition, 1960.

[41] I. V. Lindell, A. H. Sihvola, and K. Suchy. Six-vector formalism in electromagnetics of bi-anisotropic media. *J. Electro. Waves Applic.*, **9**(7/8), 887–903, 1995.

[42] J. D. Logan. *An Introduction to Nonlinear Partial Differential Equations.* John Wiley & Sons, New York, 1994.

[43] J. C. Luke. A perturbation method for nonlinear dispersive wave problems. *Proc. Roy. Soc. London Ser. A*, **292**, 403–412, 1966.

[44] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon. Experimental observation of picosecond pulse narrowing and solitons in optical fibers. *Phys. Rev. Lett.*, **45**(13), 1095–1098, September 1980.

[45] D. F. Nelson. Nonlinear electroacoustics of dielectric crystals: Fundamentals and applications. *J. Physique*, **40**(C8), C8/149–163, November 1979.

[46] M. Norgren. Optimization approaches to frequency-domain inverse problems for Kerr-like and non-Kerr-like nonlinear media. Technical Report TRITA-TET 98-15, Department of Electromagnetic Theory, S-100 44 Stockholm, Sweden, 1998.

[47] E. J. Post. *Formal Structure of Electromagnetics.* North-Holland, Amsterdam, 1962.

[48] R. E. Raab and A. H. Sihvola. On the existence of linear non-reciprocal bi-isotropic (NRBI) media. *J. Phys. A: Math. Gen.*, **30**, 1335–1344, 1997.

[49] C. Rogers, H. M. Cekirge, and A. Askar. Electromagnetic wave propagation in non-linear dielectric media. *Acta Mech.*, **26**, 59–73, 1977.

[50] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*. John Wiley & Sons, New York, 1991.

[51] Y. R. Shen. *Principles of Nonlinear Optics*. John Wiley & Sons, New York, 1984.

[52] A. Sihvola. Are nonreciprocal, bi-isotropic materials forbidden indeed? *IEEE Trans. Microwave Theory Tech.*, **43**(9), 2160–2162, 1995.

[53] J. A. Stratton. *Electromagnetic Theory*. McGraw-Hill, New York, 1941.

[54] A. Taflove. *Computational electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Boston, London, 1995.

[55] B. D. H. Tellegen. The gyrator, a new electric network element. *Philips Res. Rep.*, **3**, 81–101, 1948.

[56] R. A. Toupin and R. S. Revlin. Linear functional electromagnetic constitutive relations and plane waves in a hemihedral isotropic material. *Arch. Rational Mech. Anal.*, **6**, 188–197, 1960.

[57] V. Volterra. *Theory of functional and of integral and integro-differential equations*. Dover Publications, New York, 1959.

[58] V. H. Weston. Time-domain wave-splitting of Maxwell's equations. *J. Math. Phys.*, **34**(4), 1370–1392, 1993.

[59] G. B. Whitham. A general approach to linear and nonlinear dispersive waves using a Lagrangian. *J. Fluid Mech.*, **22**(2), 273–283, 1965.

[60] G. B. Whitham. *Linear and nonlinear waves*. John Wiley & Sons, New York, 1974.

[61] J. L. Young. Propagation in linear dispersive media: Finite difference time-domain methodologies. *IEEE Trans. Antennas Propagat.*, **43**(4), 422–426, 1995.

Paper I

Reconstruction of nonlinear material properties for homogeneous, isotropic slabs using electromagnetic waves

D. Sjöberg

Inverse Problems, **15**(2), 431–444, 1999.

Abstract

This paper addresses the inverse problem of reconstructing a medium's instantaneous, nonlinear response to electromagnetic excitation. Using reflection and transmission data for an almost arbitrary incident field on a homogeneous slab, we are able to obtain the nonlinear constitutive relations for both electric and magnetic fields, with virtually no assumptions made on the specific form of the relations. It is shown that for a nonmagnetic material, reflection data suffices to obtain the electrical nonlinear response. We also show that the algorithms are well posed. Numerical examples illustrate the analysis presented in this paper.

1 Introduction

There has been an increased interest in nonlinear electromagnetic materials recently, much due to the progresses in nonlinear optics. This is especially so for the nonlinear effects in optical fibers, *i.e.*, the experimental verification of soliton solutions [13, 14, 21], and the use of different field-dependent scattering mechanisms for amplification of a propagating signal [1]. Some chaotic effects have also been studied [10].

The research in this field is largely conducted in the frequency domain, where the nonlinearities manifest in the generation of multiple frequencies. In this paper, we study nonlinear effects in the time domain, where the nonlinearities rather cause the steepening of a propagating pulse. This steepening may ultimately turn into a shock solution, where the pulse becomes discontinuous after a finite propagation time, although we will endeavour to avoid shock solutions in this paper.

We study a material which has an instantaneous, nonlinear response, *i.e.*, we do not consider memory effects of any kind. We further assume the material to be passive, isotropic and homogeneous, and solve the problem of reconstructing the constitutive relations. Then we are able to reconstruct the nonlinear relation between $\boldsymbol{E}$ and $\boldsymbol{D}$ as well as between $\boldsymbol{H}$ and $\boldsymbol{B}$ with reflection and transmission data from a finite slab for an (almost) arbitrary input signal. Since no further assumptions have to be made regarding the specific form of the constitutive relations, the reconstruction is model independent.

Previous work in the field include the propagation of pulses in nonlinear slabs, where the paper by Kazakia and Venkataraman deserves special attention [18]. They have obtained an analytical solution for the propagation of a step function through a slab with some special constitutive functions. Reference [24] presents a method to solve the reflection and refraction problem at oblique incidence on a nonlinear half space. The wave propagation in more complicated nonlinear materials has appeared, *i.e.*, mixed nonlinearities [19], bi-anisotropic and bi-isotropic media [5], and nonlinearities in chiral media [2, 23].

Though much work has been done on the direct problem of wave propagation in nonlinear media, our solution of the inverse problem of reconstructing the material seems to be novel. It extends and improves the results in [20], where the inverse problem is solved for a nonmagnetic material, based on measurements inside the material.

In Section 2 we formulate the stratified Maxwell equations, introduce the constitutive relations for the studied materials and try to interprete the dynamics in terms known from the linear case. The main theory is contained in Section 3, where we formulate the necessary boundary conditions and state the solution to our inverse problems. Some numerical results are contained in Section 4.

2 Prerequisites

2.1 The Maxwell equations in one spatial dimension

In a source-free environment the Maxwell equations are

$$\begin{aligned} \nabla \times \boldsymbol{E}(\boldsymbol{r}, t) + \partial_t \boldsymbol{B}(\boldsymbol{r}, t) &= 0 \\ \nabla \times \boldsymbol{H}(\boldsymbol{r}, t) - \partial_t \boldsymbol{D}(\boldsymbol{r}, t) &= 0. \end{aligned}$$

Since we wish to study a homogeneous medium, it is sufficient to observe variations for only one direction. We thus assume that the fields depend on only one spatial variable, say z, in a Cartesian coordinate system (x, y, z). Then the curl operator can be written $\nabla \times = \hat{z} \times \mathbf{I}\partial_z = \mathbf{J}\partial_z$, where $\mathbf{J}$ denotes a rotation $\pi/2$ around the z-axis, and the Maxwell equations become

$$\begin{aligned} \mathbf{J} \cdot \partial_z \boldsymbol{E}(z, t) + \partial_t \boldsymbol{B}(z, t) &= 0 \\ \mathbf{J} \cdot \partial_z \boldsymbol{H}(z, t) - \partial_t \boldsymbol{D}(z, t) &= 0. \end{aligned}$$

We now assume the fields to be linearly polarized and the material to be isotropic, *i.e.*, the $\boldsymbol{D}$ and $\boldsymbol{B}$ fields are parallel to the $\boldsymbol{E}$ and $\boldsymbol{H}$ fields, respectively, which vary only in amplitude. This means we can write the Maxwell equations in a scalar form,

$$\begin{aligned} \partial_z E(z, t) + \partial_t B(z, t) &= 0 \\ \partial_z H(z, t) + \partial_t D(z, t) &= 0, \end{aligned}$$

where E and D denote an arbitrary transversal component, say x, of $\mathbf{J} \cdot \boldsymbol{E}$ and $\mathbf{J} \cdot \boldsymbol{D}$, respectively. H and B denote the corresponding component of $\boldsymbol{H}$ and $\boldsymbol{B}$, respectively. The geometry of the scattering situation studied in this paper is depicted in Figure 1.

2.2 Constitutive relations, passive materials

We consider the field strengths E and H to be the primary fields, and the flux densities D and B as effects of these. If we assume that the material responds instantaneous to excitation, we are studying the following situation:

$$\begin{aligned} D(z, t) &= \varepsilon_0 F_{\mathrm{e}}(E(z, t)) \\ B(z, t) &= \frac{1}{c_0} F_{\mathrm{m}}(\eta_0 H(z, t)), \end{aligned}$$

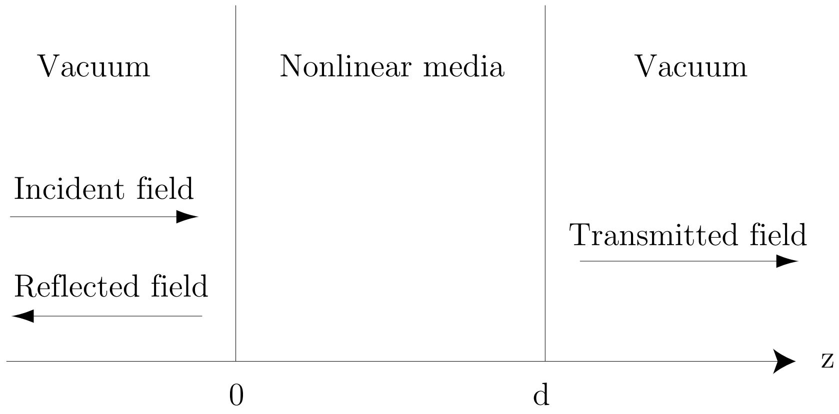

Figure 1: The scattering geometry studied in this paper.

where the constants $c_0 = 1/\sqrt{\varepsilon_0\mu_0}$ (speed of light in vacuum), ε_0 (permittivity of vacuum), and $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$ (wave impedance of vacuum) are explicit for convenience. As usual, μ_0 denotes the permeability of vacuum. The functions $F_{\rm e}(E)$ and $F_{\rm m}(\eta_0 H)$ are continuously differentiable scalar functions of one variable, and generalize the linear optical responses, $F_{\rm e}^{\rm lin}(E) = \varepsilon_r E$ and $F_{\rm m}^{\rm lin}(\eta_0 H) = \mu_r\eta_0 H$. This kind of nonlinear constitutive response with similar dynamics is investigated in [20], [3, Chap. 2], and [11, Chap. 6]. In nonlinear optics similar relations are often used, although frequently in the context of the frequency domain [1, 4].

Some thermodynamic restrictions can be put on the constitutive relations [6], but these deal mainly with the symmetry of cross terms, *i.e.*, $\frac{\partial D}{\partial H}$ and $\frac{\partial B}{\partial E}$, which we do not take into account here. Reference [20] discusses the restrictions on the functions $F_{\rm e}$ and $F_{\rm m}$ in order to model passive media; though they call it dissipative.[1] The result is that for a passive, nonmagnetic material, $F_{\rm e}'(x) \geq a > 0$ is a *sufficient* condition. In this paper we generalize this to materials which also have $F_{\rm m}'(x) \geq b > 0$, and call these *positive passive*.

When demanding isotropy, we have the implication that a change of sign in the electric and magnetic fields leads to a change of sign in the electric and magnetic fluxes, *i.e.*, $(E, H) \to (-E, -H) \Rightarrow (D, B) \to (-D, -B)$. This is also true for crystals with an *inversion symmetry*, see [4, Chap. 1] for further discussions of material properties. This property implies that the constitutive functions should be odd functions of their argument, which will be important in the following.

Eliminating the D and B fields using the constitutive relations, the scalar Maxwell equations become

$$\begin{aligned} \partial_z E + \frac{1}{c_0} F_{\rm m}' \partial_t \eta_0 H = 0 \\ \partial_z \eta_0 H + \frac{1}{c_0} F_{\rm e}' \partial_t E = 0, \end{aligned} \tag{2.1}$$

where we have dropped the arguments of the functions $F_{\rm m}'$, $F_{\rm e}'$ for simplicity.

[1] With a passive material we mean that the electromagnetic energy produced in a region is nonpositive for all times, *i.e.*, the material is not active.

2.3 The dynamics as a symmetric system, physical interpretation

Though it is possible to directly introduce the well known Riemann invariants $\frac{1}{2}(\int_0^E \sqrt{F_\mathrm{e}'(x)}\,dx \pm \int_0^{\eta_0 H} \sqrt{F_\mathrm{m}'(x)}\,dx)$ as in [3, Sec. 2.4] or [11, Sec. 6.13], we wish to follow a different approach, where we try to interprete our variables and make comparisons to the linear case. We start by formulating the dynamics as

$$\begin{pmatrix} F_\mathrm{e}'\partial_t E \\ F_\mathrm{m}'\partial_t \eta_0 H \end{pmatrix} + c_0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_z \begin{pmatrix} E \\ \eta_0 H \end{pmatrix} = 0,$$

which after division by the square root of the derivative of F_e and F_m leads to

$$\begin{pmatrix} \sqrt{F_\mathrm{e}'}\partial_t E \\ \sqrt{F_\mathrm{m}'}\partial_t \eta_0 H \end{pmatrix} + c_0 \begin{pmatrix} 0 & \frac{1}{\sqrt{F_\mathrm{e}' F_\mathrm{m}'}} \\ \frac{1}{\sqrt{F_\mathrm{e}' F_\mathrm{m}'}} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{F_\mathrm{e}'}\partial_z E \\ \sqrt{F_\mathrm{m}'}\partial_z \eta_0 H \end{pmatrix} = 0.$$

We now introduce the functions,

$$\begin{aligned} g_\mathrm{e}(E) &= \int_0^E \sqrt{F_\mathrm{e}'(x)}\,dx \\ g_\mathrm{m}(\eta_0 H) &= \int_0^{\eta_0 H} \sqrt{F_\mathrm{m}'(x)}\,dx. \end{aligned}$$

These functions can be thought of as the generalizations of the linear expressions $\sqrt{\varepsilon_r}E$ and $\sqrt{\mu_r}\eta_0 H$. The product of the derivative of the functions, $g_\mathrm{e}' g_\mathrm{m}'$, which appears in the wave speed below, can be viewed as the generalization of $\sqrt{\varepsilon_r \mu_r}$, the relative refractive index. Furthermore, for an isotropic, positive passive material, the g-functions are odd and monotone, since the integrands are always even and positive. With these new functions we can write the dynamics as

$$\partial_t \begin{pmatrix} g_\mathrm{e}(E) \\ g_\mathrm{m}(\eta_0 H) \end{pmatrix} + \frac{c_0}{g_\mathrm{e}'(E) g_\mathrm{m}'(\eta_0 H)} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_z \begin{pmatrix} g_\mathrm{e}(E) \\ g_\mathrm{m}(\eta_0 H) \end{pmatrix} = 0,$$

which in the new variables $u_1 = g_\mathrm{e}(E)$ and $u_2 = g_\mathrm{m}(\eta_0 H)$ is the symmetric system

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + c(u_1, u_2) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_z \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0, \qquad (2.2)$$

where the wave speed c is

$$c(u_1, u_2) = \frac{c_0}{g_\mathrm{e}'(g_\mathrm{e}^{-1}(u_1)) g_\mathrm{m}'(g_\mathrm{m}^{-1}(u_2))} = c_0 \left(\frac{d}{du_1} g_\mathrm{e}^{-1}(u_1)\right)\left(\frac{d}{du_2} g_\mathrm{m}^{-1}(u_2)\right). \qquad (2.3)$$

This result generalizes the nonmagnetic case given in [20].

3 Methods to solve the inverse problem

In this section we demonstrate the methods used to solve the propagation problem and to resolve the boundary conditions. We also state our inverse problems of reconstructing the materials constitutive relations.

3.1 Wave splitting

The symmetric system (2.2) can be written as a system of one-dimensional wave equations with the wave splitting [8, 9, 20],

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u^+ \\ u^- \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} u^+ \\ u^- \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

This change of variables is exactly the introduction of the Riemann invariants of the one-dimensional Maxwell equations, which was mentioned in Section 2.3. The dynamics (2.2) now becomes

$$\partial_t \begin{pmatrix} u^+ \\ u^- \end{pmatrix} + c(u^+ + u^-, u^+ - u^-) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \partial_z \begin{pmatrix} u^+ \\ u^- \end{pmatrix} = 0, \tag{3.1}$$

with c defined by (2.3). This is a system of one-dimensional wave equations, which couple only through the wave speed c.

Analytical solutions for the wave propagation have been found in [18, 22] for some special constitutive relations. These solutions could be used to benchmark an algorithm for the wave propagation, though this is not performed in this work.

3.2 Propagation along characteristics

We can solve the propagation problem of the system (3.1) via the method of characteristics. A characteristic curve for this kind of differential equation is one on which the dependent variables are constant. We study the development of the variables $u^\pm(z,t)$ on the paths $(z,t) = (\zeta^\pm(\tau), \tau)$, where $\zeta^\pm(\tau) = \zeta_0 \pm \int_0^\tau c(\boldsymbol{u}')\, d\tau'$. The notation $c(\boldsymbol{u}')$ is short hand for $c(\boldsymbol{u}(\zeta^\pm(\tau'), \tau'))$, and $\boldsymbol{u} = (u^+, u^-)$. The variation of $u^\pm(z,t)$ along these curves are

$$\frac{d}{d\tau} u^\pm(\zeta^\pm(\tau), \tau) = \frac{\partial u^\pm}{\partial t} + \frac{d\zeta^\pm(\tau)}{d\tau} \frac{\partial u^\pm}{\partial z} = \frac{\partial u^\pm}{\partial t} \pm c(\boldsymbol{u}) \frac{\partial u^\pm}{\partial z} = 0,$$

since $u^\pm$ satisfy the differential equations $u_t \pm c u_z = 0$. Thus, we conclude that u^+ is constant along the characteristic path $\zeta(\tau) = \zeta_0 + \int_0^\tau c(\boldsymbol{u}')\, d\tau'$, and u^- is constant along the characteristic path $\zeta(\tau) = \zeta_0 - \int_0^\tau c(\boldsymbol{u}')\, d\tau'$.

This means we can find the values of the fields at a point (z,t) if we can trace the characteristics to some boundary where they are known. If only one of the waves is present, it is particularly simple; then the characteristics are straight lines, with a slope given by the boundary values [20].

We see that since the slope of the characteristics is governed by the boundary values, they may cross each other if we do not choose these boundary values carefully. When two characteristics cross each other, we have two possible solutions to the wave equation, and a *shock* occurs.

Theorem 3.1 in [20] concerns the extent of the shock-free region for one-way wave propagation in a semi-infinite media with given boundary conditions. This can be used to estimate how fast the incident field may vary in order not to create a shock in the slab. The suitable boundary conditions are $u^+(z,0) = u^-(z,0) = 0$, $u^+(0,t) = h(t)$ and $u^-(0,t) = 0$, for which the theorem states that there can be no shock in the region $0 \leq z \leq d$ if $\sup\left\{-\frac{d}{dt}\frac{1}{c(u^+(0,t),u^-(0,t))}\right\} = \sup\left\{-\frac{d}{dt}\frac{1}{c(h(t),0)}\right\} \leq \frac{1}{d}$. Since $\frac{c_0}{c(u^+,0)} = g_\mathrm{e}'(g_\mathrm{e}^{-1}(u^+))g_\mathrm{m}'(g_\mathrm{m}^{-1}(u^+))$, the condition will be

$$\sup_t \left\{-\left(\frac{g_\mathrm{e}''}{g_\mathrm{e}'}g_\mathrm{m}' + g_\mathrm{e}'\frac{g_\mathrm{m}''}{g_\mathrm{m}'}\right)h'\right\} \leq \frac{c_0}{d}. \tag{3.2}$$

We see, that we can always avoid shocks by using a signal with sufficiently small variation, *i.e.*, the derivative of $h(t)$ should be small compared to $1/(\frac{g_\mathrm{e}''}{g_\mathrm{e}'}g_\mathrm{m}' + g_\mathrm{e}'\frac{g_\mathrm{m}''}{g_\mathrm{m}'})$. Also, if this quantity and h' have the same sign, there is no risk of a shock. With positive second derivatives of $g_\mathrm{e,m}$, this means that shocks can only occur when $h' < 0$, *i.e.*, on the decreasing part of a signal.

3.3 Boundary conditions

Since we want to study propagation in a nonlinear slab, we must solve the problem of satisfying the boundary conditions. In this paper, we are studying a slab imbedded in vacuum. The generalization to more general linear materials follows from the method used.

The solution is based on the wave splitting, which allows us to determine in which direction the energy of the fields are travelling. In the surrounding vacuum, the splitting corresponds to the appropriate identification of incident, reflected and transmitted field. The boundary conditions we have to satisfy are the usual, *i.e.*, continuity of the tangential electric and magnetic field strengths. Since we are assuming normal incidence, this means continuity of the total fields E and H. Inside the slab, the electric and magnetic fields can be expressed as

$$\begin{aligned} E_\mathrm{slab} &= g_\mathrm{e}^{-1}(u^+ + u^-) \\ \eta_0 H_\mathrm{slab} &= g_\mathrm{m}^{-1}(u^+ - u^-). \end{aligned}$$

In vacuum, the magnetic field strength is related to the electric field strength via $\eta_0 H^\pm = \pm E^\pm$, where the $\pm$ indicate right(left) propagating fields, *i.e.*, waves travelling towards higher(lower) z-values.

It is possible to define differential reflection and transmission coefficients relating infinitesimal changes in the incident field to infinitesimal changes in the reflected and transmitted field, respectively, *i.e.*, $dE^r = r\cdot dE^i$ and $dE^t = t\cdot dE^i$. These differential coefficients look exactly like the linear expressions, where the square roots of the permittivity and permeability $\sqrt{\varepsilon}$ and $\sqrt{\mu}$ are replaced by $\sqrt{F_\mathrm{e}'(E)}$ and $\sqrt{F_\mathrm{m}'(\eta_0 H)}$, respectively. This method is used in [18] to solve the boundary problem, but in this paper we will prefer to simply state the boundary conditions in explicit form and solve these numerically for the desired fields when implementing the forward problem.

3.3.1 The left boundary

In vacuum, $z < 0$, we have an incident field from the left E^i, and a reflected field into vacuum, E^r. In the slab two fields are present: a right propagating field u^+, and a left propagating field u^-. The continuity of electric and magnetic fields implies that

$$\begin{cases} E^i + E^r = g_{\mathrm{e}}^{-1}(u^+ + u^-) \\ E^i - E^r = g_{\mathrm{m}}^{-1}(u^+ - u^-) \end{cases} \Leftrightarrow \quad \begin{cases} g_{\mathrm{e}}(E^i + E^r) = u^+ + u^- \\ g_{\mathrm{m}}(E^i - E^r) = u^+ - u^-. \end{cases} \tag{3.3}$$

This gives two, generally nonlinear, equations from which the desired fields u^+ and E^r can be determined:

$$\begin{cases} 2E^i = g_{\mathrm{e}}^{-1}(u^+ + u^-) + g_{\mathrm{m}}^{-1}(u^+ - u^-) \\ 2u^- = g_{\mathrm{e}}(E^i + E^r) - g_{\mathrm{m}}(E^i - E^r). \end{cases}$$

The incident field is given, but also the left propagating field u^- can be thought of as known. This is because this field can be traced back in time via a characteristic curve into the slab, and is therefore, from a computational point of view, known. Since the g-functions are monotone for a positive passive material, their inverses are too. This means that the right hand sides of the equations above, treated as functions of u^+ and E^r, are invertible, and we can find all desired fields numerically.

3.3.2 The right boundary

At the right boundary, $z = d$, we have just a transmitted field in the vacuum, but we still have both right and left propagating fields in the slab. Continuity of the fields now gives

$$\begin{cases} E^t = g_{\mathrm{e}}^{-1}(u^+ + u^-) \\ E^t = g_{\mathrm{m}}^{-1}(u^+ - u^-) \end{cases} \Leftrightarrow \quad \begin{cases} g_{\mathrm{e}}(E^t) = u^+ + u^- \\ g_{\mathrm{m}}(E^t) = u^+ - u^-. \end{cases} \tag{3.4}$$

From this we get the following equations to determine u^- and E^t:

$$\begin{cases} 2u^+ = g_{\mathrm{e}}(E^t) + g_{\mathrm{m}}(E^t) \\ 2u^- = g_{\mathrm{e}}(E^t) - g_{\mathrm{m}}(E^t). \end{cases}$$

We can consider the field u^+ as known, since it can be traced back in time into the slab. The same conclusions as above about the solvability of these equations apply here.

3.4 Inverse problems

The objective of this paper is to find methods from which the material properties can be obtained from measurements outside the slab, *i.e.*, the incident, reflected, and transmitted fields.

3.4.1 Reflection

If we can ignore the left-propagating field at the left boundary, *i.e.*, $u^- = 0$, the boundary conditions (3.3) become

$$\begin{aligned} E^i + E^r &= g_{\mathrm{e}}^{-1}(u^+) \\ E^i - E^r &= g_{\mathrm{m}}^{-1}(u^+). \end{aligned}$$

A situation where this approximation applies is a half space (see [20]) or a sufficiently thick slab, where the reflection from the right boundary, $z = d$, does not appear until after some time. This delay is at least one completed roundtrip for a wave propagating with maximal speed, $2d/(\sup_{\boldsymbol{u}} c(\boldsymbol{u}))$. For the models considered in this paper, the speed is maximal for infinitesimally small fields, *i.e.*, when the right propagating field at the left boundary, $u^+(0,t)$, is equal to zero until $t = 0$, there will be no left propagating field at the left boundary, $u^-(0,t)$, separated from zero until $t > 2d/c(0)$.

In the case of the approximation $u^- = 0$, the relation between the measurable quantities $E^i + E^r$ and $E^i - E^r$ becomes

$$g_{\mathrm{e}}(E^i + E^r) = g_{\mathrm{m}}(E^i - E^r),$$

and the composite function $g_{\mathrm{e}}^{-1}(g_{\mathrm{m}}(\cdot))$ (or its inverse $g_{\mathrm{m}}^{-1}(g_{\mathrm{e}}(\cdot))$) can be determined. The fields $E = \pm g_{\mathrm{e}}^{-1}(g_{\mathrm{m}}(\eta_0 H))$ are the electric fields which combined with $\eta_0 H$ gives a right(left) propagating wave in the slab. Differentiating this relation, we get $dE = \pm \frac{g'_{\mathrm{m}}(\eta_0 H)}{g'_{\mathrm{e}}(E)} \eta_0 dH$, which lets us define a differential wave impedance relative to vacuum as $\frac{g'_{\mathrm{m}}(\eta_0 H)}{g'_{\mathrm{e}}(E)}$.

In nonlinear optics, the materials can often be considered as nonmagnetic. This implies $g_{\mathrm{m}}(x) = x$, and we can easily determine the electric response function g_{e}, from which we get F_{e} or the wave speed c. We see that the range of the input signal E^i puts bounds on the domain of the reconstructed function g_{e}. Thus, we can not gain information on how the material responds to fields greater than those we probe with, unless we extrapolate our results.

3.4.2 Transmission

If we neglect the fields that are reflected at the right boundary, $z = d$, we are considering a problem where the wave speed depends on only one variable, and the u^+-fields propagate independently of the u^--fields. This means that the characteristic curves for the right-going fields are straight lines, which can be used to our advantage. Since the left propagating wave induced by an internal reflection is in general rather small compared with the direct wave, this is an acceptable approximation.

We assume that for $z = 0+$, the right propagating field $u^+(0,t)$ considered as a function of time has a pulse shape, *i.e.*, it is continuous with finite support, and has only one extremum, *e.g.*, a maximum. This implies that there are two times for which u^+ assumes the same value, *i.e.*, $u^+(0,t_1) = u^+(0,t_1+\tau)$ for some time separation τ. Since the wave speed depends only on u^+ when we neglect the left

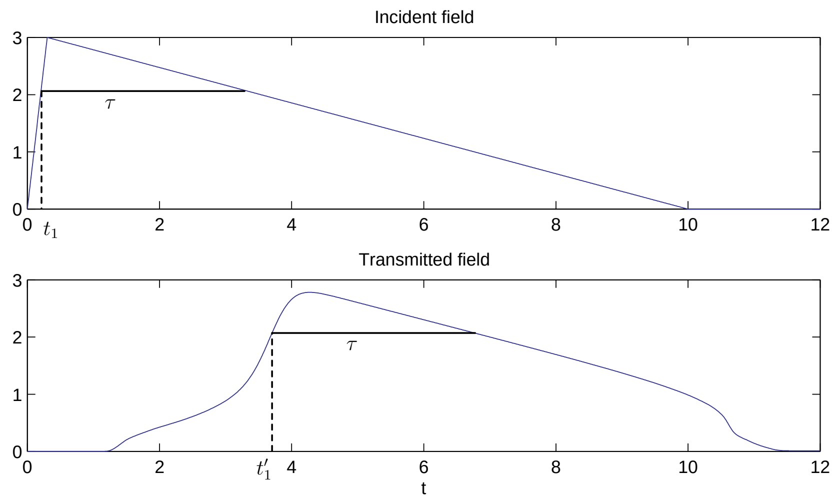

Figure 2: Method for extracting the travel time for different field amplitudes. Since equal amplitudes travel with equal speed, they arrive with the same time separation and the travel time is $t_1' - t_1$.

propagating field u^-, these two points of equal amplitude will travel with the same speed, and thus appear with the same time separation on the right side of the slab, *i.e.*, $u^+(d, t_1') = u^+(d, t_1' + \tau)$ for some time t_1'. This can be used to find the propagation time corresponding to the amplitude in question, $t_1' - t_1$, and thereby the wave speed $c(u^+) = d/(t_1' - t_1)$.

One complication is that we can only measure the fields outside the slab, but using the boundary conditions (3.3) and (3.4),

$$\begin{cases} 2u^+ = g_\mathrm{e}(E^i + E^r) + g_\mathrm{m}(E^i - E^r) \\ 2u^+ = g_\mathrm{e}(E^t) + g_\mathrm{m}(E^t), \end{cases}$$

we find that there is a one-to-one correspondence between the incident field strength and the u^+-level, and between the transmitted field strength and the u^+-level. This means that if $E^i(t_1) = E^i(t_1 + \tau)$, then there is a time t_1' for which $E^t(t_1') = E^t(t_1' + \tau)$, and we can find our transmission time $t_1' - t_1$.

In other words, we take a segment of a certain length τ of the time axis, and fit this into the curves $E^i(t)$ and $E^t(t)$. The time difference between the fits is the travel time for this particular amplitude, see Figure 2. This does not work with shock solutions, but the only consequence is that we cannot get any information on the travel time for the amplitudes over which the shock occurs.

We have the following relationships determined by reflection data and transmission time:

$$\begin{aligned} E^i + E^r &= g_\mathrm{e}^{-1}(g_\mathrm{m}(E^i - E^r)) \\ c(E^i + E^r, E^i - E^r) &= \frac{c_0}{g_\mathrm{e}'(E^i + E^r) g_\mathrm{m}'(E^i - E^r)}. \end{aligned}$$

If we denote the measurable quantities $E^i + E^r$ and $E^i - E^r$ by e and h, we have the experimentally determined functions

$$
\begin{aligned}
e(h) &= g_{\mathrm{e}}^{-1}(g_{\mathrm{m}}(h)) \\
c(e,h) &= \frac{c_0}{g_{\mathrm{e}}'(e)g_{\mathrm{m}}'(h)}.
\end{aligned}
\tag{3.5}
$$

The derivative of e with respect to h is $\frac{de}{dh} = \frac{g_{\mathrm{m}}'(h)}{g_{\mathrm{e}}'(e)}$, corresponding to the differential wave impedance. We can thus find $g_{\mathrm{e}}'(e)^2 = F_{\mathrm{e}}'(e)$ and $g_{\mathrm{m}}'(h)^2 = F_{\mathrm{m}}'(h)$ by combining these relations:

$$
\begin{aligned}
F_{\mathrm{e}}'(e) &= \frac{c_0\frac{dh}{de}}{c(e,h(e))} \quad \Rightarrow \quad F_{\mathrm{e}}(e) = \int_0^{h(e)} \frac{c_0 dh'}{c(e(h'),h')} \\
F_{\mathrm{m}}'(h) &= \frac{c_0\frac{de}{dh}}{c(e(h),h)} \quad \Rightarrow \quad F_{\mathrm{m}}(h) = \int_0^{e(h)} \frac{c_0 de'}{c(e',h(e'))}.
\end{aligned}
\tag{3.6}
$$

From these expressions we conclude that there is a one-to-one correspondence between $F_{\mathrm{e,m}}$ and $c(e,h)$ once the relation between e and h is given. Since this is given by $g_{\mathrm{e}}(e) - g_{\mathrm{m}}(h) = 0$, and $g_{\mathrm{e,m}}$ are monotone functions, this is a one-to-one relation. With shockfree propagation of a pulseshaped signal, the transmitted signal should also be pulseshaped, see *e.g.*, the example in Figure 2. Then the wavespeed $c(e,h(e)) = c(e(h),h)$ must be unique, and we conclude that the reconstructed functions are unique, always exist, and depend continuously on the data. Thus the algorithm is well posed.

3.5 Implementation of the forward problem

In order to obtain the reflected and transmitted fields from the slab, an algorithm using finite differences has been implemented in MATLAB. The algorithm is based on interpolating the wave speed and fields between two neighboring points in the grid with a linear function, and tracing the characteristics back one time step. The tracing is made by searching for the point in the grid for which the interpolated wave speed points to the new grid point. The method is described in [12, Chap. 8].

This method does not handle discontinuous solutions very well, but rather smears the discontinuity over 10-20 grid points. Since we never use shock solutions in our reconstruction algorithm, this is not a problem. When tested, the travel time for shocks seems to be correct, though.

For numerical reasons, it is advantageous to scale the problem. We have access to two different scalings; one scales the spacetime and one scales the fields. The scaling is most obvious when looking at the original Maxwell equations,

$$
\partial_z \begin{pmatrix} E \\ \eta_0 H \end{pmatrix} + \frac{1}{c_0} \begin{pmatrix} 0 & F_{\mathrm{m}}'(\eta_0 H) \\ F_{\mathrm{e}}'(E) & 0 \end{pmatrix} \partial_t \begin{pmatrix} E \\ \eta_0 H \end{pmatrix} = 0.
$$

When multiplying this equation by a factor a, and introducing the new fields $\tilde{E} = aE$ and $\tilde{H} = a\eta_0 H$, we get

$$
\partial_z \begin{pmatrix} \tilde{E} \\ \tilde{H} \end{pmatrix} + \frac{1}{c_0} \begin{pmatrix} 0 & F_{\mathrm{m}}'(\tilde{H}/a) \\ F_{\mathrm{e}}'(\tilde{E}/a) & 0 \end{pmatrix} \partial_t \begin{pmatrix} \tilde{E} \\ \tilde{H} \end{pmatrix} = 0.
$$

There is no problem incorporating the factor $\frac{1}{a}$ in the constitutive relations, *e.g.*, when $F_{\mathrm{e}}'(E) = \varepsilon_1 + \varepsilon_3 E^2$, we have $F_{\mathrm{e}}'(\tilde{E}/a) = \varepsilon_1 + \frac{\varepsilon_3}{a^2}\tilde{E}^2 = \varepsilon_1 + \tilde{\varepsilon}_3\tilde{E}^2 = \tilde{F}_{\mathrm{e}}{}'(\tilde{E})$. We see that the fields can be quite arbitrarily scaled, as long as we scale the constitutive relations as well. In our simulations, we have chosen to use a factor a such that the nonlinear terms in the constitutive relations are of the same order as the linear ones, when using a numerical field strength of a few units.

We also see that we still have the possibility to scale the spacetime, since this only effects the differential operators. Note that this is true only for homogeneous media; for inhomogenous media we would have to scale the constitutive relations once again. In our simulations, we have chosen to scale the spacetime so that the vacuum wave speed c_0 is 1, and the slab has width 1. The slab is discretized with 100 grid points in space, and the step size in time is chosen the same as that in space. This guarantees that when tracing the characteristics back in time, we stay within the nearest grid points in space.

Since we can scale the field strength and spacetime individually, and must avoid shock solutions but still have substantial nonlinear effects, our results will apply to situations with either strong fields and short propagation distances, or weak fields and long propagation distances. Of course, the concepts strong–weak and short–long, must be related to the exact physical media being modeled.

4 Numerical results

4.1 Reflection

When implementing this reconstruction, it is difficult not committing the inverse crime, *i.e.*, using the same algorithm for both simulating data and reconstructing the constitutive functions, leading to a perfect match [7, p. 121].

It is therefore meaningless to present *any* results for reconstruction with pure reflection data, unless some measured data is available, which is not the case at the present time. The reconstruction is anyway used in the transmission reconstruction, where we get good results.

4.2 Transmission

Simulations have been run, giving reflection and transmission data for a given input signal and the constitutive relations

$$F_{\mathrm{e}}(E) = 1.5E + 2\frac{E^3}{1+E^2}$$
$$F_{\mathrm{m}}(H) = H + 2\frac{H^3}{1+H^2},$$

where we have used the scaling in Section 3.5 to define dimensionless variables and functions. These constitutive relations describe a Kerr material with saturation, *i.e.*, it behaves as a material with a nonlinear behavior for weak fields, and as a linear material for strong fields.

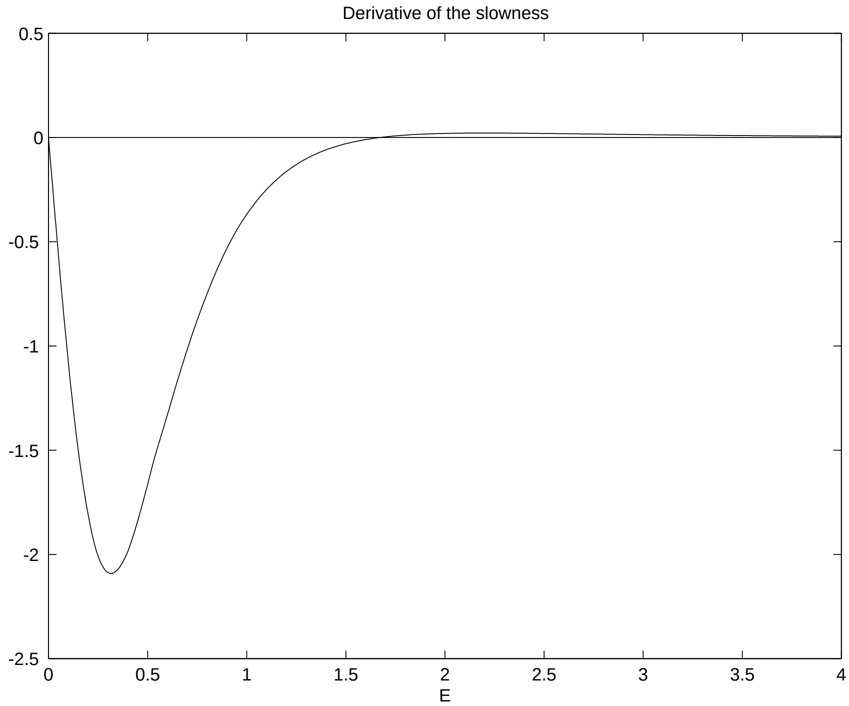

Figure 3: Graph of $-\frac{\partial}{\partial E}\frac{c_0}{c(E,H(E))}$. Observe that it is slightly positive for large field strengths, about 0.022.

We have previously stated that the condition (3.2) must be satisfied to avoid shock solutions. Figure 3 depicts the function which, when multiplied with $\frac{\partial E^i}{\partial t}$, should be less than $c_0/d = 1$. Since the function is mostly negative, we see that the greatest danger is when $\frac{\partial E^i}{\partial t} < 0$, *i.e.*, on the trailing edge of the pulse. This can be avoided by using an incident field wich decays sufficiently slow. When the derivative is positive, there is an upper limit on $\frac{\partial E^i}{\partial t}$ set by the reciprocal of the greatest positive value of the function in Figure 3, *i.e.*, $1/0.022 = 46$. Thus, we can use an incident field which rises very rapidly, but not instantly. We want it to rise fast enough so that its peak value is obtained before the reflected field at the back has returned; this gives us an exact map of the relation $g_{\mathrm{e}}(E) = g_{\mathrm{m}}(H)$, since then $g_{\mathrm{e}}(E) - g_{\mathrm{m}}(H) = 2u^- = 0$. The incident field used is depicted in Figure 4.

It should be stressed that it is not necessary to make the measurement of reflected and transmitted fields simultaneously. This is because the reflected field is only used to establish the relation between the electric and the magnetic fields necessary to create only a forward propagating field u^+, *i.e.*, $E = g_{\mathrm{e}}^{-1}(g_{\mathrm{m}}(H))$.

Figures 4 and 5 show the calculated fields and the reconstructed constitutive functions. The fields are calculated using the full forward problem, *i.e.*, the left propagating field u^- in the slab is present. The mean relative error in the reconstruction was 2.3% for F_{e} and 2.5% for F_{m}.

The algorithm is based on neglecting the field reflected from the back edge. To investigate the validity of this approximation the following test has been made. The left propagating field was neglected in the solution of the forward problem,

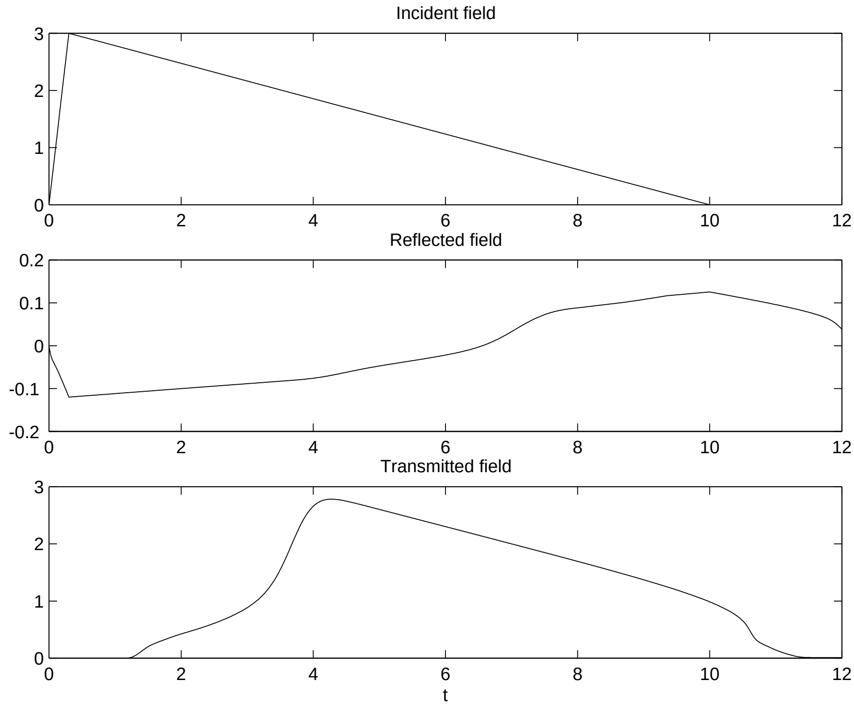

Figure 4: Incident, reflected and transmitted fields.

i.e., we used straight characteristics. Then we used the full forward problem, and compared the travel times obtained in the two cases. The mean relative difference between them was 0.19%, which shows that the approximation is good, at least for the materials studied in this paper.

In Figure 6 we have depicted the travel time as a function of the incident field strength for the two methods, as well as the difference between them. It is clearly seen that the greatest difference in travel time is for small field strengths. Remember that the expression for the slowness is $\frac{c_0}{c(u^+,u^-)} = g_{\mathrm{e}}'(u^+ + u^-)g_{\mathrm{m}}'(u^+ - u^-)$, which means that the error in travel time when neglecting u^- should be small when u^+ is relatively large.

5 Discussion and conclusions

It has been shown that it is possible to reconstruct the constitutive functions of a nonlinear slab, with the help of reflection and transmission data, not necessarily measured simultaneously. The algorithm is based on the fact that equal amplitudes travel with almost equal and constant speeds. When one of the constitutive functions is known, for instance for a nonmagnetic material, the other function is obtained with reflection data only. The algorithm seems to be robust and simple, and may be useful for measuring instantaneous nonlinear effects, with virtually no assumptions made on the specific form of the constitutive function, *i.e.*, the inverse algorithm is model independent.

Since the algorithm is based on shock free propagation, it is necessary to con-

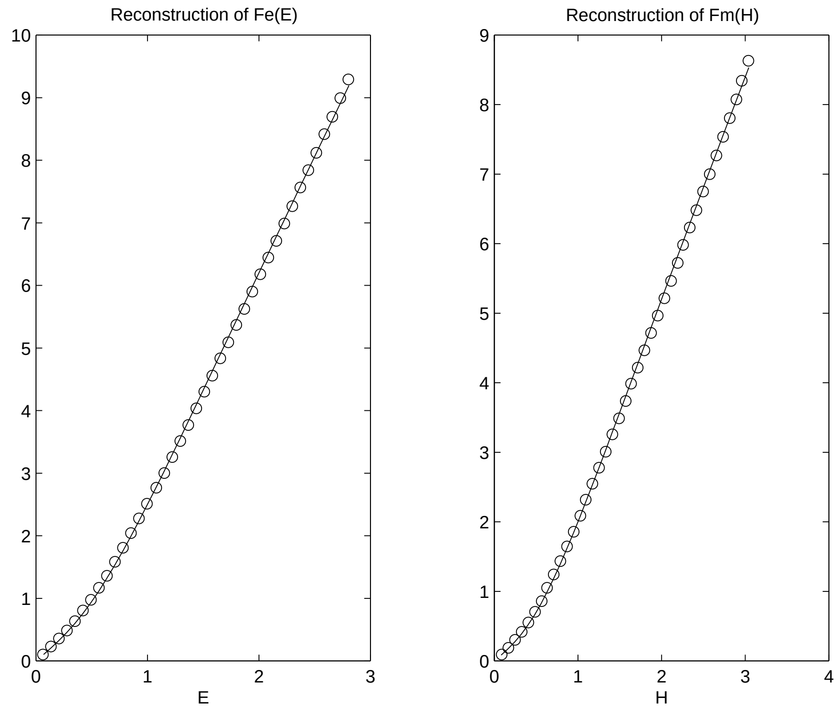

Figure 5: Reconstructed functions, from the fields in Figure 4. The circles are the reconstructed values, and the solid lines are the true functions.

struct a suitable input signal. When measuring reflected and transmitted field simultaneously, the input signal should rise fast enough so that its maximum is reached before the first reflection from the back boundary turns up, and then decrease slow enough not to create a shock in the transmitted field. This may be a difficult field to create.

The neglection of u^- in the propagation corresponds to the first term in a series expansion of the slowness at $u^- = 0$, *i.e.*, $\frac{c_0}{c(u^+,u^-)} = \frac{c_0}{c(u^+,0)} + O(u^-)$. The term $O(u^-)$ is proportional not only to u^- but also to the derivative of the slowness, which is proportional to the second derivative of the constitutive relations. A material is defined as weakly nonlinear if this second derivative is small compared to the reciprocal of the field strength. We then expect our method to work well for such materials, since the neglected term is a product of two small quantities. The series approach can in principle be used to establish a definite bound on the error in travel time, deduced directly from the constitutive relations. Though, this is a formidable problem, which is under current research. A rigorous analysis of such an expansion of the slowness may also be used to further develop the reconstruction algorithm presented in this paper, and will probably clarify which properties of the constitutive relations are important for the wave propagation.

An interesting fact is that it is conceivable to have a material with nonlinear behavior in both electric and magnetic fields. If the media changes from being dominantly electric to being dominantly magnetic, or vice versa, we may get a very small reflection for a very strong incident wave. This might have some implications

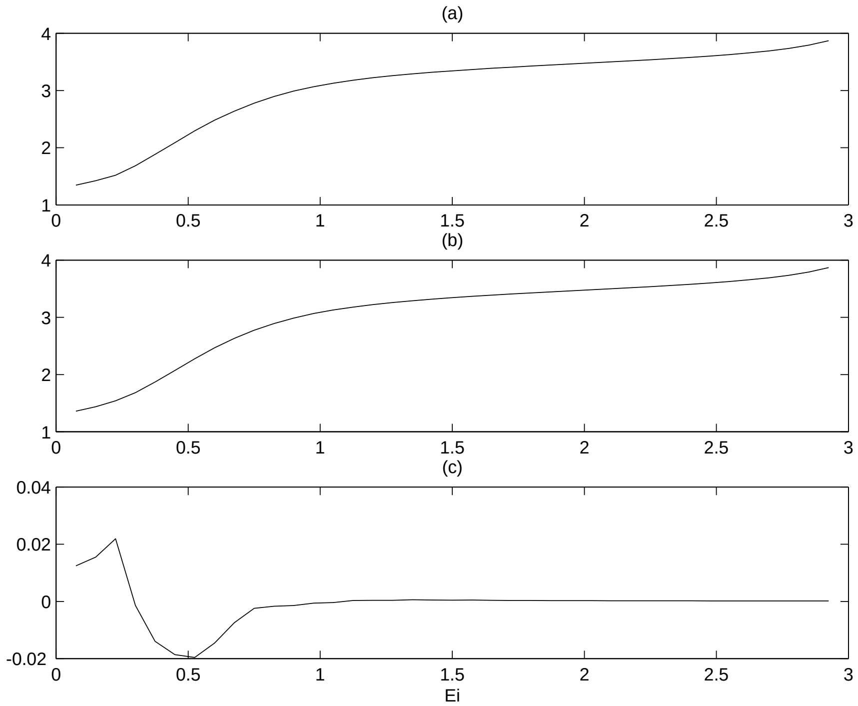

Figure 6: Amplitude-dependent travel time as a function of the incident field strength. (a) Travel time when using straight characteristics, *i.e.*, neglecting the left propagating field u^-. (b) Travel time when using the full forward problem. (c) The difference in travel time for the two methods.

on the theory of nonreflecting materials, or provide a new kind of electric shutter.

6 Acknowledgement

The author wishes to express his sincere gratitude to Prof. G. Kristensson for valuable discussions and suggestions during this project.

The work reported in this paper is partially supported by a grant from the Swedish Research Council for Engineering Sciences, and its support is gratefully acknowledged.

References

[1] G. P. Agrawal. *Nonlinear fiber optics.* Academic Press, San Diego, 1995.

[2] F. G. Bass, G. Y. Slepyan, S. A. Maksimenko, and A. Lakhtakia. Time-domain Beltrami-Maxwell solitons in certain nonlinear chiral media. *Microwave Opt. Techn. Lett.*, **9**(4), 218–221, July 1995.

[3] F. Bloom. *Mathematical problems of classical nonlinear electromagnetic theory.* Longman Scientific & Technical, Burnt Mill, Harlow, England, 1993.

[4] R. W. Boyd. *Nonlinear Optics.* Academic Press, San Diego, 1992.

[5] D. Censor and Y. Ben-Shimol. Wave propagation in weakly nonlinear bi-anisotropic and bi-isotropic media. *J. Electro. Waves Applic.*, **11**, 1763–1779, 1997.

[6] B. D. Coleman and E. H. Dill. Thermodynamic restrictions on the constitutive equations of electromagnetic theory. *Z. angew. Math. Phys.*, **22**, 691–702, 1971.

[7] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer-Verlag, Berlin, 1992.

[8] J. P. Corones, M. E. Davison, and R. J. Krueger. The effects of dissipation in one-dimensional inverse problems. In A. J. Devaney, editor, *Inverse Optics, Proceedings of the SPIE*, pages 107–114, Bellingham, WA, 1983. Proc. SPIE 413, SPIE.

[9] M. E. Davison and R. C. Winther. A general approach to splitting and invariant imbedding for linear wave equations. *J. Math. Anal. Appl.*, **188**(1), 158–181, 1994.

[10] M. de Magistris, G. Miano, L. Verolino, C. Visone, and E. Zamparelli. A numerical analysis of the behaviour of the Galerkin equations relevant to electromagnetic wave propagation in nonlinear media. *IEEE Trans. Magnetics*, **30**, 3196–3199, 1994.

[11] A. C. Eringen and G. A. Maugin. *Electrodynamics of Continua I.* Springer-Verlag, New York, 1990.

[12] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time Dependent Problems and Difference Methods.* John Wiley & Sons, New York, 1995.

[13] A. Hasegawa. *Optical Solitons in Fibers.* Springer-Verlag, Berlin Heidelberg, 1990.

[14] A. Hasegawa and Y. Kodama. *Solitons in Optical Communications.* Clarendon Press, Oxford, 1995.

[15] L. Hörmander. *Lectures on Nonlinear Hyperbolic Differential Equations.* Number 26 in Mathémathiques & Applications. Springer-Verlag, Berlin, 1997.

[16] F. John. *Partial Differential Equations.* Springer-Verlag, New York, fourth edition, 1982.

[17] R. M. Joseph and A. Taflove. FDTD Maxwell's equations models for nonlinear electrodynamics and optics. *IEEE Trans. Antennas Propagat.*, **45**, 364–374, 1997.

[18] J. Y. Kazakia and R. Venkataraman. Propagation of electromagnetic waves in a nonlinear dielectric slab. *Z. angew. Math. Phys.*, **26**, 61–76, 1975.

[19] A. Kluwick and E. A. Cox. Nonlinear waves in materials with mixed nonlinearity. *Wave Motion*, **27**, 23–41, 1998.

[20] G. Kristensson and D. J. N. Wall. Direct and inverse scattering for transient electromagnetic waves in nonlinear media. *Inverse Problems*, **14**, 113–137, 1998.

[21] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon. Experimental observation of picosecond pulse narrowing and solitons in optical fibers. *Phys. Rev. Lett.*, **45**(13), 1095–1098, September 1980.

[22] C. Rogers, H. M. Cekirge, and A. Askar. Electromagnetic wave propagation in non-linear dielectric media. *Acta Mech.*, **26**, 59–73, 1977.

[23] G. Y. Slepyan, S. A. Maksimenko, F. G. Bass, and A. Lakhtakia. Nonlinear electromagnetics in chiral media: Self-action of waves. *Phys. Rev. E*, **52**(1), 1049–1058, July 1995.

[24] J. Veldhuis and H. Blok. Reflection and refraction of electromagnetic waves by the plane boundary of a non-linear medium: application of the simple wave theory. *J. Phys. A: Math. Gen.*, **11**(8), 1663–1680, 1978.

Paper II

Simple wave solutions for the Maxwell equations in bianisotropic, nonlinear media, with application to oblique incidence

D. Sjöberg

Technical Report LUTEDX/(TEAT-7078)/1–17/(1999)

Abstract

Using simple waves and six-vector formalism, the propagation of electromagnetic waves in nonlinear, bianisotropic, nondispersive, homogeneous media is analyzed. The Maxwell equations are formulated as an eigenvalue problem, whose solutions are equivalent to the characteristic directions of the wave front. Oblique incidence of plane waves in vacuum on a half space of nonlinear material is solved, giving reflection and transmission operators for all angles of incidence and all polarizations of the incident field. A condition on Brewster angles is derived.

1 Introduction

Wave propagation in nonlinear media is a wide and quickly expanding area. In particular, the nonlinear optics field has been very prosperous [1, 3]. One of the most exciting areas is that of solitons, *i.e.*, pulses which have a very specific shape, in which the nonlinear steepening effects are precisely balanced by the dispersive broadening, thereby producing a pulse that is temporally or spatially unchanged during propagation. This delicate balance can only be understood by studying both contributing effects. In this paper, we are devoted to the nonlinear effects which occur in materials with no memory, *i.e.*, no dispersion.

Whereas the linear dispersion has been thoroughly investigated, *e.g.*, [4, 14, 20], the nonlinear properties may not have received enough attention. Some early works are summarized in [2], and especially the papers on wave propagation in nonlinear dielectrics [5, 6, 17, 21, 29] are worthy of attention. A prominent feature of nonlinear wave propagation, where the nonlinearity acts as an amplitude-dependent wave speed, is the formation of shock waves. These are discontinuous waves, which must be interpreted in a generalized way as weak solutions, see *e.g.*, [28, pp. 369–373], and the theory of these has been thoroughly studied [15, 18, 27, 31]. It is often argued that the shock waves are eliminated by the linear dispersion, see *e.g.*, [1, pp. 117–120], but since we are ignoring dispersion in this study, we expect our model to be accurate only when we are not in the vicinity of any shock formations.

An often encountered problem when studying nonlinear materials is that of finding suitable constitutive relations. In the treatise of Eringen and Maugin [9, 10], the constitutive relations for virtually every reasonable situation are presented. Some important thermodynamic restrictions are presented in [8]. The derivation of constitutive relations from a quantum mechanical point of view is presented in [3], and some theory about nonlinear dielectrics is found in [7].

This paper aims to improve the understanding of a nonlinear optical response, *i.e.*, an instantaneous nonlinear response. Earlier works, as reported above, have often made some important restrictions, such as assuming the material to be isotropic or uniaxial. Here we present a theory describing wave propagation in bianisotropic materials. We show that a generalized form of plane waves, called simple waves, can be used to analyze the wave propagation, and we reformulate the Maxwell equations as an eigenvalue problems. A brief presentation on simple waves in partial differential equations is given in [19, p. 52], and a more extensive treatment is given in [16, Chap. 3]. There are also some related results in [13, p. 47].

The paper is organized as follows: in Sections 2 and 3 we present the simple wave Ansatz and the six-vector formalism, which are the basic tools used in this paper. This is applied to the Maxwell equations in Section 4, which transforms the dynamics into an eigenvalue problem. Special notice is taken to isotropic media. In Section 5 we introduce the theory on how to classify materials. We then apply our formalism in Section 6 to the problem of a plane wave obliquely impinging on a nonlinear half space and solve the problem of finding the reflected and transmitted fields. Some results on suitable conditions on the Brewster angles are also presented, as well as a numerical example.

2 Simple wave Ansatz

Plane waves constitute a powerful tool in the analysis of wave phenomena in linear materials. The concept of plane waves transforms the problem of three spatial dimensions into a problem along the propagation direction. Simple waves are the generalization of this concept. They have previously been used in the description of nonlinear electromagnetic waves [5, 6], and are explained in basic books on partial differential equations [19, p. 52]. They also define the characteristics of the wave equation.

The simple wave Ansatz is suitable for materials which respond instantaneously to excitation, and states that the fields depend only upon a scalar parameter, which we denote ϕ. This parameter is a function of space and time. For an isotropic, linear media the simple wave Ansatz reduces to the usual phase function, $\phi(\boldsymbol{r}, t) = \boldsymbol{k}\cdot\boldsymbol{r} - \omega t$.

It is obvious that if a quantity u depends on space and time as $u(\boldsymbol{r}, t) = u(\phi(\boldsymbol{r}, t))$, the spatial gradient $\nabla\phi$ representents a propagation direction. We identify the quantity $-\frac{\nabla\phi/\phi_t}{|\nabla\phi/\phi_t|}$ as the propagation direction and $|\phi_t|/|\nabla\phi|$ as the propagation speed, where ϕ_t denotes the time derivative of ϕ. The minus sign comes from implicit differentiation of the equation $\phi(\boldsymbol{r}, t) =$ constant, which is the equation of the wave front.

3 Six-vector formalism

When describing bianisotropic phenomena, it is often advantageous to use the six-vector formalism, see *e.g.*, [24]. In this approach, we make no real distinction between the electric and magnetic fields, but rather treat them as components of a single field. We define our fields as

$$\begin{cases} \boldsymbol{e} = \begin{pmatrix} \sqrt{\epsilon_0}\boldsymbol{E} \\ \sqrt{\mu_0}\boldsymbol{H} \end{pmatrix} \\ \boldsymbol{d} = \begin{pmatrix} \frac{1}{\sqrt{\epsilon_0}}\boldsymbol{D} \\ \frac{1}{\sqrt{\mu_0}}\boldsymbol{B} \end{pmatrix} \end{cases}$$

where ϵ_0 and μ_0 denote the permittivity and permeability of vacuum, respectively. The six-vector fields now both have the same dimension, *i.e.*, $\sqrt{\text{energy/volume}}$.

The scalar product between two six-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined as $\boldsymbol{a}\cdot\boldsymbol{b} = \sum_{i=1}^{6} a_i b_i$. Operations with three-vectors on six-vectors are understood in the obvious manner, *i.e.*, the scalar and cross products are

$$\boldsymbol{v}\cdot\boldsymbol{e} = \begin{pmatrix} \boldsymbol{v}\cdot\sqrt{\epsilon_0}\boldsymbol{E} \\ \boldsymbol{v}\cdot\sqrt{\mu_0}\boldsymbol{H} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{v}\times\boldsymbol{e} = \begin{pmatrix} \boldsymbol{v}\times\sqrt{\epsilon_0}\boldsymbol{E} \\ \boldsymbol{v}\times\sqrt{\mu_0}\boldsymbol{H} \end{pmatrix}.$$

Using the operator $\mathbf{J} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}$, which is formed from the three-dimensional spatial identity operator $\mathbf{I}$, we write the source free Maxwell equations as

$$\nabla\times\boldsymbol{e} - \frac{1}{c_0}\mathbf{J}\cdot\partial_t\boldsymbol{d} = 0,$$

where c_0 denotes the wave speed in vacuum, $1/\sqrt{\epsilon_0\mu_0}$. The spatial differential operator ∇ is treated as a three-vector, and is sometimes merged with the operator $\mathbf{J}$ to form the symmetric operator $\nabla\times\mathbf{J}$, as in [12]. This approach will be beneficial later on in this work.

4 The Maxwell equations as an eigenvalue problem

The constitutive relation for a material with no memory, *i.e.*, where the fluxes $\boldsymbol{d}$ depend only upon the present values of the field strengths $\boldsymbol{e}$, can be written

$$\boldsymbol{d}(\boldsymbol{r},t) = \boldsymbol{d}(\boldsymbol{e}(\boldsymbol{r},t)). \tag{4.1}$$

We now apply the simple wave Ansatz together with the constitutive relation,

$$\begin{cases} \boldsymbol{e}(\boldsymbol{r},t) = \boldsymbol{e}(\phi(\boldsymbol{r},t)) \\ \boldsymbol{d}(\boldsymbol{r},t) = \boldsymbol{d}(\boldsymbol{e}(\phi(\boldsymbol{r},t))). \end{cases}$$

This means that the curl operator turns into a cross product, $\nabla\times\boldsymbol{e} = \nabla\phi\times\boldsymbol{e}'$, and the time derivative becomes $\partial_t\boldsymbol{d} = \phi_t[\nabla_e\boldsymbol{d}]\cdot\boldsymbol{e}'$, where the prime denotes differentiation with respect to ϕ. The operator ∇_e denotes the field gradient operator, *i.e.*, $[\nabla_e\boldsymbol{d}]_{nm} = \frac{\partial}{\partial e_m}d_n(\boldsymbol{e})$. Since we write the linear constitutive relations as $\boldsymbol{d} = \boldsymbol{\varepsilon}\cdot\boldsymbol{e}$, where $\boldsymbol{\varepsilon}$ is a six-dyadic, we denote $[\nabla_e\boldsymbol{d}]$ by $\boldsymbol{\varepsilon}(\boldsymbol{e})$, and often suppress the argument to obtain a less cumbersome notation.

With the simple wave Ansatz, the Maxwell equations contain the generic field $\boldsymbol{e}' = \frac{d}{d\phi}\boldsymbol{e}$. However, for reasons that become more obvious below we prefer to use the time derivative, $\dot{\boldsymbol{e}} = \partial_t\boldsymbol{e} = \phi_t\boldsymbol{e}'$. This choice also becomes advantageous when implementing the equations later on. Since $\phi(\boldsymbol{r},t) = \text{constant}$ is the equation for the wave front, we identify the wave slowness $1/c$ and the propagation direction $\hat{k}$ of the simple wave by the following expressions,

$$\begin{cases} \dfrac{1}{c} = \dfrac{|\nabla\phi|}{|\phi_t|} \\ \hat{k} = -\dfrac{\nabla\phi/\phi_t}{|\nabla\phi/\phi_t|} = -\dfrac{\nabla\phi}{\phi_t}c. \end{cases}$$

Using these expressions, we write the Maxwell equations as

$$\frac{1}{c}\hat{k} \times \dot{\boldsymbol{e}} + \frac{1}{c_0}\mathbf{J} \cdot \boldsymbol{\varepsilon} \cdot \dot{\boldsymbol{e}} = 0.$$

This is an eigenvalue problem, which becomes more obvious in the form

$$\frac{c}{c_0}\dot{\boldsymbol{e}} = \boldsymbol{\varepsilon}^{-1} \cdot [\hat{k} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}, \tag{4.2}$$

which follows from $\mathbf{J}^{-1} = -\mathbf{J}$ and $\mathbf{J} \cdot [\hat{k} \times \mathbf{I}] = [\hat{k} \times \mathbf{J}]$. Observe that $[\hat{k} \times \mathbf{J}]$ is a symmetric operator. The dyadic $\boldsymbol{\varepsilon}$ is postulated to be positive definite and symmetric, and is thus invertible. In the linear case, it is possible to show that $\boldsymbol{\varepsilon}$ has to be a symmetric, positive definite dyadic in order to model passive media [12]. The assumptions made on the dyadic $\boldsymbol{\varepsilon}$ is a natural generalization of the result in the linear case.

The solution to (4.2) gives conditions on the wave speed and propagation direction in terms of the fields. In the linear case, only the directions of the field will be important, but for nonlinear materials there is also a dependence on the amplitude. For an isotropic material, where

$$\boldsymbol{\varepsilon}(\boldsymbol{e}) = \begin{pmatrix} \epsilon(E)\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mu(H)\mathbf{I} \end{pmatrix},$$

the conditions are

$$c = \frac{c_0}{\sqrt{\epsilon(E)\mu(H)}} \quad \text{and} \quad \hat{k} \cdot \dot{\boldsymbol{e}} = 0 \quad \Rightarrow \quad \dot{\boldsymbol{e}} = \begin{pmatrix} \frac{1}{\sqrt{\epsilon}}\boldsymbol{v} \\ \frac{1}{\sqrt{\mu}}\hat{k} \times \boldsymbol{v} \end{pmatrix}, \tag{4.3}$$

where the three-vector $\boldsymbol{v}$ is orthogonal to $\hat{k}$. Observe that it is the direction of the *derivatives* of the fields that are important, not the fields themselves.

For a given propagation direction $\hat{k}$ the operator $\boldsymbol{\varepsilon}^{-1} \cdot [\hat{k} \times \mathbf{J}]$ has six eigenvectors $\dot{\boldsymbol{e}}_j$, $j = 1, \ldots, 6$. Since the operator is not symmetric, these solutions are not guaranteed to be mutually orthogonal. We symmetrize the operator by

$$\frac{c}{c_0}(\sqrt{\boldsymbol{\varepsilon}} \cdot \dot{\boldsymbol{e}}_j) = \left[\sqrt{\boldsymbol{\varepsilon}}^{-1} \cdot [\hat{k} \times \mathbf{J}] \cdot \sqrt{\boldsymbol{\varepsilon}}^{-1}\right] \cdot (\sqrt{\boldsymbol{\varepsilon}} \cdot \dot{\boldsymbol{e}}_j),$$

where we have used the square root of the positive definite and symmetric dyadic $\boldsymbol{\varepsilon}$, which is also positive definite and symmetric. It is concluded that the eigenvectors $\sqrt{\boldsymbol{\varepsilon}} \cdot \dot{\boldsymbol{e}}_j$ are real and orthogonal, which imply that the eigenvectors $\dot{\boldsymbol{e}}_j$ are real and linearly independent. The operator $\sqrt{\boldsymbol{\varepsilon}}^{-1}\cdot[\hat{k}\times\mathbf{J}]\cdot\sqrt{\boldsymbol{\varepsilon}}^{-1}$ is a congruence transformation (see *e.g.*, [11, p. 251]) of $[\hat{k}\times\mathbf{J}]$, which has the (double) eigenvalues -1, 0 and 1. Since the signs are preserved under congruence transforms, we conclude that for a given propagation direction $\hat{k}$ there are two modes propagating in the $+\hat{k}$-direction (positive eigenvalues) and two modes propagating in the $-\hat{k}$-direction (negative eigenvalues), while two modes do not propagate with respect to $\hat{k}$ at all (zero eigenvalue). The last two can be written explicitly as $\dot{\boldsymbol{e}}_{5,6}=\binom{\pm\hat{k}}{\hat{k}}$.

5 Classification of materials

Materials are often classified as, *e.g.*, isotropic, bi-isotropic or uniaxial depending on the invariance under symmetry transformations. In our formulation, the natural way to classify the materials is by the corresponding invariance of the dyadic $\boldsymbol{\varepsilon}(\boldsymbol{e})$. This is motivated by the following way of writing the constitutive relations (4.1):

$$\boldsymbol{d}(\boldsymbol{e})=\int_0^{\boldsymbol{e}}\boldsymbol{\varepsilon}(\boldsymbol{e}')\cdot d\boldsymbol{e}',$$

where the integral should be understood in terms of integration along a parametrized curve in $\mathbb{R}^6$. The prime is not to be confused with time differentiation, it is only denoting the integration variable. When applying a spatial transformation $\mathbf{S}$ on the field strength $\boldsymbol{e}$, we get

$$\boldsymbol{d}(\mathbf{S}\cdot\boldsymbol{e})=\int_0^{\mathbf{S}\cdot\boldsymbol{e}}\boldsymbol{\varepsilon}(\boldsymbol{e}')\cdot d\boldsymbol{e}'=\int_0^{\boldsymbol{e}}\boldsymbol{\varepsilon}(\mathbf{S}\cdot\boldsymbol{e}'')\cdot\mathbf{S}\cdot d\boldsymbol{e}'',$$

where we have made the change of variables $\boldsymbol{e}'=\mathbf{S}\cdot\boldsymbol{e}''$. Materials are classified depending on which group of transformations $\mathbf{S}$ that satisfies $\boldsymbol{d}(\mathbf{S}\cdot\boldsymbol{e})=\mathbf{S}\cdot\boldsymbol{d}(\boldsymbol{e})$, *i.e.*, which group of transformations that commutes with $\boldsymbol{\varepsilon}$.

Since this must hold for all transformations in the bi-isotropic case, we see that an $\boldsymbol{\varepsilon}(\boldsymbol{e})$ as

$$\boldsymbol{\varepsilon}(\boldsymbol{e})=\begin{pmatrix}\epsilon(E,H)\mathbf{I} & \xi(E,H)\mathbf{I}\\ \zeta(E,H)\mathbf{I} & \mu(E,H)\mathbf{I}\end{pmatrix},$$

describes a bi-isotropic material, where ϵ, ξ, ζ and μ are scalar functions of the field strengths. Common restrictions on constitutive relations, [8, 12], say that $\xi=\zeta$, and if they are equal to zero, the material is said to be isotropic.

6 Oblique incidence

To demonstrate the possible application of the simple wave approach, we analyze the problem of a plane electromagnetic wave obliquely impinging from vacuum on a nonlinear half space. The problem has been studied to some extent in [5, 6], though they specialize their treatment to a uniaxial material with nonlinearity in electric field only, where the optical axis is in a special direction.

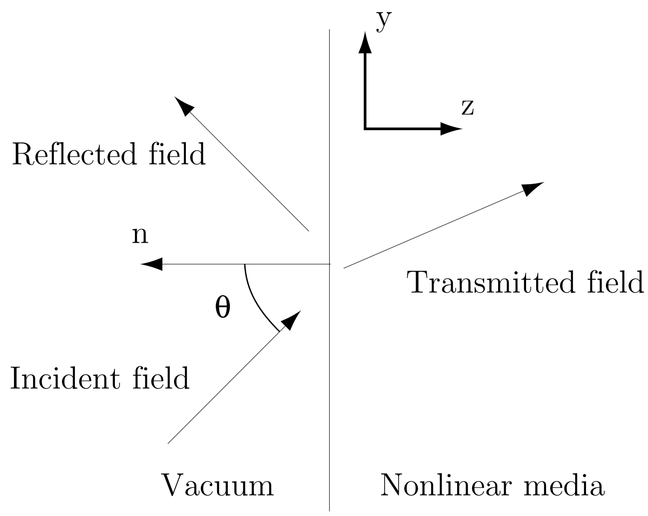

Figure 1: The geometry of the problem of oblique incidence.

6.1 Geometry and boundary conditions

The geometry of the problem is depicted in Figure 1. The incident field is a plane wave, and we make the Ansatz

$$\begin{cases} \boldsymbol{e}^i(\boldsymbol{r},t) = \boldsymbol{e}(\hat{k}^i \cdot \boldsymbol{r} - c_0 t) \\ \boldsymbol{e}^r(\boldsymbol{r},t) = \boldsymbol{e}(\hat{k}^r \cdot \boldsymbol{r} - c_0 t) \\ \boldsymbol{e}^t(\boldsymbol{r},t) = \sum \boldsymbol{e}^t_j(\phi_j(\boldsymbol{r},t)), \end{cases}$$

where c_0 denotes the wave speed in vacuum. We thus assume that the transmitted field may consist of several simple waves, as we can expect from the linear, anisotropic case. The number of those is restricted to two in Section 6.5. The usual boundary conditions apply, *i.e.*, the tangential components of the field strengths should be continuous and the normal component of the fluxes should be continuous (no sources at the interface). We write this as

$$\begin{cases} \boldsymbol{e}^i_{\parallel} + \boldsymbol{e}^r_{\parallel} = \boldsymbol{e}^t_{\parallel} \\ \hat{z} \cdot (\boldsymbol{d}^i + \boldsymbol{d}^r) = \hat{z} \cdot \boldsymbol{d}^t. \end{cases} \tag{6.1}$$

The latter condition is not used in the present analysis.

6.2 Reflection law and Snell's law

Since the boundary conditions (6.1) must hold for all times on the surface $z = 0$, we can differentiate them with respect to both time and y. The simple wave Ansatz implies that the operator ∂_y equals $\frac{1}{c} k_y \partial_t$, where $k_y = \hat{y} \cdot \hat{k}$. Using this result and $\boldsymbol{e}^t(\boldsymbol{r},t) = \sum \boldsymbol{e}^t_j(\phi_j(\boldsymbol{r},t))$ we write the time and y derivative of the tangential fields as

$$\begin{cases} \dot{\boldsymbol{e}}_{\parallel}^{i} + \dot{\boldsymbol{e}}_{\parallel}^{r} = \sum (\dot{\boldsymbol{e}}_{j}^{t})_{\parallel} \\ \dfrac{1}{c_0} k_y^i \dot{\boldsymbol{e}}_{\parallel}^{i} + \dfrac{1}{c_0} k_y^r \dot{\boldsymbol{e}}_{\parallel}^{r} = \sum \dfrac{1}{c_j(\boldsymbol{e}^t)} k_{yj}^t (\dot{\boldsymbol{e}}_{j}^{t})_{\parallel}. \end{cases}$$

These conditions are satisfied if the following holds:

$$k_y^i = k_y^r = \frac{c_0}{c_j(\boldsymbol{e}^t)} k_{yj}^t, \tag{6.2}$$

for all values of j, *cf.*, phase-matching [22, p. 104]. The quotient between the wave speeds corresponds to the refractive index, and since k_y^i and k_{yj}^t are the sines of the angles of incidence and transmission, respectively, (6.2) is the well-known Snell's law. This is a purely kinematic law, so it is not surprising that it is valid also in the nonlinear case. Note that since there are several possible values for the wave speed c_j, there are several possible angles of transmission.

Since the propagation directions are normalized and there is no propagation in the x-direction, we now also have the normal reflection law for the reflected field, *i.e.*,

$$\hat{k}^r = k_y^i \hat{y} - k_z^i \hat{z}.$$

The transmitted field is more complicated, since it involves the wave speed, which may depend on the field strength.

6.3 Decomposition of the propagation direction

It seems natural to consider a decomposition of the propagation direction $\hat{k}$ in (4.2) in a y and z part. Using Snell's law and $|\hat{k}_j^t| = 1$, we find

$$\frac{c_0}{c_j}\hat{k}_j^t = \frac{c_0}{c_j} k_{yj}^t \hat{y} + \frac{c_0}{c_j} k_{zj}^t \hat{z} = k_y^i \hat{y} + \frac{c_0}{c_j}\sqrt{1 - (\frac{c_j}{c_0} k_y^i)^2}\, \hat{z}.$$

Using the eigenvalue problem (4.2) for each simple wave in the nonlinear material, we get

$$\begin{aligned} \boldsymbol{\varepsilon} \cdot \dot{\boldsymbol{e}}_j^t &= \frac{c_0}{c_j}[\hat{k}_j^t \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_j^t \\ [\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_j^t &= \frac{c_0}{c_j} k_{zj}^t [\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_j^t \\ \frac{c_j}{c_0}\frac{1}{k_{zj}^t} \dot{\boldsymbol{e}}_j^t &= [\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\hat{z} \times \mathbf{J}] \cdot \dot{\boldsymbol{e}}_j^t. \end{aligned}$$

Since the operator $[\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\hat{z} \times \mathbf{J}]$ is independent of j, all simple waves in the nonlinear material are found from the same eigenvalue problem,

$$\lambda_j \boldsymbol{a}_j = [\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\hat{z} \times \mathbf{J}] \cdot \boldsymbol{a}_j, \tag{6.3}$$

where λ_j denotes the number $c_j/(c_0 k^t_{zj})$ and $\boldsymbol{a}_j$ is shorthand for $\dot{\boldsymbol{e}}^t_j$. The corresponding problem for the vacuum fields is easily found,

$$\pm\frac{1}{k^i_z}\dot{\boldsymbol{e}}^{i,r} = [\mathbf{I} - k^i_y\hat{y}\times\mathbf{J}]^{-1}\cdot[\hat{z}\times\mathbf{J}]\cdot\dot{\boldsymbol{e}}^{i,r}, \tag{6.4}$$

where the $\pm$ comes from $k^r_z = -k^i_z$. The operator $[\mathbf{I} - k^i_y\hat{y}\times\mathbf{J}]^{-1}$ is positive definite, since $|k^i_y| < 1$. If all eigenvalues to $\boldsymbol{\varepsilon}$ are greater than one, *i.e.*, the material is denser than vacuum, the operator $[\boldsymbol{\varepsilon} - k^i_y\hat{y}\times\mathbf{J}]^{-1}$ is also positive definite.

6.4 Properties of the eigenvectors

The eigenvalue problem (6.3) is put in a symmetric form in the same manner as in Section 4. We observe that $[\boldsymbol{\varepsilon} - k^i_y\hat{y}\times\mathbf{J}]$ is positive definite and symmetric. In this section we temporarily denote this operator $\mathbf{C}$. By multiplying (6.3) with $\sqrt{\mathbf{C}}$, which is also positive definite and symmetric, we obtain

$$\begin{aligned}\lambda_j\sqrt{\mathbf{C}}\cdot\boldsymbol{a}_j &= \sqrt{\mathbf{C}}^{-1}\cdot[\hat{z}\times\mathbf{J}]\cdot\sqrt{\mathbf{C}}^{-1}\cdot\sqrt{\mathbf{C}}\cdot\boldsymbol{a}_j\\ \lambda_j\boldsymbol{u}_j &= \left[\sqrt{\mathbf{C}}^{-1}\cdot[\hat{z}\times\mathbf{J}]\cdot\sqrt{\mathbf{C}}^{-1}\right]\cdot\boldsymbol{u}_j.\end{aligned}$$

The λ_j:s are now eigenvalues to a symmetric operator, which implies that they are real. The symmetric operator $\sqrt{\mathbf{C}}^{-1}\cdot[\hat{z}\times\mathbf{J}]\cdot\sqrt{\mathbf{C}}^{-1}$ is a congruence transformation of $[\hat{z}\times\mathbf{J}]$, which has the (double) eigenvalues -1, 0 and 1. Since the signs are preserved under congruence transformations, the eigenvalues can be characterized by

$$\begin{aligned}\lambda_{1,2} &> 0\\ \lambda_{3,4} &< 0\\ \lambda_{5,6} &= 0.\end{aligned}$$

Since the $\boldsymbol{u}_j$:s are eigenvectors to a symmetric operator, they are real and mutually orthogonal. This implies that $\boldsymbol{a}_j = \sqrt{\mathbf{C}}^{-1}\cdot\boldsymbol{u}_j$ are linearly independent vectors. The eigenvectors corresponding to $\lambda_{5,6}$ can be constructed from $\boldsymbol{a}_{5,6} = \binom{\pm\hat{z}}{\hat{z}}$, which implies that $\boldsymbol{a}_{1,2,3,4}$ are the only eigenvectors needed to form the tangential fields.

The sign of the eigenvalue indicates in which direction each mode represented by an eigenvector is propagating, *i.e.*, $\boldsymbol{a}_{1,2}$ represent waves propagating in the $+z$-direction and $\boldsymbol{a}_{3,4}$ represent waves propagating in the $-z$-direction, while $\boldsymbol{a}_{5,6}$ represent waves which do not propagate with respect to z at all.

6.5 Transmission operator

Temporarily introduce the dyadic

$$\mathbf{A} = k^i_z[\mathbf{I} - \hat{z}\hat{z}]\cdot[\mathbf{I} - k^i_y\hat{y}\times\mathbf{J}]^{-1}\cdot[\hat{z}\times\mathbf{J}].$$

From (6.4) we see that $\dot{\boldsymbol{e}}_{\|}^{i,r} = \pm \mathbf{A} \cdot \dot{\boldsymbol{e}}_{\|}^{i,r}$. By multiplying the boundary condition $\dot{\boldsymbol{e}}_{\|}^{i} + \dot{\boldsymbol{e}}_{\|}^{r} = \dot{\boldsymbol{e}}_{\|}^{t}$ with $\mathbf{A}$ we now have

$$\dot{\boldsymbol{e}}_{\|}^{i} - \dot{\boldsymbol{e}}_{\|}^{r} = \mathbf{A} \cdot \dot{\boldsymbol{e}}_{\|}^{t}.$$

In the previous section, we found that only the eigenvectors $\boldsymbol{a}_{1,2,3,4}$ involve the tangential fields. Specifically, $\boldsymbol{a}_{1,2}$ correspond to waves travelling in the $+z$-direction. To this end, the transmitted tangential field is expanded as

$$\dot{\boldsymbol{e}}_{\|}^{t} = \sum_{j=1}^{2} \alpha_j [\mathbf{I} - \hat{z}\hat{z}] \cdot \boldsymbol{a}_j, \tag{6.5}$$

provided there are no sources in the region $z > 0$, *i.e.*, no waves travelling in the $-z$-direction. We have now restricted the number of simple waves in the nonlinear material to two. From (6.3) follows

$$\begin{aligned} \mathbf{A} \cdot [\mathbf{I} - \hat{z}\hat{z}] \cdot \boldsymbol{a}_j &= k_z^i [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\hat{z} \times \mathbf{J}] \cdot [\mathbf{I} - \hat{z}\hat{z}] \cdot \boldsymbol{a}_j \\ &= \lambda_j k_z^i [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}] \cdot \boldsymbol{a}_j, \end{aligned}$$

where we have used $[\hat{z} \times \mathbf{J}] \cdot [\mathbf{I} - \hat{z}\hat{z}] = [\hat{z} \times \mathbf{J}]$. The operator

$$\begin{aligned} \mathbf{B} &= [\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\boldsymbol{\varepsilon} - k_y^i \hat{y} \times \mathbf{J}] \\ &= \mathbf{I} + [\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\boldsymbol{\varepsilon} - \mathbf{I}] \end{aligned} \tag{6.6}$$

is positive definite with eigenvalues greater than one. The boundary conditions are

$$\begin{cases} \dot{\boldsymbol{e}}_{\|}^{i} + \dot{\boldsymbol{e}}_{\|}^{r} = \displaystyle\sum_{j=1}^{2} \alpha_j [\mathbf{I} - \hat{z}\hat{z}] \cdot \boldsymbol{a}_j \\ \dot{\boldsymbol{e}}_{\|}^{i} - \dot{\boldsymbol{e}}_{\|}^{r} = \displaystyle\sum_{j=1}^{2} \alpha_j \lambda_j k_z^i [\mathbf{I} - \hat{z}\hat{z}] \cdot \mathbf{B} \cdot \boldsymbol{a}_j. \end{cases} \tag{6.7}$$

By adding these equations, we eliminate the reflected field, and obtain

$$2\dot{\boldsymbol{e}}_{\|}^{i} = \sum_{j=1}^{2} \alpha_j [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} + \lambda_j k_z^i \mathbf{B}] \cdot \boldsymbol{a}_j. \tag{6.8}$$

The only unknown quantities in this equation are the coefficients α_j. If we multiply the equation by $\boldsymbol{a}_{1,2}$ from the left, we obtain a 2×2 system, which is used to extract the coefficients $\alpha_{1,2}$:

$$\begin{cases} 2\boldsymbol{a}_1 \cdot \dot{\boldsymbol{e}}_{\|}^{i} = \alpha_1 \boldsymbol{a}_1 \cdot [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} + \lambda_1 k_z^i \mathbf{B}] \cdot \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_1 \cdot [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} + \lambda_2 k_z^i \mathbf{B}] \cdot \boldsymbol{a}_2 \\ 2\boldsymbol{a}_2 \cdot \dot{\boldsymbol{e}}_{\|}^{i} = \alpha_1 \boldsymbol{a}_2 \cdot [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} + \lambda_1 k_z^i \mathbf{B}] \cdot \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 \cdot [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} + \lambda_2 k_z^i \mathbf{B}] \cdot \boldsymbol{a}_2. \end{cases} \tag{6.9}$$

This system is always solvable provided the following determinant is non-zero:

$$\begin{aligned}\Delta =&(\boldsymbol{a}_1\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_1 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_1)(\boldsymbol{a}_2\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_2 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_2)\\ &-(\boldsymbol{a}_2\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_1 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_1)(\boldsymbol{a}_1\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_2 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_2)\\ =&(\boldsymbol{a}_1\cdot\boldsymbol{v}_1)(\boldsymbol{a}_2\cdot\boldsymbol{v}_2)-(\boldsymbol{a}_2\cdot\boldsymbol{v}_1)(\boldsymbol{a}_1\cdot\boldsymbol{v}_2)\\ =&\boldsymbol{a}_1\cdot(\boldsymbol{v}_1\boldsymbol{v}_2-\boldsymbol{v}_2\boldsymbol{v}_1)\cdot\boldsymbol{a}_2,\end{aligned}$$

where we have introduced the vectors $\boldsymbol{v}_{1,2}=[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_{1,2}k_z^i\mathbf{B}]\cdot\boldsymbol{a}_{1,2}=\mathbf{R}_{1,2}\cdot\boldsymbol{a}_{1,2}$. The operators $\mathbf{R}_{1,2}$ are obviously positive semi-definite, where the semi-definiteness comes from the projection $[\mathbf{I}-\hat{z}\hat{z}]$. It is conjectured that these properties imply $\Delta>0$.

Using the explicit inverse of a 2×2-matrix, we can write the solution to (6.9) as

$$\begin{cases}\alpha_1=\dfrac{2}{\Delta}\Big\{(\boldsymbol{a}_2\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_2 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_2)(\boldsymbol{a}_1\cdot\dot{\boldsymbol{e}}_\parallel^i)\\ \qquad -(\boldsymbol{a}_1\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_2 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_2)(\boldsymbol{a}_2\cdot\dot{\boldsymbol{e}}_\parallel^i)\Big\}\\ \alpha_2=\dfrac{2}{\Delta}\Big\{(\boldsymbol{a}_1\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_1 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_1)(\boldsymbol{a}_2\cdot\dot{\boldsymbol{e}}_\parallel^i)\\ \qquad -(\boldsymbol{a}_2\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_1 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_1)(\boldsymbol{a}_1\cdot\dot{\boldsymbol{e}}_\parallel^i)\Big\}.\end{cases}\tag{6.10}$$

This can be written as $\alpha_{1,2}=\frac{2}{\Delta}\boldsymbol{b}_{1,2}\cdot\dot{\boldsymbol{e}}_\parallel^i$ by introducing the vectors

$$\begin{cases}\boldsymbol{b}_1=(\boldsymbol{a}_2\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_2 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_2)\boldsymbol{a}_1-(\boldsymbol{a}_1\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_2 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_2)\boldsymbol{a}_2\\ \boldsymbol{b}_2=(\boldsymbol{a}_1\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_1 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_1)\boldsymbol{a}_2-(\boldsymbol{a}_2\cdot[\mathbf{I}-\hat{z}\hat{z}]\cdot[\mathbf{I}+\lambda_1 k_z^i\mathbf{B}]\cdot\boldsymbol{a}_1)\boldsymbol{a}_1.\end{cases}\tag{6.11}$$

The map between $\boldsymbol{a}_{1,2}$ and $\boldsymbol{b}_{1,2}$ has the same determinant as the map between the coefficients $\alpha_{1,2}$ and the incident field, *i.e.*, Δ, which was assumed greater than zero previous in this section. This implies that the vectors $\boldsymbol{b}_{1,2}$ are linearly independent. We now formulate the relation $\dot{\boldsymbol{e}}_\parallel^t=\sum_{j=1}^2\alpha_j[\mathbf{I}-\hat{z}\hat{z}]\cdot\boldsymbol{a}_j$ as a dyadic relation between incident and transmitted fields,

$$\begin{aligned}\dot{\boldsymbol{e}}_\parallel^t=&\frac{2}{\Delta}[\mathbf{I}-\hat{z}\hat{z}]\cdot[\boldsymbol{a}_1\boldsymbol{b}_1+\boldsymbol{a}_2\boldsymbol{b}_2]\cdot\dot{\boldsymbol{e}}_\parallel^i\\ =&\mathbf{T}_\parallel\cdot\dot{\boldsymbol{e}}_\parallel^i,\end{aligned}\tag{6.12}$$

where we have introduced the notation $\mathbf{T}_\parallel$ for the transmission operator acting on the tangential fields. Since the transmitted field consists of only the modes $\boldsymbol{a}_{1,2}$, the transmission operator extends to the total transmitted field:

$$\dot{\boldsymbol{e}}^t=\frac{2}{\Delta}[\boldsymbol{a}_1\boldsymbol{b}_1+\boldsymbol{a}_2\boldsymbol{b}_2]\cdot\dot{\boldsymbol{e}}_\parallel^i=\mathbf{T}\cdot\dot{\boldsymbol{e}}_\parallel^i.\tag{6.13}$$

Since the vectors $\boldsymbol{b}_{1,2}$ are linearly independent, they represent the two different polarizations of the incident field which generate the two possible modes $\boldsymbol{a}_{1,2}$ in the nonlinear material.

6.6 Reflection operator and Brewster angles

It is well known that at certain angles and polarizations of the incident field there is no reflected field at all — the Brewster angles [22, 25, 26]. From (6.7) we see that the reflected field can be written

$$2\dot{\boldsymbol{e}}_{\|}^{r} = \sum_{j=1}^{2} \alpha_j [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \lambda_j k_z^i \mathbf{B}] \cdot \boldsymbol{a}_j.$$

Using $\alpha_{1,2} = \frac{2}{\Delta} \boldsymbol{b}_{1,2} \cdot \dot{\boldsymbol{e}}_{\|}^{i}$ we find the following relationship between the reflected and incident field:

$$\begin{aligned} 2\dot{\boldsymbol{e}}_{\|}^{r} =& \frac{2}{\Delta} \Big[\Big\{ [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \lambda_1 k_z^i \mathbf{B}] \cdot \boldsymbol{a}_1 \Big\} \boldsymbol{b}_1 \\ & + \Big\{ [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \lambda_2 k_z^i \mathbf{B}] \cdot \boldsymbol{a}_2 \Big\} \boldsymbol{b}_2 \Big] \cdot \dot{\boldsymbol{e}}_{\|}^{i} \\ =& \frac{2}{\Delta} [\mathbf{I} - \hat{z}\hat{z}] \cdot [\boldsymbol{b}_1' \boldsymbol{b}_1 + \boldsymbol{b}_2' \boldsymbol{b}_2] \cdot \dot{\boldsymbol{e}}_{\|}^{i} \\ =& 2\mathbf{R}_{\|} \cdot \dot{\boldsymbol{e}}_{\|}^{i}. \end{aligned}$$

This is the reflection operator $\mathbf{R}_{\|}$ for the tangential fields, which is represented as a factorization in the simple dyads $\boldsymbol{b}_1' \boldsymbol{b}_1$ and $\boldsymbol{b}_2' \boldsymbol{b}_2$, where $\boldsymbol{b}_{1,2}' = [\mathbf{I} - \lambda_{1,2} k_z^i \mathbf{B}] \cdot \boldsymbol{a}_{1,2}$.

Since the vectors $\boldsymbol{b}_{1,2}$ are linearly independent, we see that the Brewster angles are characterized by

$$\begin{cases} \dot{\boldsymbol{e}}_{\|}^{i} = \beta [\mathbf{I} - \hat{z}\hat{z}] \cdot \boldsymbol{b}_j \\ 0 = [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \lambda_j k_z^i \mathbf{B}] \cdot \boldsymbol{a}_j \end{cases} \quad j = 1, 2, \tag{6.14}$$

where β is a scalar. This means that if the incident field is polarized along $\boldsymbol{b}_j$ and $\boldsymbol{a}_j$ is in the null space of $[\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \lambda_j k_z^i \mathbf{B}]$, there is no reflected field. These conditions determine the possible Brewster angles. We have

$$\begin{aligned} 0 &= [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \lambda_j k_z^i \mathbf{B}] \cdot \boldsymbol{a}_j \\ &= [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \frac{c_j}{c_0} \frac{k_z^i}{k_{zj}^t} (\mathbf{I} + [\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}]^{-1} \cdot [\boldsymbol{\varepsilon} - \mathbf{I}])] \cdot \boldsymbol{a}_j \\ &= [\mathbf{I} - \hat{z}\hat{z}] \cdot [\mathbf{I} - \frac{c_j}{c_0} \frac{k_z^i}{k_{zj}^t} (\mathbf{I} + \frac{1}{(k_z^i)^2} [\mathbf{I} + k_y^i \hat{y} \times \mathbf{J} - (k_y^i)^2 \hat{y}\hat{y}] \cdot [\boldsymbol{\varepsilon} - \mathbf{I}])] \cdot \boldsymbol{a}_j, \end{aligned}$$

where we have introduced the explicit inverse $[\mathbf{I} - k_y^i \hat{y} \times \mathbf{J}]^{-1} = \frac{1}{(k_z^i)^2} [\mathbf{I} + k_y^i \hat{y} \times \mathbf{J} - (k_y^i)^2 \hat{y}\hat{y}]$, which can be verified by straightforward calculations. The y-component of this equation is

$$\begin{aligned} 0 &= \hat{y} \cdot \boldsymbol{a}_j - \frac{c_j}{c_0} \frac{k_z^i}{k_{zj}^t} (\hat{y} + \frac{1}{(k_z^i)^2} [\hat{y} - (k_y^i)^2 \hat{y}] \cdot [\boldsymbol{\varepsilon} - \mathbf{I}]) \cdot \boldsymbol{a}_j \\ &= \hat{y} \cdot \boldsymbol{a}_j - \frac{c_j}{c_0} \frac{k_z^i}{k_{zj}^t} (\hat{y} + \hat{y} \cdot [\boldsymbol{\varepsilon} - \mathbf{I}]) \cdot \boldsymbol{a}_j \\ &= \hat{y} \cdot \boldsymbol{a}_j - \frac{c_j}{c_0} \frac{k_z^i}{k_{zj}^t} \hat{y} \cdot \boldsymbol{\varepsilon} \cdot \boldsymbol{a}_j. \end{aligned}$$

In Section 4 it was shown that a propagating field in an isotropic material is described by $\boldsymbol{a}_j = (\frac{1}{\sqrt{\epsilon}}\boldsymbol{v}_j, \frac{1}{\sqrt{\mu}}\hat{k}\times\boldsymbol{v}_j)$, where the three-vector $\boldsymbol{v}_j$ is orthogonal to $\hat{k}$, and the only possible wave speed is $\frac{c_j}{c_0} = \frac{1}{\sqrt{\epsilon\mu}}$. In the remainder of this section, we suppress the index j, and separate the two modes in the end. The Brewster angles can now be found from the y-component defined above. By explicitly considering both electric and magnetic fields we have

$$\hat{y}\cdot\begin{pmatrix}\frac{1}{\sqrt{\epsilon}}\boldsymbol{v}\\ \frac{1}{\sqrt{\mu}}\hat{k}^t\times\boldsymbol{v}\end{pmatrix} = \frac{1}{\sqrt{\epsilon\mu}}\frac{k_z^i}{k_z^t}\hat{y}\cdot\begin{pmatrix}\sqrt{\epsilon}\boldsymbol{v}\\ \sqrt{\mu}\hat{k}^t\times\boldsymbol{v}\end{pmatrix}$$

$$\begin{pmatrix}\frac{1}{\sqrt{\epsilon}}\hat{y}\cdot\boldsymbol{v}\\ \frac{1}{\sqrt{\mu}}\hat{k}^t\cdot(\boldsymbol{v}\times\hat{y})\end{pmatrix} = \frac{k_z^i}{k_z^t}\begin{pmatrix}\frac{1}{\sqrt{\mu}}\hat{y}\cdot\boldsymbol{v}\\ \frac{1}{\sqrt{\epsilon}}\hat{k}^t\cdot(\boldsymbol{v}\times\hat{y})\end{pmatrix}$$

$$\begin{pmatrix}\frac{1}{\sqrt{\epsilon}}\hat{y}\cdot\boldsymbol{v}\\ \frac{1}{\sqrt{\mu}}k_z^t\hat{z}\cdot(\boldsymbol{v}\times\hat{y})\end{pmatrix} = \frac{k_z^i}{k_z^t}\begin{pmatrix}\frac{1}{\sqrt{\mu}}\hat{y}\cdot\boldsymbol{v}\\ \frac{1}{\sqrt{\epsilon}}k_z^t\hat{z}\cdot(\boldsymbol{v}\times\hat{y})\end{pmatrix}.$$

It is now obvious that one of the following sets of conditions have to be satisfied in order to satisfy the Brewster angle criterion.

$$\begin{cases}\hat{y}\cdot\boldsymbol{v} = 0\\ \sqrt{\epsilon}k_z^t = \sqrt{\mu}k_z^i\end{cases} \quad\text{or}\quad \begin{cases}\hat{z}\cdot(\boldsymbol{v}\times\hat{y}) = 0\\ \sqrt{\mu}k_z^t = \sqrt{\epsilon}k_z^i.\end{cases}$$

Observe that $\hat{z}\cdot(\boldsymbol{v}\times\hat{y}) = 0$ is equivalent to $\hat{x}\cdot\boldsymbol{v} = 0$, *i.e.*, the first set of conditions corresponds to TE-polarization and the second to TM-polarization. Remember that $k_z^i = \cos\theta^i$ and $k_z^t = \cos\theta^t$, where $\theta^{i,t}$ denote the angles of incidence and transmission, respectively, and we have recovered the well known results for linear isotropic materials. Since we in general have $\theta^t < \theta^i$, only one of the above possible Brewster angles is feasible.

An interesting question is whether it always suffices to study the y-component of our original Brewster-angle-condition in (6.14). This is a problem that goes beyond the scope of this paper.

6.7 Algorithm for the direct problem

In this section we summarize the algorithm for solving the direct problem of propagating the incident field through a boundary between vacuum and a nonlinear, nondispersive, homogeneous, bianisotropic halfspace.

We have to calculate the eigenvectors $\boldsymbol{a}_{1,2}$, the eigenvalues $\lambda_{1,2}$ and the operator $\mathbf{B}$ to obtain the reflection and transmission dyadics. These quantities are determined from the relations

$$\begin{cases}\lambda_j\boldsymbol{a}_j = [\boldsymbol{\varepsilon} - k_y^i\hat{y}\times\mathbf{J}]^{-1}\cdot[\hat{z}\times\mathbf{J}]\cdot\boldsymbol{a}_j\\ \mathbf{B} = [\mathbf{I} - k_y^i\hat{y}\times\mathbf{J}]^{-1}\cdot[\boldsymbol{\varepsilon} - k_y^i\hat{y}\times\mathbf{J}],\end{cases}$$

i.e., we have to solve an eigenvalue problem (first row), extract the eigenvectors corresponding to positive eigenvalues, and calculate the operator $\mathbf{B}$. These calculations are evaluated at the transmitted field values at a specific time. The operators are supposed to act on time derivatives of the fields. We discretize the problem with central differences in time, and use the previously calculated values for the transmitted fields in the solution of the eigenvalue problem.

Once we have calculated the tangential fields, it is an easy task to obtain the normal components of the fields. For the transmitted fields these are already given by the transmission operator, see (6.13), and for the reflected field they are given by the relation $\hat{k}^r \cdot \dot{\boldsymbol{e}}^r = 0$, which implies $\dot{\boldsymbol{e}}^r_z = -\frac{k^r_y}{k^r_z}\dot{\boldsymbol{e}}^r_y$.

The algorithm can be summarized as follows, where the indices denote at which time level the different quantities are to be evaluated.

$$
\begin{aligned}
&(\text{eigenvalue problem})_j \Rightarrow (\lambda_{1,2})_j, (\boldsymbol{a}_{1,2})_j\\
(\mathbf{B})_j &= \mathbf{B}((\boldsymbol{e}^t)_j)\\
(\mathbf{T})_j &= \mathbf{T}((\lambda_{1,2})_j, (\boldsymbol{a}_{1,2})_j, (\mathbf{B})_j)\\
(\mathbf{R}_{\|})_j &= \mathbf{R}_{\|}((\lambda_{1,2})_j, (\boldsymbol{a}_{1,2})_j, (\mathbf{B})_j)\\
(\dot{\boldsymbol{e}}^i_{\|})_j &= \frac{(\boldsymbol{e}^i_{\|})_{j+1} - (\boldsymbol{e}^i_{\|})_{j-1}}{2\Delta t}\\
(\boldsymbol{e}^t)_{j+1} &= (\boldsymbol{e}^t)_{j-1} + 2\Delta t (\mathbf{T})_j \cdot (\dot{\boldsymbol{e}}^i_{\|})_j\\
(\boldsymbol{e}^r_{\|})_{j+1} &= (\boldsymbol{e}^r_{\|})_{j-1} + 2\Delta t (\mathbf{R}_{\|})_j \cdot (\dot{\boldsymbol{e}}^i_{\|})_j\\
(\boldsymbol{e}^r_z)_{j+1} &= -\frac{k^r_y}{k^r_z}(\boldsymbol{e}^r_y)_{j+1}
\end{aligned}
$$

6.8 Numerical example

The algorithm in the previous section has been implemented for a nonlinear, anisotropic material, and the result is depicted in Figure 2. We have scaled the fields to obtain dimensionless field strengths and substantial nonlinearities for field strengths of a few units, see *e.g.*, [23, 30]. The constitutive relation is characterized by the dyadic $\boldsymbol{\varepsilon}$, which is represented in the xyz-coordinate system as

$$
\boldsymbol{\varepsilon} = \begin{bmatrix}
2+E^2 & 0 & 0 & 0 & 0 & 0\\
0 & 3+E^2 & 0 & 0 & 0 & 0\\
0 & 0 & 4+E^2 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.
$$

Thus, the material is non-magnetic, anisotropic with principal axis in the xyz-directions, and has a nonlinear permittivity depending on the square of the electric field strength. The angle of incidence is 70°, and the incident field has the magnetic field perpendicular to the plane of incidence, *i.e.*, in the x-direction,

$$
\boldsymbol{e}^i(\boldsymbol{r}, t) = f(t - \hat{k}^i \cdot \boldsymbol{r}/c_0)\begin{pmatrix} -\hat{k}^i \times \hat{x} \\ \hat{x} \end{pmatrix}, \quad f(t) = e_0\sqrt{|t|}.
$$

The time dependence of the amplitude of the incident field is chosen so that its square, which is proportional to the field energy in vacuum, depends linearly on time.

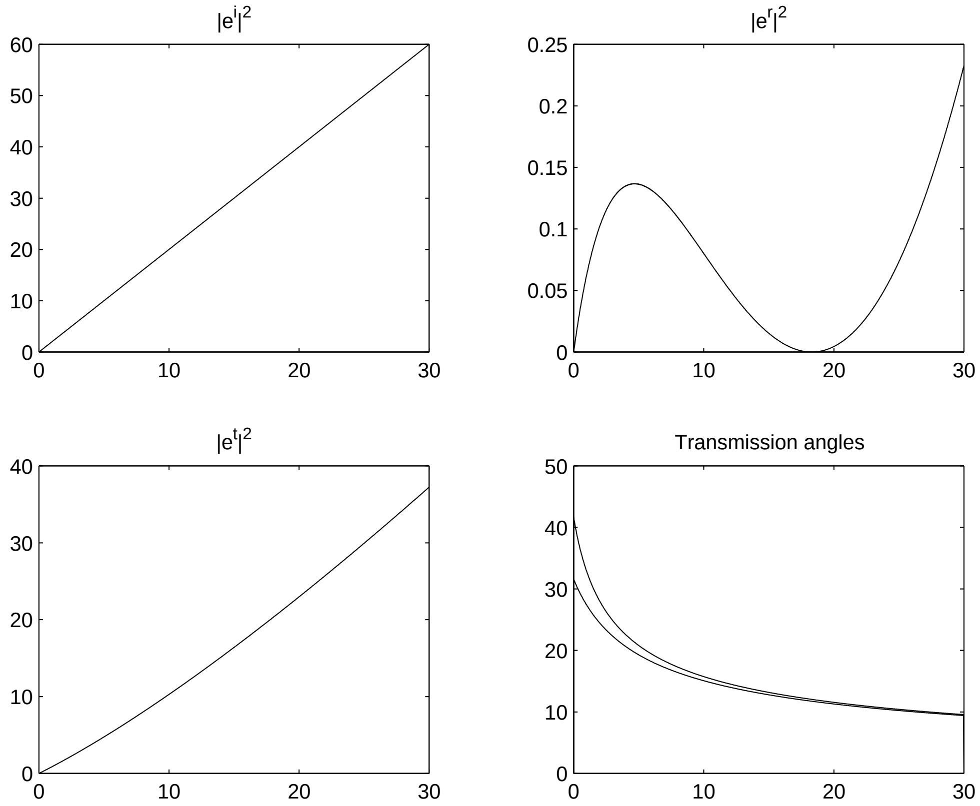

Figure 2: Oblique incidence on an anisotropic Kerr material. Observe that the horizontal scales can be used both as time and energy of the incident field. The diagrams show the squares of the incident, reflected and transmitted fields, and the two possible transmission angles.

This implies that the horizontal scales in Figure 2 can be used both as time and energy. We see that the reflected field displays a strong dependence on the incident field energy, whereas the transmitted field has a more moderate dependence.

It is clearly seen that the Brewster angle occurs when the incident energy is approximately 18. Had the principle axis of the material not been in the xyz-directions, we would have needed another polarization of the incident field to obtain a reflected field that is zero.

The possible transmission angles start off as clearly separated, as can be expected for an anisotropic material, but become more and more equal as the incident energy increases. This can be interpreted from the material dyadic: when the electric field strength grows, the diagonal elements become essentially E^2. Thus the material becomes more and more isotropic, *i.e.*, it has only one possible angle of transmission. Observe that due to our choice of polarization of the incident field, only one of the modes is excited.

7 Conclusions

In this paper, we have introduced the concept of simple waves, as a means to analyze wave propagation problems in nonlinear materials with instantaneous response. We have applied the method to the problem of oblique incidence of a plane electromagnetic wave on a nonlinear material, and found that the direct problem can be solved for all materials and all possible polarizations of the incident wave.

The drawback of the simple wave solutions, is that they do not apply to materials with dispersion, *i.e.*, materials with memory. Our mathematical model with instantaneous nonlinearity, predicts that *all* reasonable waves eventually turn into shocks. It is often argued that the presence of linear dispersion eliminates these shocks, see *e.g.*, [1, pp. 117–120]. Therefore, we can expect our model to be accurate only when there is no shock-like behaviour and the dispersion effects are small, *i.e.*, for sufficiently smooth and slowly varying pulses. It is possible to calculate what propagation distances are necessary for the shock to develop, which means we can estimate the region of validity for our model.

The methods presented in this paper may be useful to propagate the wave front when studying wave propagation in more advanced materials. Temporal dispersion and inhomogeneous media may appear as lower order terms in the Maxwell equations, and can be treated as sources to the fields treated here.

8 Acknowledgement

The author wishes to express his sincere gratitude to Professor Gerhard Kristensson for invaluable support and encouragement during the work presented in this paper. The author's colleagues at the Department of Electromagnetic Theory are also acknowledged for valuable discussions and suggestions.

The work reported in this paper is partially supported by a grant from the Swedish Research Council for Engineering Sciences, and its support is gratefully acknowledged.

References

[1] G. P. Agrawal. *Nonlinear fiber optics.* Academic Press, San Diego, 1995.

[2] F. Bloom. *Mathematical problems of classical nonlinear electromagnetic theory.* Longman Scientific & Technical, Burnt Mill, Harlow, England, 1993.

[3] R. W. Boyd. *Nonlinear Optics.* Academic Press, San Diego, 1992.

[4] L. Brillouin. *Wave propagation and group velocity.* Academic Press, New York, 1960.

[5] L. J. F. Broer. Wave propagation in non-linear media. *Z. angew. Math. Phys.*, **16**, 18–26, 1965.

[6] L. J. F. Broer and P. H. A. Sarluy. On simple waves in non-linear dielectric media. *Physica*, **30**, 1421–1432, 1964.

[7] A. Chelkowski. *Dielectric Physics.* Elsevier Science Publishers, Amsterdam, 1980.

[8] B. D. Coleman and E. H. Dill. Thermodynamic restrictions on the constitutive equations of electromagnetic theory. *Z. angew. Math. Phys.*, **22**, 691–702, 1971.

[9] A. C. Eringen and G. A. Maugin. *Electrodynamics of Continua I.* Springer-Verlag, New York, 1990.

[10] A. C. Eringen and G. A. Maugin. *Electrodynamics of Continua II.* Springer-Verlag, New York, 1990.

[11] H. Goldstein. *Classical Mechanics.* Addison-Wesley, Reading, MA, USA, second edition, 1980.

[12] M. Gustafsson. Time domain theory of the macroscopic maxwell equations. Technical Report LUTEDX/(TEAT-7062)/1–24/(1997), Lund Institute of Technology, Department of Electromagnetic Theory, P.O. Box 118, S-211 00 Lund, Sweden, 1997.

[13] L. Hörmander. *Lectures on Nonlinear Hyperbolic Differential Equations.* Number 26 in Mathémathiques & Applications. Springer-Verlag, Berlin, 1997.

[14] J. D. Jackson. *Classical Electrodynamics.* John Wiley & Sons, New York, second edition, 1975.

[15] A. Jeffrey. Wave propagation and electromagnetic shock wave formation in transmission lines. *J. Math. and Mech.*, **15**(1), 1–13, 1966.

[16] A. Jeffrey. *Quasilinear hyperbolic systems and waves.* Research Notes in Mathematics. Pitman, London, 1976.

[17] A. Jeffrey. Non-dispersive wave propagation in nonlinear dielectrics. *Z. angew. Math. Phys.*, **19**, 741–745, 1968.

[18] A. Jeffrey and V. P. Korobeinikov. Formation and decay of electromagnetic shock waves. *Z. angew. Math. Phys.*, **20**, 440–447, 1969.

[19] F. John. *Partial Differential Equations.* Springer-Verlag, New York, fourth edition, 1982.

[20] A. Karlsson and G. Kristensson. Constitutive relations, dissipation and reciprocity for the Maxwell equations in the time domain. *J. Electro. Waves Applic.*, **6**(5/6), 537–551, 1992.

[21] J. Y. Kazakia and R. Venkataraman. Propagation of electromagnetic waves in a nonlinear dielectric slab. *Z. angew. Math. Phys.*, **26**, 61–76, 1975.

[22] J. A. Kong. *Electromagnetic Wave Theory.* John Wiley & Sons, New York, 1986.

[23] G. Kristensson and D. J. N. Wall. Direct and inverse scattering for transient electromagnetic waves in nonlinear media. *Inverse Problems*, **14**, 113–137, 1998.

[24] I. V. Lindell, A. H. Sihvola, and K. Suchy. Six-vector formalism in electromagnetics of bi-anisotropic media. *J. Electro. Waves Applic.*, **9**(7/8), 887–903, 1995.

[25] I. V. Lindell, A. H. Sihvola, S. A. Tretyakov, and A. J. Viitanen. *Electromagnetic Waves in Chiral and Bi-isotropic Media.* Artech House, Boston, London, 1994.

[26] I. V. Lindell, A. H. Sihvola, and A. J. Viitanen. Plane-wave reflection from a bi-isotropic (nonreciprocal chiral) interface. *Microwave Opt. Techn. Lett.*, **5**(2), 79–81, 1992.

[27] J. D. Logan. *An Introduction to Nonlinear Partial Differential Equations.* John Wiley & Sons, New York, 1994.

[28] R. D. Richtmyer. *Principles of Advanced Mathematical Physics*, volume 1. Springer-Verlag, New York, 1978.

[29] C. Rogers, H. M. Cekirge, and A. Askar. Electromagnetic wave propagation in non-linear dielectric media. *Acta Mech.*, **26**, 59–73, 1977.

[30] D. Sjöberg. Reconstruction of nonlinear material properties for homogeneous, isotropic slabs using electromagnetic waves. Technical Report LUTEDX/(TEAT-7070)/1–16/(1998), Lund Institute of Technology, Department of Electromagnetic Theory, P.O. Box 118, S-211 00 Lund, Sweden, 1998.

[31] G. B. Whitham. *Linear and nonlinear waves.* John Wiley & Sons, New York, 1974.