
Andreas P. Mayer 1,*, Elena A. Mayer 1, Markus Mayer 2, Werner Ruile 2,†, Vikrant Chauhan 2, Thomas Forster 2,‡ and Karl C. Wagner 2

1 Department B+W, Offenburg University of Applied Sciences, Klosterstr. 14, 77723 Gengenbach, Germany
2 RF360 Europe GmbH, Anzinger Strasse 13, 81671 Munich, Germany
* Correspondence: andreas.mayer@hs-offenburg.de
† Retired author.

Abstract: In the framework of electro-elasticity theory and the finite element method (FEM), a model is set up for the computation of quantities in surface acoustic wave (SAW) devices accounting for nonlinear effects. These include second-order and third-order intermodulations, second and third harmonic generation and the influence of electro-acoustic nonlinearity on the frequency characteristics of SAW resonators. The model is based on perturbation theory, and requires input material constants, e.g., the elastic moduli up to fourth order for all materials involved. The model is two-dimensional, corresponding to an infinite aperture, but all three Cartesian components of the displacement and electrical fields are accounted for. The first version of the model pertains to an infinite periodic arrangement of electrodes. It is subsequently generalized to systems with a finite number of electrodes. For the latter version, a recursive algorithm is presented which is related to the cascading scheme of Plessky and Koskela and strongly reduces computation time and memory requirements. The model is applied to TC-SAW systems with copper electrodes buried in an oxide film on a LiNbO3 substrate. Results of computations are presented for the electrical current due to third-order intermodulations and the displacement field associated with the second harmonic and second-order intermodulations, generated by monochromatic input tones. The scope of this review is limited to methodological aspects with the goal to enable calculations of nonlinear quantities in SAW devices on inexpensive and easily accessible computing platforms.

Keywords: finite element analysis; intermodulation distortions; nonlinear acoustics; nonlinear systems; surface acoustic waves

1. Introduction

Signal processing devices based on surface acoustic waves (SAW), like frequency filters, delay lines and others, have, nowadays, far-reaching applications in various fields, especially in the growing market of mobile communication. The working principle of almost all these devices makes use of the piezoelectric effect with interdigital electrode structures and relies on linear behavior. New standards like 5G further raise the linearity requirements. This has led to an urgent need for gaining an in-depth understanding of the influence of nonlinearity on the propagation of SAWs and for means to reduce this influence.

Major undesired nonlinear effects that occur in SAW devices and need to be counteracted are harmonic generation and intermodulations, which can deteriorate signal quality in mobile communication. Of similar importance is the power dependence of the resonance frequencies. All these effects will be addressed in this review.

So far, modelling of nonlinear effects in SAW devices has mostly been carried out phenomenologically by extending coupling of modes (COM) and P-matrix theory [1–4] or by nonlinear extensions of an equivalent circuit model (especially the Mason model) [5].
latter were successfully applied to bulk acoustic wave (BAW) devices [6,7] (and references therein). Parameters containing information on nonlinearity were obtained by fitting to experimental data. Calculations, relating nonlinear effects like third-order intermodulational distortions (IMD3) directly to linear and nonlinear material constants of the elastic media involved, use the finite element method (FEM) [8–15]. In contrast to the phenomenological models mentioned above, finite element calculations do not rely on restrictive model assumptions. They yield reliable results without the need for fit parameters, if the relevant material constants are known to a sufficient extent and accuracy. In addition, they provide full-wave fields, which can be very useful for the designer of SAW devices and allow for detailed insights into the nonlinear behavior of a device.

To the authors’ knowledge, detailed investigations of nonlinear propagation effects in connection with SAWs began in the mid-sixties of the previous century (for a review see [16]). The field was boosted by the development of the elastic convolver, an analog, nonlinear, signal-processing device based on SAWs, which, however, did not reach the market (see the reviews [17,18]). Here, electro-acoustic nonlinearity was a desired effect. Nonlinearity in the context of SAWs has gained renewed attention in connection with further miniaturization, leading to higher power densities, and tightened linearity requirements for signal processing devices, as well as in the field of non-destructive evaluation [19].

The major “classical” sources of nonlinearity are the third- and higher-order terms in an expansion of the materials’ potential energy per unit volume in powers of displacement gradients and electrical field components. The theory of nonlinear electro-elasticity was presented in detail in classical texts [20–23]. Theoretical studies of effects of this classical nonlinearity on the propagation of surface acoustic waves were first carried out with the help of coupled amplitude equations, derived by Reutov [24], or evolution equations in the form of integral equations ([25] and references therein). They focused on homogeneous media and layered structures, including piezoelectric materials, with planar surfaces and interfaces. They allowed for simulations and, hence, for quantitative predictions of nonlinear propagation effects directly on the basis of material constants. The investigations included intermodulations at an early stage [26,27].

The complexity of the displacement field and the spatial distribution of electrostatic potential in SAW-based devices with their arrangements of electrodes on the surface or buried in a dielectric film calls for the use of numerical techniques, such as FEM for quantitative predictions of nonlinear effects. The goal of this review paper is to outline, in detail, a two-dimensional FEM-based approach, treating the nonlinearity within perturbation theory. To demonstrate its applicability to quantitative analysis of electrical currents and mechanical displacements generated by the electro-elastic nonlinearities of the constituent materials of a SAW device, a few numerical results are shown, which refer to a temperature-compensated SAW system (TC-SAW system) consisting of copper electrodes on a LiNbO$_3$ 128°YX substrate and a silicon oxide film. The presentation focuses on methods and algorithms rather than on numerical results. The complexity of the simulation methods and their theoretical basis in nonlinear electro-elasticity induced us to confine this review to methodological aspects concerning theory rather than including experimental results or giving a detailed overview of the field’s historical development.

The paper is structured in the following way:

In Section 2, the part of nonlinear elasticity theory underlying the FEM model is briefly summarized, and the notation used in the following sections is introduced. In particular, we introduce a four-component generalized displacement field, its fourth component being the electrostatic potential. In this context, the material constants used as input data in the numerical calculations for the considered TC-SAW system are mentioned and their sources provided.

Section 3 presents the FEM model itself for an infinite periodic arrangement of electrodes and for a finite number of electrodes.

In Section 4, two different methods of calculating the current in the electrodes due to third-order intermodulations are introduced. One of them allows the calculation of the
nonlinear current generated by third-order intermodulations without first determining the displacement field and electrostatic potential associated with the third-order intermodulations.

In Section 5, we show how the influence of nonlinearity on the resonance curve and phase response of a SAW resonator can be efficiently determined by combining FEM with perturbation theory.

In Section 6, reflections of bulk waves from the bottom of the substrate are considered, which are part of the (generalized) displacement field of the second harmonic of a SAW excited in an interdigital transducer.

Section 7 is concerned with FEM calculations for systems with a finite number of equidistantly arranged electrodes. A recursive algorithm is introduced that enables calculations of nonlinear quantities (mechanical and electrical displacement fields associated with intermodulations and higher harmonics) for a large number of electrodes with considerably reduced numerical effort as compared to the standard approach. The paper ends with a short conclusion.

2. Nonlinear Electro-Elasticity

2.1. General Theory

The theory of nonlinear electro-elasticity underlying our FEM calculations was developed in detail in the nineteen-sixties to eighties (see, e.g., [20–23,25]). Here, only a brief summary is provided. Further references to earlier work may be found in [16,23].

As degrees of freedom, we consider the three Cartesian components of the displacement field, \( u_1, u_2, u_3 \), referring to the spatial frame [20] and the electrical potential. The latter will be treated as the fourth component, \( u_4 \), of a four-component generalized displacement field. This concept was introduced for linear piezoelectricity in [28] and was applied to nonlinearity in [25,29–32]. The components of the generalized displacement field are labelled by lower-case Latin indices. They depend on time \( t \) and on the Cartesian coordinates \( X_1, X_2, X_3 \), referring to the material frame [20] and being the Cartesian components of the position vector \( X \). We invoke a summation convention that implies summation over repeated lower-case Latin indices from 1 to 4 and summation over repeated upper-case Latin indices from 1 to 3. These indices are always subscripts. (We note that there is no implicit summation over any other kind of index.).

The equations of motion and, partly, the boundary conditions for the four-component generalized displacement field can be derived from a Lagrangian,

\[
L = \int (T - \Phi) \, d^3X, \tag{1}
\]

via Hamilton’s principle (see for example [20,23,33]). In (1), \( d^3X \) denotes the volume element in material coordinates, and the integral extends over the entire space filled with elastic material. The kinetic energy per unit volume of undeformed elastic medium is

\[
T = \frac{1}{2} \rho(X) \left( \frac{\partial}{\partial t} u_m(X,t) \right) \Delta_{mn} \left( \frac{\partial}{\partial t} u_n(X,t) \right). \tag{2}
\]

Consequently, \( \rho(X) \) is the mass density at position \( X \) of the undeformed medium. Furthermore, \( \Delta_{mn} \) is a generalized Kronecker symbol defined as \( \Delta_{mn} = 1 \), if \( m = n \) and \( m \neq 4 \), and \( \Delta_{mn} = 0 \) otherwise.

For the systems we are considering here, which consist of elastic and partially piezoelectric media, it is assumed that the potential energy per unit volume of the undeformed medium, \( \Phi \), may be expanded in powers of gradients of the generalized displacement field components \( u_{i,M}(X,t) = \partial u_i(X,t) / \partial X_M \),

\[
\Phi = \frac{1}{2} C_{mM \, nN} u_m u_M u_n u_N + \frac{1}{4} S_{mM \, nN \, pP} u_m u_M u_n u_P u_P + \frac{1}{4} S^{(3)}_{mM \, nN \, pP \, qQ} u_m u_M u_n u_P u_P u_q u_Q + O(u^5). \tag{3}
\]
The expansion coefficients on the right-hand side of (3) may be regarded as components of generalized tensors $C$, $S^{(2)}$, $S^{(3)}$, which may depend on position $X$. In order to avoid confusion, we note that the superscripts $(2)$ and $(3)$ in the above equation refer to second-order (i.e., lowest-order) and third-order nonlinearity, respectively, and do not correspond to the order in the expansion of $\Phi$ in powers of generalized displacement gradient components. For simplicity, we only treat cases where the system is in static equilibrium, where no external static stress or electric field is applied to the system. In such cases, the expansion coefficients in the first term on the right-hand side of (3) have the following meaning:

For the indices $m$, $n$ ranging from 1 to 3, $C_{mM \, nN}$ are second-order elastic constants (SOECs). $C_{4M \, lN} = C_{lN \, 4M} = \delta_{1L} \, C_{M \, NN}$, $l = 1, 2, 3$; $M, N, L = 1, 2, 3$, are the linear piezoelectric constants, and $C_{4M \, 4N} = -\varepsilon_{NM}$, apart from the minus-sign, are the static dielectric constants of the corresponding medium.

In the second term on the right-hand side of (3), $S^{(2)}_{mM \, nN \, pp}$ are second-order and third-order elastic constants (TOECs), if the indices $m, n, p$ range between 1 and 3. The coefficients $S^{(2)}_{4M \, nN \, pp}$ with $n, p$ ranging between 1 and 3 are linear combinations of electro-elastic and piezoelectric constants; $S^{(2)}_{4M \, 4N \, pp}$ with $p = 1, 2, 3$, are related to the electrostriction constants, and $S^{(2)}_{4M \, 4N \, 4p}$ to the static second-order electrical permittivities.

In the third term on the right-hand side of (3), $S^{(3)}_{mM \, nN \, pp \, pq}$ are combinations of second-order, third-order and fourth-order elastic constants (FOECs), if the lower-case indices $m, n, p, q$ range between 1 and 3 [34]. If at least one of the indices $m, n, p, q$ is equal to 4, the corresponding coefficient is related to the electrical potential. All fourth-order material constants entering the generalized eighth-rank tensor $S^{(3)}$ are specified in [23]. (For further details, see [16,20,23,25,34].) Explicit expressions for the elements of the tensors $S^{(2)}$, $S^{(3)}$ in terms of the material constants are provided in the Appendix A.

We note that the index pairs, consisting of a lower-case and an upper-case index, may be permuted at the quantities $C_{kM \, lN}$, $S^{(2)}_{mM \, nN \, pp}$, and $S^{(3)}_{mM \, nN \, pp \, pq}$.

2.2. Input Data for the TC-SAW System

In order to obtain reliable results in FEM simulations, all material constants entering the generalized tensors $C$, $S^{(2)}$ and $S^{(3)}$ must be known for the materials that constitute the device to be modelled. For the example system of a TC-SAW interdigital transducer considered here, these are 128° rotated Y-cut LiNbO$_3$ as substrate material (Euler angles: $\lambda = 0.0^\circ$, $\mu = 37.85^\circ$, $\theta = 0.0^\circ$), copper electrodes and a temperature-compensating silicon oxide film, which covers the electrodes and the substrate surface. For simplicity, the surface of the SiO$_2$ film is taken to be planar in our model. A complete set of material constants occurring in $C$, $S^{(2)}$ and $S^{(3)}$ is hardly available for any material. The situation for the three materials in our example system will now be discussed.

2.2.1. Material Constants for LiNbO$_3$

The second-order elastic constants, piezoelectric constants and dielectric constants entering the generalized tensor $C$ were taken from [35]. For the third-order material constants (TOECs, electro-elastic constants, electrostriction constants and second-order electrical permittivities) a complete set of measured data is available [36] and was used in the generalized tensor $S^{(2)}$, along with the linear material constants of [35]. Since no fourth-order material constants for LiNbO$_3$ are known to us, we have equated the generalized tensor $S^{(3)}$ to zero in most of our calculations of IMD3. To demonstrate the role of the material constants of second and third order in the tensor $S^{(3)}$, we compare, in Section 4, the results for IMD3 and the third harmonic obtained when equating $S^{(3)}$ to zero, with those obtained with $S^{(3)}$ containing certain second-order and third-order, but no fourth-order material constants.
As a caveat, we would like to mention here that substantial compensations of material constants of different orders can occur in the tensors $S^{(2)}$ and $S^{(3)}$. In order to account for the unknown FOECs in a rough approximate way, an isotropic approximation was introduced in [15], and the four fourth-order Lamé constants were treated as fit parameters. For the mass density of LiNbO$_3$, we chose the value of $4628$ kg/m$^3$.

2.2.2. Material Constants for SiO$_2$

The two second-order elastic constants and mass density of the isotropic SiO$_2$ film are provided in [37] (second row in Table II). Our sources for the three TOECs are [38,39] and for the four FOECs [40]. For the static relative dielectric constant, the value $3.75$ was chosen. All third-order and fourth-order material constants related to the electrostatic potential either vanish or have been neglected. We note that the material constants of the SiO$_2$ film may be sensitive to its fabrication process. Yost and Breazeale [39] point out that residual stresses at this surface (open-circuit boundary condition).

2.2.3. Material Constants for Copper Electrodes

As mentioned above, subtle compensations between elastic constants of different order may occur in the generalized tensors $S^{(2)}$ and $S^{(3)}$. Therefore, a consistent data set of material constants is desirable, which is obtained from the same source. In the case of crystalline copper, such a data set is available from ab-initio calculations based on density functional theory [41]. Good agreement for second-order and third-order elastic moduli was found between the ab-initio results and measured data. All FOECs are provided in [41], too. These constants have been isotropized by averaging using the Voigt assumption [42,43]. The mass density value chosen for our calculations was $8920$ kg/m$^3$.

3. FEM Modeling of Nonlinear SAW Systems

3.1. General Considerations

Our FEM modeling of SAW systems was confined to two dimensions, assuming an infinite aperture of the electrodes and no dependence of the displacement field and electric potential on $X_2$ in the coordinate system introduced in Figure 1. However, all three Cartesian components of the displacement field are taken into account. On the planar surface of a piezoelectric substrate, metal electrodes are placed, which are embedded in a dielectric film covering the substrate surface and have a trapezoidal cross-section. An extension to substrates consisting of several layers stratified parallel to the $(X_1,X_2)$-plane is straight-forward.

![Figure 1. Geometries treated by FEM. Infinite periodic (a) and finite (b) arrangement of electrodes. Perfectly matched layers I–VII.](image)

Two types of systems are distinguished: Infinite periodic (I) and finite (II) arrangements of electrodes on the substrate surface (Figure 1). The surface of the dielectric film is traction-free, and the normal component of the material electric displacement vector vanishes at this surface (open-circuit boundary condition).
free, and the normal component of the material electric displacement vector vanishes at this surface (open-circuit boundary condition).

When reflections must be avoided, a perfectly matched layer (PML) is placed underneath the substrate part of the simulated system (Figure 1) and—in case (II)—at the right and left sides of the finite system, too. When reflections from the bottom of the substrate are to be investigated, an impedance boundary condition has been applied that mimics certain prescribed boundary conditions (e.g., zero traction and zero normal component of the material electric displacement field, or a prescribed value of the electric potential) at the bottom surface of a wafer with realistic thickness. These boundary conditions will be described in more detail below. In addition, periodic boundary conditions along the $X_1$-direction are imposed in case (I).

Following the iso-parametric concept, the four-components of the generalized displacement field, $u_j(X,t)$, are represented in each element as a linear combination of test functions $N_k(r,s)$ with $k = 1, \ldots, 8$, corresponding to rectangular 8-node elements [44]. Here, $r(X_1, X_3), s(X_1, X_3)$ are non-dimensional internal variables. When inserting this representation into the Lagrangian (1) with (2) and (3), Hamilton’s principle yields nonlinear coupled equations for the four degrees of freedom $\hat{u}_j^{(\ell)}(t)$ with $j = 1, \ldots, 4$, which are the three Cartesian components of the node displacement vector and the electrical potential at the corresponding node position in the un-deformed medium,

$$\sum_{\ell'} M_{jj'}^{(\ell\ell')} \frac{\partial^2}{\partial t^2} \hat{u}_j^{(\ell')}(t) + \frac{1}{2} \sum_{\ell'\ell''} K_{jj'}^{(\ell\ell\ell'\ell'')} \hat{u}_j^{(\ell')}(t) \hat{u}_j^{(\ell'')}(t) + \frac{1}{3!} \sum_{\ell'\ell''\ell'''} K_{jj'}^{(\ell\ell\ell'\ell''\ell''')} \hat{u}_j^{(\ell')}(t) \hat{u}_j^{(\ell'')} \hat{u}_j^{(\ell''')}(t) + O(\hat{u}^4) = \delta_{jj'} V^{(\ell)}(t).$$

The superscript $\ell$ labels the nodes of the total mesh. In (4), $M$ is the global mass matrix and $K^{(2)}$ is the linear global stiffness matrix. $K^{(3)}, K^{(4)}, \ldots$ are higher-order global stiffness matrices, where $K^{(3)}$ depends linearly on $S^{(2)}$, and likewise, $K^{(4)}$ depends linearly on $S^{(3)}$. The drive term $V^{(\ell)}(t)$ on the right-hand side of (4) results from the time-dependent electrical potential prescribed at the electrodes. It is non-zero only if $\ell$ refers to a node of an element in an electrode.

A dimensionless quantity characterizing the size of the nonlinearity in the systems considered here is the dynamic strain associated with a SAW. Since this quantity is much smaller than one, the application of perturbation theory is justified. When expanding the node variables in powers of a dimensionless expansion parameter $\varepsilon$,

$$\hat{u}_j^{(\ell)} = \varepsilon \hat{u}_j^{(\ell,1)} + \varepsilon^2 \hat{u}_j^{(\ell,2)} + \varepsilon^3 \hat{u}_j^{(\ell,3)} + O(\varepsilon^4),$$

a hierarchy of systems of linear inhomogeneous ordinary differential equations is obtained from (4).

Among the nonlinear phenomena relevant for SAW devices, we shall focus at first on third-order intermodulations (IMD3). The electrical potential of the electrodes is fixed and chosen as a superposition of two input tones $f_1, f_2$,

$$u_4(X,t) = \pm \left( \phi_1 e^{i\omega_1 t} + \phi_2 e^{i\omega_2 t} + \text{c.c.} \right),$$

where $\omega_1 = 2\pi f_1, \omega_2 = 2\pi f_2, \phi_1, \phi_2$ are complex potential values and c.c. stands for the conjugate complex. The potential alternates between neighboring electrodes. Equation (6) implies for the time-dependence of the right-hand side of (4):

$$V^{(\ell)}(t) = \pm \varepsilon \left( \hat{V}_1^{(\ell)} e^{i\omega_1 t} + \hat{V}_2^{(\ell)} e^{i\omega_2 t} + \text{c.c.} \right),$$
where $\hat{V}_1^{(t)}$, $\hat{V}_2^{(t)}$ are non-zero only for nodes belonging to elements in the electrodes. For the time dependence of the node variables at first order up to third order of $\varepsilon$ we may assume a superposition of time-harmonic functions.

At first order of $\varepsilon$, we may decompose

$$\hat{U}_j^{(1)}(t) = \hat{U}_j^{(1,F1)} e^{i\omega_1 t} + \hat{U}_j^{(1,F2)} e^{i\omega_2 t} + c.c.$$  \hspace{1cm} (8)

with time-independent node variables $\hat{U}_j^{(1,F1)}$, $\hat{U}_j^{(1,F2)}$. The symbols $F1$ and $F2$ denote the first and second fundamental (input tone), respectively. The time-independent node variables satisfy the linear inhomogeneous equations

$$\sum_{\ell'} \hat{M}_j^{(\ell'\ell)}(\omega_n) \hat{U}_j^{(\ell',\ell_n)} = \mp \hat{V}_n^{(\ell)},$$  \hspace{1cm} (9)

where $n = 1, 2$ labels the two input tones. In order to account for the constant pre-defined electrical potential on the electrodes, we define two additional matrices $\tilde{M}$ and $\hat{M}$ as follows:

$$\tilde{M}_j^{(\ell'\ell)}(\omega) = M_j^{(\ell'\ell)} \omega^2 - K_j^{(\ell'\ell)}. \hspace{1cm} (10)$$

The matrix $\tilde{M}_j^{(\ell'\ell)}(\omega)$ differs from $M_j^{(\ell'\ell)}(\omega)$ in the following components: For nodes $\ell$ belonging to an electrode, the components $\tilde{M}_j^{(\ell'\ell)}(\omega)$ are zero for $j = 1, \ldots, 4$ and all nodes $\ell'$ except for the diagonal elements $\tilde{M}_{\ell\ell}^{(\ell'\ell)}(\omega)$ which are equal to 1.

At second order of $\varepsilon$, it is sufficient to account for the frequencies $2f_1$ (second harmonic of first input tone, 2H) and $f_1 - f_2$ (low-frequency component, LF) for the computation of the IMD3 generalized displacement field. Decomposing

$$\hat{U}_j^{(2)}(t) = \hat{U}_j^{(2,2H)} e^{2i\omega_1 t} + \hat{U}_j^{(2,LF)} e^{i(\omega_1 - \omega_2) t} + c.c. + \Delta \hat{U}_j^{(2)}(t), \hspace{1cm} (11)$$

where $\Delta \hat{U}_j^{(2)}(t)$ stands for the remaining terms which oscillate with frequencies $2f_2$ and $f_1 + f_2$. The corresponding systems of linear algebraic equations are:

$$\sum_{\ell'} \tilde{M}_j^{(\ell'\ell)}(2\omega_1) \hat{U}_j^{(\ell',2H)} = \frac{1}{2} \sum_{\ell'',\ell'''} \tilde{K}_{\ell j}^{(\ell'\ell''\ell''')} \hat{U}_j^{(\ell',F1)} \hat{U}_j^{(\ell''',F1)}, \hspace{1cm} (12)$$

$$\sum_{\ell'} \tilde{M}_j^{(\ell'\ell)}(\omega_1 - \omega_2) \hat{U}_j^{(\ell',LF)} = \frac{1}{2} \sum_{\ell'',\ell'''} \tilde{K}_{\ell j}^{(\ell'\ell''\ell''')} \hat{U}_j^{(\ell',F1)} \hat{U}_j^{(\ell''',F2)}, \hspace{1cm} (13)$$

modified such that $\hat{U}_j^{(2,2H)}$ and $\hat{U}_j^{(2,LF)}$ vanish at nodes of elements in the electrodes.

At third order of $\varepsilon$, we only consider third-order intermodulations with frequency $2f_1 - f_2$. Therefore, we write

$$\hat{U}_j^{(3)}(t) = \hat{U}_j^{(3,IM3)} e^{i(2\omega_1 - \omega_2) t} + c.c. + \Delta \hat{U}_j^{(3)}(t), \hspace{1cm} (14)$$

where $\Delta \hat{U}_j^{(3)}(t)$ contains all other frequency components of the complete solution at third order of $\varepsilon$.

The system of linear algebraic equations to be solved in order to obtain the generalized displacement field of IMD3 is the following:
\[\sum_{\ell} M_{ij}^{(\ell)} (2\omega_1 - \omega_2) \hat{U}_\nu^{(\ell,IM3)} = \sum_{\ell,\ell'} K^{(3,\ell,\ell')}_{ij} Q^{(\ell,2H)}_{\nu} Q^{(\ell',F2)}_{\nu} + \sum_{\ell,\ell'} K^{(3,\ell,\ell')}_{ij} Q^{(\ell,LF)}_{\nu} Q^{(\ell',F1)}_{\nu} + \frac{1}{2} \sum_{\ell,\ell',\ell''} K^{(4,\ell,\ell',\ell'')}_{ij} Q^{(\ell,F1)}_{\nu} Q^{(\ell',F1)}_{\nu} Q^{(\ell'',F2)}_{\nu}\]

The first two terms on the right-hand side of (15) represent an effective third-order nonlinearity as a result of cascading the second-order nonlinearity, having a quadratic dependence on the generalized tensor \(S^{(2)}\), whereas the third term on the right-hand side depends linearly on \(S^{(3)}\).

It is straightforward to apply this scheme to other nonlinear processes such as higher harmonic generation and nonlinear mixing of three input tones.

As an alternative to the presented derivation of the above equations, one may apply the expansion in powers of \(\epsilon\) directly to the generalized displacement field,

\[u_i(X_1, X_3, t) = \epsilon u_i^{(1)}(X_1, X_3, t) + \epsilon^2 u_i^{(2)}(X_1, X_3, t) + \epsilon^3 u_i^{(3)}(X_1, X_3, t) + O(\epsilon^4),\]

insert this expansion into the Lagrangian (1), and apply the variational principle to each individual order of \(\epsilon\). For the investigation of IMD3 with the two input tones \(f_1, f_2\), the functions \(u_i^{(n)}(X_1, X_3, t), n = 1, 2, 3\), have the form

\[u_i^{(1)}(X_1, X_3, t) = e^{i\omega_1 t} U_i^{(F1)}(X_1, X_3) + c.c.,\]

\[u_i^{(2)}(X_1, X_3, t) = e^{i2\omega_1 t} U_i^{(2H)}(X_1, X_3) + e^{i(\omega_1 - \omega_2)t} U_i^{(LF)}(X_1, X_3) + c.c. + \Delta u_i^{(2)}(X_1, X_3, t),\]

\[u_i^{(3)}(X_1, X_3, t) = e^{i(2\omega_1 - \omega_2)t} U_i^{(IM3)}(X_1, X_3) + c.c. + \Delta u_i^{(3)}(X_1, X_3, t),\]

where \(\Delta u_i^{(2)}(X_1, X_3, t), \Delta u_i^{(3)}(X_1, X_3, t)\) are irrelevant for the determination of the IMD3 generalized displacement field with frequency \(2f_1 - f_2\) at third order of \(\epsilon\).

For systems of type (I) (infinite periodic arrangement of electrodes), the generalized displacement field, associated with the various frequency components, can be set up as Bloch functions, \(e^{-i k x_1} U_j(X_1, X_3)\), \(j = 1, \ldots, 4\), where \(U_j(X_1, X_3) = U_j(X_1 + p, X_3)\) is a \(p\)-periodic function. This allows to confine the calculations to a unit cell with extension along the \(x_1\)-direction equal to the mechanical period \(p\) (pitch) of the system. In this case, the node variables \(\hat{U}_j^{(i,F1)}, \hat{U}_j^{(i,F2)}, \hat{U}_j^{(i,2H)}, \hat{U}_j^{(i,LF)}, \hat{U}_j^{(i,IM3)}\) are the values of the corresponding periodic function \(U_j(X_1, X_3)\) at the position of node \(i\) in the undeformed state of the system. The Bloch–Floquet exponent \(k\) may be chosen to be equal to \(\pi/p\) plus an integer multiple of \(2\pi/p\) for the fields of the fundamentals \((F1, F2)\) and of the third-order intermodulation \((IM3)\). For the fields associated with the second harmonic \((2H)\) and the low-frequency contribution \((LF)\), \(k\) has to be chosen as zero plus an integer multiple of \(2\pi/p\). The stiffness matrices \(K^{(n)}, n = 2, 3, 4, \ldots\) will then depend on the Bloch–Floquet exponents. Viscous damping can be introduced in a straightforward way, leading to the additional term \(-i\omega G_{ij}^{(\ell)}\) on the right-hand side of (10). The matrix \(G\) depends on the viscosity tensor \(\eta\) of each material in the same way as the linear stiffness matrix \(K^{(2)}\) on the tensor of second-order elastic moduli.

3.2. Nonlinear Perfectly Matched Layers

Unless effects of reflections from the bottom of a substrate of finite thickness are the topic of investigation, a semi-infinite substrate is assumed, and bulk-wave reflections are suppressed by placing a perfectly matched layer (PML) below the part of the substrate.
that is included in the simulation domain (Figure 1a,b). When simulating finite arrays of electrodes, additional PMLs are applied as shown in Figure 1b.

Our implementation of linear PMLs largely follows [45]. Complex functions \( h_L = p_L + i q_L, L = 1, 3 \), are defined in each PML with real (imaginary) parts \( p_L, q_L \). The argument of the real functions \( p_L, q_L \) is the Cartesian coordinate \( X_L \). In our implementation we have chosen \( p_1(X_1) = p_3(X_3) = 1 \) in all PMLs and \( q_1(X_1) = 0, q_3(X_3) = \sigma_3 (X_3 - X_3,0)^2 \) for the PML in Figure 1a and for the PML domain VII in Figure 1b, \( q_1(X_1) = \sigma_1 (X_1 - X_1,\text{left})^2, q_3(X_3) = 0 \) for the PML domains I, III, \( q_1(X_1) = -\sigma_1 (X_1 - X_1,\text{right})^2, q_3(X_3) = 0 \) for the PML domains II, IV, \( q_1(X_1) = \sigma_1 (X_1 - X_1,\text{left})^2, q_3(X_3) = \sigma_3 (X_3 - X_3,0)^2 \) for the PML domain V, \( q_1(X_1) = -\sigma_1 (X_1 - X_1,\text{right})^2, q_3(X_3) = \sigma_3 (X_3 - X_3,0)^2 \) for the PML domain VI in Figure 1b with two real parameters \( \sigma_1, \sigma_3 \). Their values must be determined in tests such that reflections are minimized over a sufficiently wide frequency range. For each PML domain, a \( 3 \times 3 \) matrix \( H \) is defined via

\[
H(X_1, X_3) = \begin{pmatrix}
\frac{1}{h_1(X_1)} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & \frac{1}{h_3(X_3)}
\end{pmatrix}
\]

and the quantity \( \bar{h}(X_1, X_3) = h_1(X_1) h_3(X_3) \).

In linear PMLs, the mass density \( \bar{\rho}(X_1, X_3) \) and the components of the generalized fourth-rank tensor \( \bar{C}(X_1, X_3) \) containing the second-order elastic, the piezoelectric and dielectric constants, are defined as

\[
\bar{\rho} = \bar{\rho}_j, \quad \bar{C}_{kMN} = \bar{h} H_{MM'} H_{NN'} C_{kMN'},
\]

where \( \rho \) and \( C \) are the corresponding quantities in the neighboring (piezo-)elastic medium.

In order to avoid reflections at the interface between the PML and the neighboring medium in the presence of nonlinearity (nonlinear PML), generalized tensors \( \tilde{S}^{(n)}(X_1, X_3), n = 2,3,\ldots \), are introduced and defined as

\[
\tilde{S}^{(n)}_{1,1M_1,1,2M_2,\ldots,1,M_n} = \tilde{h} H_{M_1M_1'} H_{M_2M_2'} \cdots H_{M_nM_n'} S^{(n)}_{1,1M_1',1,2M_2',\ldots,1,M_n'},
\]

where \( S^{(0)} \) is the corresponding generalized tensor in the neighboring medium.

In calculations of the IMD2, IMD3, second or third harmonics in an infinitely periodic arrangement of electrodes (Figure 1a) according to the perturbation scheme outlined above, this extension of the PML to nonlinearity is not necessary as long as the generalized displacement fields of the fundamentals are localized at the surface. However, it is needed if at least one of the fundamental waves corresponds to a leaky surface mode or the input frequency of a fundamental is much larger than the resonance frequency. In this case, a linear PML is insufficient, and the Extension (23) does in fact eliminate reflections in the limit of infinitely small element sizes.

### 4. Calculation of Nonlinear Currents

#### 4.1. Line Integrals

The main goal of the FEM simulations discussed here is the computation of the alternating electrical current per unit length of aperture, for each electrode, generated by the third-order intermodulation. This has been achieved via two different approaches. The first consists of carrying out an integral on a closed line in the undeformed media around the electrode. The integrand is the component normal to this line of the electric displacement field, referred to the material frame [20] (called material electric displacement field for short). Within the perturbation scheme applied here, the material electric displacement field \( D_M(X, t), M = 1, \ldots, 3 \), may be written as

\[
D_M(X, t) = e^{(2\omega_1 - \omega_2) t} D_M^{(IM3)}(X) + c.c. + \Delta D_M(X, t),
\]

(24)
\[ D^{(1M3)}_M(X) = C_{4M} nN U_{n,N}^{(1M3)}(X) + S_{4M}^{(2)} nN \frac{\partial}{\partial p} U_{n,N}^{(2H)}(X) U_{p,p}^{(2)}(X) + U_{n,N}^{(LF)}(X) U_{p,p}^{(F1)}(X) \]

(25)

Here, as in (3), the subscript \((n, N)\) denotes the partial derivative of the \(n\)-th component with respect to \(X_N\).

The IMD3 current per unit length of aperture through an electrode, \(I^{(1M3)}(t) = \exp[i(2\omega_1 - \omega_2)t] I^{(1M3)} + c.c.,\) is calculated by evaluating with FEM the integral

\[ \tilde{I}^{(1M3)}(t) = i(2\omega_1 - \omega_2) \int D^{(1M3)}_M(X) \tilde{N}_M(X) \, ds \]

over a closed path around the electrode. Here, \(\tilde{N}_M(X), M = 1, 2, 3,\) are the three Cartesian components of a unit vector, pointing in the outward direction normal to the closed path at point \(X\).

4.2. Overlap Integrals

The following alternative approach does not require the computation of the field \(U_n^{(1M3)}(X), m = 1, \ldots, 4,\) but only involves the fields occurring at first and second order of \(\varepsilon\) in our perturbation scheme, and a reference field that may be regarded as a third fundamental corresponding to the IMD3 frequency \(f_3 = \omega_3/(2\pi),\) where \(\omega_3 = 2\omega_1 - \omega_2.\)

We first consider the case of a finite number \(N_E\) of electrodes. The reference field is a solution of the linear field equations and boundary conditions of the form \(e^{i\omega_3 t} \phi_3^{(l)}(X_1, X_3)\) generated by a set of time-dependent potential values \(e^{i\omega_3 t} \phi_3^{(l)}, l = 1, \ldots, N_E.\) This means that \(U_4^{(F3)}(X_1, X_3) = \phi_3^{(l)}\) for spatial points \((X_1, X_3)\) in the \(l\)-th electrode.

The four-component IMD3 field \(U_m^{(1M3)}(X), m = 1, \ldots, 4,\) has to satisfy the nonlinear equation of motion and Poisson’s equation,

\[-\rho \Delta_{mm} \omega_3^2 U_n^{(1M3)} + \frac{\partial}{\partial X_M} (C_{mm} nN + i\omega_3 \eta_{mm} nN) U_n^{(1M3)} = -\frac{\partial}{\partial X_M} K_{mm}.\]

(27)

For the nonlinear driving term on the right-hand side of (27), we have introduced the abbreviation

\[ K_{mm}(X) = S_{mm}^{(2)} nN \frac{\partial}{\partial p} \left[ U_n^{(2H)}(X) U_p^{(2)}(X) + U_n^{(LF)}(X) U_p^{(F1)}(X) \right] + \frac{1}{2} S_{mm}^{(3)} nN \frac{\partial}{\partial p} U_n^{(F1)}(X) U_p^{(F2)}(X) U_n^{(F2)}(X).\]

(28)

Both sides of (27) are now multiplied by \(U_m^{(F3)}(X_1, X_3),\) summed over \(m\) and integrated over the area \(A\) in the \((X_1, X_3)\)-plane filled by elastic media (Figure 1b). Applying integration by parts, the resulting equation may be transformed into

\[ \omega_3^2 \int_A \rho U_m^{(F3)} \Delta_{mn} U_n^{(1M3)} dX_1 dX_3 + \int_A U_m^{(F3)} (C_{mm} nN + i\omega_3 \eta_{mm} nN) U_n^{(1M3)} dX_1 dX_3 = -I_{NL} + I_Q.\]

(29)

The right-hand side of (29) consists of the overlap integral

\[ I_{NL} = \int_A U_m^{(F3)} K_{mm} dX_1 dX_3,\]

(30)
which extends over all media of the total system, and a sum of line integrals over the boundaries of all electrodes

\[
I_Q = \sum_{l=1}^{N_L} \oint_{L_l} U_4^{(F3)} \left[ C_{4M} nN U_{n,N}^{(1M3)} + K_{4M} \right] \hat{N}_M^{(I)} ds. \tag{31}
\]

Here, \( \hat{N}_M^{(I)} \) is the \( M \)-th Cartesian component of a unit vector normal to the boundary \( L_l \) of the \( l \)-th electrode, pointing into the outward direction. All other boundary terms occurring in the integration by parts vanish because of the boundary conditions satisfied by the fields \( U_m^{(1M3)} \) and \( U_m^{(F3)} \), \( m = 1, \ldots, 4 \). These are the continuity of these fields at interfaces; the traction

\[
T_m^{(1M3)}(X) = \hat{N}_M(X) \left[ (C_{mM} nN + i\omega_3 \eta_{mM} nN) U_{n,N}^{(1M3)}(X) + K_{mM}(X) \right], \tag{32}
\]

\( m = 1, 2, 3 \), and the normal component of the material electric displacement field

\[
\hat{N}_M(X) D_M^{(1M3)}(X) = \hat{N}_M(X) \left[ C_{4M} nN U_{n,N}^{(1M3)}(X) + K_{4M}(X) \right] \tag{33}
\]

are continuous at interfaces and vanish at surfaces, except for the interfaces between the metal electrodes and the piezoelectric substrate and the dielectric film. Here the discontinuity of the normal component of \( D^{(1M3)} \) leads to the boundary term (31). Also, we use the fact that the field gradients vanish at infinity because of the presence of damping, in particular due to the PMLs. In (32) and (33), \( \hat{N}(X) \) is a vector normal to the corresponding interface or surface at position \( X \).

Another integration by parts of the second integral on the left-hand side of (29) transforms (29) into

\[
\prod_{\Lambda} U_n^{(1M3)} \left[ \frac{\omega_3^2 \rho^2 \Delta_{nm} U_m^{(F3)}}{\omega^2_{F3}} - \frac{\partial}{\partial x_n} (C_m nN + i\omega_3 \eta_{mM} nN) U_m^{(F3)} \right] dX_1 dX_3 = -I_{NL} + I_Q. \tag{34}
\]

From the equation of motion and Poisson’s equation, satisfied by the generalized displacement field \( U_m^{(F3)} \), we may conclude that the left-hand side of (34) vanishes. In the second integration by parts, no boundary terms remain because of the boundary conditions satisfied by \( U_m^{(1M3)} \) and \( U_m^{(F3)} \), \( m = 1, \ldots, 4 \). Especially, use is made of the fact that

\[
T_m^{(F3)}(X) = \hat{N}_M(X) (C_{mM} nN + i\omega_3 \eta_{mM} nN) U_{n,N}^{(F3)}(X), \tag{35}
\]

\( m = 1, 2, 3 \), and the normal component of the material electric displacement field

\[
\hat{N}_M(X) D_M^{(F3)}(X) = \hat{N}_M(X) C_{4M} nN U_{n,N}^{(F3)}(X) \tag{36}
\]

are continuous at interfaces and vanish at surfaces, apart from the interfaces between the metal electrodes and the surrounding media. Here, \( \hat{N}_M D_M^{(F3)} \) has a discontinuity. However, this does not generate any additional boundary terms since the discontinuity is multiplied by \( U_4^{(1M3)} \), which vanishes on the metal electrodes according to the boundary conditions we have imposed on our FEM solution.

Since the electric potential is spatially constant within an electrode, the line integral around the \( l \)-th electrode in (31) becomes

\[
\phi_3^{(I)} \oint_{L_l} \left[ C_{4M} nN U_{n,N}^{(1M3)} + K_{4M} \right] \hat{N}_M^{(I)} ds, \tag{37}
\]
which is equal to the product of $\phi_3^{(l)}$ and the charge per unit length of aperture, $Q_{l}^{(IM3)}$, on the $l$-th electrode, associated with IMD3. Equation (34) together with (30) finally yields

$$\sum_{l=1}^{N_e} \phi_3^{(l)} Q_{l}^{(IM3)} = \int_A U_{m,M}^{(F3)} K_{mM} \, dX_1 dX_3.$$  (38)

The current $\tilde{I}_l^{(IM3)}$ per unit length of aperture flowing into or out of the $l$-th electrode can be determined with the relation $\tilde{I}_l^{(IM3)} = i\omega_3 Q_{l}^{(IM3)}$.

In order to obtain the total IMD3 current per unit length of aperture, flowing into one busbar, in one computation, a linear field $U_{m}^{(F3)}$, $m = 1, \ldots, 4$, in (38) has to be generated by a potential distribution with potential amplitudes $\phi_3^{(l)}$ having the same constant value on all electrodes connected with one of the two bus bars. The potential amplitudes on the electrodes connected to the other bus bar have to be zero.

On the other hand, if one wants to calculate the IMD3 current in one single selected electrode $l_0$, the potential amplitude $\phi_3^{(l_0)}$ must be different from zero, while $\phi_3^{(l)} = 0$ for $l \neq l_0$.

Charge neutrality is guaranteed as long as open circuit boundary conditions are chosen at the surfaces of the system. This follows from (38) and the fact that the displacement components and the frequency of the reference field do not change if a time-dependent, but spatially constant term is added to its electric potential. An advantage of determining the total IMD3 or H3 current in a finite system with a periodic sequence of electrodes by using the overlap integral method rather than the standard line integral approach is the following. Unlike the IMD3 field $U^{(IM3)}$ or the H3 field $U^{(H3)}$, the reference field $U^{(F3)}$ is periodic in the periodic part of the finite system. Therefore, it can be computed by fully cascading the periodic cells in the same way as the input fields, which leads to an additional speed-up of the calculations.

The overlap integral approach to the determination of the IMD3 current is applied to the infinite periodic system of Figure 1a in a straightforward way. Making use of the fact that both the IMD3 field and the reference field are Bloch functions with Bloch–Floquet wavenumber $k = \pi / p$, the resulting overlap integral has to be extended over one mechanical period of the system, only.

4.3. IMD3 and Third-Harmonic Current for the Infinite Periodic System

Figure 2 shows the IMD3 current in a TC-SAW system with an infinite number of periodically arranged electrodes. The distance between the two input tones $f_1$ and $f_2$ was kept fixed at 6 MHz. The plots exhibit the typical triple-peak. In Figure 3, the current in the electrodes due to the third harmonic (H3) is shown. The mechanical period (pitch $p$) is 2 $\mu$m, the thickness of the SiO$_2$ film is 1.385 $\mu$m, the metallization ratio is 50% and the electrode height is 287 nm. The linear admittance of this system is shown in Figure 4 for two different sets of second-order elastic moduli (Lamé constants) of the SiO$_2$ film. The constants provided in [37] were measured in films, whereas those of [39] refer presumably to bulk material. This leads to a difference of about 7 MHz in the resonance frequency and also a shift of the feature at about 910 MHz, which corresponds to a mode of predominantly shear-horizontal polarization.
Figure 2. IMD3 current as function of frequency $f_3 = 2f_1 - f_2$. Tone distance $f_1 - f_2 = 6$ MHz. Blue: Second-order and third-order elastic constants for SiO$_2$ film from [39], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], $S^{(3)}$ for LiNbO$_3$ equal to zero. Black dotted: Second-order and third-order elastic constants for SiO$_2$ film from [37,38], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], $S^{(3)}$ for LiNbO$_3$ equal to zero. Red: Second-order and third-order elastic constants for SiO$_2$ film from [37,38], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], in $S^{(3)}$ for LiNbO$_3$ only second-order and third-order elastic constants accounted for [35,36]. Green: Second-order and third-order elastic constants for SiO$_2$ film from [37,38], $S^{(2)}$ and $S^{(3)}$ for LiNbO$_3$ equal to zero.

Figure 3. H3 current as function of third-harmonic frequency $3f_1$. Blue (approx. equal to red curve): Second-order and third-order elastic constants for SiO$_2$ film from [39], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], $S^{(3)}$ for LiNbO$_3$ equal to zero. Black: Second-order and third-order elastic constants for SiO$_2$ film from [37,38], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], $S^{(3)}$ for LiNbO$_3$ equal to zero. Red: Second-order and third-order elastic constants for SiO$_2$ film from [37,38], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], in $S^{(3)}$ for LiNbO$_3$ only second-order and third-order elastic constants accounted for [35,36]. Blue crossed: Second-order and third-order elastic constants for SiO$_2$ film from [39], material constants for LiNbO$_3$ in $S^{(2)}$ from [35,36], in $S^{(3)}$ for LiNbO$_3$ all second-order and third-order material constants [35,36] accounted for. Green: Second-order and third-order elastic constants for SiO$_2$ film from [39], $S^{(2)}$ and $S^{(3)}$ for LiNbO$_3$ equal to zero.
In order to investigate the sensitivity of the nonlinear current to the material constants as input data, results of calculations are compared that were carried out with the TOECs of the isotropic SiO$_2$ film taken from two different sources [38,39]. Together with the TOECs of [39], we used the second-order elastic constants provided in the same publication, whereas the second-order constants of [37] were used together with the TOECs of [38]. For both IMD3 and H3, the differences of nonlinear currents computed with the two different sets of constants for fused quartz are mainly related to the shift in the linear resonance frequency seen in the linear admittance curves. In these calculations, the generalized tensor $S^{(3)}$ was set equal to zero for the substrate material LiNbO$_3$ (blue and black curves in Figures 2 and 3), because a large number of material constants entering $S^{(3)}$ are unknown for LiNbO$_3$ (blue curves in Figures 2 and 3).

In addition, the IMD3 and H3 currents were calculated with a generalized tensor $S^{(3)}$ for LiNbO$_3$ that only contains the contributions of the second-order and third-order elastic constants, while all other material constants in the expressions for $S^{(3)}$ (given explicitly in the Appendix) were set equal to zero (red solid curves in Figures 2 and 3). In Figure 3, referring to H3, the blue and red curves are virtually indistinguishable. In the case of H3, a further calculation was carried out with the second-order constants of [35] and all third-order constants [36] of LiNbO$_3$ entering $S^{(3)}$, including those associated with the electric field (dashed-dotted red curve in Figure 3). The results suggest that the generalized tensor of third-order nonlinearity ($S^{(3)}$) for LiNbO$_3$ seems to play a largely negligible role for IMD3 and H3 currents. In fact, the black, solid red and dashed red curves in Figure 3 are not distinguishable without strong magnification. When setting the tensor of second-order nonlinearity, $S^{(2)}$, and of third-order nonlinearity, $S^{(3)}$, both equal to zero for LiNbO$_3$ (green curves in Figures 2 and 3), the resulting IMD3 current differs insignificantly from the cases discussed before in the neighborhood of the resonance peaks (triple-peak). This applies even more to the H3 current and suggests that in these frequency ranges, the nonlinearity of LiNbO$_3$ is largely negligible. However, at frequencies much below the resonance peaks, corresponding to excited fields of predominantly electrical nature, differences between the green and red curves are visible. In this frequency range, the unknown fourth-order material constants of LiNbO$_3$ associated with the electric field may also play an appreciable role.

5. Nonlinear Amplitude and Phase Response of a SAW Resonator

An important, normally undesired effect occurring in oscillators with effective third-order nonlinearity is the bending of the resonance curve, often termed the Duffing effect,
which can lead to bistability. It may be demonstrated with the help of the Duffing equation [46]
\[ \frac{\partial^2}{\partial t^2}u + r \frac{\partial}{\partial t}u + \omega_0^2 u + \beta u^3 = g \cos(\omega t), \] (39)
where \( \omega_0/(2\pi) \) is the resonance frequency in the linear limit without damping, \( r \) is a damping constant, \( \beta \) is the pre-factor of the third-order nonlinearity, and \( \omega/(2\pi) \) and \( g \), respectively, are the frequency and strength of the drive.

Using asymptotic perturbation theory (see for example [47]), it is straightforward to derive a first-order differential equation that describes the behavior of the Duffing oscillator for drive frequencies in the neighborhood of the linear resonance frequency. For this purpose, an expansion parameter \( \epsilon \) and a “stretched” time coordinate \( \tau = \epsilon^2 t \) are introduced. Also, quantities \( \gamma, F \) and \( \Omega \) are defined via \( r = 2\epsilon^2 \gamma, g = 4\epsilon^3 \omega_0 F, \omega = \omega_0 + \epsilon^2 \Omega \), and are assumed to be of order \( O(1) \) in the expansion parameter \( \epsilon \). Expanding \( u \) in powers of \( \epsilon \),
\[ u(t) = \epsilon \left[ A(\tau) e^{i\omega_0 \tau} + \text{c.c.} \right] + O(\epsilon^2), \] (40)
the following evolution equation is obtained for the complex amplitude \( A(\tau) \):
\[ i \left( \frac{d}{d\tau} + \gamma \right) A(\tau) + N |A(\tau)|^2 A(\tau) = F e^{i\Omega \tau} \] (41)
with the coefficient \( N = 3\beta/(2\omega_0) \) in the nonlinear term. (For more details of this approach see Chapter 6 in [16] and references therein).

The evolution equation (41) also holds for the complex amplitude \( A(\tau) \) of the generalized displacement field \( u_j(X_1, X_3, t) \), \( j = 1, \ldots, 4 \), in a SAW resonator in the neighborhood of the resonance frequency \( f_0 = \omega_0/(2\pi) \). To show this, we write this field as an asymptotic expansion (16) and introduce, as before, the “stretched” time coordinate \( \tau = \epsilon^2 t \). On the \( l \)-th electrode we impose the electric potential
\[ u_4(X_1, X_3, t) = \epsilon^3 \phi^{(l)} e^{i\omega_0 \tau} + \text{c.c.} = \epsilon^3 \phi^{(l)} e^{i\Omega \tau} e^{i\omega_0 \tau} + \text{c.c.} \] (42)
with drive frequency \( f = \omega/(2\pi) \) close to the resonance frequency \( f_0, \omega - \omega_0 = \epsilon^2 \Omega \).

The quantities \( \Omega \) and \( \phi^{(l)} \) are of order \( O(1) \) in the expansion parameter \( \epsilon \). For the viscosity tensor, we apply the scaling \( \eta = \epsilon^2 \tilde{\eta} \) with the tensor \( \tilde{\eta} \) being of order \( O(1) \).

The first-order term,
\[ u_j^{(1)}(X_1, X_3, t) = \epsilon^3 \phi^{(l)} U_j^{(F)}(X_1, X_3) A(\tau) + \text{c.c.} \] (43)
is a solution of the linear boundary value problem following from the Lagrangian (1) obeying the condition that the electric potential on the electrodes vanishes. It is easily determined by FEM by solving a generalized eigenvalue problem, where \( \omega_0^2 \) is the eigenvalue and \( U_j^{(F)}(X_1, X_3) \) is the corresponding eigenvector, normalized in a suitable way. A convenient normalization adopted in the following is
\[ 2\omega_0 \int_A \rho(X_1, X_3) U_j^{(F)*}(X_1, X_3) \Delta_{jj'} U_{j'}^{(F)}(X_1, X_3) dX_1 dX_3 = 1 \] (44)

In the same way as in Section 4, the double integrals in (44) and in the following have to be carried out over the area \( A \) filled with material in the finite system (Figure 1b) or in a unit cell of the infinite periodic system (Figure 1a). To ensure the convergence of the integrals occurring in this derivation, we require that the field \( U^{(F)} \) is fully surface-localized. In the case of a finite system of electrodes, we assume, in addition, that the geometry of the system is such that \( U^{(F)} \) decays exponentially into the positive and negative \( X_1 \)-direction or is spatially confined to a stripe with finite extension along the \( X_1 \)-axis. In the case of an
infinite periodic arrangement of electrodes (Figure 1a), area integrals of the kind (44) need to be extended over a unit cell only, as mentioned above.

The second-order field consists of two parts, the second harmonic \((2H)\) and a static contribution \((S)\),

\[
u_j^{(2)}(X_1, X_3, t) = e^{i\omega_0 t} U_j^{(2H)}(X_1, X_3) \{A(\tau)\}^2 + c.c. + U_j^{(S)}(X_1, X_3) |A(\tau)|^2. \tag{45}
\]

Both contributions can be determined by applying FEM and solving an inhomogeneous system of linear equations.

The third-order field \(u_j^{(3)}(X_1, X_3, t), j = 1, \ldots, 4\), contains a contribution of the form

\[
e^{i\omega_0 t} U_j^{(3)}(X_1, X_3, \tau). \tag{46}
\]

The equation of motion and Poisson’s equation, obeyed by this contribution, has the following form:

\[
\rho \Delta_m \omega_0^2 U_n^{(3)} = \frac{\partial}{\partial X_M} C_{mM} nN U_{nN}^{(3)} U_{nN}^{(3)} + i \omega_0 \frac{\partial}{\partial X_M} \tilde{\eta}_{mM} nN U_{nN}^{(3)} A(\tau) + \frac{\partial}{\partial X_M} \tilde{\eta}_{mM} |A(\tau)|^2 A(\tau)
\]

where

\[
\tilde{\eta}_{mM}(X) = \sum_{\phi}^{(2)} nN_{pP} \left[ U_{nN}^{(2H)}(X) U_{pP}^{(2)}(X) + U_{nN}^{(S)}(X) U_{pP}^{(S)}(X) \right] + \sum_{\phi}^{(3)} nN_{pP} qQ U_{nN}^{(1)}(X) U_{pP}^{(1)}(X) U_{qQ}^{(1)}(X).
\]

Following [48], (46) is projected on the field \(U^{(F)}\); i.e., both sides of (46) are multiplied by \(U_m^{(f)}(X_1, X_3)\), summed over \(m\) and integrated over \(X_1, X_3\). In the same way as in the derivation of (38) in the previous section, two integrations by parts are subsequently carried out and the boundary conditions satisfied by the fields \(U^{(3)}\) and \(U^{(F)}\) are made use of. Especially, one has to account for the jump of the component of the material electric displacement field \(D^{(F)}\) normal to the electrode boundaries, where

\[
D_M^{(F)} = C_{4M} nN U_{nN}^{(F)}, \tag{48}
\]

and the imposed electric potential (42). In the integrations by parts, all boundary terms vanish except for the line integrals

\[
\oint_{L_l} \phi_3^{(l)} D_M^{(F)} N_M^{(l)} ds = \phi_3^{(l)} Q_l^{(F)}, \tag{49}
\]

\(l = 1, \ldots, N_E\), over the boundaries \(L_l\) of the \(N_E\) electrodes. Here, \(Q_l^{(F)}\) is the charge per unit length of aperture on the \(l\)-th electrode, associated with the field \(U^{(F)}\).

This procedure yields the evolution Equation (41) for the amplitude \(A(\tau)\) with the following expressions for the coefficients \(F, \gamma\) and \(N\) in terms of the fields \(U^{(F)}, U^{(2H)}, U^{(S)}\), which are evaluated with our FEM code. The strength of the drive term is provided by

\[
F = \sum_{l=1}^{N_E} \phi_3^{(l)} Q_l^{(F)}, \tag{50}
\]

with (49) and (48), and the damping constant by

\[
\gamma = \omega_0 \int_A U_{mM}^{(F)} \tilde{\eta}_{mM} nN U_{nN}^{(F)} dX_1 dX_3. \tag{51}
\]
The coefficient \( N \) in front of the nonlinear term in (41) may be decomposed into three parts,
\[
N = N^{(2H)} + N^{(S)} + N^{(D)},
\]
which all have the form of overlap integrals. The first term results from the second harmonic,
\[
N^{(2H)} = \int_A S^{(2)}_{mM nN pP} U^{(F)}_{m,M} \overline{U^{(F)}_{n,N}} \overline{U^{(F)}_{p,P}} dX_1dX_3.
\]
(53)

The second term is related to the static part of the second-order field,
\[
N^{(S)} = \int_A S^{(2)}_{mM nN pP} U^{(F)}_{m,M} U^{(S)}_{n,N} U^{(F)}_{p,P} dX_1dX_3.
\]
(54)

The third term is the direct contribution of third-order nonlinearity,
\[
N^{(D)} = \frac{1}{2} \int_A S^{(3)}_{mM nN pP qQ} U^{(F)}_{m,M} \overline{U^{(F)}_{n,N}} U^{(F)}_{p,P} U^{(F)}_{q,Q} dX_1dX_3.
\]
(55)

While \( N^{(S)} \) and \( N^{(D)} \) are real, the contribution \( N^{(2H)} \) of the second harmonic is ordinarily complex due to radiation of the second harmonic into the bulk. We shall denote the real part of the total effective nonlinearity parameter \( N \) by \( N' \) and its imaginary part by \( N'' \). In an infinite periodic system, bulk waves are radiated at the second harmonic frequency into the substrate with wave-vector vertical to the surface. The reason for their generation is explained in Figure 5. In an infinite periodic structure, the Bloch–Floquet exponent of the generalized displacement field, associated with the second harmonic of a Brillouin zone boundary excitation, is equal to zero in the reduced zone scheme. Consequently, it consists of bulk waves with wave-vectors vertical to the surface (blue arrow in the right subfigure of Figure 5). In the case of the third harmonic, the magnitude of the Bloch–Floquet exponent in the reduced zone scheme is equal to \( \pi / p \). Therefore, the generalized displacement field associated with the third harmonic contains bulk waves with wave-vectors inclined to the surface normal (red arrows in the right subfigure of Figure 5). The inclination angles of these wave-vectors may be obtained from the sections of the \((X_1,X_3)\)-plane and the corresponding sheets of the slowness surface of bulk waves in the substrate material.

**Figure 5.** Mechanism of nonlinear bulk wave generation by a surface wave at the second- and third-harmonic frequency.

Bulk-wave radiation at the second and third harmonic frequencies in a SAW resonator was verified experimentally in [49]. For an infinite periodic system, it may be shown in a way similar to the derivation of Equation (6.71) in [16] that the imaginary part of \( N^{(2H)} \) is related to the power flow of the second harmonic away from the electrodes and that it must be positive. Unless the bottom of the substrate reflects the bulk waves at the second harmonic frequency back to the electrodes and their damping is negligible, the imaginary part of \( N^{(2H)} \) is larger than zero. In Table 1, the real part of the effective nonlinearity coefficient \( N \) and the relative sizes of the second harmonic contribution (both real and
imaginary parts), of the static and of the direct contributions are provided for a TC-SAW system. The normalization

$$\int_A U_m^{(F)} \Delta_{mn} U_n^{(F)} dX_1 dX_3 = 1$$

(56)

was chosen. The data show that there is a significant dependence of the nonlinearity coefficient on the metallization ratio. The direct contribution via the generalized tensor $S^{(3)}$ plays a dominant role.

**Table 1.** Real part ($N'$) of coefficient $N$ in the nonlinear term of Equation (41), and relative contributions of second harmonic ($N^{(2H)}$), static field ($N^{(S)}$) and third-order nonlinearity ($N^{(D)}$) to $N$. $\eta$: metallization ratio, $f_r$: resonance frequency. a: determined with generalized tensor $S^{(3)}$ for LiNbO$_3$ that contains the contributions of the second-order and third-order elastic constants, while the FOECs were set equal to zero. b: determined with $S^{(3)} = 0$ for LiNbO$_3$.

<table>
<thead>
<tr>
<th>$N'$ (a.u.)</th>
<th>$N^{(2H)}/N'$</th>
<th>$N^{(S)}/N'$</th>
<th>$N^{(D)}/N'$</th>
<th>$\eta$</th>
<th>$f_r$(MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.21</td>
<td>$-0.44 + 1.30i$</td>
<td>$-1.19$</td>
<td>$2.64$</td>
<td>$0.375$</td>
<td>885</td>
</tr>
<tr>
<td>4.08</td>
<td>$-0.81 + 2.13i$</td>
<td>$-1.68$</td>
<td>$3.48$</td>
<td>$0.5$</td>
<td>875</td>
</tr>
<tr>
<td>3.09</td>
<td>$-0.98 + 2.90i$</td>
<td>$-2.14$</td>
<td>$4.12$</td>
<td>$0.625$</td>
<td>870</td>
</tr>
<tr>
<td>2.33</td>
<td>$-1.11 + 3.46i$</td>
<td>$-2.84$</td>
<td>$4.95$</td>
<td>$0.75$</td>
<td>869</td>
</tr>
<tr>
<td>7.04</td>
<td>$-0.47 + 1.23i$</td>
<td>$-0.97$</td>
<td>$2.44$</td>
<td>$0.5$</td>
<td>875</td>
</tr>
</tbody>
</table>

We note that the above considerations also apply to BAW resonators. For solidly mounted resonators with the piezoelectric layer between upper and lower electrode placed on a Bragg reflector, the coefficient $N$ can acquire an imaginary part if the second harmonic of the linear resonance does not fall into a stop band of the Bragg reflector.

For frequencies in the neighborhood of the resonance frequency, the shape of the resonance curve and the phase response follows from (41). Inserting the stationary Ansatz $A(\tau) = A_0 \exp(i\Omega\tau)$ in (41) yields a cubic equation in $|A_0|^2$, which may be brought into the form

$$\Omega = \pm \sqrt{\left(|F|^2 / |A_0|^2\right) - \left(\gamma + N'' |A_0|^2\right)^2 + N'|A_0|^2}.$$ 

(57)

From (57), one readily deduces that the sign of the real part $N'$ of the complex coefficient $N$ decides on whether the resonance curve is inclined to negative ($N' < 0$) or positive ($N' > 0$) frequencies. The amplitude $|A_0|$ reaches its maximum when the argument of the square root in (57) vanishes. Consequently, the presence of a positive imaginary part $N''$ of $N$ reduces this maximum in comparison to the linear case. The effect of the real and imaginary parts of $N$ on the amplitude and phase response of a resonator is shown schematically in Figure 6.

The phase response of the resonator is affected by both real and imaginary parts of the coefficient $N$, too. Choosing $F$ to be real and writing $A_0 = |A_0| \exp(i\alpha)$, we obtain for the phase angle $\alpha$:

$$\tan \alpha = \left(\gamma + N'' |A_0|^2\right) / \left(\Omega - N'|A_0|^2\right).$$

(58)
6. Reflection of Bulk Waves from the Bottom of the Substrate

As mentioned in the previous subsection, even an infinite periodic arrangement of electrodes generates bulk waves as part of second and third harmonics of fundamentals that are surface waves. If the bottom surface of the substrate is sufficiently smooth, these bulk waves undergo specular reflection, travel back to the upper surface and influence the current in the electrodes. This has been verified experimentally by Solal et al. [48]. In order to simulate this phenomenon, it is desirable to restrict computations to an area in the (X1, X3)-plane that is not larger in depth than the area used in FEM simulations of linear SAWs (h in Figure 7). However, in practice, the distance between the two surfaces of the sample (H in Figure 7) is much larger, H >> h. In order to bypass this problem, one can take advantage of the fact that below a certain depth, the generalized displacement field at the second-harmonic or third-harmonic frequency of a fundamental surface wave is approximately that of pure bulk waves, and apply an impedance boundary condition at the bottom of a simulation area of thickness h. This is demonstrated in the following for bulk wave radiation of the second harmonic frequency in a system with an infinite periodic arrangement of electrodes.

Figure 6. Amplitude and phase response of a SAW resonator, schematically. Black: linear limit, blue: real effective third-order nonlinearity parameter N, red: complex N.

Figure 7. Application of impedance boundary condition at lower edge of simulation area in order to account for reflections from bottom of sample.
In this system, the generalized displacement field associated with the second harmonic, \( \exp(2i\omega_1 t)U^{(2H)}_j(X_1, X_3) \), is a \( p \)-periodic function of \( X_1 \) and may therefore be written as a Fourier series,

\[
U^{(2H)}_j(X_1, X_3) = \sum_{n=-\infty}^{\infty} \left[ e^{2i\pi n X_1 / p} V^{(n)}_j(X_3) \right].
\]  

(59)

In the substrate, the functions \( V^{(n)}_j(X_3) \) are of the form

\[
V^{(n)}_j(X_3) = \sum_{r=1}^{N_r} w^{(n,r)}_j e^{iq^{(n,r)}(X_3+h)} A^{(n,r)} + \delta_m \delta_4 (\varphi_0 + \varphi_1(X_3 + h))
\]

(60)

with generalized four-component polarization vectors \( w^{(n,r)}_j \), normalized in a suitable way, wave-vector components \( q^{(n,r)} \) and coefficients \( A^{(n,r)}, \varphi_0, \varphi_1 \). For \( n \neq 0 \) the number \( N_r \) of partial waves in the sum on the right-hand side of (60) is eight, while in the case \( n = 0 \) it is equal to six. The polarization vectors and wave-vector components follow from the Christoffel equation for the substrate, extended to account for piezoelectricity. The coefficients \( A^{(n,r)}, \varphi_0, \varphi_1 \) are unknown.

Depending on the value of \( \omega_1 \), the functions \( V^{(n)}_j(X_3) \) are localized at the surface, i.e., they decay exponentially into the substrate, or they contain plane wave components that radiate energy into the bulk of the substrate and from the bottom of the substrate back to the top surface. For frequencies \( \omega_1 / (2\pi) \) smaller or close to the resonance frequency, it is only the Fourier component \( n = 0 \) in (60) that contains bulk wave contributions. At the bottom of the substrate \( (X_3 = -H) \), boundary conditions must be specified. As an example, we require that the bottom surface is traction-free and the electric potential is zero. This yields the four equations

\[
\sum_{r=1}^{6} C_m k_3 w^{(0,r)}_k \exp \left( i q^{(0,r)}(h - H) \right) i q^{(0,r)} A^{(0,r)} + C_m 43 \varphi_1 = 0,
\]

\[ m = 1, 2, 3, \]

(61)

\[
\sum_{r=1}^{6} w^{(0,r)}_4 \exp \left( i q^{(0,r)}(h - H) \right) A^{(0,r)} + \varphi_0 + \varphi_1(h - H) = 0
\]

(62)

At \( X_3 = -h \), the generalized four-component displacement field of the second harmonic and its first derivative with respect to \( X_3 \) is continuous, which yields the eight equations

\[
U^{(2H)}_j(X_1, -h) = \sum_{r=1}^{N_r} w^{(0,r)}_j A^{(0,r)} + \delta_j \delta_4 \varphi_0
\]

(63)

\[
\left[ \frac{\partial}{\partial X_3} U^{(2H)}_j(X_1, X_3) \right]_{X_3 = -h} = \sum_{r=1}^{N_r} w^{(0,r)}_j i q^{(0,r)} A^{(0,r)} + \delta_j \delta_4 \varphi_1
\]

(64)

for \( j = 1, \ldots, 4 \). The depth \( h \) was chosen to be sufficiently large, such that the \( X_1 \)-dependence of \( U^{(2H)}_j \) may be neglected. Solving (61)–(63) for eight variables \( A^{(0,r)}, \varphi_0, \varphi_1 \) and inserting this solution in (64) yields an effective boundary condition at \( X_3 = -h \) of the form

\[
\left[ \frac{\partial}{\partial X_3} U^{(2H)}_j(X_1, X_3) \right]_{X_3 = -h} = M_{jk}(H - h) U^{(2H)}_k(X_1, -h)
\]

(65)

with a frequency-dependent matrix \( M \), which follows from the solution of a system of eight linear equations.

This procedure can be extended to Fourier components \( n \neq 0 \) in a straightforward way. Figure 8 shows the three components of the displacement field associated with the
second harmonic at a fixed depth. They were computed with the boundary condition (65). For two different thicknesses of the substrate, typical Fabry–Perot type resonances are visible. For a rough estimate of the periodicity of the resonance peaks in the case of the displacement component $U_3$, one may argue as follows. The distance between neighboring frequencies of plate vibrations with displacements predominantly vertical to the surfaces is approximately $v_L/(2L)$, where $L$ is the thickness of the plate and $v_L$ is the phase velocity of quasi-longitudinal bulk waves with wave vector vertical to the surfaces of the plate. In the plots of the displacement field associated with the second harmonic as a function of the fundamental frequency $f_1$, we therefore expect a distance $\Delta f_1 \approx v_L/(4H)$ between neighboring resonance peaks of the same kind for the displacement component $U_3$. With the velocity of quasi-longitudinal waves in the substrate (LiNbO$_3$-rot128YX) with wave-vector in the $X_3$-direction $v_L = 7091$ m/s, we expect $\Delta f_1 \approx 30$ MHz for $H = 59.24$ µm and $\Delta f_1 \approx 8.5$ MHz for $H = 209.24$ µm. This corresponds well with the data in Figure 8. The reflected quasi-transverse waves, polarized predominantly in the $X_2$-direction, and purely transverse waves, polarized in the $X_1$-direction, together with the quasi-longitudinal waves give rise to complex interference patterns, since all three components are coupled in the spatial domain close to the surface.

The Fabry–Perot resonances in the generalized displacement field of the second harmonic influences the IMD3 current as shown in Figure 9.

Figure 8. Displacement field of second harmonic at fixed depth ($X_3 = -8.3$ µm). Blue: $U_1$; green: $U_2$; red: $U_3$. Substrate wafer thickness 58.3 µm (top) and 208.3 µm (bottom).

Figure 9. IMD3 current as function of IMD3 frequency $f_3 = 2f_1 - f_2$. Distance between input tones $f_1$, $f_2$: 2 MHz. Blue: total IMD3 current; red: contribution of the second harmonic of input tone $f_1$ for substrate thickness 208.3 µm; green: total IMD3 current for semi-infinite substrate.
At frequencies $f_1$ much larger than the resonance frequency, where terms of the Fourier series (59) with $n \neq 0$ are not localized at the surface, the corresponding bulk waves propagate away from the surface at an angle different from 90°. After being reflected at the bottom of the substrate, they will travel back and reach the top surface outside the area covered by electrodes in the case of a finite device. Nevertheless, they may influence the IMD3 current in this device.

7. Recursive Procedure for Finite, Truncated Periodic Systems

In this section, a method is presented for an efficient calculation of the non-linearly generated fields and currents for a system consisting of a periodically arranged, but finite number of electrodes (Figure 1b). In the case of third-order intermodulations, this concerns the fields $\mathbf{U}^{2H}$, $\mathbf{U}^{LF}$, $\mathbf{U}^{IM}$ and the IMD3 current. The basic task is to numerically solve the inhomogeneous systems of linear Equations (9), (12), (13) and (15) with the structure of the matrix $\mathbf{M}(\omega)$ and the inhomogeneities (right-hand sides of the linear equations) resulting from the geometry of the system.

Recently, Koskela and Plessky [50,51] introduced a method for the FEM calculation of linear fields and admittances in such systems, which leads to an enormous reduction of computing time as compared to the standard FEM approach. This method is based on the elimination of the internal degrees of freedom in a unit cell via the Schur complement and the principle of cascading, well-known in $\mathbf{P}$-matrix theory [52]. For the computations of nonlinear fields within the perturbation scheme used here, cascading is not directly applicable. This is because the inhomogeneities differ between different cells unlike the linear case, where the inhomogeneities result from the electrodes’ electric potential and, consequently, there are only two different cell inhomogeneities.

Here, we present a simple scheme which makes use of recursion relations and which is described in the following. It is applied to the system shown in Figure 10, which is divided in $N+2$ cells, $N$ of them being identical. The mesh is chosen such that no boundary between neighboring cells divides a finite element into two parts. In addition, we impose the following requirements on the arrangement of the mesh. The finite elements on the left side of a boundary between two neighboring cells (blue vertical line in Figure 10), having common nodes with this boundary, form a vertical layer. We require that these vertical layers are identical in the cells with numbers 0 to $N$. Likewise, the vertical layers of finite elements on the right of a boundary having common nodes with this boundary are identical in the cells with numbers 1 to $N+1$. These additional requirements to the mesh are not essential, but simplify the calculations. The sub-meshes in the cells 1 to $N$ are identical.

![Figure 10](image-url)

**Figure 10.** Truncated periodic system consisting of $N+2$ cells. Grey: piezoelectric substrate, red: metal electrodes, yellow: dielectric film, white: PML. Cells 1, ..., $N$ are identical with alternating potential on their electrodes. Vertical blue lines: boundaries between neighboring cells. $\nu$: internal degrees of freedom; $\kappa$: internal inhomogeneities; $\mathbf{U}$: degrees of freedom on cell boundaries; $\mathbf{K}$: inhomogeneities on cell boundaries.
7.1. Derivation of the Recursion Relation
The systems of linear Equations (9), (12), (13) and (15) are of the general form

\[ \sum_{\ell} M_{j\ell}^{(n)}(\omega) \hat{U}_{\ell}^{(n)} = \hat{K}_{j}^{(n)}. \]  

(66)

We now introduce a cell index \( n \) and group the degrees of freedom \( \hat{U}_{\ell}^{(n)} \) on nodes \( \ell \) in the interior of the \( n \)-th cell in a column vector \( \nu^{(n)} \). The degrees of freedom \( \hat{U}_{\ell}^{(n)} \) on nodes \( \ell \), which are situated on the boundary between the \( n \)-th and \( (n-1) \)-th cell, are grouped in a column vector \( U^{(n)} \). In the same way, with the quantities \( \hat{K}_{j}^{(n)} \) on the right-hand side of (66) with node number \( \ell \) referring to the interior of the \( n \)-th cell, a column vector \( \kappa^{(n)} \) is formed. The quantities \( \hat{K}_{j}^{(n)} \) with node number \( \ell \) referring to the boundary between the \( n \)-th and \( (n-1) \)-th cell form the column vector \( K^{(n)} \).

The interior degrees of freedom may be eliminated by using the following relation:

\[ M^{(n,ii)} \nu^{(n)} = \kappa^{(n)} - M^{(n,j<)} U^{(n-1)} - M^{(n,j>)} U^{(n)} \]  

(67)

for \( n = 1, \ldots, N \). For \( n = 0 \), the second term on the right-hand side and for \( n = N + 1 \) the third term on the right-hand side of (67) is absent. \( M^{(n,ii)}, M^{(n,j<)}, M^{(n,j>)} \) are submatrices of the coefficient matrix on the left-hand side of (66) with the left node index \( \ell \) of its matrix elements \( M_{j\ell}^{(n)}(\omega) \) referring to all nodes in the interior of the \( n \)-th cell. The matrix elements of \( M^{(n,ii)} \) connect two nodes in the interior of cell \( n \) (these two nodes may be identical). The right node indices of the matrix elements of \( M^{(n,j<)} \) (\( M^{(n,j>)} \)) refer to nodes on the left (right) boundary of cell \( n \).

The system of Equation (67) is solved for \( \nu^{(n)} \) and the result is inserted in

\[ M^{(n,<)} \nu^{(n)} + M^{(n-1,>)} \nu^{(n-1)} = \kappa^{(n)} - M^{(n,<)} U^{(n)} \]  

(68)

for \( n = 1, \ldots, N + 1 \), where \( M^{(n,<)}, M^{(n,>)} \) are submatrices of the coefficient matrix in (66). The right node indices of \( M^{(n,<)} \) and \( M^{(n,>)} \) refer to nodes in the interior of cell \( n \). The left node indices of the matrix elements of \( M^{(n,<)} \) (\( M^{(n,>)} \)) refer to nodes on the left (right) boundary of cell \( n \), whereas the matrix elements of \( M^{(n,<)} \) connect two (possibly identical) nodes on the left boundary of cell \( n \).

We note that with the conditions we have imposed on the mesh, the matrices \( M^{(n,ii)}, M^{(n,j<)}, M^{(n,>)} \) are independent of the cell number \( n \) for the cells 1 to \( N \). In addition, \( M^{(N+1,<)} = M^{(N：<)} \). Therefore, we shall drop the cell number in the notation of these matrices and keep this number only in \( M^{(0,ii)}, M^{(N+1,ii)}, M^{(0,>)} \), \( M^{(N+1,>)} \).

Having eliminated \( \nu^{(n)}, n = 0, \ldots, N + 1 \), the following recursion relation results for the vectors \( U^{(n)} \),

\[ AU^{(n-1)} + BU^{(n)} + CU^{(n+1)} = f^{(n)}, \]  

(69)

for \( n = 2, \ldots, N \). Furthermore

\[ B^{(0)} U^{(1)} + CU^{(2)} = f^{(1)}, \]  

(70)

\[ AU^{(N)} + B^{(N+1)} U^{(N+1)} = f^{(N+1)}. \]  

(71)

The matrices \( A, B, C, B^{(0)}, B^{(N+1)} \) are related to the quantities introduced before via

\[ A = -M^{(>)} M^{(i<)} - M^{(j<)} M^{(i>)} + M^{(i<)} \]  

(72)

\[ B = -M^{(j<)} M^{(i<)} - M^{(j<)} M^{(i>)} + M^{(i<)} \]  

(73)

\[ C = -M^{(i<)} M^{(i<)} \]  

(74)
\[ B^{(0)} = -M^{0,i,j} M^{0,i,j-1} M^{0,i,j} - M^{i,j} M^{i-1,j} M^{j,i} + M^{<i,j}, \]
\[ B^{(N+1)} = -M^{(i,j)} M^{(i,j-1)} M^{(i,j)} - M^{(N+1,i,j)} M^{(N+1,i,j-1)} M^{(N+1,i,j)} + M^{<i,j}. \]

The column vectors \( f^{(n)} \), \( n = 1, \ldots, N \), are defined as:
\[ f^{(n)} = -M^{(i,j)} M^{(i)j-1} \kappa^{(n)} - M^{i,j} M^{i-1,j} \kappa^{(n)} + K^{(n)}, \]

and
\[ f^{(1)} = -M^{(0,i,j)} M^{(0,i,j-1)} \kappa^{(0)} - M^{i,j} M^{i-1,j} \kappa^{(1)} + K^{(1)}, \]
\[ f^{(N+1)} = -M^{(i,j)} M^{(i)j-1} \kappa^{(N)} - M^{(N+1,i,j)} M^{(N+1,i,j-1)} \kappa^{(N+1)} + K^{(N+1)} . \]

From (69)–(71), the degrees of freedom on the cell boundaries, \( U^{(n)} \), \( n = 1, \ldots, N + 1 \), are determined in the following way.

First, we compute matrices \( \hat{B}_n \), \( n = 1, \ldots, N - 1 \), and vectors \( \hat{j}_n \), \( n = 1, \ldots, N \), recursively via
\[ \hat{B}_n = B - A \hat{B}_{n-1} C, \]
\[ \hat{j}_{n+1} = f^{(n+1)} - A \hat{B}_{n-1} \hat{j}_n, \]

starting the recursions with
\[ \hat{B}_0 = B^{(0)}, \]
\[ \hat{j}_1 = f^{(1)}, \]

and defining
\[ \hat{B}_N = B^{(N+1)} - A \hat{B}_{N-1} C. \]

With the help of these matrices and vectors, which have been stored, the vectors \( U^{(n)} \), \( n = 1, \ldots, N + 1 \) are subsequently computed by making use of the following downwards recursion:
\[ U^{(n)} = \hat{B}_{n-1}^{-1} \left( \hat{j}_n - C U^{(n+1)} \right), \]

for \( n = N, \ldots, 1 \), and starting with
\[ U^{(N+1)} = \hat{B}_N^{-1} \hat{j}_{N+1}. \]

### 7.2. Results for IMD2 Displacements

The recursion scheme outlined in the previous subsection has been implemented for the computation of displacements and electrical currents, generated by second-order nonlinearity in TC-SAW systems consisting of finite numbers of electrodes. In Figure 11, results are shown for the displacements at the surface of the SiO₂ film at positions above the centers of the electrodes and halfway between the surface points above the centers of neighboring electrodes (Figure 11b). The pitch, metallization ratio, electrode height and film thickness were the same as those in TC-SAW systems consisting of finite numbers of electrodes. In Figure 11, results are shown for the displacements at the surface of the SiO₂ film at positions above the centers of the electrodes and halfway between the surface points above the centers of neighboring electrodes (Figure 11b). The pitch, metallization ratio, electrode height and film thickness were the same as those in the infinite periodic system that Figures 2–4 refer to. The linear admittance of this system is shown in Figure 11a. The displacement patterns for two input frequencies as linear response to a periodic electric potential of ±1 V (peak value) are shown in Figure 11c,d. The input frequency \( f_1 \) (860 MHz) is about 14 MHz below, and \( f_2 \) (890 MHz) about 16 MHz above the resonance frequency. Figure 11e shows the displacement pattern for the second harmonic (H2) of the input tone \( f_2 \). It varies more strongly along the length of the device than the displacement associated with its fundamental (Figure 11d). The displacement pattern of the second order intermodulation (IMD2) at frequency \( f_1 + f_2 \) is shown in Figure 11f. Similar to the pattern associated with the input tone \( f_1 \) (Figure 11c), the displacements in the center of the device are small as compared to those near the 30th and 70th cells.
with reasonable numerical effort and which are directly based on material properties. For perturbation theory is justified by the smallness of typical strains occurring in SAW devices. Periodic as well as finite systems with spatial periodicities can be treated. The application of quantities in SAW devices generated by second-order and third-order nonlinearity. These patterns in this system (Figure 11) show above the centers of the electrode and between the electrodes (b). Linear fields at excitation frequencies 860 MHz (c) and 890 MHz (d). Field of the second harmonic at 2 × 890 MHz (e) and of IMD2 at frequency (860 + 890) MHz (f).

8. Conclusions

In summary, a perturbation-theoretical approach, based on nonlinear electro-elasticity theory and the finite element method, has been presented, which allows for calculations of quantities in SAW devices generated by second-order and third-order nonlinearity. These include displacements and electric fields as well as electrical currents associated with intermodulations of second and third order, second and third harmonic generation and the dependence of the resonance curve and phase response of a SAW resonator on input voltage. The FEM approach works with two-dimensional elements, but fully accounts for all three Cartesian components of the mechanical and electric displacement field. Infinite periodic as well as finite systems with spatial periodicities can be treated. The application of perturbation theory is justified by the smallness of typical strains occurring in SAW devices. It leads to a hierarchy of linear boundary value problems. By this approach, complex systems become accessible to numerical simulations of nonlinear effects, which would be hard to deal with by numerical schemes solving directly nonlinear partial differential equations.

A major goal of our approach was to enable computations of third-order intermodulations on a laptop computer. This opens the possibility for the designer of SAW devices to minimize undesired nonlinear effects with the help of simulations, which can be carried out with reasonable numerical effort and which are directly based on material properties. For this purpose, algorithms were developed that increase the efficiency of certain parts of the computations. These include an alternative way of calculating IMD2 and IMD3 currents in the electrodes via overlap integrals, which avoid the computation of the IMD2/IMD3 field. In addition, an impedance boundary condition is presented to account for reflections of bulk waves from the bottom of a thick substrate and, in particular, a recursive algorithm.
for finite periodic systems which is related to the cascading method introduced by Koskela and Plessky for FEM computations of linear quantities in SAW devices [50,51].

The recursive scheme, introduced here in the context of electro-elastic nonlinearity, may also be applied to simulations of thermal effects in SAW devices to reduce computational effort; for example, the effect of intrinsic heating on the resonance frequency, giving rise to additional nonlinearity (e.g., [53] for the case of BAW resonators). The variation of the temperature field between different unit cells of the interdigital transducers or reflectors precludes cascading, but still allows for application of the recursive algorithm.

Prerequisite of quantitative simulations of nonlinear effects is the availability of reliable nonlinear material constants up to fourth order for all materials that fill those parts of the SAW device where comparatively large strains or electric fields occur. In the case of thin films, the required material constants may differ from those of the corresponding bulk material and may depend on details of the fabrication process.

Author Contributions: Conceptualization, A.P.M. and W.R.; formal analysis, A.P.M., E.A.M., M.M., W.R., V.C. and T.F.; writing—original draft preparation, A.P.M.; writing—review and editing, A.P.M., with (A.P.M. is indebted to Vladimir Mozhaev for having brought relevant literature

Acknowledgments: All authors have read and agreed to the published version of the manuscript.

Conceptualization, A.P.M. and W.R.; formal analysis, A.P.M., E.A.M., M.M., W.R., V.C., T.F. and K.C.W.; supervision and project administration, K.C.W. and W.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available in the article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The generalized tensors $\tilde{S}^{(2)}$ and $\tilde{S}^{(3)}$ may be decomposed into a part due to the mechanical properties or electrical polarizability of the material they refer to, $\tilde{S}^{(2)}_1$, $\tilde{S}^{(3)}_1$, plus a part related to effects of the electric field that are also present in vacuum, like the Maxwell stresses, $\tilde{S}^{(2)}_2$, $\tilde{S}^{(3)}_2$.

The explicit expressions for the elements of the generalized tensors $\tilde{S}^{(2)}$, $\tilde{S}^{(3)}$ in terms of the material constants of second, third and fourth order are obtained by comparing the expansion of the density of potential energy $\rho_\Phi$ terms in the components of the field $W$ and Green–Lagrange strain tensor $E$ in Equation (1.85) of [23] with the expansion (3) of the potential energy density $\Phi$ in powers of generalized displacement gradients. Identifying the $M$-th component of $W$ with $-u_{4,M}$ and the component $E_{MN}$ of the tensor $E$ with $(1/2)(\delta_{Mn}u_{m,N} + \delta_{Nn}u_{m,M} + \delta_{Lj} \delta_{lj} u_{j,M}u_{k,N})$, this comparison yields,

$$\tilde{S}^{(2)}_{KL \, mN \, pQ} = \delta_{kK} \delta_{mM} \delta_{pp} C_{KL \, MN \, PQ} + \delta_{pp} \delta_{kn} C_{LN \, PQ} + \delta_{mn} \delta_{kp} C_{MN \, LQ}, \quad k, m, p = 1, 2, 3, \quad (A1)$$

$$\tilde{S}^{(2)}_{mN \, jP \, 4L} = \delta_{MM} \delta_{jj} e_{KL \, MN \, jP} + \delta_{nj} e_{LN \, NP}, \quad m, j = 1, 2, 3, \quad (A2)$$

$$\tilde{S}^{(2)}_{mN \, 4K \, 4L} = -\delta_{mn} e_{KL \, MN}, \quad m = 1, 2, 3, \quad (A3)$$

$$\tilde{S}^{(2)}_{4K \, 4L \, 4M} = \epsilon_0 \chi_{KL\, MN}. \quad (A4)$$

Equation (A4) corrects a sign error in Equation (2.30) in [16].

$$\tilde{S}^{(3)}_{KL \, mN \, pQ \, rS} = \delta_{kK} \delta_{mM} \delta_{pp} \delta_{rr} C_{KL \, MN \, PQ \, RS} + \delta_{kK} \delta_{mp} C_{KL \, MN \, QS} + \delta_{pp} \delta_{rr} C_{KL \, MN \, JS} + \delta_{mm} \delta_{rr} C_{MN \, PQ \, LS} + \delta_{mn} \delta_{rr} C_{MN \, PQ \, RS}, \quad k, m, r, p = 1, 2, 3, \quad (A5)$$
\begin{align}
S^{(3)}_{KL mN \, pQ \, 4J} &= \delta_{KK}\delta_{mM}\delta_{pP} \varepsilon_{J \, KL \, MN \, PQ} + \delta_{KK}\delta_{mp} \varepsilon_{J \, KL \, NQ} + \delta_{mM}\delta_{kp} \varepsilon_{J \, MN \, LQ} \\
&+ \delta_{pp}\delta_{km} \varepsilon_{J \, LN \, PQ}, \quad k, m, p = 1, 2, 3, \tag{A6}
\end{align}

\begin{align}
S^{(3)}_{KL mN \, 4I \, J} &= -\delta_{KK}\delta_{mM} I^{(s)}_{J \, KL \, MN} - \delta_{km} I^{(s)}_{J \, LN}, \quad k, m = 1, 2, 3, \tag{A7}
\end{align}

\begin{align}
S^{(3)}_{mN \, 4I \, J \, 4K} &= \varepsilon_0 \delta_{mM} I^{(s)}_{J \, KMN}, \quad m = 1, 2, 3. \tag{A8}
\end{align}

\begin{align}
\tilde{S}^{(3)}_{mN \, 4I \, J \, 4K \, 4L} &= -\varepsilon_0 \chi_{JKL}. \tag{A9}
\end{align}

The contributions $\tilde{S}^{(2)}$, $\tilde{S}^{(3)}$ to the generalized tensors $S^{(2)}$ and $S^{(3)}$ are given explicitly by

\begin{align}
S^{(2)}_{mN \, 4K \, 4L} &= \varepsilon_0 \delta_{mM} (\delta_{MK}\delta_{NL} + \delta_{ML}\delta_{NK} - \delta_{MN}\delta_{KL}), \quad m = 1, 2, 3, \tag{A10}
\end{align}

\begin{align}
S^{(3)}_{JP \, RS \, 4K \, 4L} &= \varepsilon_0 \delta_{JL} \delta_{RS} \delta_{PK} \delta_{JL} + \delta_{RS} \delta_{PL} \delta_{JL} - \delta_{RP} \delta_{KS} \delta_{JL} - \delta_{RP} \delta_{KS} \delta_{JL} \\
&- \delta_{JL} \delta_{RS} \delta_{PK} \delta_{JL} + \delta_{PR} \delta_{SJ} \delta_{KL} + \delta_{KR} \delta_{JS} \delta_{PL} + \delta_{LR} \delta_{KS} \delta_{PJ} - \delta_{KR} \delta_{PS} \delta_{LJ} - \delta_{LR} \delta_{KS} \delta_{KJ} - \delta_{JS} \delta_{PK} \\
&- \delta_{KR} \delta_{JS} \delta_{PL}), \quad k, m = 1, 2, 3. \tag{A11}
\end{align}

In Equations (A1)–(A11) we have largely followed the notation for the various material constants used in [20,23]. In the case of the electrostriction constants $I^{(s)}_{J \, KL, MN}$ and the even electroelasticity coefficients of fourth order $I^{(s)}_{J \, KL, MN}$, the superscript $(s)$ was added to indicate that these constants do not contain terms related to effects of the electric field in vacuum, i.e., they vanish in a non-polarizable medium.

References


Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.