



LUND UNIVERSITY

High-frequency switching and Kerr effect - nonlinear problems solved with nonstationary time domain techniques

Åberg, Ingegerd

1996

[Link to publication](#)

Citation for published version (APA):

Åberg, I. (1996). *High-frequency switching and Kerr effect - nonlinear problems solved with nonstationary time domain techniques*. (Technical Report LUTEDX/(TEAT-7052)/1-39/(1996); Vol. TEAT-7052). [Publisher information missing].

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

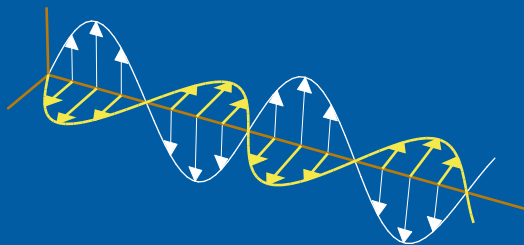
LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

High-frequency switching and Kerr effect—nonlinear problems solved with nonstationary time domain techniques

Ingegerd Åberg

Department of Electrosience
Electromagnetic Theory
Lund Institute of Technology
Sweden



Ingegerd Åberg

Department of Electromagnetic Theory

Lund Institute of Technology

P.O. Box 118

SE-221 00 Lund

Sweden

Editor: Gerhard Kristensson

© Ingegerd Åberg, Lund, December 10, 1996

Abstract

A time domain method building on the concept of wave splitting is used to study direct wave propagation phenomena in weakly nonlinear media. The starting point is the linear wave equation with time-dependent coefficients. This means that the studied nonlinear medium in some sense has to be approximated with a nonstationary medium which changes while the wave passes through. For the nonstationary equation homogeneous as well as particular solutions can be obtained. Two different iterative procedures to find the nonlinear solutions are discussed. They are illustrated by two problems fetched from different research fields of current interest. In the first case, the nonlinear term is linearized using the Fréchet derivative. This leads into a truly nonstationary, mixed initial boundary value problem with a linear equation characterized by both time-dependent coefficients and source terms. In this example a semiconductor device used for switching in high-frequency applications is considered. It can be described as a coplanar waveguide loaded with distributed resonant tunnel diodes. In the other example, wave propagation in Kerr media is considered. Then Taylor expansion transforms the nonlinear equation into a linear one with nonstationary source terms. In this case the nonlinearity does not lead to time-depending coefficients in the equation. The way to obtain the solution is a nonlinear variant of the Born approximation.

1 Introduction

The interest in nonlinear wave phenomena arose after an incident in August 1834. The Scottish engineer John Scott Russell then observed a very special water wave, a solitary wave, rolling forwards for at least a couple of miles with great velocity and without change of shape in the Union Canal between Edinburgh and Glasgow, see Ref. 22. Some 50 years later, the event was explained by the mathematicians Korteweg and de Vries, who derived a model equation describing the unidirectional propagation of long waves in water of relatively shallow depth. The KdV equation has a localized solution, a solitary wave. This was the “hump” discovered by Scott Russell.

It took more than a century after this initial event, until the significance bestowed upon solitary waves in modern technology was anticipated. The interest in solitons prevailed among mathematicians, other people regarded them an unimportant mathematical curiosity. Other important equations derived in those early days, exhibiting soliton-type solutions are the Boussinesq equation, the Sine-Gordon equation, the Born-Infeld equation and the nonlinear Schrödinger equation [21, 22, 26]. In 1955, the solitary waves were rediscovered among physicists and engineers. That year E. Fermi, J. Pasta and S. Ulam made a discovery during early computer simulations of a heat transfer problem [22]. It took some time before they realized, that the phenomenon they had run into was a solitary wave. By the way, here, as in many other papers, the words soliton and solitary wave will be used as synonyms. To some authors [22], solitons are exact traveling solutions of wave equations which have a very stable shape and speed that do not alter by collisions with other solitons. A

solitary wave, on the other hand, is the soliton-like localized wave, which propagates in nature at long distances in one direction only, without being deformed.

During the last 20 years research regarding nonlinear wave phenomena has attracted great interest. Scientists have given high priority to investments in this field, urged by the rapid development and high demands within modern telecommunications. There has been an exceptional growth in the number of users, asking for increased accessibility and high fidelity of the transmitted signal. Today digital signals are of increased importance. They have developed from the first telegraphy pulses invented by S. Morse in 1838 [19]. However, the problem of intersymbol interference, discovered in 1843 when the American Congress granted money for a telegraph line between Washington and Baltimore, has not been completely resolved. Dispersion made the pulses reshape and interfere under propagation, so upon arrival in Baltimore the telegraphed messages could not be interpreted correctly. The immediate remedy was to dig up the telegraph cables, which had been buried in the ground. With telegraph lines in the air the effects of dispersion diminished and the velocity of transmission could be increased.

Although the first wireless transmissions, performed by Marconi and Popov independently in 1897, were pulsed, analogous radio transmissions soon had the monopoly. The low frequency signal containing the information was imposed on a high-frequency carrier, first by means of amplitude modulation, then frequency modulation. There was a quick growth in the number of radio stations, and in order to meet with the demands on increase of the capacity, the carrier waves were chosen with higher and higher frequencies.

Following the introduction of optical fibers, there was a breakthrough for nonlinear wave research [6, 10, 11, 17]. The boom started with the introduction of reliable semiconductor lasers. In optical fibers the information is carried by light. Investigations of analogous signals (optical beams) as well as digital signal (pulse propagation) research show, that the compression of the beam (pulse) introduced by the nonlinearity in the medium is counteracted by diffraction and dispersion. This allows solitary beams (solitary pulses) to come about. When carried by solitary waves, information can be transferred long distances without being regenerated. In this paper the focus is on pulse propagation. Traditionally, in the search for solitary waves in optical fibers, the nonlinear Schrödinger equation with the pulse envelop as the dependent variable has played a central part. In this study it is replaced by the hyperbolic time domain wave equation with the total electric field as the dependent variable.

The demand for large amounts of information to be transmitted simultaneously has given rise to TDM (time division multiplexing). The trend is from picosecond to femtosecond pulses. By shortening the duration of each pulse, pulses from many sources (e.g. different telephone calls) can be stored and transmitted together. At the receiver the pulses are assorted and the original messages are recreated. In this perspective it becomes increasingly important that the pulses do not reshape and interfere with neighboring pulses. Although research keeps pushing the frontier forwards, there are many problems that remain to be solved.

The trend towards shorter pulse durations necessitates a similar development of

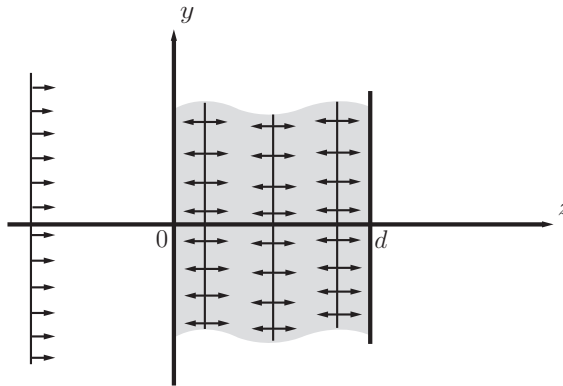


Figure 1: Physical background.

the high-frequency and optoelectronic equipment used for amplification and for measurements and testing of the transmitting devices [23]. At analogous transmissions, when the signal containing the information is modulating a high-frequency carrier, the resulting wave is continuous in time. Those signals have a limited bandwidth, so the testing device only needs to work within a limited frequency range. In contrast, in digital signal transmission where pulses of short duration are used, the frequency content of the signal is infinite. Therefore large bandwidths are required in the measuring apparatus. By distributing nonlinear electronic components, e.g. Schottky diodes, along a transmission line, it has been possible to attain bandwidths around 500 GHz. This means that test pulses can be generated with rise times ~ 1 ps.

2 The method of solution

The full details of the underlying theory can be found in the Refs 2 and 4. A recapitulation of the main characteristics is given below.

A common feature of the studied nonlinear scattering problems is, that the homogeneous hyperbolic equation(s) modeling the problem is solved iteratively, either by using a linearized version, which needs to be currently updated, or by treating a linear equation with the nonlinearity entering as a source function, which has to be reevaluated for each iteration. How these intermediate, iterative equations are found is discussed in Section 3.

The physical background is illustrated in Figure 1. The scattering medium, which is inhomogeneous and nonlinear, is confined to a slab with thickness d . The slab is surrounded by a lossless, linear, homogeneous and time-invariant medium. All fields are assumed quiescent before a fixed time. In this study the fields are evoked by external sources to the left of the slab. They generate uniform plane waves, which impinge on the slab at $z = 0$, $t \geq 0$.

After transforming the original problem into one, which can be solved iteratively, the original, nonlinear medium in every step is replaced by an intermediate linear model medium, that may be inhomogeneous and non-stationary. As the wave

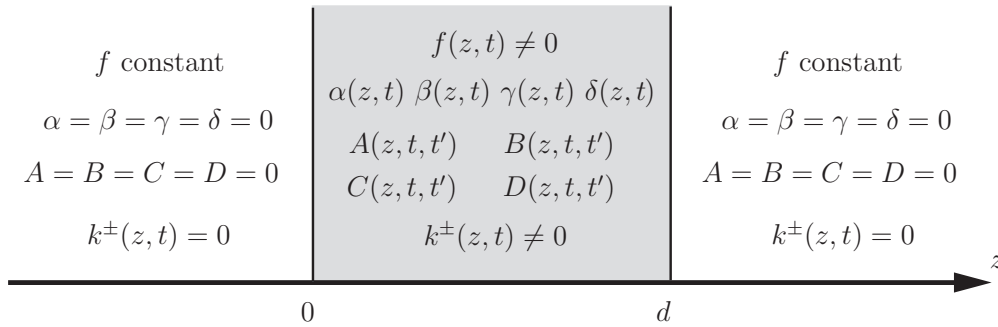


Figure 2: Geometry of the problem.

originating from the external sources penetrates the original nonlinear medium, the nonlinear reaction in every point reached by the wave is to evoke radiating sources, here modeled as planar, distributed sources, transmitting both left- and right-going waves. The direction of propagation of all waves is perpendicular to the slab.

Before starting the iteration, each intermediate equation is subject to a transformation of the dependent variable, i.e. the wave splitting operation. How this is done is discussed in Appendix A. The reason for the operation is to obtain a set of natural dependent coordinates, u^\pm , better adapted to the scattering situation. Outside the slab, in the homogeneous and time-invariant medium, the u^+ -wave travels to the right, and the u^- -wave to the left. Inside the slab, u^\pm cannot be interpreted as physical waves, there they are just mathematical wave components. The result of the wave splitting is a set of first order hyperbolic equations, that can be described by the general form

$$\begin{aligned} \frac{\partial}{\partial z} \begin{pmatrix} u^+(z, t) \\ u^-(z, t) \end{pmatrix} - f(z, t) \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} u^+(z, t) \\ u^-(z, t) \end{pmatrix} - \begin{pmatrix} \alpha(z, t) & \beta(z, t) \\ \gamma(z, t) & \delta(z, t) \end{pmatrix} \begin{pmatrix} u^+(z, t) \\ u^-(z, t) \end{pmatrix} \\ - \int_{-\infty}^t \begin{pmatrix} A(z, t, t') & B(z, t, t') \\ C(z, t, t') & D(z, t, t') \end{pmatrix} \begin{pmatrix} u^+(z, t') \\ u^-(z, t') \end{pmatrix} dt' = \begin{pmatrix} k^+(z, t) \\ k^-(z, t) \end{pmatrix} \end{aligned} \quad (2.1)$$

One immediate conclusion is that this equation models wave propagation in a non-stationary medium. This can be seen from the slowness function f , which in general is a function of both time t and space z . (The numerical treatment only permits the z -dependence so far.) The multiplicative coefficients α , β , γ and δ and the source functions k^\pm depend on z and t as well. The integral kernels A , B , C and D have a double time dependence. The variable t describes the current time, whereas the variable t' is an integral variable, related to the starting time of the excitation.

Some information about the coefficient functions can be found in Figure 2. Also note that the positive function f is a continuous, bounded function of the variables z and t everywhere. It is assumed to be constant outside the slab and continuously differentiable, with bounded derivatives, in z and t everywhere inside the slab. The functions α , β , γ and δ are equal to zero outside the slab, and they are continuous, bounded functions inside the slab (not necessarily continuous at the edges of the slab).

The functions A , B , C and D are always zero outside of the slab $(0, d)$. Due to causality, these functions vanish identically inside the slab provided $t < t'$. For simplicity, the functions A , B , C and D are assumed continuous and bounded as functions of the variables z , t and t' in the region $t > t'$, $0 < z < d$.

The source functions $k^\pm(z, t)$ are integrable functions in the domain $(z, t) \in (0, d) \times (-\infty, \infty)$. Outside the slab, they vanish identically.

The solution of the general problem, characterized by (2.1), with both exterior and interior sources, can be obtained after superposition of waves originating from different kinds of sources. The wave solutions originating from external sources are discussed first. They are the homogeneous solutions of (2.1). Then, the particular solutions of (2.1) are treated. They model waves, generated by the internal sources.

External sources

The underlying set of equations is given by (2.1) with $k^+(z, t) = k^-(z, t) = 0$. Only external sources to the left of the slab, in $z < 0$, are considered. The analysis can easily be extended to include sources in the region $z > d$. The solution of the external source problem can be found either by use of an imbedding argument, or by a Green functions technique. Only the Green functions alternative is pursued here. The details of the imbedding technique can be found in Ref. 4.

The waves, generated by the sources in $z < 0$, impinge on the slab at $z = 0$. In the analysis, they are introduced as the boundary value $u^+(0, t)$. From Duhamel's integral [4, 8], an explicit mapping of the excitation $u^+(0, t)$ to the internal fields $u^\pm(z, t)$ can be obtained. This means, that the boundary waves $u^+(0, t)$ can be propagated to all locations $z > 0$. The mapping is

$$\begin{cases} u^+(z, \tau^+(z; 0, t)) = u^+(0, t)p(0, z, t) + \int_{-\infty}^t G^+(0, z, t, t')p(0, z, t')u^+(0, t') dt' \\ u^-(z, \tau^+(z; 0, t)) = \int_{-\infty}^t G^-(0, z, t, t')p(0, z, t')u^+(0, t') dt' \end{cases} \quad (2.2)$$

The wave front factor p of the u^+ -wave is given by

$$p(0, z, t) = \exp \left\{ \int_0^z \alpha(\zeta, \tau(\zeta; 0, t)) d\zeta \right\}$$

The u^+ -wave front proceeds along the characteristic traces $\tau^+(z; 0, t)$ of the first equation in (2.1). The two propagator kernels $G^\pm(0, z, t, t')$ are also known as the Green functions. A first order 2×2 system of hyperbolic, partial differential equations with G^\pm as the dependent variables, i.e. the Green functions equations, can be derived by performing the calculation of

$$\frac{\partial}{\partial z} \begin{pmatrix} u^+ \\ u^- \end{pmatrix}$$

in two different ways. The first way is obtained through explicit differentiation of the definition of the Green functions in (2.2), and the second way is obtained by

differentiating (2.2) with respect to t , and then inserting the resulting expressions into the general dynamics (2.1). The two alternative expressions are then compared and balanced. The result is the two Green functions equations, together with the initial condition

$$G^-(0, z, t, t')|_{t'=t} = -\frac{1}{2} \frac{\gamma(z, \tau(z; 0, t))}{f(z, \tau(z; 0, t))} \frac{\partial \tau}{\partial t}(z; 0, t)$$

and the boundary conditions

$$\begin{cases} G^+(0, 0, t, t') = 0 \\ G^-(0, d, t, t') = 0 \end{cases}$$

The Green functions equations are complicated coupled integro-differential first order equations, see Refs 2,4. The two Green functions $G^\pm(0, z, t, t')$ can be calculated numerically from an algorithm, based upon these propagator equations.

Internal sources

In this part of the analysis, the relevant boundary conditions are

$$\begin{cases} u^+(0, t) = 0 \\ u^-(d, t) = 0 \end{cases}$$

This means that there are no external sources. Therefore, only the particular solutions of (2.1) are of interest. They model the waves, generated by the internal sources.

In Ref. 2, the solutions of (2.1) are derived in two similar steps. Only one of them is presented here. The emphasis is then put on waves generated by the source function $k^+(z, t)$, so $k^-(z, t) = 0$. The solutions of (2.1) are given by

$$\begin{cases} u_1^+(z, t) = \int_0^z \left\{ \int_{-\infty}^{\tau^+(z_0; z, t)} E_1^+(z, t; z_0, t_0) k^+(z_0, t_0) dt_0 \right\} dz_0 \\ \quad + \int_z^d \left\{ \int_{-\infty}^{\tau^-(z_0; z, t)} E_1^+(z, t; z_0, t_0) k^+(z_0, t_0) dt_0 \right\} dz_0 \\ u_1^-(z, t) = \int_0^z \left\{ \int_{-\infty}^{\tau^+(z_0; z, t)} E_1^-(z, t; z_0, t_0) k^+(z_0, t_0) dt_0 \right\} dz_0 \\ \quad + \int_z^d \left\{ \int_{-\infty}^{\tau^-(z_0; z, t)} E_1^-(z, t; z_0, t_0) k^+(z_0, t_0) dt_0 \right\} dz_0 \end{cases} \quad (2.3)$$

Here E_1^\pm are the solutions of the fundamental set of equations, which is obtained if the right side matrix

$$\begin{pmatrix} k^+(z, t) \\ k^-(z, t) \end{pmatrix}$$

is replaced by the point source matrix

$$\begin{pmatrix} \delta(z - z_0)\delta(t - t_0) \\ 0 \end{pmatrix} \quad (2.4)$$

The waves observed to the right of the source point $z = z_0$ have wave fronts propagating along the characteristic traces $t = \tau^+(z; z_0, t_0)$. Vice versa, waves observed in the region $z < z_0$ have wave fronts traveling along the characteristic curves $t = \tau^-(z; z_0, t_0)$ of the second equation in (2.1).

In the second step, waves generated by the source function $k^-(z, t)$ are calculated, so $k^+(z, t) = 0$. This means that in the source point matrix (2.4), the two elements are interchanged. The solutions $u_2^\pm(z, t)$ are obtained from (2.3) if subscript 1 is replaced by 2 and k^+ is replaced by k^- . Finally, the solutions of the complete system (2.1) with both source functions $k^\pm(z, t) \neq 0$ are obtained after superimposing the two sets of solutions $u_1^\pm(z, t)$ and $u_2^\pm(z, t)$, respectively.

The fundamental solutions $E_1^\pm(z, t; z_0, t_0)$ can be understood as the waves generated in the slab by a right-going impulse introduced at the space-time point (z_0, t_0) . Again, the derivation includes two steps. First, the responses $E_1^\pm(z_0 \pm 0, t; z_0, t_0)$ are calculated. Observe that only waves at the points $z = z_0 \pm 0$ are of interest in this part of the calculation. All waves at $z = z_0$ have their origin in the reaction of the medium to the impulse, and these waves, with exception of the initial δ -pulse, can be represented by regular functions $H_1^\pm(z_0, t, t_0)$. Then, in a second step, the fundamental wave expression $E_1^+(z_0 + 0, t; z_0, t_0)$ can be propagated to the right into the region $z > z_0$ using propagator expressions similar to (2.2). Likewise, the propagation of the fundamental wave expressions $E_1^-(z_0 - 0, t; z_0, t_0)$ to the left into $z < z_0$ can be described by expressions with propagator kernels denoted Q^\pm . The wave front factor is p^- . In a similar manner, that has already been discussed for the propagator kernels G^\pm , the kernels Q^\pm can be calculated from a mixed initial boundary value problem of a set of coupled first order integro-differential equations, see Refs 2, 4.

The final expressions of the propagated fundamental solutions $E_1^\pm(z, t; z_0, t_0)$ are for $t \geq \tau^+(z; z_0, t_0)$ and $z_0 \leq z \leq d$

$$\left\{ \begin{array}{l} E_1^+(z, t; z_0, t_0) = p^+(z_0, z, \tau^+(z_0; z, t)) \delta(\tau^+(z_0; z, t) - t_0) \\ \quad + p^+(z_0, z, \tau^+(z_0; z, t)) H_1^+(z_0, \tau^+(z_0; z, t), t_0) \\ \quad + p^+(z_0, z, t_0) G^+(z_0, z, \tau^+(z_0; z, t), t_0) \\ \quad + \int_{t_0}^{\tau^+(z_0; z, t)} G^+(z_0, z, \tau^+(z_0; z, t), t') p^+(z_0, z, t') H_1^+(z_0, t', t_0) dt' \\ E_1^-(z, t; z_0, t_0) = p^-(z_0, z, t_0) G^-(z_0, z, \tau^-(z_0; z, t), t_0) \\ \quad + \int_{t_0}^{\tau^-(z_0; z, t)} G^-(z_0, z, \tau^-(z_0; z, t), t') p^-(z_0, z, t') H_1^-(z_0, t', t_0) dt' \end{array} \right.$$

and for $t \geq \tau^-(z; z_0, t_0)$ and $0 \leq z < z_0$

$$\left\{ \begin{array}{l} E_1^-(z, t; z_0, t_0) = p^-(z_0, z, \tau^-(z_0; z, t))H_1^-(z_0, \tau^-(z_0; z, t), t_0) \\ \quad + \int_{t_0}^{\tau^-(z_0; z, t)} Q^-(z_0, z, \tau^-(z_0; z, t), t')p^-(z_0, z, t')H_1^-(z_0, t', t_0) dt' \\ E_1^+(z, t; z_0, t_0) = \\ \quad \int_{t_0}^{\tau^-(z_0; z, t)} Q^+(z_0, z, \tau^-(z_0; z, t), t')p^-(z_0, z, t')H_1^-(z_0, t', t_0) dt' \end{array} \right.$$

Due to causality, $E_1^\pm(z, t; z_0, t_0) = 0$ for $t < \tau^\pm(z; z_0, t_0)$ in $z_0 \leq z \leq d$, and $E_1^\pm(z, t; z_0, t_0) = 0$ for $t < \tau^-(z; z_0, t_0)$ in $0 \leq z < z_0$.

A few words should be inserted about the derivation of the expressions for $H_1^\pm(z_0, t, t_0)$, i.e. the regular parts of $E_1^\pm(z_0 \pm 0, t; z_0, t_0)$. They can be obtained, if the fundamental system of equations obtained from (2.1), with the k^\pm -function matrix replaced by the δ -function matrix (2.4), is integrated from $z = z_0 - 0$ to $z = z_0 + 0$. This yields two equations in the variable t . These equations can be combined with two other equations, which result if the propagator expressions for G^- and Q^+ are evaluated at $z = z_0 \pm 0$, respectively. The resulting system of equations is

$$\left\{ \begin{array}{l} E_1^+(z_0 + 0, t; z_0, t_0) - E_1^+(z_0 - 0, t; z_0, t_0) = \delta(t - t_0) \\ E_1^-(z_0 + 0, t; z_0, t_0) - E_1^-(z_0 - 0, t; z_0, t_0) = 0 \\ E_1^-(z_0 + 0, t; z_0, t_0) = \int_{t_0}^t G^-(z_0, z_0, t, t')E_1^+(z_0 + 0, t'; z_0, t_0) dt' \\ E_1^+(z_0 - 0, t; z_0, t_0) = \int_{t_0}^t Q^+(z_0, z_0, t, t')E_1^-(z_0 - 0, t'; z_0, t_0) dt' \end{array} \right.$$

From these equations $E_1^\pm(z_0 \mp 0, t; z_0, t_0)$ can be eliminated, and the remaining two equations can be solved. One of them is a Volterra equation of the second kind. The unique solution can be constructed using the concept of resolvent kernels. This is further explained in Ref. 2.

Numerical implementation

The major part of the numerical treatment, i.e. the calculation of the contributions from the internal sources to the u^\pm -waves electric fields, has been treated thoroughly in Ref. 1. In that part, the main task is to evaluate the integrals (2.3). Throughout this investigation, integrals are computed using the trapezoidal method. The calculation of the internal source contribution is illustrated in Figure 3. The input data to the program contains information about the medium properties. The source functions k^\pm always must be updated between successive iterations, and in the type A solution, see Section 3, this is also true about the coefficient functions α , β , γ and δ and the integral kernels A , B , C and D . The calculation and accessibility of the propagator kernels are crucial, not only in the internal source part. The algorithms build on the method of characteristics. The propagator kernels G^\pm and

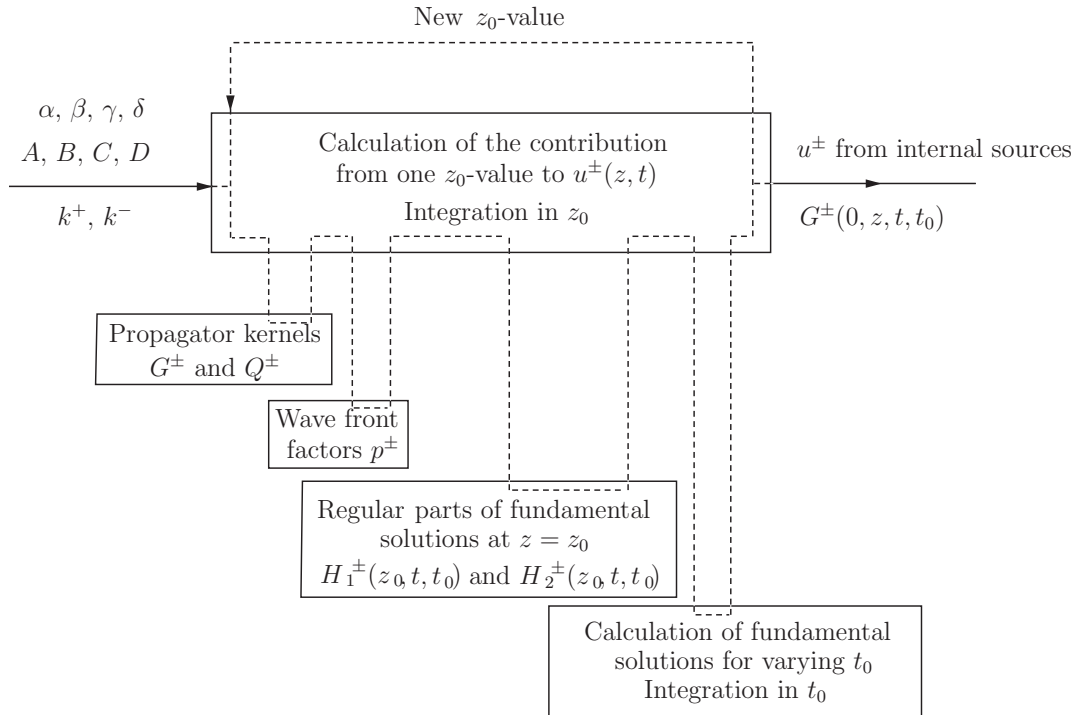


Figure 3: The internal source contributions.

Q^\pm as well as the fundamental solutions E_1^\pm and E_2^\pm are 4-parameter functions, but 4-dimensional data fields require too much computer memory! A way to circumvent the problem is to calculate the contributions to the u^\pm -variables at all field points (z, t) from one source space point z_0 at a time. The start is from $z_0 = d$ and at first 3-dimensional data fields are filled with Q^\pm -kernel data. As the z_0 -value is decreased, the G^\pm -kernel data gradually takes over, and after the $z_0 = 0$ -calculation is finished, the $G^\pm(0, z, t, t')$ -data is ready for use in the second part of the numerical calculation, i.e. the one focused on external source contributions. That part only contains one step, i.e. the evaluation of the integrals (2.2).

When the theory of this section is used to solve nonlinear wave propagation problems, a co-ordinating iterative procedure must be appended. That discussion is saved until Section 3.

The main objective of this presentation is to demonstrate two alternative ways to evaluate wave propagation in weakly nonlinear media. Yet, the hope is that some hidden aspects of the studied physical and technological phenomena are illuminated. One restriction should be recognized. In the present computer code, the scattering medium is confined to a slab of finite length. Therefore, for a given length, the duration and frequency variations of the pulse are limited by the finest possible discretization. The pulse excitation should be sampled with at least three or four mesh points in order to resolve the propagation of the pulse. In the Kerr effect example, a semi-infinite waveguide might have been a better choice. This is within the reach of the method, and a tempting plan for the future.

3 Two nonlinear wave propagation problems

It is clear from Section 2, that only problems modeled by hyperbolic equations are studied in this paper. Other presumptions are that the model equations are quasi-linear, and that the nonlinearities are weak [21, 26].

A quasi-linear equation is one which is linear in the highest order derivatives of the dependent variable, but nonlinear in the variable itself, and (or) the lowest order derivatives. A general, quasi-linear, second order wave equation is given by

$$\frac{\partial^2 U}{\partial z^2} - F\left(U, \frac{\partial U}{\partial z}, \frac{\partial U}{\partial t}, z, t\right) \frac{\partial^2 U}{\partial t^2} - G\left(U, \frac{\partial U}{\partial z}, \frac{\partial U}{\partial t}, z, t\right) = 0 \quad (3.1)$$

The second prerequisite, i.e. that the nonlinearity is weak, implies that in some sense it is possible to expand the nonlinear terms in a Taylor series. Two ways to obtain a solution are suggested. First the two alternatives are introduced in a more general way starting from the second order wave equation (3.1). Then two specific examples are given. As can be seen from the first one, the starting point must not necessarily be the second order equation. The alternative is a set of first order equations. The second example actually starts from the second order wave equation.

Type A solution

If the function $F(U, \frac{\partial U}{\partial z}, \frac{\partial U}{\partial t}, z, t)$ depends only on z , equation (3.1) can be transformed by a travel time coordinate transformation into an equation of the independent normalized wave variables x (space) and s (time) [3, 20]:

$$\frac{\partial^2 u}{\partial x^2}(x, s) - \frac{\partial^2 u}{\partial s^2}(x, s) - g\left(u(x, s), \frac{\partial u}{\partial x}(x, s), \frac{\partial u}{\partial s}(x, s), x, s\right) = 0 \quad (3.2)$$

The nonlinearity of the problem appears in the last term

$$g\left(u(x, s), \frac{\partial u}{\partial x}(x, s), \frac{\partial u}{\partial s}(x, s), x, s\right) \quad (3.3)$$

If the nonlinearity is weak, i.e. the wave field u departs only from a function $u^i(x, s)$ slightly, the nonlinear term can be linearized by use of the Fréchet derivative in a two terms Taylor expansion about the function $u^i(x, s)$, see Ref. 25. The result is

$$\begin{aligned} & \frac{\partial^2 u}{\partial x^2}(x, s) - \frac{\partial^2 u}{\partial s^2}(x, s) - g^i(x, s) - \partial_1 g^i(x, s) (u(x, s) - u^i(x, s)) \\ & - \partial_2 g^i(x, s) \left(\frac{\partial u}{\partial x}(x, s) - \frac{\partial u^i}{\partial x}(x, s) \right) \\ & - \partial_3 g^i(x, s) \left(\frac{\partial u}{\partial s}(x, s) - \frac{\partial u^i}{\partial s}(x, s) \right) = 0 \end{aligned} \quad (3.4)$$

The notation $g^i(x, s)$ simply means that the function (3.3) is evaluated at the center function $u(x, s) = u^i(x, s)$. Similarly, $\partial_k g^i(x, s)$, $k = 1, 2, 3$, are the partial

derivatives of g with respect to u , $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial s}$, respectively, evaluated at $(u^i, \frac{\partial u^i}{\partial x}, \frac{\partial u^i}{\partial s})$. Equation (3.4) can be used iteratively to find a solution to the quasi-linear equation (3.2). The iteration scheme, which is equivalent to the well-known Newton's method is the following [28], $i=0,1,2,\dots$

$$\begin{aligned} & \frac{\partial^2 u^{i+1}}{\partial x^2}(x, s) - \frac{\partial^2 u^{i+1}}{\partial s^2}(x, s) - \partial_1 g^i(x, s) u^{i+1}(x, s) \\ & \quad - \partial_2 g^i(x, s) \frac{\partial u^{i+1}}{\partial x}(x, s) - \partial_3 g^i(x, s) \frac{\partial u^{i+1}}{\partial s}(x, s) \\ & = g^i(x, s) - \partial_1 g^i(x, s) u^i(x, s) \\ & \quad - \partial_2 g^i(x, s) \frac{\partial u^i}{\partial x}(x, s) - \partial_3 g^i(x, s) \frac{\partial u^i}{\partial s}(x, s) \end{aligned} \quad (3.5)$$

Observe, that the coefficient functions $\partial_k g^i(x, s)$, $k = 1, 2, 3$, through the dependence on u^i and its partial derivatives $\frac{\partial u^i}{\partial x}$ and $\frac{\partial u^i}{\partial s}$, depend on both space and time. The solution to this linear nonstationary problem is obtained by the methods of Section 2. The first step is to transform equation (3.5) into the set (2.1) by performing a wave splitting operation, i.e. a transformation of the dependent variable. That way general expressions for the coefficient functions α , β , γ and δ and the integral kernels A , B , C and D , and the source functions k^\pm result. The wave splitting is demonstrated in Appendix A.

Roughly, the iterative scheme proceeds as follows: First, choose $i = 0$ and guess a function $u^0(x, s)$. Then, from (3.5) with $u^i = u^0$, evaluate the coefficient and source functions of (2.1) and solve (2.1) for u^\pm . Note that the solution is a sum of solutions originating from both external and internal sources. Finally, the solution $u^1(x, s)$ of equation (3.5) is obtained as the sum $u^1 = u^+ + u^-$. The next step is to choose $i = 1$ and solve equation (3.5) again. Continue this procedure until a converged solution has been obtained. It can be seen from equation (3.4), that if $u(x, s) = u^i(x, s)$, then $u^i(x, s)$ is actually a solution of the nonlinear equation (3.2), thus not only a solution of a linearized version of (3.2).

One difficult question remains. When does the solution to the iteration scheme (3.5) actually converge to a solution of (3.2)? Regarding the criterion of convergence, the guideline has been to demand a decreasing trend of errors between successive iterations. In addition, the maximum absolute value $|u^{i+1}(x, s) - u^i(x, s)|$ of the last step should be less than 0.005 of the amplitude of the incident wave u_{in} . In one case, this ultimate goal was hard to reach. One reason may be that the chosen model did not properly represent the underlying physical principle. Details of convergence will also be discussed in connection with the specific examples.

Type B solution

In this case, the start may eventually be from equation (3.2) again. Another possibility, illustrated by the Kerr example below, is that the G -function of the third term of (3.1) is linear or vanishes, so the nonlinearity appears only in the F -function of the second term

$$F\left(U, \frac{\partial U}{\partial z}, \frac{\partial U}{\partial t}, z, t\right)$$

Say, that due to the weakness of the nonlinearity, this function by means of a Taylor expansion can be approximated with

$$F_1(z) + F_2\left(U, \frac{\partial U}{\partial z}, \frac{\partial U}{\partial t}, z, t\right)$$

Then, after a travel time coordinate transformation, equation (3.1) can be expressed as

$$\frac{\partial^2 u}{\partial x^2}(x, s) - \frac{\partial^2 u}{\partial s^2}(x, s) - \frac{\partial^2}{\partial s^2} f_2\left(u(x, s), \frac{\partial u}{\partial x}(x, s), \frac{\partial u}{\partial s}(x, s), x, s\right) = 0 \quad (3.6)$$

eventually extended with additional lower order, linear terms. Equation (3.6) can be solved iteratively. The procedure is the nonlinear variation of the Born approximation, see Ref. 24. Equation (3.6) is first rewritten as an iteration scheme, $i=0,1,2, \dots$

$$\frac{\partial^2 u^{i+1}}{\partial x^2}(x, s) - \frac{\partial^2 u^{i+1}}{\partial s^2}(x, s) = \frac{\partial^2}{\partial s^2} f_2\left(u^i(x, s), \frac{\partial u^i}{\partial x}(x, s), \frac{\partial u^i}{\partial s}(x, s), x, s\right) \quad (3.7)$$

The procedure is partially the same as in the type A solution. First, equation (3.7) is cast on the form of equation (2.1) through a wave-splitting operation. Observe, that since the nonlinearity only affects the source term of (3.7), the coefficient functions α , β , γ and δ and the integral kernels A , B , C and D of (2.1) are the same throughout the iterations. The source functions, $k^\pm(x, s)$, on the other hand, in every new iteration depend on the solution of the previous step, and therefore they need be updated. First choose $i = 0$ and an initial guess $u^0(x, s)$. Evaluate the source functions $k^\pm(x, s)$. Thereafter, the solution of (2.1), which is a sum of solutions originating from external and internal sources, is obtained with the methods outlined in Section 2. Again, $u^1(x, s)$ is obtained as $u^1 = u^+ + u^-$. Then choose $i = 1$ and solve equation (3.7) another time. The criterion for convergence is the same as in the type A solution. Experience shows that the convergence is quicker if in some sense the size of the right side of equation (3.7) is small.

3.1 The nonlinear transmission line switching device

Background

The background to the demonstration of the type A solution can be found in Refs 13, 15, 16. The demands for high information transmission rates have brought about a dramatic increase in the operating frequencies of electronic systems. Today, research within ultrafast electronics and optoelectronics explores how far the limits of high-frequency technology can be pushed. Important issues are how to generate and detect transient electrical signals with risetimes of 0.1–5 ps, and sinusoidal signals with frequencies ranging up to several THz.

Many radio and radar signals operate at very high frequencies, $f > 100$ GHz. An advantage is that they are narrowbanded. Therefore, electronic equipment dealing with that kind of signals, only needs to operate within a narrow frequency

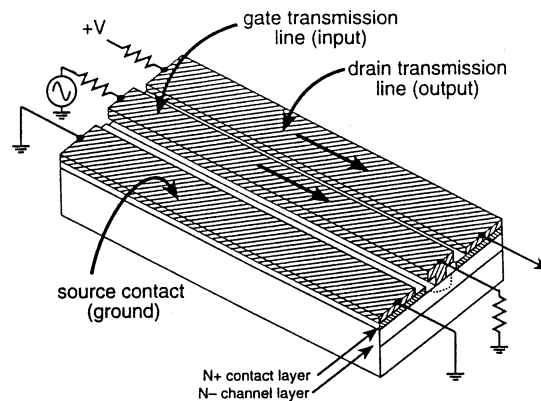


Figure 4: The fully distributed traveling-wave transistor amplifier. The original figure is from Ref. 23.

band. Circuits with a single nonlinear lumped element suffice. One example is the monolithic resonant-matched HEMT-amplifier (high-electron-mobility field-effect-transistor), which provides a 6 dB gain over a 112–115 GHz bandwidth.

In pulsed systems, the operating frequencies are lower, instead they require broad bandwidths. The pulse-code-modulation signal transmitted by a 100 Gb/s fiberoptic transmission link, shows significant power over a ~ 1 MHz to 100 GHz bandwidth. For pulsed signals, distributed circuits permit efficient coupling to semiconductor devices. The monolithic HEMT distributed amplifier, which incorporates a set of lumped transistor elements into a transmission line, is then a better choice than the one transistor amplifier. Other distributed circuits used for signal generation and detection, have Schottky diodes as active elements.

The general information on distributed circuits can be found in a paper by Rodwell et al. [23]. In this section a brief summary of the properties of distributed circuits is given. The concept of distributed circuits can be pursued even further [23]. In the fully distributed semiconductor device, the distribution is applied to the semiconductor device itself. In the fully distributed HEMT-TWA (traveling wave amplifier), one single long transistor is used. The extended source, gate and drain electrodes are used as transmission line conductors. Figure 4 shows the fully distributed amplifier [23]. So far, the lumped element distributed amplifier has proven more efficient than the fully distributed amplifier. The main difficulties are semiconductor losses together with the large disparity in the dimensions required for high-frequency semiconductor devices versus transmission lines. When merged into one common, fully distributed structure, the transmission line and semiconductor device must be of the same size. The result is either a badly transforming transistor, not attaining its maximum gain–bandwidth product, or a transmission line characterized by large losses.

A soliton propagation device, containing distributed Schottky diodes, used for generation of large amplitude electrical impulses of pico-second duration, attracts interest as an eventual application for the present method. However, the nonlinear

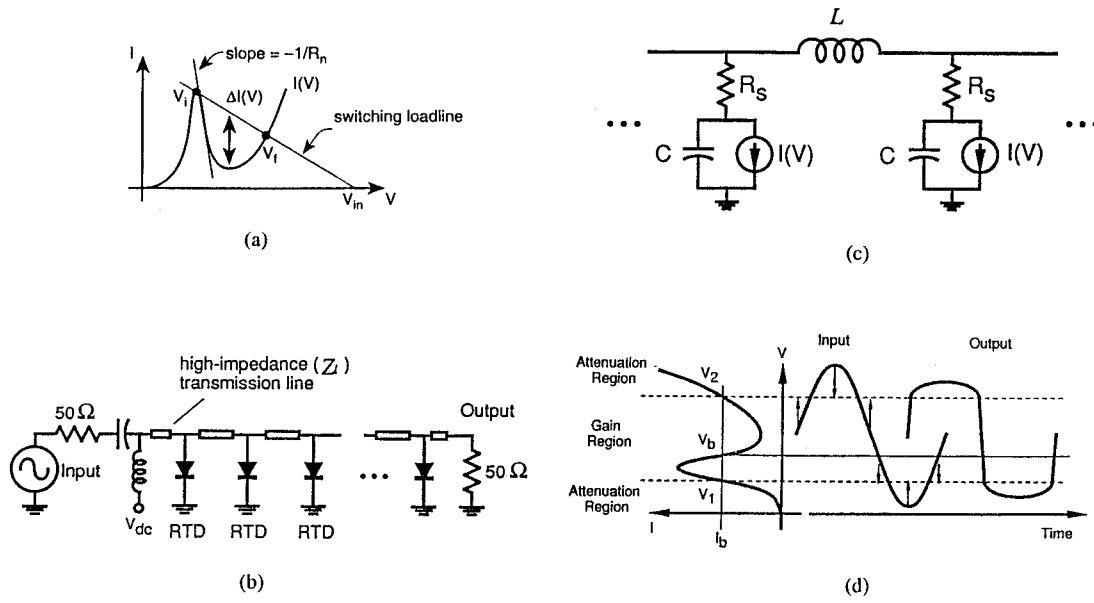


Figure 5: Characteristic features of the lumped RTD distributed pulse generator. The original figures are from Refs 23 and 27.

effect shows as a voltage dependence in the capacitance, leading into a nonlinear wave front velocity. Unfortunately, this nonlinearity is strong. It cannot be expanded in a Taylor series, and therefore it does not lead into the type B solution.

An application more suitable for the present method, the traveling-wave resonant tunnel diode pulse generator (TWRTD), has been fabricated and described by R.Y. Yu et al. [27], and it is also reported in Ref. 23. The ability to produce extremely fast transitions makes this device an attractive candidate for switching and harmonic applications. The TWRTD pulse generator consists of a CPW (coplanar waveguide) transmission line loaded by several lumped GaAs/AlAs resonant tunnel diodes.

For switching circuits containing a single GaAs/AlAs RTD, risetimes of 6 ps have been reported [9]. The bistability arises from the negative resistance of the RTD. Figure 5a shows the I-V curve of the RTD. As the slowly varying input voltage increases, the load line of the device shifts until the current supplied to the RTD exceeds the peak current. The RTD then switches abruptly to the stable state defined by the intersection of the I-V curve and the load line. The rise time is inversely proportional to the magnitude of the negative slope $-1/R_n$.

The circuit diagram of the pulse generator is shown in Figure 5b. It consists of a high impedance transmission line of impedance $Z_l = 75 \Omega$, periodically loaded with 27 RTDs. The RTD spacing is $47 \mu\text{m}$. The approximative equivalent circuit of the RTD is illustrated in Figure 5c. The inductance in each section of the transmission line is $L = 3.13 \cdot 10^{-11} \text{ H}$ and the capacitance $C = 9.63 \cdot 10^{-15} \text{ F}$. Note that C is the sum of the capacitance C_d of the RTD and the capacitance C_l of the section. The impedance of one period of the structure is $Z_0 = 57 \Omega$. Also observe, the

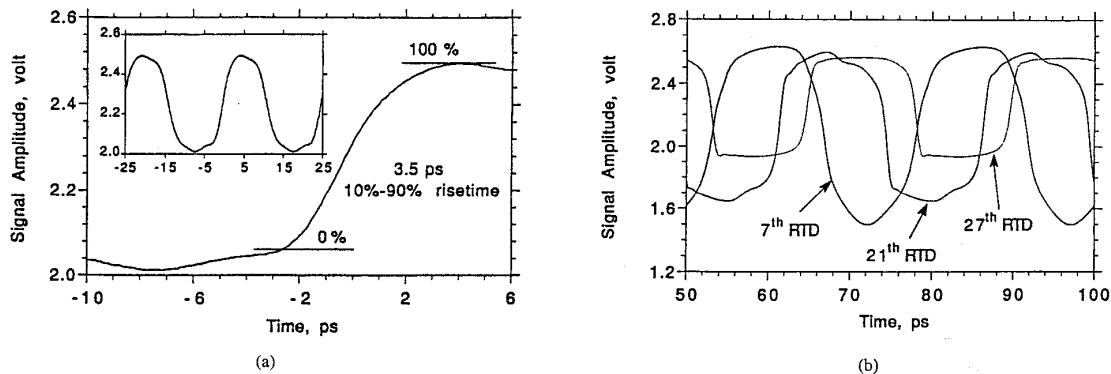


Figure 6: Measured and SPICE simulated waveforms. The original figures are from Ref. 27.

voltage-controlled current source $I(V)$, modeling the I-V curve of the RTD.

Each RTD is reported to have an average negative conductance of -2 mS, the capacitance $C_d = 4.0 \cdot 10^{-15}$ F, and the series resistance $R_s = 12.5 \Omega$ [7]. The series resistance has its main importance in setting the limit for the maximum frequency of oscillations for the single lumped element RTD. Theoretically, in the limit of very small R_s , the cut-off frequency of the device becomes infinite. In practise, however, the charging time of the capacitance sets the limit.

Periodic structures like the one in Figure 5c are undesired low-pass filters with cut-off frequency $\omega_{per} \approx 2/\sqrt{LC}$. In the 27 diode pulse generator the cut-off frequency is $f_{per} = 580$ GHz.

The underlying principle of the RTD switch can be understood from Figure 5d. The I-V curve of the diode is shown to the left. Due to the direction of the current emitted from the current sources when the voltage is close to the input voltage peak values, the capacitors discharge. This leads to a voltage decay. In the low voltage areas on the other hand, the voltage over the capacitors is built up by the current sources. The net influence from the RTDs on the sinusoidal input voltage, is therefore a diminuation and a change into an almost square wave.

With a sinusoidal input of $3 V_{p-p}$ and the frequency 40 GHz, the input signal has a 10% –90% risetime of 7.38 ps over a 2.4 V transition. The output signal which has been measured and reported by Yu et al. [27] is shown in Figure 6a. After deconvolution of the active probe step response, the risetime of the TWRTD pulse generator is 2.5 ps over a 400 mV transition.

One aspect of using active components has not yet been mentioned. In the ideal case, a bias of 2.2 V is applied uniformly along the transmission line. However, there is a 2Ω per section DC resistance causing a bias voltage gradient along the transmission line. With bias voltages applied at both ends of the waveguide, the bias variation is cut in half. In the present computations, two cases are considered. One has a constant bias voltage of 2.2 V, and one has a bias voltage of 2.2 V at the edges and a linear decrease towards a minimum at the center of the device.

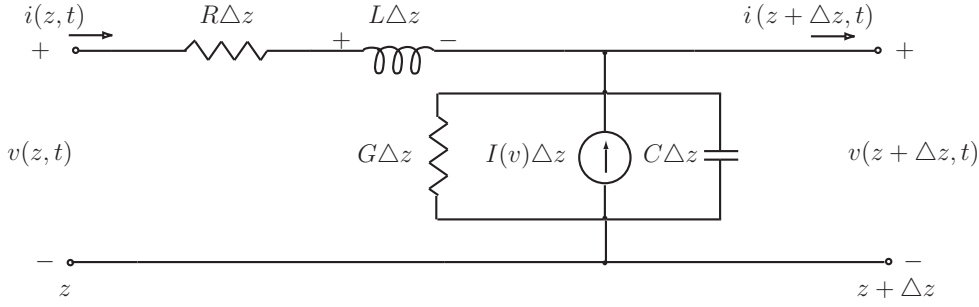


Figure 7: Transmission line model.

Modeling equations

The background of the general transmission line equations and their coupling to the set (2.1) have been thoroughly treated in Ref. 4. Figure 7 shows a model for the transmission line circuit segment Δz . Note, that the direction of the current source is reversed compared to the one in Figure 5c. The corresponding nonlinear set of transmission line equations is

$$\begin{pmatrix} 0 & C \\ L & 0 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} i \\ v \end{pmatrix} + \frac{\partial}{\partial z} \begin{pmatrix} i \\ v \end{pmatrix} + \begin{pmatrix} 0 & G \\ R & 0 \end{pmatrix} \begin{pmatrix} i \\ v \end{pmatrix} = \begin{pmatrix} I(v) \\ 0 \end{pmatrix} \quad (3.8)$$

The suggested model for a type A solution is followed. First, the nonlinear term $I(v)$ is expanded about a center function $v = v^i$

$$I(v) \approx I(v^i) + I'(v^i) (v - v^i)$$

The linearized version of the set (3.8) is given by

$$\begin{pmatrix} 0 & C \\ L & 0 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} i \\ v \end{pmatrix} + \frac{\partial}{\partial z} \begin{pmatrix} i \\ v \end{pmatrix} + \begin{pmatrix} 0 & G - I'(v^i) \\ R & 0 \end{pmatrix} \begin{pmatrix} i \\ v \end{pmatrix} = \begin{pmatrix} I(v^i) - I'(v^i) \cdot v^i \\ 0 \end{pmatrix} \quad (3.9)$$

The second step is to transform the set (3.9) into the dynamics (2.1). This is done through a transformation of the dependent variables, i.e. the wave splitting, demonstrated in Appendix A. In normalized coordinates $x = z/d$ (space) and $s = t/l$ (time), where $l = d\sqrt{LC}$ is the time it takes the wave front to propagate one way through the transmission line, the coefficient functions of (2.1) are

$$\begin{cases} f(x, s) = 1 \\ \alpha(x, s) = -\frac{d}{2}\sqrt{\frac{L}{C}} [G - I'(v^i(x, s))] - \frac{d}{2}\sqrt{\frac{C}{L}}R \\ \beta(x, s) = +\frac{d}{2}\sqrt{\frac{L}{C}} [G - I'(v^i(x, s))] - \frac{d}{2}\sqrt{\frac{C}{L}}R \\ \gamma(x, s) = -\beta(x, s) \\ \delta(x, s) = -\alpha(x, s) \end{cases}$$

and

$$A = B = C = D = 0$$

The source functions are given by

$$k^\pm(x, s) = \frac{d}{2} [I(v^i(x, s)) - I'(v^i(x, s)) \cdot v^i(x, s)]$$

Choice of parameters

In the present calculations, the input signal is a $3 V_{p-p}$ sinusoidal wave of frequency 40 GHz. The length of the transmission line is $d = 1.269$ mm. The approach to the TWRTD pulse generator is here the fully distributed circuit of Figure 4. However, this time the active distributed component is an RTD, not a transistor. From the transmission line model in Figure 7, it is apparent, that the inherent properties R , G , L and C are the resistance, conductance, inductance and capacitance per unit length of the transmission line, respectively. From Yu et al. [27] the DC resistance is known to be 2Ω per section. With high-frequency signals, a rise in resistance can be expected due to skin effect. In this study, the skin effect is disregarded. After a straightforward recalculation of the per section R and L values given by Yu et al., the per unit length values $R = 43 \text{ k}\Omega/\text{m}$ and $L = 0.666 \text{ }\mu\text{H}/\text{m}$ are obtained.

The evaluation of G and C is a bit more complicated. From the pulse generator circuit model of Figure 5c it is seen that a resistance $R_s = 12.5 \Omega$ is connected in series with the capacitance. For a given frequency, the series circuit can be transformed into a parallel circuit. If the fundamental frequency $f = 40$ GHz is chosen, one gets $R_p = 13.7 \text{ k}\Omega$ per section and $C_p = C = 9.63 \cdot 10^{-15}$ F per section. The corresponding per unit length values are $G = 1.556 \text{ S}/\text{m}$ and $C = 0.205 \text{ nF}/\text{m}$, respectively. At the cut-off frequency $f = 580$ GHz, the corresponding values would be $G = 275 \text{ S}/\text{m}$ and $C = 0.171 \text{ nF}/\text{m}$, respectively. The main implication is an increase in the high-frequency contents of the output voltage and a shortening of the transition time compared to a more perfect model. However, as will soon be seen, of even greater importance to the output voltage transition time is the choice of model for the I-V curve. With different I-V models a large spread of transition times are obtained.

The RTD pulse generator considered in the present paper, differs from the original one not only through the transformation of the lumped RTDs into fully distributed RTDs. In Figure 5b the resistance of the signal generator and the load resistance are both 50Ω . With $Z_0 = 57 \Omega$ (thereby disregarding R and G), the aim is an almost perfect match with nearly zero reflection at both transmission line edges. However, in reality, with high microwave losses due to the skin effect, there is mismatch implying an increase in reflection. In the present method, the load at both ends is the impedance of free space $\eta = \sqrt{\mu_0/\epsilon_0} \approx 377 \Omega$.

Computations with the cubic polynomial model

The traveling wave RTD has been studied extensively by Il'ina et al. [14]. They suggest that the I-V curve is approximated with a cubic polynomial in V . The

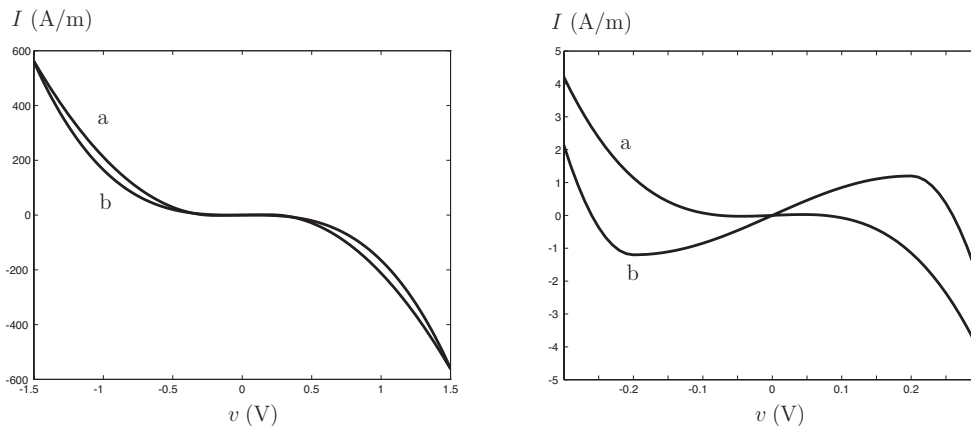


Figure 8: The RTD models. Curve *a* illustrates the cubic polynomial model and curve *b* the sinusoidal parabolic model. The right is a close-up of the details.

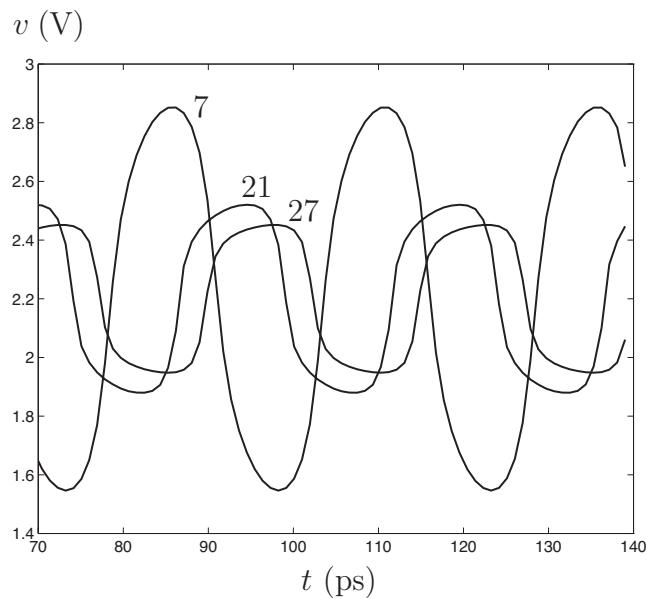


Figure 9: The voltage after 7, 21 and 27 RTDs—cubic polynomial model.

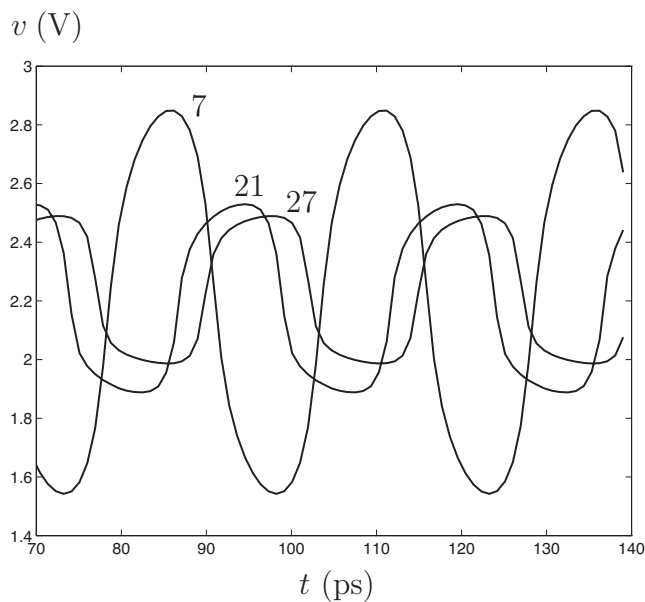


Figure 10: The voltage after 7, 21 and 27 RTDs—cubic model with bias voltage gradient.

curves denoted a in Figure 8 agree with that idea. The curves are given by the cubic polynomial

$$I(v) \approx -166.28 v^3 + 0.94 v$$

Observe, that the current sources $I(v)$ are distributed, so the correct unit is A/m. From the measured waves in Figure 6a, and also from Figure 6b, showing results obtained by a SPICE simulation performed by Yu et al. [27], it can be seen that the attenuation from $3 V_{p-p}$ to $\sim 500 \text{ mV}_{p-p}$ is considerable. Tests show that in order to obtain a large attenuation, the peak cutting domains of Figure 5d must dominate. For the I-V curve, this means that the current values of the peaks are small (in curve a of Figure 8 the peak values are $\pm 0.027 \text{ A/m}$) compared with the current values ($\sim 600 \text{ A/m}$) at the larger voltage magnitudes. Unfortunately, in a cubic polynomial with small current peak values, the magnitude of the derivative $I'(v)$ corresponding to the negative resistance area of the RTD, is always small, which means that the RTD average negative conductance is only $\sim 2\%$ of the -2 mS valid for the real RTD.

Calculations with the cubic polynomial model have been performed with a discretization net of 32×300 meshes in the (x, s) -space. The voltage waves of the Figures 9–10 are representative of the results. They can be compared directly to the measured and SPICE-simulated waves of Figure 6. In Figure 9, the assumption is a constant 2.2 V bias throughout the transmission line. This means that the working point of the RTD in Figure 8 coincides with the point given by $v = 0 \text{ V}$. Compared to the SPICE-simulations, the voltage wave is somewhat larger after the passage of 7 RTDs, then decreasing more rapidly. At the end of the transmission line the output voltage is smaller than the SPICE-simulated wave, but similar to the measured wave in Figure 6a. A 3.7 ps risetime can be measured over a 400 mV transition. This is close to the result in Figure 6a. However, remember that the present result should be compared to the *deconvolved* measured response with a rise time of 2.5 ps !

In Figure 10, the result of a similar calculation is shown. The assumption is now a constant bias negative voltage gradient. The bias decrease is from the edges towards the center of the line. The working point of the RTD then moves from $v = 0 \text{ V}$ in Figure 8 linearly with the distance from the transmission line edge towards the point $v = -0.10 \text{ V}$ at the center of the transmission line. Close examination shows that the voltage curves are not quite symmetric. As the working point moves to the left, the voltage minima are cut more than the maxima, pushing the average of the voltage upwards. At the end of the transmission line, the voltage average is $\sim 2.23 \text{ V}$, not 2.2 V which is the value of the bias. The result is in accordance with the measured wave in Figure 6a.

The reason that the calculations are performed with a relatively coarse mesh size is the wish to show the evolution of the wave during a relatively long period of time, i.e. up to $\sim 140 \text{ ps}$. Limits set by the available computer memory and the time it takes to perform a calculation (a 32×300 points calculation needs 24 hours on a SPARCstation 20) imply that if a calculation with twice as many space points is performed, the wave evolution can only be studied for approximately 45 ps . Such calculations have been performed and compared with the results of the coarser

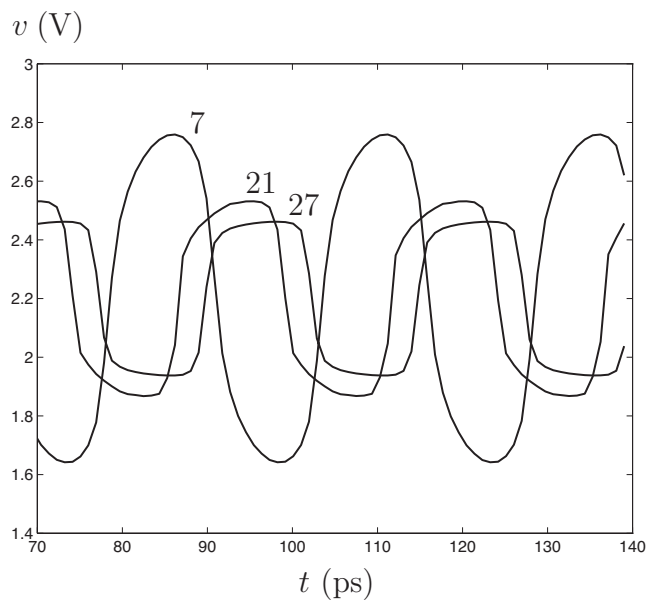


Figure 11: The voltage after 7, 21 and 27 RTDs—sinusoidal-parabolic model.

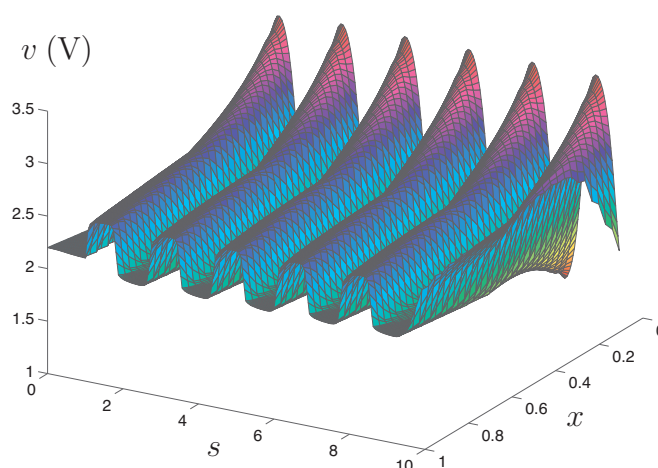


Figure 12: Evolution of the voltage wave in a case with bias voltage gradient.

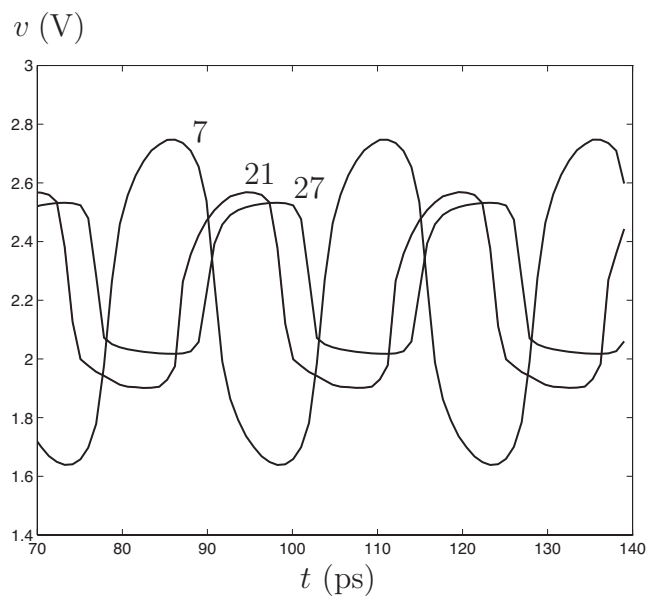


Figure 13: The voltage after 7, 21 and 27 RTDs—Sinusoidal-parabolic model with bias voltage gradient.

calculations. Tests of convergence show that for the cubic polynomial model and the coarser mesh, there is some discrepancy in the front edge at $x = 0$. However, the curves displayed in the Figures 9–10 show almost perfect agreement with the corresponding finer mesh results.

Computations with the sinusoidal parabolic model

As mentioned, the reason for the calculated output voltage not being close enough to the measured response is that the third degree polynomial does not allow a sufficiently steep slope about $v = 0$ V in curve *a* of Figure 8. A better model is suggested by curve *b*. It is given by

$$\begin{cases} I(v) = -\frac{1000}{3}(v - 0.2)^2 + 1.2 \quad (\text{A/m}), & 0.2 \text{ V} \leq v \\ I(v) = 1.2 \sin\left(\frac{\pi v}{0.4}\right) \quad (\text{A/m}), & -0.2 \text{ V} < v < 0.2 \text{ V} \\ I(v) = +\frac{1000}{3}(v + 0.2)^2 - 1.2 \quad (\text{A/m}), & v < -0.2 \text{ V} \end{cases}$$

The advantage of this sinusoidal-parabolic model is that it can easily be manipulated into an arbitrarily chosen slope at $v = 0$ V. The average negative conductance of the RTD modeled by curve *b* is $\sim 20\%$ of the corresponding conductance for the real RTD.

The convergence of 32×300 points calculations with the sinusoidal-parabolic model now stays completely within the given frames. About 10 iterations are needed. Calculations with a finer mesh show a small disagreement at the front edge $x = 0$ compared to the coarser results in the overview of Figure 12, but for all other curves the agreement is good.

Figure 11 illustrates the result of a calculation with constant 2.2 V bias. The change of model is the only difference from the calculation presented in Figure 9. Again, the voltage after a distance comparable to 7 RTD sections is larger than that of the SPICE-simulation (but smaller than in the cubic polynomial model calculation), then decreases quickly. The output voltage compares well with the measured one, demonstrating a 2.7 ps, 400 mV transition.

Figure 12 illustrates the evolution of the same sinusoidal input wave as in Figure 11, but this time the bias voltage has a 0.18 V dip at the transmission line center. It is of interest to mention that the corresponding calculation with the cubic polynomial model is disturbed by overflow. Remember that the space coordinate $x = 1$ corresponds to the end of the line $z = 1.269$ mm, so the view is from the output. The normalized end time $s = 10$ corresponds to ~ 140 ps. The voltage waves of Figure 13 are studied at positions corresponding to 7, 21 and 27 RTD sections. They are details taken from the same calculation. Note that as the working point moves from the center towards the left in Figure 8, the voltage minima are cut more than the maxima, so in spite of the waves being superimposed on a bias less than 2.2 V, the average voltage is pushed upwards towards values greater than 2.2 V. The shape of the wave becomes more asymmetric with flatter minima and smoother

peaks compared to the wave in Figure 11. The risetime suffers from a small increase $\sim +0.15$ ps, but on the other hand there is a similar decrease of the falltime.

In the sinusoidal-parabolic model, by choosing the sinusoidal middle part of curve b in Figure 8 even steeper, thereby bringing the negative conductance of the model closer to the real conductance, the transition time can be made even shorter. In fact, tests hint at a 1.5 ps transition time if the negative conductance of the model approaches the real -2 mS. As the slope of curve b grows, the calculations approach the limit of the numerical method, so the results become more uncertain. Yet, although keeping in mind that the model is not perfect, the eventual conclusion is that, provided the fully distributed TWRTD pulse generator can be manufactured, it may give even faster responses than its lumped element distributed counterpart!

Comparison of type A and B results

Before leaving this first example, the results of an attempt to perform the same calculations with the type B method, should be reported. In attempts where the size of current injected by the current sources into the transmission line is small, $|I(v)| < 15$ A/m, the type B method works well. It converges to approximately the same solution as the type A solution. The number of iterations needed for convergence are then ~ 10 in the type A solution, while only ~ 5 iterations are needed for the type B solution. However, as soon as the range of the injected current is increased further, the type B solution starts to oscillate from one iteration to the other, in most cases soon exceeding the dynamical range of the computer. All calculations illustrated in the Figures 9–13 are impossible to perform with the type B method. The difference between the two methods in this example is seen to be the two extra terms $I'(v^i) \cdot v$ and $I'(v^i) \cdot v^i$, one at each side of the first equation in (3.9). It seems like these terms help stabilizing the solution, leading it into the right direction.

From the modest experience obtained so far, it is impossible to point out which method works the best. The truth is probably that the type A and B solutions complement each other. Sometimes one of the methods works better, sometimes the other. An advantage is that with the present algorithm, solving both time-invariant and nonstationary equations, it is easy to switch between the methods, choosing the one which in a specific case works the best!

3.2 Pulse propagation in Kerr media

Background

One reason for the intense interest in the Kerr effect during the last decades, is due to modeling of nonlinear effects in silicon dioxide SiO_2 and other materials used in optical fibers. The Kerr effect has some negative consequences for the propagating signal, since four different kinds of dispersion, i.e. self phase modulation (SPM), cross phase modulation (XPM), four wave mixing (FWM) and modulational instability, originate from it. However, the great interest follows from the discovery, that during special circumstances (anomalous propagation) the Kerr effect is causing optical

beams to self-focus and pulses to compress, thus counteracting the broadening effects of diffraction and linear dispersion. In this way, if the power is high enough, self-trapped beams and picosecond pulses with the characteristic features of solitary waves are created. The optical Kerr effect causing solitary waves is self-induced. This means that the phase velocity of the wave depends on the wave's own intensity. Other interesting applications for the Kerr effect are electro-optic modulators and all-optical switches. In the electro-optical Kerr effect an external static voltage or low frequency electric field controls the refractive index of the medium. One example is the traditional Kerr cell, where the modulating voltage, applied to plate electrodes, rotates the plane of polarization of an incident low amplitude light wave [12]. The active media is nitrobenzene $C_6H_5NO_2$ or carbon disulfide CS_2 . The Kerr cell can be used as an optical modulator or alternatively a shutter for high speed photography.

The Kerr effect originates from non-harmonic motion of electrons bound in molecules, caused by an external electric field. A typical time scale of the response is 10^{-15} s. In this study, as in most others, the response is presumed instantaneous, which is true if the variation of the electric field is considerably slower than the Kerr effect response time. The Kerr medium is also assumed isotropic, which is true for gases, liquids and various solid media. However, in general this is a simplification, since most Kerr media are anisotropic.

Now consider a nonlinear isotropic medium, initially nonpolarized. Suppose, that a plane electromagnetic wave propagates in the z -direction with the electric field vector

$$\mathbf{E} = E(z, t) \mathbf{e}_y \quad (3.10)$$

and the electric displacement field

$$\mathbf{D} = \epsilon_0 \epsilon(E) E \mathbf{e}_y \quad (3.11)$$

Here \mathbf{e}_y is a unit vector in the y -direction. Note, that the permittivity $\epsilon_0 \epsilon(E)$ depends nonlinearly on the electric field magnitude E . Usually externally applied electric fields are small in comparison with interatomic or crystalline fields. This is true even with focused laser light. In all such cases of weak nonlinearity, it is possible to expand $\epsilon(E)$ in a Taylor's series about $E = 0$ [24]:

$$\epsilon(E) = \epsilon(E)|_{E=0} + \epsilon'(E)|_{E=0} E + \frac{1}{2} \epsilon''(E)|_{E=0} E^2 + \frac{1}{6} \epsilon'''(E)|_{E=0} E^3 + \dots$$

In media like SiO_2 used in optical fibers, there is inversion symmetry at the molecular level [24], which means that the electric displacement \mathbf{D} and the electric field \mathbf{E} change sign simultaneously. The conclusion is that the permittivity $\epsilon(E)$ is an even function, implying that $\epsilon'(E)|_{E=0} = 0$, so the dominant nonlinear term is quadratic in E . Higher order terms are insignificant. Therefore, if the refractive index of the nonmagnetic and nonexcited medium $n_0 = \sqrt{\epsilon(E)|_{E=0}}$ is introduced together with the notation $\frac{1}{2} \epsilon''(E)|_{E=0} = \mathcal{S} n_0^4$, then

$$\mathbf{D} = (\epsilon_0 n_0^2 + \epsilon_0 \mathcal{S} n_0^4 E^2) \mathbf{E} \quad (3.12)$$

Typical values of the Kerr coefficient \mathcal{S} range from 10^{-23} to 10^{-14} m^2/V^2 . The larger values are typical of organic compounds with large molecules, e.g. Chinese

tea. Since electrons bound in large molecules are less mobile, the response time of the nonlinear effect increases with \mathcal{S} . In SiO_2 , measurements show that the value of the Kerr coefficient is $\mathcal{S} = 1.3 \cdot 10^{-22} \text{ m}^2/\text{V}^2$.

Kerr coefficients can be positive or negative. In most investigations, especially those concerning optical fibers, the Kerr coefficients are chosen positive [24]. The reason is that media with positive Kerr coefficients are self-focusing, which is a prerequisite for the appearance of solitary beams. In an optical beam, the intensity is larger in the central part. The nonlinear effect causes a self-induced increase of the refraction index, higher in the center of the beam. The trajectories of the optical rays follow from Fermat's principle. The paths bend towards areas of higher indices of refraction, which causes the self-focusing. Negative Kerr coefficient media, on the other hand, have lower indices of refraction in the central part, which lead into self-defocusing.

In this investigation the Kerr coefficients are necessarily large ($\mathcal{S} = \pm 10^{-14} \text{ m}^2/\text{V}^2$). The reason is the restriction enforced by the finest discretization mesh possible. For pulses of picosecond duration the typical length of the slab is 1 mm, so the time it takes the pulse to propagate from one side to the other is ~ 5 ps. For the nonlinear effect to develop and the pulse to reshape during that short a time, \mathcal{S} must be large. It is conjectured that the same pulse traveling through a slab of larger depth but with a smaller degree of nonlinearity, at the end of the slab would show similar characteristics as the pulse in the present calculation.

In the optical fibers with small Kerr coefficients used today, short pulses of high power, which are hard to generate, are necessary for nonlinear effects like solitary waves to develop. To overcome this drawback, the search for media with high Kerr coefficients and short response times is essential [18]. Potential candidates are expected among polymers or semiconductors.

Modeling equations

Now return to the constitutive relation (3.12). If dispersion is present in the scattering medium, this is expressed as an extra term

$$\epsilon_0 \int_{-\infty}^t \chi_e(t-t') E(z, t') dt' \quad (3.13)$$

added to the right side of (3.12). Note that, although not necessary with the present method, the susceptibility kernel χ_e is assumed homogeneous and time invariant.

The relations (3.10), (3.11) and (3.12) are substituted into Maxwell's equations

$$\begin{cases} \nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}}{\partial t}(\mathbf{r}, t) \\ \nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\partial \mathbf{D}}{\partial t}(\mathbf{r}, t) \end{cases}$$

For a nonmagnetic material a second constitutive relation is

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{H}(\mathbf{r}, t)$$

which is used to eliminate the magnetic induction \mathbf{B} and the magnetic field \mathbf{H} from Maxwell's equations. The resulting wave equation is then transformed by a travel time coordinate transformation. The normalized space coordinate is $x = z/d$ and the normalized time coordinate is $s = t/l$. Note that d is the thickness of the scattering medium, and $l = n_0 d \sqrt{\epsilon_0 \mu_0}$ is the period of time needed for the wave front to travel one way through the medium. In the normalized variables x and s , the nonlinear wave equation is

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2}(x, s) - \frac{\partial^2 u}{\partial s^2}(x, s) - \sqrt{\epsilon_0 \mu_0} \frac{d}{n_0} \frac{\partial^2}{\partial s^2} \int_{-\infty}^s \chi(s - s') u(x, s') ds' \\ - \mathcal{S} n_0^2 \frac{\partial^2}{\partial s^2} u^3 = 0 \end{aligned} \quad (3.14)$$

where

$$u(x, s) = E(xd, sl) \quad \text{and} \quad \chi(s - s') = \chi_e(sl - s'l)$$

respectively. Equation (3.14) has the form of equation (3.6). Therefore, it can be solved by a type B iterative solution, see above. The details of how to transform equation (3.14) into the set (2.1) can be found in Appendix A. Also note that in all calculations of pure Kerr effect without dispersion, the coefficient functions α , β , γ and δ and the integral kernels A , B , C and D of (2.1) vanish.

Observe, that in this time domain approach the magnitude of the total electric field is studied. Theoretically, this permits any expression for the electric field. In most other investigations [6, 11], the electric field has the form

$$E(z, t) = A(z, t) \cos(\omega_0 t - k_0 z) \quad (3.15)$$

The function $A(z, t)$ is the pulse envelop, ω_0 and k_0 are the carrier frequency and wave number, respectively. This ansatz leads to the nonlinear Schrödinger equation with the envelop function $A(z, t)$ as the dependent variable. It can be derived by substituting the Fourier transform $\tilde{E}(z, \omega)$ of (3.15) into the Fourier transformed equation (3.14).

In all Kerr effect calculations performed here, the refractive index of the unexcited medium is $n_0 = 1.5$. For comparison, two different types of Kerr media are discussed, one negative and one positive, with Kerr coefficients ($\mathcal{S} = \mp 10^{-14} \text{ m}^2/\text{V}^2$), respectively. The thickness of the medium is $d = 1 \text{ mm}$. The Figures 14–17 show results for the negative Kerr coefficient medium, and the Figures 18–21 illustrate the positive Kerr coefficient medium. In all these figures, the exciting wave is one half period of a sinusoidal wave of frequency $f = 6 \cdot 10^{11} \text{ Hz}$, impinging perpendicularly on the slab. There are two ways to produce the nonlinear effects, one has already been mentioned, i.e. by choosing a large magnitude Kerr coefficient. The other possibility is to increase the amplitude of the impinging wave. In the negative Kerr coefficient medium, the amplitude of the incoming sinusoidal electric field is $E_0 = 2.0 \cdot 10^6 \text{ V/m}$, and in the positive Kerr coefficient medium it is $E_0 = 1.9 \cdot 10^6 \text{ V/m}$. On the whole, Kerr effect calculations in negative media seem to degenerate at lower input amplitudes than those of positive media. The amplitude $E_0 = 2.0 \cdot 10^6 \text{ V/m}$ is chosen

close to the upper limit. Further increase of the amplitude results in very high peak values, not manageable with the present numerical code.

One main concern is how to model dispersion. The high \mathcal{S} -values imply large molecules, often found in polar liquids, where the Debye model is appropriate. Therefore, the susceptibility kernel is given by

$$\chi(t) = H(t)\chi_0 e^{-t/\tau} \quad (3.16)$$

Here, $H(t)$ is the Heaviside step function, χ_0 is an alignment frequency, and τ is the relaxation time. Unless otherwise stated, the choice of relaxation time in these calculations is $\tau = l = 0.5 \cdot 10^{-11}$ s. In Figure 17 the pulses b are calculated with $\chi_0 = 10.0 \cdot 10^{11}$ s $^{-1}$, and the soliton-like pulses b in Figure 19 have $\chi_0 = 5.0 \cdot 10^{11}$ s $^{-1}$. In the saturated medium studied in Figure 21, soliton-like pulses b are achieved with $\chi_0 = 4.5 \cdot 10^{11}$ s $^{-1}$, and strongly attenuated pulses c have $\chi_0 = 10.0 \cdot 10^{11}$ s $^{-1}$.

In a calculation with dispersion, the coefficient functions of (2.1) are

$$\alpha = \beta = -\gamma = -\delta = -\frac{1}{2} \frac{d}{c_0 n_0} \chi_0$$

and the integral kernels are

$$A = B = -C = -D = \frac{1}{2} \frac{d^2}{c_0^2} \frac{\chi_0}{\tau} H(t)$$

The source functions $k^\pm(x, s)$ were calculated from the nonlinear term of equation (3.14). They are the same in all calculations performed here:

$$k^\pm(x, s) = \mp \frac{3}{2} \mathcal{S} n_0^2 u^2(x, s) \frac{\partial u}{\partial s}(x, s)$$

The differentiation of u is performed by a 3-point Lagrangian differentiation, see Ref. 5.

Numerical calculations

In all Kerr calculations, the number of mesh points is chosen as 64×144 points in the (x, s) -space. The starting function $u^0(x, s)$ of each iteration is chosen identically zero, and the numbers of necessary iterations range between 6 and 33. Pure Kerr effect calculations with high peak values need the higher number in order to converge. With an increasing degree of dispersion, fewer iterations are needed. On the other hand, dispersion requires time-consuming calculation of integrals, so each iteration takes longer time. Therefore, the total computer time needed for Kerr calculations with or without dispersion is almost the same, about 12 hours on a SPARCstation 20.

Before actually discussing the pictures, a couple of remarks must be inserted. Although the calculations have been performed with specific values of the frequency f , the thickness of the slab d , and (if dispersion is present) the alignment frequency χ_0 , then, under the assumption that all other values remain unaffected, the same results also apply to other choices of f , d and χ_0 as long as the products $f \cdot d$ and

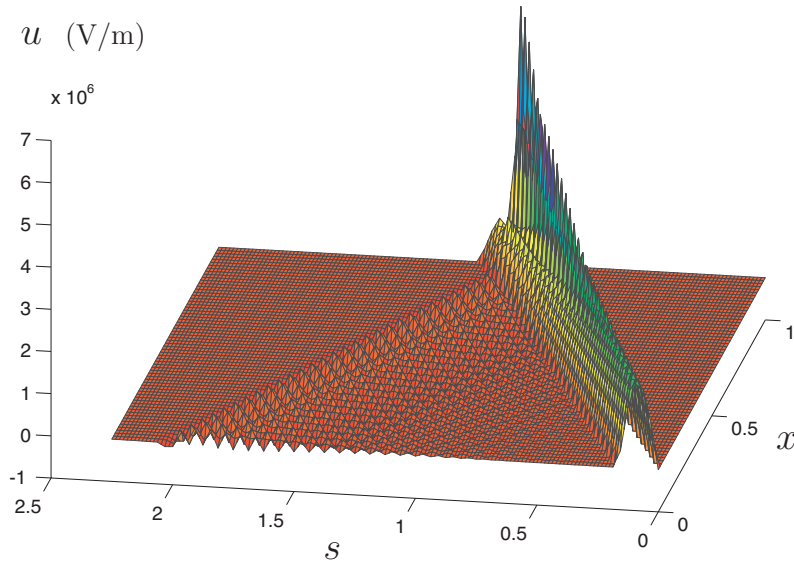


Figure 14: A half period sinusoidal pulse in a negative Kerr coefficient medium. Pulse a in the figures 15–17 are details of the pulse in this figure. Data: $\mathcal{S} = -10^{-14} \text{ m}^2/\text{V}^2$, $d = 1 \text{ mm}$, $n_0 = 1.5$, $f = 600 \text{ GHz}$, $E_0 = 2.0 \text{ MV/m}$, $\chi_0 = 0 \text{ s}^{-1}$.

$d \cdot \chi_0$ are kept constant. A warning must be given! At too high frequencies there is a risk to get into unphysical situations. The pulse duration must always be long compared to the Kerr response time!

The second remark is that in all figures, the electric field of the wave is shown. The reason is that the magnetic field graphs are similar to their electric counterparts. The inverse wave splitting shows that the electric field is the sum of the positive and negative wave components $u^+ + u^-$, whereas the magnetic field is proportional to the difference $u^+ - u^-$, see Appendix A. Calculations confirm that in the direct (primary) wave, the u^+ -contents dominates. In the secondary field, on the other hand, the major part of the field can be related to u^- . Therefore, the approximate magnetic field is obtained from the electric field by division with the wave impedance $\sqrt{\mu_0/n_0^2\epsilon_0}$ followed by a change of sign of the secondary field!

Results

In Figure 14 it is shown, how the half period sinusoidal wave, impinging on the negative Kerr coefficient medium, evolves in space and time. The effect of the nonlinearity can be seen as radiating sources being created by the incoming wave along its path of propagation. These sources emit waves moving to the right (higher x -values), thus adding to the primary wave, and moving to the left (smaller x -values), giving rise to a secondary field. Another peculiarity is, that as the pulse moves along, there is a change of shape in the primary wave. It splits into one sharp peak and one low amplitude part. At the far end of the slab, the peak gain is ~ 3.5 .

A minor part of the primary wave is reflected in the rear edge of the slab. This part is called the reflected wave below. In Figure 14, at a close look, the reflected field can be seen as a little ridge rising above the secondary field emitted by the internal sources. Naturally, as can be seen if the calculated data is studied carefully, the reflected wave and the secondary field waves give rise to new radiating sources along their path. However, this emitted field is too small to appear in the Figures 14 and 16.

Figure 15 shows a snapshot of the calculated wave at $s = \frac{62}{64}$. For comparison, the sinusoidal pulse traveling through a medium with $\mathcal{S} = 0$ is included. The disagreement at the front ends of the pulses is unphysical. As mentioned, the amplitude of the incoming wave has been chosen as high as possible in order to exaggerate the nonlinear effect. The disagreement probably points at a numerical inefficiency in depicting the sharp peak at the front end of the Kerr pulse. In spite of the deficiency, Figure 15 is of interest, since it hints at a possible break-down of the pulse, if the nonlinearity is strong enough.

Figure 16 shows the time evolution of the pulse at position $x = \frac{49}{64}$. Notice, the evolution of the primary wave, the secondary field waves and finally the reflected wave.

In Figure 17 the upper curves (*a*) are snapshots of the propagating Kerr pulse at $s = \frac{14}{64}, \frac{36}{64}$ and $\frac{62}{64}$. They are shown together with snapshots (*b*) of the same initial pulse, propagating through a dispersive Kerr medium, modeled with the same negative Kerr coefficient and a susceptibility kernel $\chi(t)$. In spite of the high degree of attenuation, the pulse *b* in the dispersive medium shows a strong influence from the negative Kerr nonlinearity as it approaches $x = 1$.

The remaining figures illustrate the positive Kerr coefficient medium. Remember, that the amplitude of the impinging wave is a little smaller, $E_0 = 1.9 \cdot 10^6$ V/m compared to $E_0 = 2.0 \cdot 10^6$ V/m in the negative Kerr coefficient medium! These calculations show a faster rate of convergence than the negative ones. The peak gain is less, and no noticeable disagreements at the pulse edges occur.

Figure 18 illustrates the evolution of the pulse in space and time. Observe, that the view is chosen differently than in Figure 14. The reason is, that the pulse shape is changed in a reverse order to what was found in the negative Kerr coefficient medium. This can also be seen from the upper curves *a* in Figure 19, which are snapshots of the pulse in the pure (nondispersive) Kerr medium, taken at $s = \frac{14}{64}, \frac{38}{64}$ and 1.0, respectively and Figure 20 showing the time evolution of the same pulse at position $x = \frac{20}{64}$. Contrary to what is seen in the negative Kerr coefficient medium, the peak in the positive Kerr coefficient medium appears in the rear edge, and the low amplitude part develops in the front end of the pulse. In Figure 19, the lower pulses *b* are the results of an attempt to create a solitary wave! They were formed by including a carefully balanced amount of dispersion in the medium characteristics. Notice, that all pulses in the figures of the positive Kerr coefficient medium have the same length. The reason for this will be discussed below.

Note also that in the Figures 18–21 showing pulse propagation in positive Kerr media, the sharp increase and peak of the electric field in the rear edge, imply a concentration of high-frequency components in that edge. This is due to a blue

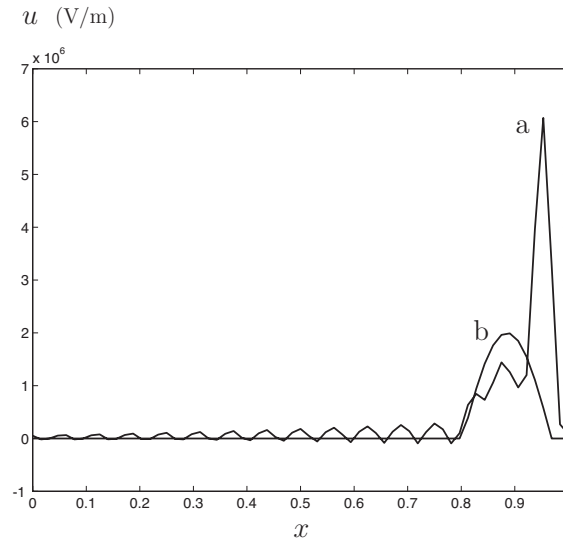


Figure 15: Snapshots at $s = 62/64$. Change of data from Figure 14: Pulse b has $\mathcal{S} = 0$.

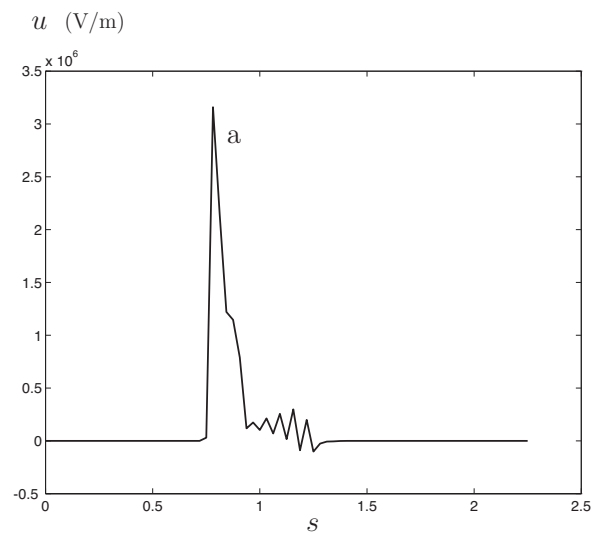


Figure 16: Time evolution of the pulse in Figure 14 ($x = 49/64$).

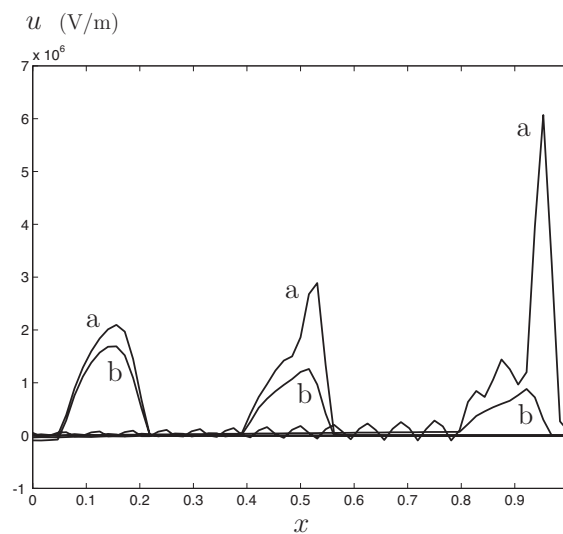


Figure 17: Snapshots at $s = 14/64$, $36/64$ and $62/64$. Change of data from Figure 14: Pulse b has $\chi_0 = 10.0 \cdot 10^{11} \text{ s}^{-1}$ and $\tau = 0.5 \cdot 10^{-11} \text{ s}$.

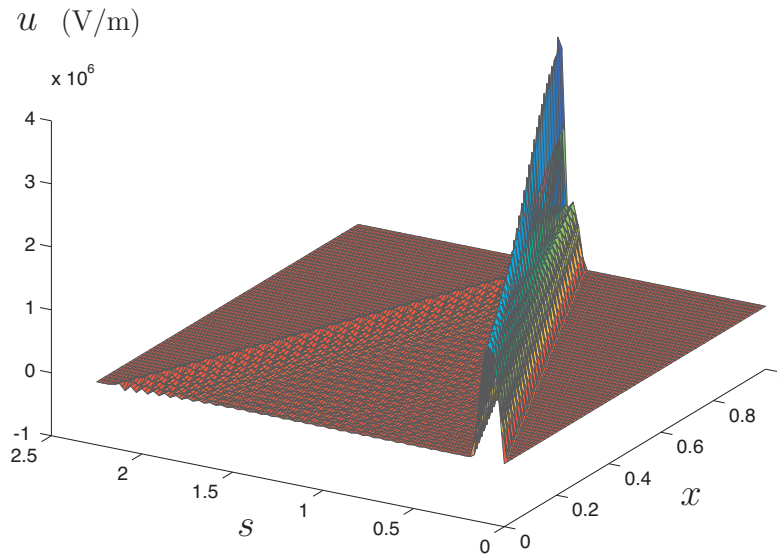


Figure 18: A half period sinusoidal pulse in a positive Kerr coefficient medium. Pulse a in the figures 19–20 are details of the pulse in this figure. Data: $\mathcal{S} = 10^{-14}\text{m}^2/\text{V}^2$, $d = 1\text{mm}$, $n_0 = 1.5$, $f = 600\text{ GHz}$, $E_0 = 1.9\text{ MV/m}$, $\chi_0 = 0\text{ s}^{-1}$.

shift [17]. The smoothly curved low amplitude part of the front pulse edge similarly shows an accumulation of low frequency components, i.e. a red shift [17]. In negative Kerr media, the situation is the opposite, there is a red shift at the rear edge and a blue shift in front.

As can be seen from the pure (nondispersive, $\chi(t) = 0$) Kerr effect pictures, e.g. Figures 14 and 18, the nonlinear effect gives rise to deformed pulses with narrow peaks of very high electric field strengths. In Ref. 18 a similar phenomenon is discussed. When optical beams propagate in positive Kerr media then theoretically beams with zero width and infinite electric field amplitude can exist. In reality, however, there is an upper limit to the magnitude of the refractive index. In analogy to what was suggested in Ref. 18, the parenthesis in (3.12) can be replaced by the constitutive relation

$$\epsilon(E) = \epsilon_0 n_0^2 + \frac{\epsilon_0 \mathcal{S} n_0^4 E^2}{1 + \frac{\mathcal{S} n_0^4 E^2}{n_{sat}^2 - n_0^2}}$$

In Figure 21, the saturated index of refraction was chosen $n_{sat} = 2.0$. A comparison of the saturated pulses a in Figure 21 to the nonsaturated pulses a in Figure 19 (the latter pulses are modeled with the constitutive relation (3.12)) shows, that saturation helps keeping the electric field strength down. The effect of saturation in this specific example is similar to the effect of a small amount of dispersion in the medium. It can also be concluded that if the medium is saturated and the alignment frequency χ_0 is high, like the medium generating the pulses c in Figure 21, the resulting pulses are almost identical to pulses developing in a nonsaturated medium modeled with the same χ_0 .

It has already been mentioned, that in the figures discussed so far, there is no

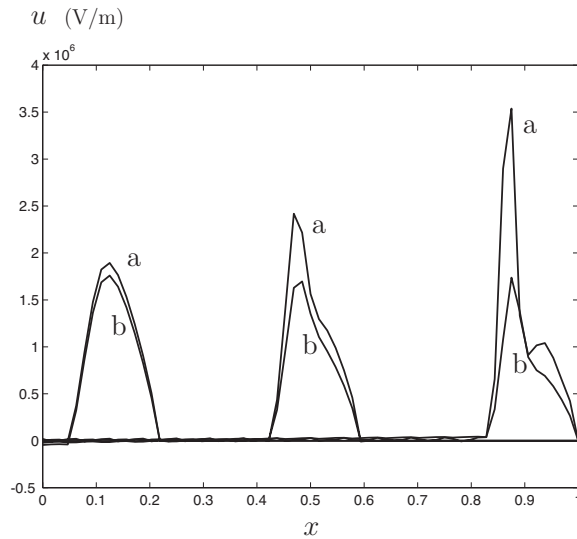


Figure 19: Snapshots at $s = 14/64, 38/64$ and 1.0 . Change of data from Figure 18: Pulse b has $\chi_0 = 5.0 \cdot 10^{11} \text{ s}^{-1}$ and $\tau = 0.5 \cdot 10^{-11} \text{ s}$.

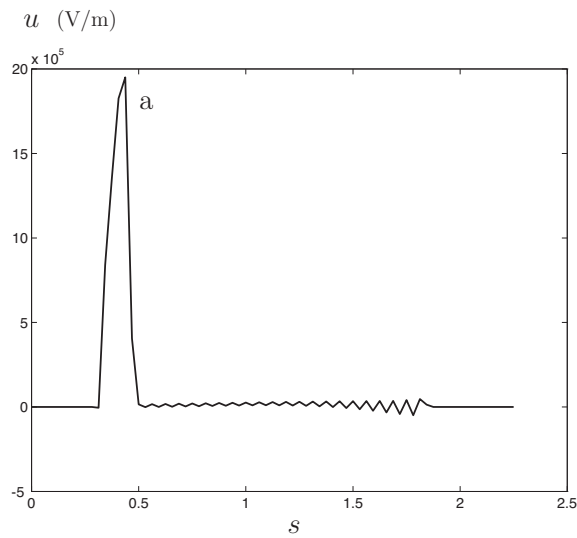


Figure 20: Time evolution of the pulse in Figure 18 ($x = 20/64$).

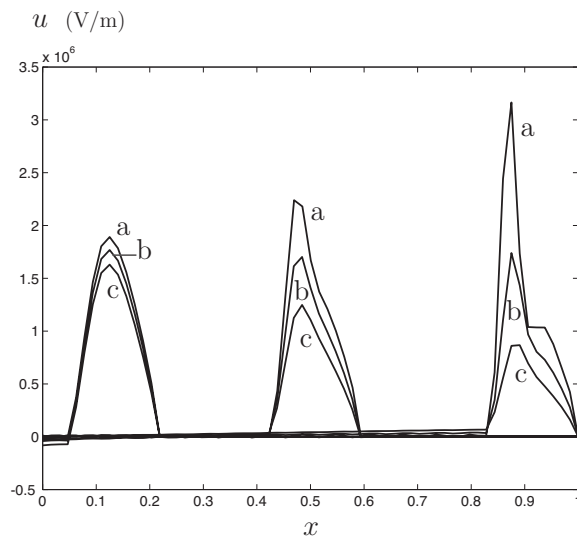


Figure 21: Saturated medium ($n_{sat} = 2.0$). Snapshots at $s = 14/64, 38/64$ and 1.0 . Pulse a : $\chi_0 = 0 \text{ s}^{-1}$. Pulse b : $\chi_0 = 4.5 \cdot 10^{11} \text{ s}^{-1}$. Pulse c : $\chi_0 = 10.0 \cdot 10^{11} \text{ s}^{-1}$. All pulses have $\tau = 0.5 \cdot 10^{-11} \text{ s}$. Data not mentioned is identical to that of Figure 18.

evidence that the pulse length is influenced by dispersion. It is known that with positive Kerr coefficient and normal dispersion (the phase velocity is a decreasing function of frequency) the pulse broadens, since the lower frequencies at the front edge travel at a higher speed, than the higher frequencies at the rear edge. Normal dispersion occurs when the 2nd derivative of the phase constant $\beta''(\omega) > 0$. At anomalous dispersion, $\beta''(\omega) < 0$, the high frequencies travel faster than the low, and the Kerr pulse is compressed. In negative Kerr media, where the frequency shift is reversed, the pulse is compressed at normal dispersion and broadened at anomalous dispersion. Linear dispersion, on the other hand, works differently. The reason is that frequencies sent out from the same position simultaneously, travel at different speeds. Therefore, at a later point of time, they are found at different positions. This means that the pulse always undergoes a linear broadening. When the nonlinear and linear dispersive effects counteract, a soliton can be obtained! Then, what type of dispersion is present in the pulses studied here?

The explanation can be obtained from the dispersion relation

$$\epsilon(\omega) = n_0^2 + \int_0^\infty \chi(t) e^{-i\omega t} dt$$

If the susceptibility kernel (3.16) is inserted, one gets

$$\epsilon(\omega) = n_0^2 + \frac{\chi_0 \tau}{1 + \omega^2 \tau^2} - i \frac{\omega \chi_0 \tau^2}{1 + \omega^2 \tau^2}$$

From this relation the propagation constant

$$\gamma(\omega) = i\omega \sqrt{\epsilon_0 \mu_0 \epsilon(\omega)}$$

can be obtained. The attenuation constant $\alpha(\omega)$ and the phase constant $\beta(\omega)$ are the real and imaginary parts of the propagation constant $\gamma(\omega)$, respectively.

In Figure 22, $\beta''(\omega)$ is shown for $\chi_0 = 5.0 \cdot 10^{11} \text{ s}^{-1}$ and $\tau = 0.5 \cdot 10^{-11} \text{ s}$. The dispersion is anomalous for low frequencies $\omega < 0.51 \cdot 10^{12} \text{ s}^{-1}$ and normal, but very close to zero dispersion, for all higher frequencies. Remember, that the frequencies of interest here are $\sim 4 \cdot 10^{12} \text{ s}^{-1}$! The attenuation constant is an increasing function of frequency. For the dominating frequencies, the value is almost constant, $\alpha(\omega) \sim 700 \text{ Np} \cdot \text{m}^{-1}$. Tests show that pulses, very similar to the ones shown above, can be obtained if dispersion with $\chi_0 = 5.0 \cdot 10^{11} \text{ s}^{-1}$ and $\tau = 0.5 \cdot 10^{-11} \text{ s}$ is exchanged for dissipation with a conductivity $\sigma = 4.42 \text{ S/m}$.

What happens to the pulse, if the dispersive influence is increased by a reduction of the relaxation time τ ? During the calculations, this can be noticed as a corresponding gain in the weight of the integral terms in (2.1). Roughly the $\beta''(\omega)$ -graph stays the same, showing anomalous dispersion in the lower frequency range, and almost zero dispersion at higher frequencies. However, the upper limit for anomalous dispersion moves higher up in frequency, from $\omega \approx 0.51 \cdot 10^{12} \text{ s}^{-1}$ to $\omega \approx 1.3 \cdot 10^{12} \text{ s}^{-1}$, so although the magnitude of β'' in the anomalous region is reduced by a factor 10, this type of dispersion becomes more dominant.

The two curves in Figure 23 show pulse propagation in a negative Kerr coefficient medium, both with $\chi_0 = 5.0 \cdot 10^{11} \text{ s}^{-1}$, the lower curve *a* has $\tau = 0.5 \cdot 10^{-11} \text{ s}$ and

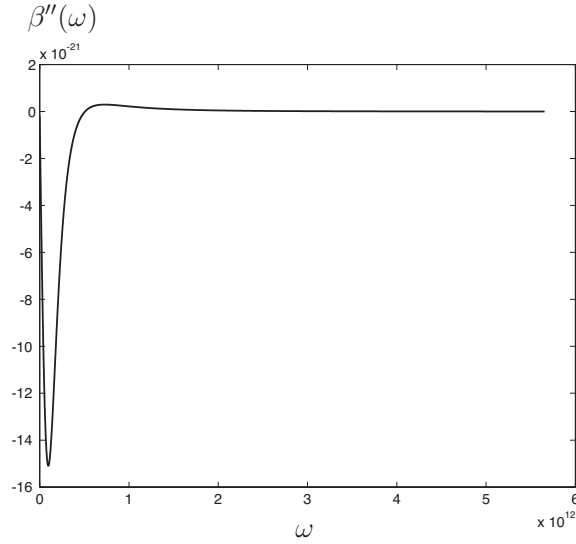


Figure 22: The second derivative of the phase constant $\beta(\omega)$ for a Debye model with $\chi_0 = 5.0 \cdot 10^{11} \text{ s}^{-1}$ and $\tau = 0.5 \cdot 10^{-11} \text{ s}$.

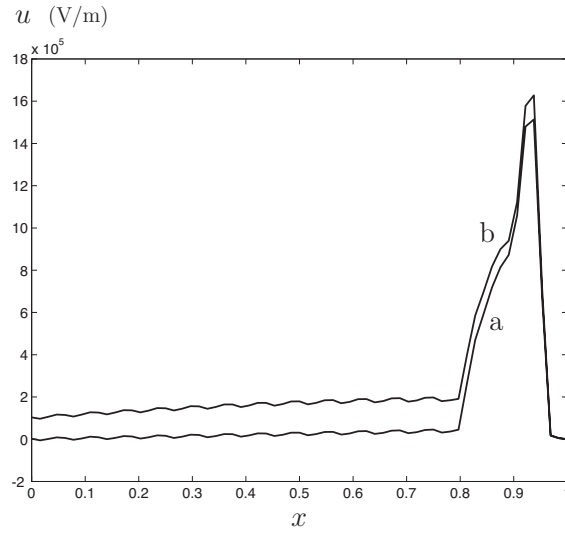


Figure 23: Snapshots at $s = 62/64$. Data: $\mathcal{S} = -10^{-14} \text{ m}^2/\text{V}^2$, $d = 1 \text{ mm}$, $n_0 = 1.5$, $f = 600 \text{ GHz}$, $E_0 = 1.9 \text{ MV/m}$, $\chi_0 = 5.0 \cdot 10^{11} \text{ s}^{-1}$. Pulse *a*: $\tau = 0.5 \cdot 10^{-11} \text{ s}$. Pulse *b*: $\tau = 0.15 \cdot 10^{-11} \text{ s}$.

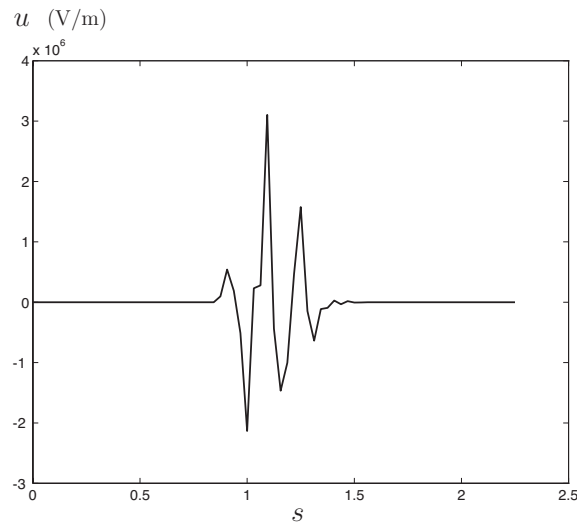


Figure 24: Time evolution of pulse in a nondispersive Kerr medium ($x = 55/64$). Data: $\mathcal{S} = 10^{-14} \text{ m}^2/\text{V}^2$, $d = 2 \text{ mm}$, $n_0 = 1.5$, $f_0 = 600 \text{ GHz}$, $E_0 = 1.9 \text{ MV/m}$.

the upper curve b has $\tau = 0.15 \cdot 10^{-11}$ s. Notice, the enhancement of the noise level, which can easily be understood from the time domain point of view: the medium “remembers” that a pulse has just passed by. Also note the broadening of the pulse! There are two possible explanations. One is that the negative Kerr pulses broaden due to nonlinear effects. The other possibility is that the broadening is due to linear dispersion, normal or anomalous. What explanation is valid here?

If a similar graph is studied for a medium with positive Kerr coefficient, the increase in noise is about the same. Furthermore, it is discovered again that the pulse broadens, particularly at the rear edge, but not as much as for the negative Kerr coefficient medium. If τ is decreased even more the trend is the same; pulses in both kinds of Kerr media broaden, but in the negative Kerr coefficient medium more than in the positive one.

A guess is that anomalous dispersion in the lower frequency range tends to compress the positive pulse somewhat, but the compression is counteracted and overtaken by linear dispersion. In the negative Kerr coefficient medium on the other hand, the two effects cooperate, so the broadening becomes more pronounced!

In all calculations discussed so far, the incoming pulse has been the same, one half period of a sinusoidal wave. That uncomplicated shape was chosen to make the nonlinear influence appear clearly, especially considering the restrictions imposed by the discretization. The intention of Figure 24 is to show, that many other pulse shapes can be treated. Here, the incident wave can be expressed on the form (3.15). It is given by

$$E(z, t) = E_0 \sin\left(\frac{1}{3}\omega_0 t\right) \sin(\omega_0 t), \quad 0 < t < \frac{3\pi}{\omega_0}$$

and it is zero at all other times. The amplitude is $E_0 = 1.9 \cdot 10^6$ V/m and the frequency $f_0 = 6 \cdot 10^{11}$ Hz. The scattering medium is the same as above, with the positive Kerr coefficient $\mathcal{S} = 10^{-14}$ m²/V², $n_0 = 1.5$ and the thickness of the scattering medium $d = 2$ mm. In Figure 24 it is shown how the pulse evolves in time at the normalized position $x = \frac{55}{64}$. The duration of the incident pulse is $s = \frac{1}{2}$, so in the figure the primary wave is mixed with some secondary field noise, excited by the internal sources, and parts of the returning, reflected primary wave. Notice, the amplification of the central part of the pulse, and the characteristic Kerr jerk also observed in the figures above.

4 Summary

Two different techniques to obtain solutions of wave propagation problems in weakly nonlinear media are discussed. Both methods are accessible from the same algorithm building on time domain analysis of nonstationary mixed initial boundary value problems. After a short summary of the underlying nonstationary mathematical analysis and its numerical implementation, the two techniques are presented. Problems from two interesting areas of research are introduced. The type A technique is applied to a generator of signals for high frequency switching, and the type

B technique is used to test how ps-pulses are influenced by the nonlinear Kerr effect. There is some discussion of the observed technical and physical notions. A limited comparison of the two techniques suggests that they are complementary.

Acknowledgment

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

Appendix A The wave splitting

The second order wave equation

Consider the generalized form of the inhomogeneous, second order wave equation expressed in normalized coordinates x and s . Note, that the value of the slowness function f in (2.1) is $f(x, s) = 1$.

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2}(x, s) - \frac{\partial^2 u}{\partial s^2}(x, s) + \mathcal{A}(x, s) \frac{\partial u}{\partial x}(x, s) + \mathcal{B}(x, s) \frac{\partial u}{\partial s}(x, s) \\ + \mathcal{C}(x, s)u(x, s) + \int_{-\infty}^s \mathcal{D}(x, s, s')u(x, s') ds' = g(x, s) \end{aligned} \quad (\text{A.1})$$

In order to see how equation (A.1) is related to the general hyperbolic wave equation (2.1), the concept of wave splitting is introduced. Wave splitting of the fields is a way to obtain a natural set of dependent coordinates that is adapted to the scattering situation. The wave splitting can be defined in several different ways. The definition adopted here uncouples the u^+ - and u^- -equations for the principal part

$$\frac{\partial^2 u}{\partial x^2}(x, s) - \frac{\partial^2 u}{\partial s^2}(x, s) = 0$$

of equation (A.1). Thus, proceeding formally, the wave splitting is defined by the following transformation of the dependent variables [3, 4, 20]:

$$\begin{pmatrix} u^+ \\ u^- \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -\partial_s^{-1} \\ 1 & \partial_s^{-1} \end{pmatrix} \begin{pmatrix} u \\ \frac{\partial u}{\partial x} \end{pmatrix}$$

with inverse

$$\begin{pmatrix} u \\ \frac{\partial u}{\partial x} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -\frac{\partial}{\partial s} & \frac{\partial}{\partial s} \end{pmatrix} \begin{pmatrix} u^+ \\ u^- \end{pmatrix}$$

where the anti-derivative ∂_s^{-1} is defined as

$$\partial_s^{-1} f(s) = \int_{-\infty}^s f(s') ds'$$

The new fields $u^\pm(x, s)$ satisfy a first order 2×2 system of hyperbolic partial differential equations, which is identical to the generalized u^\pm -dynamics in (2.1). The explicit expressions of the coefficients of (2.1) are

$$\begin{cases} \alpha(x, s) = -\frac{1}{2}\mathcal{A}(x, s) + \frac{1}{2}\mathcal{B}(x, s) \\ \beta(x, s) = +\frac{1}{2}\mathcal{A}(x, s) + \frac{1}{2}\mathcal{B}(x, s) \\ \gamma(x, s) = +\frac{1}{2}\mathcal{A}(x, s) - \frac{1}{2}\mathcal{B}(x, s) \\ \delta(x, s) = -\frac{1}{2}\mathcal{A}(x, s) - \frac{1}{2}\mathcal{B}(x, s) \end{cases}$$

and

$$\begin{cases} A(x, s, s') = \frac{1}{2} \left[+\frac{\partial \mathcal{A}}{\partial s'}(x, s') - \frac{\partial \mathcal{B}}{\partial s'}(x, s') + \mathcal{C}(x, s') + \int_{s'}^s \mathcal{D}(x, s'', s') ds'' \right] \\ B(x, s, s') = \frac{1}{2} \left[-\frac{\partial \mathcal{A}}{\partial s'}(x, s') - \frac{\partial \mathcal{B}}{\partial s'}(x, s') + \mathcal{C}(x, s') + \int_{s'}^s \mathcal{D}(x, s'', s') ds'' \right] \\ C(x, s, s') = -A(x, s, s') \\ D(x, s, s') = -B(x, s, s') \end{cases}$$

The source functions are given by

$$k^\pm(x, s) = \mp \frac{1}{2} \int_{-\infty}^s g(x, s') ds'$$

The transmission line equations

The starting point is from the set of linearized transmission line equations

$$\begin{pmatrix} 0 & C \\ L & 0 \end{pmatrix} \frac{\partial}{\partial t} \begin{pmatrix} i \\ v \end{pmatrix} + \frac{\partial}{\partial z} \begin{pmatrix} i \\ v \end{pmatrix} + \begin{pmatrix} 0 & G - I'(v^i) \\ R & 0 \end{pmatrix} \begin{pmatrix} i \\ v \end{pmatrix} = \begin{pmatrix} I(v^i) - I'(v^i) \cdot v^i \\ 0 \end{pmatrix}$$

For shortness, the (z, t) -dependence of the variables i and v and the function v^i has been suppressed. In the general case, the transmission line properties R , G , L and C also depend on z and t . These equations are easily transformed into the general first order 2×2 system of hyperbolic equations (2.1). The following wave splitting diagonalizes the set of transmission line equations:

$$\begin{pmatrix} u^+ \\ u^- \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & \sqrt{\frac{C}{L}} \\ 1 & -\sqrt{\frac{C}{L}} \end{pmatrix} \begin{pmatrix} i \\ v \end{pmatrix}$$

with inverse

$$\begin{pmatrix} i \\ v \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \sqrt{\frac{L}{C}} & -\sqrt{\frac{L}{C}} \end{pmatrix} \begin{pmatrix} i \\ v \end{pmatrix} \begin{pmatrix} u^+ \\ u^- \end{pmatrix}$$

The explicit expressions of the coefficients are:

$$\begin{cases} f(z, t) = \sqrt{LC} \\ \alpha(z, t) = -\frac{1}{2}\sqrt{\frac{L}{C}} [G - I'(v^i)] - \frac{1}{2}\sqrt{\frac{C}{L}}R \\ \beta(z, t) = +\frac{1}{2}\sqrt{\frac{L}{C}} [G - I'(v^i)] - \frac{1}{2}\sqrt{\frac{C}{L}}R \\ \gamma(z, t) = -\beta(z, t) \\ \delta(z, t) = -\alpha(z, t) \end{cases}$$

and

$$A = B = C = D = 0$$

The source functions are given by

$$k^\pm(z, t) = \frac{1}{2} [I(v^i) - I'(v^i) \cdot v^i]$$

Further details can be found in Ref. 4.

References

- [1] I. Åberg. Transient waves propagating from internal sources in non-stationary media—Numerical implementation. Technical Report LUTEDX/(TEAT-7048)/1-29/(1996), Lund Institute of Technology, Department of Electromagnetic Theory, P.O. Box 118, S-211 00 Lund, Sweden, 1996.
- [2] I. Åberg and A. Karlsson. “The source problem”—transient waves propagating from internal sources in non-stationary media”. *Wave Motion*, 1997. (in press).
- [3] I. Åberg, G. Kristensson, and D.J.N. Wall. Propagation of transient electromagnetic waves in time-varying media—direct and inverse scattering problems. *Inverse Problems*, **11**(1), 29–49, 1995.
- [4] I. Åberg, G. Kristensson, and D.J.N. Wall. Transient waves in non-stationary media. *J. Math. Phys.*, **37**(5), 2229–2252, 1996.
- [5] M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions*. Applied Mathematics Series No. 55. National Bureau of Standards, Washington D.C., 1970.
- [6] G.P. Agrawal. *Nonlinear fiber optics*. Academic Press, San Diego, 1995.
- [7] S.T. Allen, M. Reddy, M.J.W. Rodwell, R.P. Smith, S.C. Martin, J. Liu, and R.E. Muller. Submicron Schottky-Collector AlAs/GaAs resonant tunnel diodes. In *Tech. Dig.*, Washington D.C., 1993. Int. Electron Device Meet.

- [8] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 2. Interscience, New York, 1962.
- [9] S.K. Diamond, E. Ozbay, M.J.W. Rodwell, and D.M. Bloom. Fabrication of resonant tunneling diodes for switching applications. In *Proc. Conf. Picosecond Electronics and Optoelectronics*, Salt Lake City, Utah, 1989.
- [10] A. Hasegawa. *Optical Solitons in Fibers*. Springer-Verlag, Berlin Heidelberg, 1990.
- [11] A. Hasegawa and Y. Kodama. *Solitons in Optical Communications*. Clarendon Press, Oxford, 1995.
- [12] E. Hecht. *Optics*. Addison-Wesley, Reading, second edition, 1987.
- [13] R. Hülsewede, U. Effing, I. Wolff, and D. Jäger. CAD of pulse compression on nonlinear transmission lines. In *MIOP'95*, Sindelfingen, Germany, 1995.
- [14] T.M. Il'ina and R.V. Khoklov. Wave processes in lines with nonlinear shunt resistance. *Radiotekhnika i Elektronika*, **8**, 1864–1872, 1963.
- [15] D. Jäger and R. Hülsewede. Microwave propagation on nonlinear transmission lines. In *XXVth General Assembly of the international union of the radio science*, Lille, France, 1996. International Union of Radio Science.
- [16] D. Jäger, R. Kremer, and A. Stöhr. Travelling-wave optoelectronic devices for microwave applications. *IEEE MTT-S Digest*, **1**, 163–166, 1995.
- [17] M. Joindot. Lightwave communications. In J. Hamelin, editor, *Modern Radio Science 1996*. Oxford University Press, 1996.
- [18] M. Karlsson. *Nonlinear Propagation of Optical Pulses and Beams*. PhD thesis, Chalmers University of Technology, S-412 96 Göteborg, Sweden, 1994.
- [19] R.S. Kirkby, S. Withington, A.B. Darling, and F.G. Kilgour. *Engineering in history*. McGraw-Hill, New York, 1956.
- [20] G. Kristensson and R.J. Krueger. Direct and inverse scattering in the time domain for a dissipative wave equation. Part 1: Scattering operators. *J. Math. Phys.*, **27**(6), 1667–1682, 1986.
- [21] J.D. Logan. *Nonlinear Partial Differential Equations*. John Wiley & Sons, New York, 1994.
- [22] M. Remoissenet. *Waves Called Solitons*. Springer-Verlag, Berlin Heidelberg, 1994.
- [23] M.J.W. Rodwell, S.T. Allen, R.Y. Yu, M.G. Case, U. Bhattacharya, M. Reddy, E. Carman, M. Kamegawa, Y. Konishi, J. Pusch, and R. Pallela. Active and nonlinear wave propagation devices in ultrafast electronics and optoelectronics. *Proc. IEEE*, **82**(7), 1037–1059, 1994.

- [24] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics*. John Wiley & Sons, New York, 1991.
- [25] I. Stakgold. *Green's Functions and Boundary Value Problems*. John Wiley & Sons, New York, 1979.
- [26] G.B. Whitham. *Linear and nonlinear waves*. John Wiley & Sons, New York, 1974.
- [27] R.Y. Yu, Y. Konishi, S.T. Allen, M. Reddy, and M.J.W. Rodwell. A traveling-wave resonant tunnel diode pulse generator. *IEEE Microwave and Guided Wave Letters*, **4**(7), 220–222, 1994.
- [28] E. Zeidler. *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*. Springer, New York, 1986.