

Efficient modeling and optimization of surface second-harmonic generation from nanophotonic components

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Rat der Physikalisch-Astronomischen Fakultät
der Friedrich-Schiller-Universität Jena

von Master of Science in Angewandte Mathematik
und Physik Anton Pakhomov
geboren am 1. Oktober 1991 in Samara, UdSSR

Gutachter:

1. Prof. Thomas Pertsch, Friedrich-Schiller-Universität Jena, Germany
2. Prof. Costantino De Angelis, University of Brescia, Italy
3. Prof. Yuri Kivshar, Australian National University, Canberra, Australia

Tag der Disputation: 29.06.2021

Contents

1	Introduction, state of the art and goals	1
1.1	Surface and bulk SHG	5
1.2	Simulation of SHG in nanooptics: overview	6
1.3	Goals of the PhD thesis	8
2	Transfer matrix method for modeling of surface SHG from multilayer structures	11
2.1	Model and problem statement	13
2.2	Calculation of the emitted SHW by TMM	17
2.3	Calculation of the surface nonlinear polarization	20
2.4	Computational procedure with the proposed TMM	22
2.5	Multilayer structure of ultrathin layers	23
2.6	Effective surface nonlinear tensor of the stack	26
2.7	Simulation results and discussion	29
2.8	Summary	33
3	Adaptive finite-element method for modeling of surface SHG from nanostructures	35
3.1	Overview of the finite-element methods	37
3.2	Treatment of the surface SHG	40
3.3	Mathematical formalism of the surface nonlinear polarization	41
3.4	Implementation of generalized boundary conditions	43
3.5	Extension to 2D materials	45
3.6	Test results for benchmark problem	46
3.7	Summary	49
4	Signatures of surface nonlinearity in far-field patterns of SHG from noncentrosymmetric semiconductors	50
4.1	Model and analytical solution	52
4.2	SHG patterns for plane-wave illumination	58
4.3	Representation of tightly-focused Gaussian beam	63
4.4	Far-field SHG patterns under TFGFB illumination	68
4.5	Summary	74

5 Optimization of surface SHG from multilayer structures	76
5.1 Nonperiodic stacks	77
5.2 Periodic stack of SiO ₂ /TiO ₂ layer pairs	80
5.3 Periodic stack of SiO ₂ /Al ₂ O ₃ /TiO ₂ layer triples	86
5.4 Summary	88
Conclusion and outlook	93
Deutschsprachige Zusammenfassung	97
List of own contributions	99
Bibliography	101
Appendix A Generalized boundary conditions for an anisotropic medium	114
Appendix B First-order correction terms for the effective surface nonlinear tensor $\chi_{\Sigma,ijk}^{(2)}$	118
Appendix C Surface SHG from a dielectric sphere under VSH illumination	121
Abbreviations and conventions	125
Acknowledgement	126
Ehrenwörtliche Erklärung	127
Curriculum Vitae	128

Chapter 1

Introduction, state of the art and goals

Second-harmonic generation (SHG) is the simplest and most widely studied nonlinear optical process, when the incident electromagnetic wave at the fundamental frequency ω is converted in a medium into the wave at the doubled frequency 2ω [1, 2] (Fig. 1.1). The observation of SHG usually requires high-power light beams due to the weakness of the nonlinear effects and correspondingly low intensity of the generated wave at the second-harmonic (SH) frequency. Due to this reason SHG has been only experimentally detected after the invention of first lasers in 1960s [3], which made it possible to produce intense coherent light beams [4, 5]. SHG has found a number of practical applications, especially due to its sensitivity to the symmetry properties of the medium. Among them are second-harmonic imaging microscopy [6, 7, 8, 9], frequency doubling [10, 11, 12], spectroscopy [13, 14], probing of interfaces [15, 16], characterization of crystalline materials [17] and ultrashort pulse measurement with autocorrelators [18].

Second-harmonic generation can be considered as a particular case of sum-frequency generation (SFG), when two photons with frequencies ω_1 and ω_2 are converted inside a nonlinear medium into a single photon with the sum frequency $\omega_3 = \omega_1 + \omega_2$. The relation between the frequencies implies, that the energy is conserved during this process. Second-harmonic generation is thus the sum-frequency generation with $\omega_1 = \omega_2$. However, in contrast to SFG, when at least two light beams with frequencies ω_1 and ω_2 have to overlap inside the volume of a nonlinear medium, for the SHG to occur only a single optical beam is needed. Besides, SHG is the simplest case of the harmonic generation phenomena, i.e. generation of the light waves with frequencies $n\omega$ for different positive integers $n \geq 2$. SHG corresponds to $n = 2$ and the second harmonic is often the most intense harmonic in the output of a nonlinear crystal. Sometimes, however, higher-order harmonics could prevail. For instance, this happens in inversion-symmetric media, since inversion symmetry leads to the greatly reduced conversion into the second-harmonic wave (SHW). In this case the third-harmonic generation (THG) with $n = 3$ becomes the major nonlinear process.

In order to describe the property of media to produce SHG, the so-called nonlinear polarization of the media \vec{P}^{NL} is used, i.e. the extra contribution to the medium polarization beyond the usual linear one. As the result, the electric flux density \vec{D} in a nonlinear medium can be written as [2]:

$$\vec{D} = \epsilon_0 \hat{\epsilon} \vec{E} + \vec{P}^{\text{NL}},$$

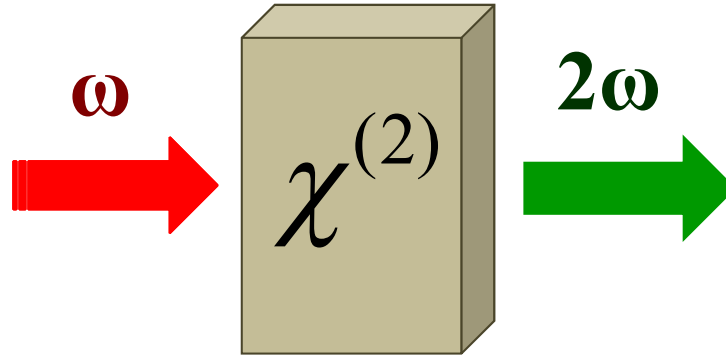


Figure 1.1: General scheme of the second-harmonic generation.

where \vec{E} is the applied electric field and $\hat{\epsilon}$ is the tensor of the linear relative dielectric permittivity.

In general the description of nonlinear optical properties of a medium would then need extra equations for the dynamics of \vec{P}^{NL} . In the simplest case of a non-resonant excitation, the nonlinear polarization can be expanded into the series of the electric field in the pump wave. The corresponding general expression for the polarization of a nonlinear medium can be written as:

$$P_j^{\text{NL}} = \epsilon_0 \epsilon_{jk} E_k + \epsilon_0 \chi_{jkl}^{(2)} E_k E_l + \epsilon_0 \chi_{jklm}^{(3)} E_k E_l E_m + \dots \quad (1.1)$$

The first term here is the usual relative dielectric permittivity, which has in general tensor form. The second term contains the third-rank tensor $\chi_{jkl}^{(2)}$, which is called the *second-order nonlinear susceptibility tensor* and is responsible for the second-order nonlinear processes. Similarly the third term incorporates the fourth-rank tensor $\chi_{jklm}^{(3)}$ called third-order nonlinear susceptibility tensor. Further higher-order terms can be added in the Eq. (1.1) to treat other harmonic generation processes.

The expansion Eq. (1.1) yields the following term for the second-order nonlinear polarization:

$$\vec{P}^{\text{NL}} = \epsilon_0 \hat{\chi}^{(2)} \vec{E} \vec{E}, \quad (1.2)$$

where the tensor notation for $\hat{\chi}^{(2)}$ is introduced. However, Eq. (1.2) does not always suffice and has to be generalized. Indeed, the term on the right-hand side of Eq. (1.2) is the so-called electric dipole term and represents just the first term of the expansion of the second-order nonlinear polarization. The full expansion into multipole components has the following form:

$$\vec{P}^{\text{NL}} = \epsilon_0 \hat{\chi}_D^{(2)} \vec{E} \vec{E} - \epsilon_0 \vec{\nabla} \cdot \left(\hat{\chi}_Q^{(2)} \vec{E} \vec{E} \right) + \frac{\epsilon_0 \mu_0}{i\omega} \vec{\nabla} \cdot \left(\hat{\chi}_M^{(2)} \vec{E} \vec{E} \right) + \dots \quad (1.3)$$

Here the second-order nonlinear susceptibility tensor in the electric-dipole approximation from Eq. (1.2) is denoted as $\hat{\chi}_D^{(2)}$. The second and the third terms in Eq. (1.3) contain the electric quadrupole tensor

$\hat{\chi}_Q^{(2)}$ and the magnetic dipole tensor $\hat{\chi}_M^{(2)}$ and are proportional to the spatial derivatives of the pump field. The relative amplitudes of the terms in Eq. (1.2) are sensitive to the symmetry properties of the material. In media without inversion symmetry the electric dipole tensor $\hat{\chi}_D^{(2)} \neq 0$ and the corresponding contribution prevails. By contrast, in media with inversion symmetry the electric dipole term vanishes $\hat{\chi}_D^{(2)} = 0$ and only the electric quadrupole and magnetic dipole terms are responsible for SHG [19]. It should be also noted, that in the regime of the extreme nonlinear optics [20] SHG can occur in a bulk inversion-symmetric material through, e.g., the conical second-harmonic generation [21] or the photon-drag-effect [22], but the consideration of these topics goes far beyond the scope of this thesis.

The wave equation for the electric field in the presence of a nonlinear medium attains the form:

$$\frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} + \vec{\nabla} \times \vec{\nabla} \times \vec{E} + \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}^L}{\partial t^2} = -\frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}^{NL}}{\partial t^2}. \quad (1.4)$$

Eq. (1.4) in general has to be supplemented with the expression for the nonlinear polarization, like Eq. (1.1). If the second-order nonlinearity only is considered and if the pump wave at the fundamental frequency can be treated as quasi-monochromatic, the electric field can be represented as a dual-frequency one as follows:

$$\vec{E}(\vec{r}, t) = \vec{E}_\omega(\vec{r}) e^{i\omega t} + \vec{E}_{2\omega}(\vec{r}) e^{2i\omega t}. \quad (1.5)$$

Upon substituting Eq. (1.5) into the wave equation Eq. (1.4), one obtains the following two coupled equations for the amplitudes of the electric field components \vec{E}_ω and $\vec{E}_{2\omega}$:

$$-\vec{\nabla} \times \vec{\nabla} \times \vec{E}_\omega + \frac{\hat{\epsilon}_\omega \omega^2}{c^2} \vec{E}_\omega = -\frac{2\hat{\chi}^{(2)} \omega^2}{c^2} \vec{E}_{2\omega} \vec{E}_\omega^*, \quad (1.6)$$

$$-\vec{\nabla} \times \vec{\nabla} \times \vec{E}_{2\omega} + \frac{4\hat{\epsilon}_{2\omega} \omega^2}{c^2} \vec{E}_{2\omega} = -\frac{4\hat{\chi}^{(2)} \omega^2}{c^2} \vec{E}_\omega \vec{E}_\omega. \quad (1.7)$$

The system of equations Eqs. (1.6)-(1.7) is the central one for both analytical and numerical studies of SHG effects. It must be stated that Eqs. (1.6)-(1.7) are not exact and are not equivalent to the wave equation Eq. (1.4). The substitution Eq. (1.5) implied only two frequency components involved. In fact, the fundamental wave (FW) at the frequency ω interacts with the second-harmonic wave through the nonlinear polarization Eq. (1.1) to produce third-harmonic, fourth-harmonic and so on. Eqs. (1.6)-(1.7) inherently neglect the generation of these higher-order harmonics. This approximation is not only well justified in many real-world situations, where second-order nonlinearity is present, but also largely reduces the complexity of the theoretical treatment. At the same time, the inclusion of the extra harmonics into Eq. (1.5) would result in a much more complex system of nonlinear equations instead of Eqs. (1.6)-(1.7).

The system of equations Eqs. (1.6)-(1.7) allows for an exact analytical solution for a homogeneous medium in 1D geometry. The result exhibits the consecutive energy transfer between the FW and SHW upon propagation in a nonlinear crystal, while the fraction of the converted FW is determined

by the value [2]:

$$\Delta k = k_{2\omega} - 2k_{\omega} = \frac{2\omega}{c} (\sqrt{\varepsilon_{2\omega}} - \sqrt{\varepsilon_{\omega}}). \quad (1.8)$$

If $\Delta k = 0$, the full transfer of the FW into SHW is possible. This condition is called *phase-matching condition* (PMC) and means that the phase velocities for the FW and SHW are the same. The violation of PMC, i.e. a phase mismatch $\Delta k \neq 0$, leads to the much less efficient nonlinear conversion into the SHW. In real nonlinear crystals it is very challenging to achieve PMC due to the significant dispersion. In uniaxial crystals PMC can be exactly reached at specific propagation angles with respect to the crystal axis. In practice, however, the perfect phase-matching is often not a realistic condition and the PMC needs to be relaxed. There are several ways proposed to circumvent the strict PMC. One of the most widely used is Quasi-phase-matching (QPM), which uses a periodically poled crystal to ensure the phase shift between two optical waves (FW and SHW) is always less than 180° along the length of the crystal [23, 24]. The PMC is not exactly fulfilled in this case, but periodic poling guarantees the constant energy flow from the pump FW to the signal (SHW).

Even though SHG phenomena are studied for a long time already after the creation of first lasers, there is now growing interest to this topic in the field of nonlinear nanooptics. Nanostructures exhibit strong resonances throughout the visible (VIS) and mid-infrared (MIR) spectral ranges. Excitation of these resonances results in the large confinement of the energy of the electromagnetic field in the tiny volume of the nanostructures. Due to the quadratic dependence of the nonlinear polarization on the pump electric field, the efficiency of the energy conversion into a SHW depends on the pumping field. Specifically, the conversion efficiency is quadratically proportional to the pumping field. Therefore the field enhancement in the nanostructures leads to the drastic increase of the conversion efficiency and reduces the excitation intensity needed to produce a required SH intensity. Both plasmonic [25, 26] and all-dielectric [27, 28] nanostructures and metasurfaces are considered for SHG applications. All-dielectric optical nanostructures, which are usually made of semiconductor materials, have several advantages, when compared to the metallic ones. In particular, those include small dissipative losses, low Joule heating and high-Q Mie resonances [29, 30]. Hence, all-dielectric nanostructures can serve as alternatives to the plasmonic nanostructures in a number of applications. Specifically, in the area of nonlinear nanophotonics the usage of the metal-less nanostructures could bring great benefits [27, 28].

As stated above, in traditional nonlinear optics the energy conversion efficiency to the SHW is mainly enhanced, when the PMC is fulfilled. It provides the unidirectional energy flow from the FW into the SHW upon propagation through a nonlinear crystal. At the same time, the characteristic propagation distance required for this energy transfer to occur is orders of magnitude larger than the light wavelength. Nanostructures, on the other hand, have typically subwavelength sizes. As the result, PMC ceases to be crucial for efficient SHG in the nanoscale components and totally different approaches are needed. The most promising approach seems to be the excitation of high-Q resonances in all-dielectric nanopillars usually made of III-V semiconductors [31, 32]. The strong resonances lead to a huge field enhancement inside the nanostructures. Together with the large bulk nonlinearity in III-V semiconductor materials it allows to achieve high conversion efficiency for experimentally accessible pump intensities. For example, a conversion efficiency $\sim 10^{-4}$ was obtained for the exci-

tation of the electric dipole (ED) and magnetic dipole (MD) resonances in AlGaAs nanodiscs [31]. More recently, the conversion efficiency in GaAs nanodiscs was increased by at least 2 orders of magnitude by means of the so-called *bound states in continuum* (BIC) with even higher Q-factors [32]. Alternative approaches include, for instance, the multiple-quantum-well structures, which suppress the band-gap absorption, but the achieved conversion efficiency in this way is still significantly lower [33].

1.1 Surface and bulk SHG

One typically assumes two sources of SHG: bulk and surface nonlinearity [2]. Bulk nonlinearity is the standard one, arising in the volume of the nonlinear medium. In contrast, surface SHG (SSHG) arises at the interfacial layer in the vicinity of the interface between two adjacent media due to the discontinuity of both the medium structure and the normal component of the electric field across the interface thus resulting in the huge electric field gradient nearby the surface [34, 35, 36]. While the bulk nonlinearity in centrosymmetric media disappears in the electric-dipole approximation, the surface nonlinearity does not. Indeed, the inversion symmetry is broken close to the surface of the centrosymmetric media as well, what leads to the strong electric-dipole second-order nonlinearity in the interfacial region. Such independence of the surface SHG on the symmetry properties of the bulk of the medium makes surface SHG an efficient tool for probing interfaces [16].

SHG from different materials is mainly driven by different nonlinear contributions. In centrosymmetric optical materials, like silicon and different metals, the bulk nonlinearity is prohibited in the electric-dipole approximation, therefore the surface SHG is typically dominating [37, 38, 39]. Moreover, in plasmonic nanostructures the electric field can only penetrate into the medium for the skin depth for the frequencies below the ultraviolet (UV) transparency threshold. Therefore the electric field gets concentrated close to the metal surface, which increases the nonlinear polarization in the surface region and the overall nonlinear response [25]. In all-dielectric nanostructures made of centrosymmetric materials, e.g. widely used Si, SiO₂, TiO₂, the field absorption is low below the fundamental band-gap and the field is not confined in the vicinity of the interface.

In noncentrosymmetric semiconductors the situation becomes different. Since these materials possess the strong bulk nonlinearity, the surface SHG is often assumed to play just a minor role. For example, SHG from nanostructures made of III-V semiconductor materials (like GaAs, GaP, InAs, AlGaAs and others) could be described assuming only the bulk nonlinearity and neglecting the surface nonlinearity [40, 41, 42]. However, further studies have demonstrated that this is not always correct and even in nanostructures made of noncentrosymmetric semiconductors the surface SHG can also play a significant role and must be treated accordingly [43, 44, 45]. Specifically, noticeable or even prevailing surface contributions were found at frequencies close to the surface resonance ones [46], dielectric metasurfaces for SHG wavelengths above the fundamental band-gap [44] as well as in the opaque region of the semiconductor [47, 45]. In general the surface and bulk SHG in noncentrosymmetric semiconductors exhibit a quite nontrivial interplay, which has not been comprehensively studied to date and deserves more detailed consideration.

1.2 Simulation of SHG in nanooptics: overview

Modeling of any nonlinear optical phenomena requires correct treatment of the nonlinear optical polarization. In the simplest case of the nonresonant nonlinearity one can use Eq. (1.1), which relates the instantaneous electric field strength with the instantaneous value of the nonlinear polarization. It is worth noting that the instantaneous value of the nonlinear polarization would in general also depend on the earlier values of the electric field strength, then one would need instead of Eq. (1.1) a differential equation for the temporal dynamics of \vec{P}^{NL} in the external electric field \vec{E} . Eq. (1.1) is still well justified, when the frequency of the pump wave is far away from any of the natural resonant frequencies of the nonlinear medium [2].

One of the most crucial issues in the modeling of SHG is how one treats the interaction between the generated second-harmonic wave and incident pump wave. The point is that the expression for the second-order nonlinear optical polarization in general includes the term:

$$\varepsilon_0 \chi^{(2)} E_{2\omega} E_{\omega}^* \sim e^{i\omega t},$$

i.e. oscillating with the pump frequency (see the right-hand side of Eq. (1.6)). If one has to take this term into account, one has to solve the system of nonlinearly-coupled equations Eqs. (1.6)-(1.7). Such approach is needed, if the conversion efficiency into the SHW is high enough, so that this term gets comparable with the linear medium polarization at the pump frequency. In nanooptics, however, the situation is usually just the opposite. Low conversion efficiency allows to reliably neglect the right-hand side of Eq. (1.6). This case is called *undepleted-pump approximation* (UPA), since it assumes the pump wave to keep constant amplitude upon propagation in the nonlinear medium. When UPA is justified, the solution of Eqs. (1.6)-(1.7) largely simplifies. Now one gets two uncoupled linear equations, which can be solved separately. Namely, one first solves the linear homogeneous Eq. (1.6) for the pump wave, then calculates the nonlinear polarization term on the right-hand side of Eq. (1.7) and finally solves the linear inhomogeneous Eq. (1.7) for the SH field.

As potentially the most suitable method for modeling SHG serves finite-difference time-domain method (FDTD) [48, 49]. The FDTD method is based on the discretization of the Maxwell's equations both in time-domain and over spatial dimensions and inherently allows generalization to the 3D fully-vectorial case. The main property of the FDTD method is the usage of two rectangular meshes for the values of the electric and magnetic field, which are shifted both in the spatial and temporal dimensions with respect to each other. Introducing two different meshes results in the second-order accuracy of the FDTD method over both the spatial and temporal steps. In the presence of the second-order nonlinearity as in Eq. (1.2), the calculation of the electric field values at a certain time step $(N+1)\Delta t$ from the electric field values at the time step $N\Delta t$ would lead in 3D vectorial case to a system of three coupled quadratic equations at each spatial point. The solution of this system of equations would give the values of E_x, E_y, E_z at the time step $(N+1)\Delta t$ at every spatial point and would allow to calculate the field at the time step $(N+2)\Delta t$ and so on. For an arbitrary form of the second-order nonlinear susceptibility tensor $\hat{\chi}_D^{(2)}$ the respective system of three coupled quadratic equations for E_x, E_y, E_z cannot be solved analytically and extra numerical efforts are needed.

The FDTD method has several significant advantages. As it directly follows from its name, the simulations in the FDTD method run in the time domain. This property is especially needed for addressing applications in ultrafast optics. For example, the modeling of the propagation of an ultrashort pulse, up to few-cycle or even subcycle, can be easily performed with the FDTD method. In such a manner different types of optical nonlinearities could be considered, like resonant, Kerr or Raman nonlinearities [50, 51, 52, 53, 54]. Also, the described algorithm of the FDTD method can be easily extended to higher-order nonlinearities. The main limitations are related to the regular rectangular mesh used in the FDTD method. This drawback may look inessential in traditional nonlinear optics, but becomes important in the nonlinear nanophotonics. Indeed, the accurate simulation of the light propagation through subwavelength nanoparticles requires fine discretization of the volume of the nanoparticle, what leads to the fine mesh in the whole computational domain. Therefore one would prefer to use some non-uniform mesh, which is finer within the volume of the nanoparticle and coarser in the outer region with much smaller field gradients. Although several subgridding algorithms for FDTD, which allow for the local mesh refinement, have been proposed to date [55, 56, 57], they require complicated meshing schemes as well as complex code management, what limits the performance of the FDTD method. Moreover, the nanoparticles of non-rectangular shape can be just loosely covered with the rectangular meshes and more flexibility in the accessible mesh geometry is needed. Such flexibility is naturally provided in the finite-element methods (FEM) [58].

The finite-element methods currently represent the often used standard for simulations of Maxwell's equations in the frequency domain. In FEM the electric field components are approximated by polynomial functions, while the computational domain is divided into tetrahedra or prisms. FEMs therefore allow to easily tune the mesh sidelength in a certain part of the computational domain, where the higher accuracy of the simulations is needed. Another degree of freedom for adjusting the computational accuracy is provided by the degree of the approximating polynomials. The simulations of SHG with the finite-element methods without UPA would be increasingly complicated, since the finite-element methods in contrast to the FDTD method cannot be so easily extended to the nonlinear case. However, when UPA is applicable, the SHG simulations split into two consecutive linear simulations, where FEM can be naturally used. Since UPA is usually well justified in nanophotonic applications and FEM provides great flexibility in dealing with the complex geometries and non-uniform meshes, FEM is widely used nowadays for SHG simulations in the nanostructures.

However, the reasoning above applies to the modeling of the bulk nonlinearity only. The accurate treatment of the surface nonlinearity needs the mesh nodes to be densely located on the surface of the nanostructure. As the result, the simulations of the surface SHG would be greatly sensitive to the ability of a numerical method to deal with the curved surfaces of the arbitrary-shape nanostructures. Particularly, the FDTD method turns out to be barely applicable, since the treatment of the curved surfaces by the uniform rectangular meshes is hardly possible. At the same time, as far as I know, modeling of the surface SHG with FEM has not been reported so far.

Several methods have been proposed to date for the modeling of the surface SHG. Those include volume integral formulations [59, 60], boundary element methods [61] and different kinds of the surface integral method [62, 63, 64, 65]. These methods usually rely on considering the surface non-

linear polarization as the surface current density and solving the integral form of Maxwell's equations. Therefore the problem gets reduced to the boundary integral equation, which then needs to be solved numerically. The main advantage of such an approach is that volume discretization is not needed, and only the surfaces have to be discretized. This approach was found to yield quite good performance and relatively low computational costs. At the same time, there are also several limitations, like the complex incorporation of the bulk nonlinearity in the media, relatively slow convergence rates (especially compared to the exponential convergence of FEM simulations) and limited flexibility in handling the surface discretization.

1.3 Goals of the PhD thesis

The major question that this thesis addresses is the development of efficient numerical techniques for the simulation of the surface SHG phenomena from nanostructures and metasurfaces. There are a number of limitations that currently existing numerical methods possess. Among them are difficulty of dealing with the complex geometries and inability to easily incorporate both the bulk and surface nonlinear properties of the nanostructures. The geometry issues arise due to the subwavelength sizes of the nanostructures and the necessity to simulate the surface nonlinear polarization for an arbitrary curved surface of the realistic nanostructures. Therefore the appropriate numerical methods should be well suited to treat surfaces, rather than volumes inside the computational domain. On the other hand, as discussed in the previous sections, interplay of the bulk and surface nonlinearities is quite typical in nanophotonic applications. It means that proper numerical methods have to provide the possibility to simulate the bulk nonlinear effects in the volume of the nanostructures as well. These two requirements in general seem to be in certain contradiction with each other. Hence, one needs to find a trade-off between the efficient treatment of the surface effects and the ability to equally account for the bulk effects. Finite-element methods seem to be the proper candidate for this role. To the best of my knowledge, finite-element methods have not been applied so far for the modeling of the SSHG. Thus, the development of such numerical method is one of the main goals of the thesis.

Besides that, one of the geometries, which was proposed for the consideration in the thesis and looks promising in the context of boosting the surface nonlinear effects, is the multilayer structure. One can intuitively expect the possibility of the constructive interference of the SH signals from the multiple interfaces inside such multilayer structure, leading to a large enhancement of the SSHG. Moreover, as some previous studies showed, for thin layers (already ~ 100 nm or smaller) the bulk nonlinearity can be neglected even in noncentrosymmetric semiconductors [66]. Therefore the development of proper numerical algorithms for the modeling of the SSHG in such geometry is in demand. The inherent 1D geometry in this case makes the application of both FDTD or FEM methods for the modeling of the SHG process overcomplicated as long as one does not deal with complex illuminating beams. Particularly, this is the case for the plane-wave illumination. To the best of my knowledge, there have been no methods proposed so far for the modeling of the SSHG in such multilayer structures. As a result, the development of this numerical method was also taken as one of the goals of the thesis.

Besides the issues related with the development of the numerical algorithms themselves, the main question that this thesis addresses is the finding of the optimal conditions for the detection of the surface SH signal and finding ways to enhance the surface nonlinear response. In order to address this question, this work focuses on the following several issues in this context. Firstly, I aim to find a reliable way to identify the surface SHW against the background of bulk SHW, when such separation is nontrivial, namely in noncentrosymmetric semiconductors with a strong bulk nonlinearity. Studies on this question are lacking, so I intend to find out the typical signatures of the surface optical nonlinearity for different available illumination geometries. Secondly, the enhancement of the conversion efficiency of the SSHG is to be examined with an appropriate optical device, namely a multilayer structure. The choice of the multilayer arrangement is inspired by the expected constructive role of the interference effects in this geometry. The main goal here is to optimize the layer composition of a multilayer structure to obtain possibly high conversion efficiency into the surface-driven SHW. As far as I know, such problem has not been suggested and analysed before.

To summarize, this work is intended to reach the following goals:

- development of a transfer matrix method for the modeling of the surface second-harmonic generation from arbitrary multilayer structures under the plane-wave illumination, with the possible extension to other more complex illuminating beams;
- development of an adaptive finite-element method for the modeling of the surface second-harmonic generation from plasmonic or dielectric nanostructures of an arbitrary shape;
- investigation of the optimal illumination parameters for the reliable identification and quantitative estimation of the strength of the surface nonlinearity in noncentrosymmetric semiconductors;
- numerical optimization of the conversion efficiency of the surface second-harmonic generation in periodic and nonperiodic multilayer structures of different composition.

The thesis is structured as follows.

Chapter 2 is dedicated to the development of a numerical approach based on the transfer matrix method for modeling of surface SHG from multilayer structures. The influence of the surface sources of the nonlinear polarization is implemented through the generalized boundary conditions for the tangential components of the electric and magnetic field, imposed at the boundaries between the media in multilayer structures. The performance of the developed transfer matrix method is tested for several exemplary stacks.

Chapter 3 is devoted to the development of a fully finite-element numerical method for efficient modeling of surface SHG effects. This method treats the surface nonlinear polarization as the delta-surface source embedded into the respective interfaces between two media. It is shown that the proposed method allows modeling of surface nonlinear effects from arbitrary-shaped all-dielectric and plasmonic isolated nanoparticles, as well as metasurfaces.

Chapter 4 focuses on the numerical investigation of the symmetry properties of the bulk and surface nonlinear tensors in noncentrosymmetric semiconductors, which result in specific signatures in the polarization-resolved far-field SHG pattern. It is demonstrated that under the plane-wave illumination of a slab of a noncentrosymmetric III-V semiconductor the differences arise in the angular dependence of the emitted second-harmonic radiation on the angle of the crystal axis rotation.

The results are then extended to the case of an illumination by a linearly polarized tightly-focused Gaussian beam and the specific signatures of surface and bulk nonlinearities are also obtained in the polarization-resolved far-field SHG patterns.

Finally, Chapter 5 is devoted to the numerical studies of the surface SHG from multilayer structures of different possible compositions. Using the transfer matrix method developed earlier in Chapter 2 the optimal geometry of the multilayer structures is sought for the efficient surface-driven SHG. Both nonperiodic and periodic multilayer structures with different number of layers in a single period are considered and analysed with respect to the achievable SHG conversion efficiency.

Chapter 2

Transfer matrix method for modeling of surface SHG from multilayer structures

Surface nonlinearity is manifested only inside a very thin interfacial layer nearby the medium interface due to the discontinuity of both the medium and the normal component of the electric field across the interface leading to the large electric field gradient at the surface [34, 35, 36]. The symmetry of the medium is always broken nearby the surface resulting in the pronounced dipole second-order nonlinearity in the surface region. Because of that, the surface SHG is dominating in the nonlinear response of centrosymmetric materials, including plasmonic structures [37, 38, 39]. Many centrosymmetric materials are widely used in nanophotonics, e.g., Si, SiO₂, TiO₂, Ge and others. Therefore the enhancement of SHG in structures made of centrosymmetric materials requires boosting the surface nonlinear response. Along with that, in III-V semiconductors with strong bulk nonlinearity the surface contribution to the SHG can also be important, for instance, in the opaque region of the medium [47, 45], close to the surface resonances [46] and for SH frequencies above the fundamental bandgap [44]. In order to enhance the SSHG one needs both larger surface-to-volume ratio and strong pump field at the surface. The first condition is fulfilled in nanostructures or nanometer-thick layers. The enhancement of the pump field can be also achieved in nanostructures, which can concentrate the energy of the electromagnetic field inside their volume due to the excitation of resonances in the VIS or MID-IR spectral ranges [67]. It should be mentioned that the PMC is not expected to play a major role for the enhancement of the SSHG. The PMC is crucial for gaining the conversion efficiency of SHG in bulk nonlinear crystals since it guarantees the constant energy flow from FW to the SHW upon propagation. In nanophotonics in general the nanoscale structures possess subwavelength dimensions and the energy transfer can not efficiently occur over so small distances, greatly reducing the effect of the PMC.

In nanostructures made of III-V semiconductors the SHG conversion efficiency was mainly enhanced thanks to the excitation of high-Q resonances [31, 32], which results in an intense fundamental field inside the volume of the nanostructure. However, when dealing with the surface nonlinearities only, the excitation of resonances does not in general provide the gaining of the field nearby the interface as well as the large overlap of the surface nonlinear polarization with the nanostructure resonant modes at the SH frequency. Therefore I aim here to investigate an alternative approach, namely the

interference of the sources of the surface nonlinear polarization inside a multilayer structure. Combining many interfaces can bring several advantages. On one hand, if the constructive interference of the surface sources is achieved, their nonlinear responses can sum up resulting in the enhancement of the SSHG. On the other hand, one-dimensional photonic crystals (PhC) are known to exhibit different resonances, which can also enhance the pump field inside the stack [68]. As the result, multilayer stacks, even composed of centrosymmetric semiconductors or dielectrics, can be expected to yield beneficial performance for the frequency conversion at the nanoscale. Existing fabrication techniques nowadays make it possible to grow layered structures of different materials with high precision and material purity with variable thickness down to a few nanometers [69, 70, 71]. It should be noted that third-order nonlinear effects can arise at some semiconductor interfaces due to the presence of the constant electric field in the surface region because of the charge trapping or interface charging. These third-order nonlinear phenomena are eventually manifested as the surface-like second-order nonlinear effects [72, 73, 74]. All these and other similar effects can be formally described in the same way as the usual surface second-order nonlinearity by adding them to the usual surface nonlinear tensor. It is therefore assumed that all effectively second-order nonlinear phenomena inside the surface region are accounted for in the treatment below.

For modeling of SSHG from multilayer structures it seems especially convenient to apply the transfer-matrix formalism. The transfer matrix method (TMM) is widely used for numerical simulations of the light propagation in different one-dimensional structures, like PhCs, Bragg mirrors, defect structures or other layered media. Some papers have addressed the extension of the TMM for the modeling of nonlinear effects, like SHG or third-harmonic generation (THG). THG in PhC structures was analyzed using the transfer matrix method under the UPA in Ref. [75]. Then, the TMM was elaborated and generalized in a number of works to second-order nonlinear processes as well and to account for the pump depletion [76, 77, 78, 79, 80, 81, 82]. However, as far as I know, the modeling of the SHG in multilayer structures has been limited so far to the bulk nonlinearities only, while surface nonlinearities were always neglected. In several works [83, 84, 85, 86] both nonlinearities were assumed thanks to the solving of the dynamical equations in the time domain for coupled Drude-Lorentz oscillators describing free and bound charges in the medium. Such approach inherently incorporates both surface and bulk contributions simultaneously, though without their explicit separation, and was able to demonstrate the enhancement of the conversion efficiency from a stack due to the larger number of surfaces and the localization of the pump field. Still such modeling by solving the dynamical equations for the media is much more complicated and resource-consuming and is mainly interesting to shed a clearer light on the origin of the nonlinear optical properties and their relation with the crystal properties.

This chapter is organized as follows. The basic theoretical framework of the proposed TMM and the problem statement are given in Section 2.1. In Section 2.2 transfer matrices for TE- and TM-polarized fields at the SH frequency are presented and the expressions relating the emitted fields at the SH frequency with the source vectors and the transfer matrices are derived. In Section 2.3 it is discussed, how the source terms are calculated from the fundamental field, by solving the linear scattering problem at the fundamental frequency with the TMM. Section 2.4 summarizes all steps of the

proposed numerical method and provides its general computational procedure as well as discussion of its applicability and possible extension. Section 2.5 is devoted to the specific case of ultrathin layers in the stack and it is shown how the analytical treatment can be simplified for this case. In Section 2.6 the effective nonlinear parameters of the whole stack for the case of ultrathin layers are provided. The results of the numerical simulations with the proposed numerical method are given in Section 2.7, where the conversion efficiency of SSHG from an exemplary multilayer stack is compared with a single layer of GaAs of similar thickness as well as with a single interface with effective surface nonlinearity of the whole stack. Finally, the summary and some concluding remarks are presented in Section 2.8.

2.1 Model and problem statement

The typical structure to be considered represents a stack of N dielectric layers located on top of a semi-infinite substrate, as schematically shown in Fig. 2.1. The m -th layer in the stack possesses the thickness h_m and the relative permittivity ε_m . The semi-infinite substrate has the dielectric permittivity ε_{sub} , and the medium in the upper half-space is assumed to be air. The stack is illuminated from the top side by a wave at the fundamental frequency. I mainly examine here the simplest case, when the illuminating field is a plane wave with the wavevector \vec{k}^{FW} and the angle of incidence θ .

In the following, it is assumed that the surface nonlinearity is dominating in the nonlinear response of the stack and the bulk nonlinearity can be thus neglected. This assumption is well justified in the case of centrosymmetric semiconductors or dielectrics even for thick layers, because the bulk nonlinearity due to the higher-order multipoles is much weaker than the surface nonlinearity. However, for thin enough layers of noncentrosymmetric materials it is also reasonable to keep the surface nonlinearity only. The general condition, when the bulk nonlinearity is to be neglected in the layer of the thickness h_m , can be written as:

$$\chi_{\text{bulk}}^{(2)} h_m \ll \chi_{\text{surf}}^{(2)}, \quad (2.1)$$

where $\chi_{\text{bulk}}^{(2)}$ and $\chi_{\text{surf}}^{(2)}$ are the bulk and surface second-order nonlinear coefficients respectively. The prevailing contribution of the surface nonlinearity in nanoscale structures made of noncentrosymmetric semiconductors has been found in experiments. For instance, in Refs. [43, 87] it was experimentally measured that the surface contribution becomes dominating in GaP nanopillars for diameters below ~ 170 nm. In Ref. [66] the surface contribution was experimentally found to prevail in the SHG response from 50-nm and 100-nm-thick layers of GaAs. These values can be used as the order-of-magnitude estimates of the threshold thickness in Eq. (2.1). One can also expect these estimates to hold for different noncentrosymmetric materials, including other III-V semiconductors.

In the presence of the surface nonlinear polarization the standard boundary conditions for the components of the electric and magnetic fields have to be replaced with the generalized boundary conditions. The generalized boundary conditions, when the source of the medium polarization is

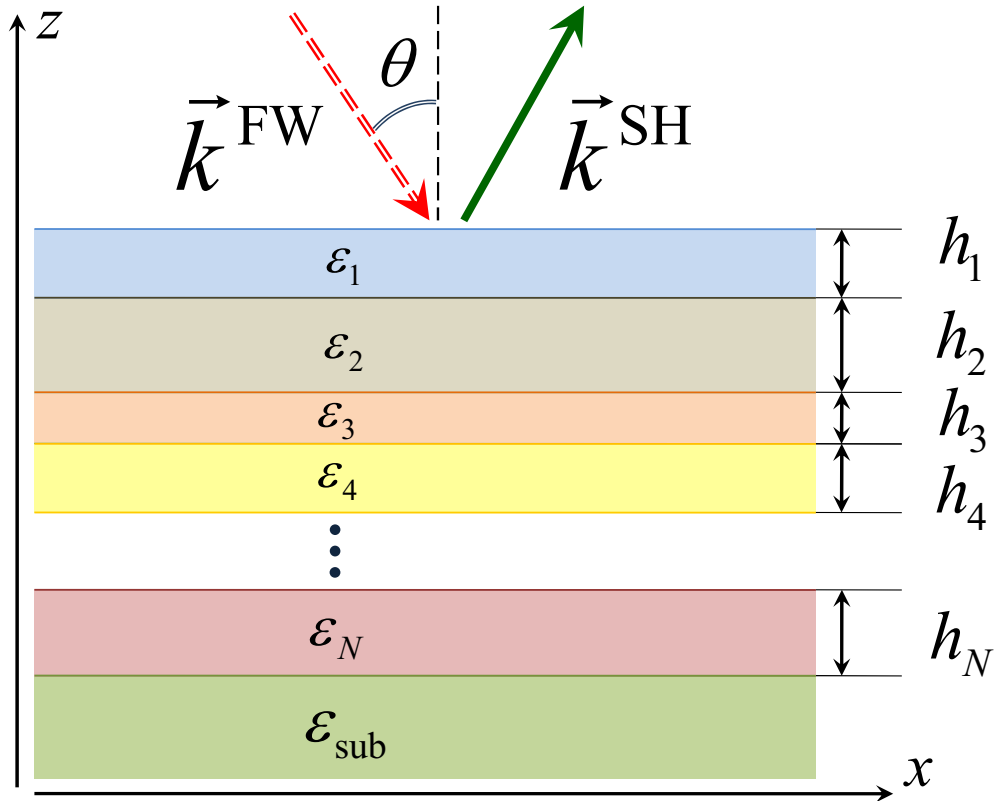


Figure 2.1: Considered multilayer structure consisting of N layers of thicknesses h_m and relative permittivities ϵ_m and placed on the top of a semi-infinite substrate of relative permittivity ϵ_{sub} .

embedded into the interface, were derived in Ref. [88] as:

$$\begin{aligned}\Delta \vec{E}_{\parallel} &= -\frac{1}{\epsilon_0 \epsilon'} \vec{\nabla}_{\parallel} P_{S,\perp}^{\text{NL}}, \\ \Delta \vec{H}_{\parallel} &= -2i\omega \vec{P}_S^{\text{NL}} \times \vec{r}_{\perp},\end{aligned}\quad (2.2)$$

where \vec{P}_S^{NL} is the vector of the surface nonlinear polarization, subindices \perp and \parallel stand for the orthogonal and in-plane components of the corresponding vector, Δ is the jump of the corresponding quantity across the interface and $\vec{\nabla}_{\parallel}$ denotes the two-dimensional gradient calculated in the plane of the interface. The relative permittivity ϵ' corresponds to the medium, where the sheet of the surface nonlinear polarization is located. Boundary conditions Eq. (2.2) are derived for the surface of an isotropic medium. However, they can be easily extended for the case of the surface of an anisotropic optical material. The corresponding derivation and the respective form of Eq. (2.2) are provided in the Appendix A.

Several remarks must be added here regarding the application of the generalized boundary conditions Eq. (2.2) to the multilayer structures. As it was stated, the sheet of the surface nonlinear polarization is assumed to be placed inside the medium with the relative permittivity ϵ' . For any interface inside the stack two options are possible, when the permittivity ϵ' corresponds either to the medium above or below the interface. I will follow hereafter the standard approach, when the com-

ponents of the vector of the surface nonlinear polarization are calculated using the pump electric field just below the interface and the sheet of the surface nonlinear polarization is placed just above the interface. Hence, at the interface between m -th and $m + 1$ -th layer inside the multilayer stack the value $\varepsilon' = \varepsilon_m$ is accepted. Such choice is done for convenience and potentially one can also place the sheet of the surface nonlinear polarization below the respective interface and/or use the components of the pump electric field above the interface. In this case the surface nonlinear tensor has to be rescaled in a proper way to guarantee the consistency of the calculations regardless of the specific choice [38].

The most important case of the illuminating field is a plane wave at the fundamental frequency. For the plane-wave illumination the generalized boundary conditions Eq. (2.2) allow obtaining certain conclusions for the generated SH field. Specifically, for the normal component of the surface nonlinear polarization at each interface one finds:

$$P_{S,\perp}^{\text{NL}} \sim e^{2i\vec{k}_{\parallel}^{\text{FW}}\vec{r}}, \quad (2.3)$$

where $\vec{k}_{\parallel}^{\text{FW}}$ is the tangential component of the wavevector of the incident plane wave. Following the standard boundary conditions for the electric field, which are still valid at the fundamental frequency, the vector $\vec{k}_{\parallel}^{\text{FW}}$ is constant throughout the whole stack. Therefore the following equalities hold:

$$\Delta\vec{E}_{\parallel} \parallel \vec{\nabla}_{\parallel} P_{S,\perp}^{\text{NL}} \parallel \vec{k}_{\parallel}^{\text{FW}}, \quad (2.4)$$

i.e., the jump of the tangential components of the electric field has a fixed direction at all interfaces inside the considered multilayer stack. From Eq. (2.3) one can also conclude that all SHWs inside the stack obey the following equation:

$$\vec{k}_{\parallel}^{\text{SH}} = 2\vec{k}_{\parallel}^{\text{FW}}. \quad (2.5)$$

Eqs. (2.4) and (2.5) determine which second-harmonic plane waves can be emitted. For example, a TE-polarized plane wave with the electric field in the direction of $\Delta\vec{E}_{\parallel}$ cannot be radiated, because in this case one would get a plane wave with $\vec{E} \cdot \vec{k}^{\text{SH}} \neq 0$. For the same reason a TM-polarized plane wave with the magnetic field in the direction of $\vec{k}_{\parallel}^{\text{FW}}$ cannot be generated, even though $\Delta\vec{H}_{\parallel}$ in general can possess a nonzero component in this direction at each interface.

Having obtained the equalities above, one is ready now to proceed with the analytical formalism for the transfer matrix method. Let us assume for definiteness, that the electric field at SH frequency is directed along the x -axis and the magnetic field is directed along the y -axis. Using the transfer matrix for the m -th layer in the stack [68], one arrives to the following relation between the fields in the m -th and $m + 1$ -th layers:

$$\begin{bmatrix} E_x \\ H_y \end{bmatrix}_{m+1,+} = \hat{M}_m \cdot \begin{bmatrix} E_x \\ H_y \end{bmatrix}_{m,+} + \hat{S}_{m(m+1)}, \quad (2.6)$$

where \hat{M}_m is the transfer matrix of the m -th layer, and the source vector is

$$\hat{S}_{m(m+1)} = \begin{bmatrix} -\Delta E_{||,m(m+1)} \\ -\Delta H_{||,m(m+1)} \end{bmatrix}, \quad (2.7)$$

where the minus sign appears because the field below the interface is expressed through the field above. The subscript "+" in Eq. (2.6) denotes the field just below the top interface of the respective layer, while the subscript "-" would stand for the field just above the bottom interface of the layer. Eq. (2.6) differs from the matrix equations in the conventional transfer-matrix analysis by the extra source vector in the right-hand side, arising thanks to the presence of the surface nonlinear polarization at the interfaces. Combining together Eq. (2.6) for each single interface in the stack, one arrives to an equation connecting the SHWs emitted into the upper medium and into the substrate:

$$\begin{bmatrix} E_x \\ H_y \end{bmatrix}_{N+1,+} = \hat{M}_\Sigma \cdot \begin{bmatrix} E_x \\ H_y \end{bmatrix}_{0,-} + \hat{S}_\Sigma \quad (2.8)$$

with matrices

$$\begin{aligned} \hat{M}_\Sigma &= \prod_{m=1}^N \hat{M}_m, \\ \hat{S}_\Sigma &= \hat{S}_{N(N+1)} + \hat{M}_N \times \hat{S}_{N(N-1)} + \hat{M}_N \times \hat{M}_{N-1} \times \hat{S}_{(N-1)(N-2)} + \dots + \hat{M}_\Sigma \times \hat{S}_{01}. \end{aligned} \quad (2.9)$$

Every term in the sum for \hat{S}_Σ in Eq. (2.9) describes the partial contribution of the interface between m -th and $m+1$ -th layers into the overall SHW from the whole stack. Eq. (2.9) attains a particularly simple form in the case of only one interface, i.e., in the absence of any layers on the top of the substrate:

$$\hat{M}_\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \hat{S}_\Sigma = \begin{bmatrix} -\Delta E_{||,01} \\ -\Delta H_{||,01} \end{bmatrix}.$$

Here it becomes obvious, how Eqs. (2.8) and (2.9) connect the fields above and below the single interface.

The emitted SH field in general has mixed polarization, i.e. contains TE- and TM-polarized components. The generalized boundary conditions Eq. (2.2), on the other hand, provide the jump of the total field only. Therefore it is necessary to split Eq. (2.2) into two terms corresponding to the jumps of the electric and the magnetic fields for different polarizations of the SHW. This separation can be easily done for the illumination by the plane wave at the fundamental frequency, while for an arbitrary incident field this can be a challenging task and requires extra efforts. Eq. (2.4) states that the jumps of the tangential component of the electric field $\Delta E_{||,m(m+1)}$ have a fixed direction at each interface inside the stack, namely the direction of $\vec{k}_{||}^{\text{FW}}$. Besides that $\vec{k}_{||}^{\text{SH}} = 2\vec{k}_{||}^{\text{SH}}$ in each layer. Let us assume the x -axis to be directed along $\vec{k}_{||}^{\text{FW}}$, then both $\vec{k}_{||}^{\text{SH}}$ and all the jumps of the electric field $\Delta E_{||,m(m+1)}$ are also directed along the x -axis. The electric field in a plane wave has to be orthogonal to the wavevector and thus cannot possess a non-zero x -component. This contradiction implies that

a TE-polarized SHW with electric field along the x -axis cannot be generated at any interface inside the stack. Therefore, the jumps $\Delta E_{||,m(m+1)}$ can only produce the TM-polarized SHWs in every layer inside the stack with the vector of the electric field in the xz -plane and the vector of the magnetic field along y -axis. The respective form of the boundary conditions from Eq. (2.2) for TM-polarized SHW is:

$$\begin{cases} \Delta \vec{E}_{||,TM} = \Delta \vec{E}_{||} \\ \Delta \vec{H}_{||,TM} = \Delta \vec{H}_{||}|_y. \end{cases} \quad (2.10)$$

Since the jumps of the tangential components of the electric and the magnetic field in Eq. (2.2) are independent, $\Delta H_{||,m(m+1)}$ at every interface would in general possess both x - and y -components. While its y -component contributes to the emission of the TM-polarized SHW, the x -component would generate the TE-polarized SHW with magnetic field in the xz -plane and the electric field along the y -axis. Due to the equality $\Delta E_{||,m(m+1)} \parallel \vec{e}_x$ for any m , the boundary conditions for TE-polarized SHW are given as:

$$\begin{cases} \Delta \vec{E}_{||,TE} = 0 \\ \Delta \vec{H}_{||,TE} = \Delta \vec{H}_{||}|_x. \end{cases} \quad (2.11)$$

The transfer matrices for TE- and TM-polarized fields are different, therefore both polarizations of the FW have to be treated separately. The corresponding expressions for the implementation of the proposed TMM are derived in the next sections both for TE- and TM-polarized FWs and SHWs.

2.2 Calculation of the emitted SHW by TMM

In this section both possible polarizations of the emitted second-harmonic wave are dealt with separately. I begin with the case of a TE-polarized SHW. Following the notation above, it is assumed that the electric field is directed along the y -axis, so that Eq. (2.8) yields:

$$\begin{bmatrix} E_{TE,y} \\ H_{TE,x} \end{bmatrix}_{N+1,+} = \hat{M}_{\Sigma}^{\text{TE}} \cdot \begin{bmatrix} E_{TE,y} \\ H_{TE,x} \end{bmatrix}_{0,-} + \hat{S}_{\Sigma}^{\text{TE}}. \quad (2.12)$$

For TE-polarization the transfer matrix of the m -th layer \hat{M}_m^{TE} has the following form:

$$\hat{M}_m^{\text{TE}} = \begin{bmatrix} \cos k_{m,z}^{\text{SH}} h_m & iZ'_m \sin k_{m,z}^{\text{SH}} h_m \\ \frac{i}{Z'_m} \sin k_{m,z}^{\text{SH}} h_m & \cos k_{m,z}^{\text{SH}} h_m \end{bmatrix}, \quad (2.13)$$

where

$$Z = \sqrt{\frac{\mu_0 \mu}{\varepsilon_0 \varepsilon}} \quad (2.14)$$

is the wave impedance of the material of the m -th layer and:

$$Z' = Z \cdot \frac{k_m}{k_{m,z}}. \quad (2.15)$$

The transfer matrix of the whole stack for TE-polarized SHW attains the form:

$$\hat{M}_{\Sigma}^{\text{TE}} = \prod_{m=1}^N \hat{M}_m^{\text{TE}}. \quad (2.16)$$

I would like to state here, that another implementation of the TMM exists, where the matrices are written for forward- and backward-propagating waves inside each layer. This approach is often used for simulations of bulk SHG in multilayer structures [75, 76, 78, 79, 80, 81, 82]. In this work the more common approach is used instead and the transfer matrices are written for the tangential components of the electric and magnetic fields inside the stack [68]. There are several reasons for this choice. Firstly, the generalized boundary conditions Eq. (2.2) provide the jump of the total electric and magnetic fields. When decomposing the fields inside each layer into forward- and backward-propagating waves, the boundary conditions Eq. (2.2) have to be split into two terms respectively, what can be difficult for complex illuminating fields. Secondly, the surface nonlinear polarization at each interface inside the stack depends on the components of the total field, i.e. the sum of the fields of the forward- and backward-propagating waves. Hence, the transfer matrices for the tangential components of the electric and magnetic fields (without their separation into forward- and backward-propagating components) appear to be more convenient for simulations.

Eq. (2.12) allows direct calculation of the amplitudes of the outgoing SHW for the plane-wave FW. Indeed, both above (in the upper space) and below the stack (in the substrate) one gets an outgoing plane wave at the SH frequency. Let us denote respective amplitudes of the electric field strength in both waves as A^+ and A^- . The electric and magnetic fields in the plane wave are related with each other by:

$$\begin{bmatrix} E_{\text{TE},y} \\ H_{\text{TE},x} \end{bmatrix}_{N+1,+} = \begin{bmatrix} A^- \\ A^-/Z'_{\text{sub}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} E_{\text{TE},y} \\ H_{\text{TE},x} \end{bmatrix}_{0,-} = \begin{bmatrix} A^+ \\ -A^+/Z'_0 \end{bmatrix}, \quad (2.17)$$

where $H_{\Sigma} = h_1 + h_2 + \dots + h_N$ is the overall thickness of the multilayer structure. Now one can substitute the obtained expressions Eq. (2.13)-(2.17) into Eq. (2.12) and arrives to the system of linear equations for the unknown amplitudes A^+ and A^- :

$$\begin{bmatrix} A^- \\ A^-/Z'_{\text{sub}} \end{bmatrix} = \hat{M}_{\Sigma}^{\text{TE}} \cdot \begin{bmatrix} A^+ \\ -A^+/Z'_0 \end{bmatrix} + \hat{S}_{\Sigma}^{\text{TE}}. \quad (2.18)$$

From Eq. (2.18), for instance, the field in the upper space A^+ is readily found as:

$$A^+ = \frac{S_{\Sigma,2}^{\text{TE}} Z'_0 Z'_{\text{sub}} - S_{\Sigma,1}^{\text{TE}} Z'_0}{M_{11} Z'_0 + M_{22} Z'_{\text{sub}} - M_{12} - M_{21} Z'_0 Z'_{\text{sub}}}. \quad (2.19)$$

where the entries of the total transfer matrix $M_{ij} = \hat{M}_{\Sigma}^{\text{TE}}(i, j)$ are used.

Next, let us move on to the case of TM-polarization of the emitted SHW. The electric field in the TM-polarized SHW is assumed to be located in the xz -plane with the magnetic field directed along the y -axis, therefore Eq. (2.8) is given as:

$$\begin{bmatrix} E_{\text{TM},x} \\ H_{\text{TM},y} \end{bmatrix}_{N+1,+} = \hat{M}_{\Sigma}^{\text{TM}} \cdot \begin{bmatrix} E_{\text{TM},x} \\ H_{\text{TM},y} \end{bmatrix}_{0,-} + \hat{S}_{\Sigma}^{\text{TM}}. \quad (2.20)$$

For TM-polarization the transfer matrix of m -th layer \hat{M}_m^{TM} has the following form:

$$\hat{M}_m^{\text{TM}} = \begin{bmatrix} \cos k_{m,z}^{\text{SH}} h_m & -iZ_m'' \sin k_{m,z}^{\text{SH}} h_m \\ -\frac{i}{Z_m''} \sin k_{m,z}^{\text{SH}} h_m & \cos k_{m,z}^{\text{SH}} h_m \end{bmatrix}, \quad (2.21)$$

where it is denoted

$$Z_m'' = Z_m \cdot \frac{k_{m,z}}{k_m}. \quad (2.22)$$

The transfer matrix of the whole stack for TM-polarized SHW is:

$$\hat{M}_{\Sigma}^{\text{TM}} = \prod_{m=1}^N \hat{M}_m^{\text{TM}}. \quad (2.23)$$

Again one gets both above and below the stack outgoing plane waves at the SH frequency, and the amplitudes of the x -component of the electric field strength are denoted as A^+ and A^- respectively. The relations between electric and magnetic field in the plane wave yield:

$$\begin{bmatrix} E_{\text{TM},x} \\ H_{\text{TM},y} \end{bmatrix}_{N+1,+} = \begin{bmatrix} A^- \\ -A^-/Z_{\text{sub}}'' \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} E_{\text{TM},x} \\ H_{\text{TM},y} \end{bmatrix}_{0,-} = \begin{bmatrix} A^+ \\ A^+/Z_0'' \end{bmatrix}. \quad (2.24)$$

The expressions Eqs. (2.21)-(2.24) now have to be substituted into Eq. (2.20), which gives the following system of equations for the unknown amplitudes A^+ and A^- :

$$\begin{bmatrix} A^- \\ -A^-/Z_{\text{sub}}'' \end{bmatrix} = \hat{M}_{\Sigma}^{\text{TM}} \cdot \begin{bmatrix} A^+ \\ A^+/Z_0'' \end{bmatrix} + \hat{S}_{\Sigma}^{\text{TM}}. \quad (2.25)$$

Now for the field in the upper space A^+ one finds:

$$A^+ = \frac{-S_{\Sigma,2}^{\text{TM}} Z_0'' Z_{\text{sub}}'' - S_{\Sigma,1}^{\text{TM}} Z_0''}{M_{11} Z_0'' + M_{22} Z_{\text{sub}}'' + M_{12} + M_{21} Z_0'' Z_{\text{sub}}''}, \quad (2.26)$$

with the entries of the total transfer matrix $M_{ij} = \hat{M}_{\Sigma}^{\text{TM}}(i, j)$.

2.3 Calculation of the surface nonlinear polarization

Eqs. (2.19) and (2.26) for the amplitudes of the emitted SHW contain the source vectors $\hat{S}_{\Sigma}^{\text{TE}}, \hat{S}_{\Sigma}^{\text{TM}}$. In order to calculate them, one needs to know the surface nonlinear polarization at every interface in the stack expressed through the parameters of the illuminating FW.

The surface nonlinear polarization is in general given by an expression of the form Eq. (1.2). Here for simplicity, the surfaces are assumed to be isotropic. At the isotropic surface the general expression reduces to only three nonzero terms as follows :

$$\begin{aligned} \vec{P}_S^{\text{NL}}(2\omega, \vec{r}) &= \varepsilon_0 \chi_{\perp\perp\perp}^{(2)} E_{\perp}^{(\omega)}(\vec{r}) E_{\perp}^{(\omega)}(\vec{r}) \cdot \vec{n} + \varepsilon_0 \chi_{\perp\parallel\parallel}^{(2)} E_{\perp}^{(\omega)}(\vec{r}) E_{\parallel}^{(\omega)}(\vec{r}) \cdot \vec{n} \\ &+ \varepsilon_0 \chi_{\parallel\perp\parallel}^{(2)} E_{\perp}^{(\omega)}(\vec{r}) E_{\parallel}^{(\omega)}(\vec{r}) \cdot \vec{\tau}, \end{aligned} \quad (2.27)$$

where \vec{n} is the outward normal unit vector and $\vec{\tau}$ is the surface tangent unit vector at the point \vec{r} , pointing in the direction of $\vec{E}_{\parallel}^{(\omega)}(\vec{r})$. If one takes a noncentrosymmetric semiconductor material, the surface would not be isotropic and its symmetry would depend on the crystal symmetry and the orientation of the crystal axes with respect to the surface normal at each specific point on the surface. This means that other nonzero terms would correspondingly arise in Eq. (2.27). The form of the surface nonlinear tensor for noncentrosymmetric III-V semiconductors will be discussed later in chapter 4, but in this chapter this complex case is not considered.

One can see from Eq. (2.27), that one has to calculate the field components $E_{\perp}^{(\omega)}, E_{\parallel}^{(\omega)}$ at every interface inside the multilayer stack in order to obtain the source vectors $\hat{S}_{\Sigma}^{\text{TE}}, \hat{S}_{\Sigma}^{\text{TM}}$. It should be also emphasized that the standard notation for the surface nonlinearity is used in the following, with the electric field of the FW in Eq. (2.27) taken just below the corresponding interface, and the sheet of the surface nonlinear polarization located just above the interface.

First, the case of TE-polarization of the illuminating plane wave is considered. Let E_I be the amplitude of the electric field strength in the FW and R_{TE} the amplitude reflection coefficient of the whole stack at fundamental frequency for TE-polarized incident plane wave. According to the notation in the previous section, vectors $\vec{k}_{\parallel}^{\text{FW}}$ and $\vec{k}_{\parallel}^{\text{SH}}$ are assumed to be directed along x -axis. Given that the incident plane wave is TE-polarized, its electric field therefore has to be directed along the y -axis and the magnetic field lies in the xz -plane. The transfer matrices from Eqs. (2.13)-(2.16) allow calculation of the electric field at the fundamental frequency just below the m -th layer:

$$\begin{aligned} \begin{bmatrix} E_{\text{TE},y} \\ H_{\text{TE},x} \end{bmatrix}_{m+1,+} &= \hat{M}_{m,\text{FW}}^{\text{TE}} \times \hat{M}_{m-1,\text{FW}}^{\text{TE}} \times \dots \times \hat{M}_{2,\text{FW}}^{\text{TE}} \times \hat{M}_{1,\text{FW}}^{\text{TE}} \times \begin{bmatrix} E_{\text{TE},y} \\ H_{\text{TE},x} \end{bmatrix}_{0,-} \\ &= \prod_{j=1}^m \hat{M}_{j,\text{FW}}^{\text{TE}} \times \begin{bmatrix} E_I(1 + R_{\text{TE}}) \\ E_I \frac{1 - R_{\text{TE}}}{Z_{0,\text{FW}}} \end{bmatrix}. \end{aligned} \quad (2.28)$$

Next, TM-polarization of the incident plane wave is considered in the same manner. Let us denote as E_I the amplitude of the electric field strength in the FW and as R_{TM} the amplitude reflection coefficient of the whole stack at the fundamental frequency for TM-polarized incident plane wave. The vectors $\vec{k}_{\parallel}^{\text{FW}}$ and $\vec{k}_{\parallel}^{\text{SH}}$ are assumed to be directed along the x -axis, so that the electric field in the

TM-polarized FW lies in the xz -plane and the magnetic field is directed along the y -axis. With the transfer matrices Eqs. (2.21)-(2.23) one obtains for the electric field at the fundamental frequency just below the m -th layer:

$$\begin{aligned} \begin{bmatrix} E_{\text{TM},x} \\ H_{\text{TM},y} \end{bmatrix}_{m+1,+} &= \hat{M}_{m,\text{FW}}^{\text{TM}} \times \hat{M}_{m-1,\text{FW}}^{\text{TM}} \times \dots \times \hat{M}_{2,\text{FW}}^{\text{TM}} \times \hat{M}_{1,\text{FW}}^{\text{TM}} \times \begin{bmatrix} E_{\text{TM},x} \\ H_{\text{TM},y} \end{bmatrix}_{0,-} \\ &= \prod_{j=1}^m \hat{M}_{j,\text{FW}}^{\text{TM}} \times \begin{bmatrix} E_{\text{I},x}(1+R_{\text{TM}}) \\ -E_{\text{I},x} \frac{1-R_{\text{TM}}}{Z'_{0,\text{FW}}} \end{bmatrix}. \end{aligned} \quad (2.29)$$

This expression gives only the tangential component of the electric field inside the stack. The normal component $E_{\text{TM},\perp}$ can be calculated using the following relation between the electric and the magnetic field in a plane wave inside $m+1$ -th layer (i.e. just below m -th layer):

$$E_{\text{TM},z} = -\frac{k_{m+1,x}^{\text{FW}}}{k_{m+1}^{\text{FW}}} \cdot Z_{m+1,\text{FW}} \cdot H_{\text{TM},y}, \quad (2.30)$$

where the tangential component of the magnetic field $H_{\text{TM},y}$ is provided by Eq. (2.29).

Lastly, I would like to present the equations for the amplitude reflection coefficients at the fundamental frequency R_{TE} and R_{TM} through the respective transfer matrices. If the incident field on the stack at the fundamental frequency and the transmitted field for TE-polarization are denoted as $E_{\text{I,TE/TM}}$ and $E_{\text{T,TE/TM}}$ respectively, both fields can be related through the transfer matrix of the stack similar to Eqs. (2.18) and (2.28) as follows:

$$\begin{bmatrix} E_{\text{T,TE}} \\ E_{\text{T,TE}}/Z'_{\text{sub,FW}} \end{bmatrix} = \hat{M}_{\Sigma,\text{FW}}^{\text{TE}} \times \begin{bmatrix} E_{\text{I}}(1+R_{\text{TE}}) \\ E_{\text{I}} \frac{1-R_{\text{TE}}}{Z'_{0,\text{FW}}} \end{bmatrix}. \quad (2.31)$$

The indices "FW" here are written in order to state that matrix $\hat{M}_{\Sigma,\text{FW}}^{\text{TE}}$ as well as all values Z'_m are taken at the fundamental frequency, while in all other equations they are assumed to be taken at the SH frequency. Eq. (2.31) yields for the amplitude reflection coefficient R_{TE} :

$$R_{\text{TE}} = \frac{M_{22}Z'_{\text{sub,FW}} - M_{11}Z'_{0,\text{FW}} + M_{21}Z'_{0,\text{FW}}Z'_{\text{sub,FW}} - M_{12}}{M_{22}Z'_{\text{sub,FW}} + M_{11}Z'_{0,\text{FW}} - M_{21}Z'_{0,\text{FW}}Z'_{\text{sub,FW}} - M_{12}}, \quad (2.32)$$

where the entries of the total transfer matrix $M_{ij} = \hat{M}_{\Sigma,\text{FW}}^{\text{TE}}(i, j)$ are used. An analogous derivation for the TM-polarization gives for the amplitude reflection coefficient R_{TM} :

$$R_{\text{TM}} = \frac{M_{22}Z''_{\text{sub,FW}} - M_{11}Z''_{0,\text{FW}} - M_{21}Z''_{0,\text{FW}}Z''_{\text{sub,FW}} + M_{12}}{M_{22}Z''_{\text{sub,FW}} + M_{11}Z''_{0,\text{FW}} + M_{21}Z''_{0,\text{FW}}Z''_{\text{sub,FW}} + M_{12}}, \quad (2.33)$$

with the entries of the total transfer matrix $M_{ij} = \hat{M}_{\Sigma,\text{FW}}^{\text{TM}}(i, j)$.

2.4 Computational procedure with the proposed TMM

Having derived in the previous sections all the necessary equations, I would like now to put them together and provide a general computational procedure for the proposed transfer matrix method. The application of the proposed TMM should be done according to the following steps:

1. Finding the transfer matrix for the whole multilayer structure at the fundamental frequency using Eq. (2.9) and the amplitude reflection coefficient $R_{TE/TM}$ from Eqs. (2.32) and (2.33).
2. Calculation of the normal and tangential components of the electric field at the fundamental frequency at each interface inside the stack using Eqs. (2.28)-(2.30).
3. Calculation of the vector of the surface nonlinear polarization with Eq. (2.27) at each interface using the electric fields obtained at step (2).
4. Finding the jumps of the electric and magnetic fields given by the generalized boundary conditions Eq. (2.2). The boundary conditions then have to be split into two terms, corresponding to the TE- and TM-polarized SHWs according to Eqs. (2.10)-(2.11)
5. Construction of the matrices $\hat{M}_\Sigma, \hat{S}_\Sigma$ for the whole multilayer structure at the SH frequency using Eqs. (2.9).
6. Calculation of the amplitude of the outgoing SHW from the multilayer structure using Eq. (2.19) or (2.26) for TE- and TM-polarized SHW respectively.

Eqs. (2.19) or (2.26) are derived for the upwards emitted SHW. The amplitudes of the SHW transmitted into the substrate can be found from Eqs. (2.18) or Eqs. (2.25). The described computational procedure allows calculation of the surface-driven SHG from an arbitrary multilayer structure. The proposed method by design accounts for the multiple reflections both at the fundamental and the second-harmonic frequency and is valid for arbitrary angles of incidence of the illuminating field. It should be also noted that in the treatment above UPA is used, so the reverse action of the SHW on the FW is neglected, what is typically the case in nanophotonics.

Several remarks have to be made regarding the applicability of the proposed computational method. All the treatment above dealt with the plane-wave illumination and the isotropic surfaces inside the stack. Therefore the extension of the method beyond these simplifications has to be addressed. For example, in order to treat an arbitrary illuminating field, this illuminating field at fundamental frequency has to be expanded into Fourier components. Afterwards the steps (1)-(2) of the procedure above have to be implemented for every single Fourier component of the illuminating field. Then the inverse Fourier transform has to be run at each interface to calculate the corresponding vectors of the surface nonlinear polarization Eq. (2.27) and the generalized boundary conditions Eq. (2.2) at the steps (3)-(4) of the procedure. Finally, the jumps of the tangential components of the electric and magnetic fields have to be expanded again into Fourier series at the SH frequency and steps (5)-(6) should be done for each obtained Fourier component separately. It should be noted that this procedure can be also directly applied to the case of the sum-frequency generation (SFG).

The extension to the case of anisotropic surfaces can be also readily done. In this case extra terms must be added to the expression for the surface nonlinear polarization in Eq. (2.27). Such anisotropic optical properties are inherently present at the surfaces of noncentrosymmetric semiconductors, but may also arise at the surfaces of centrosymmetric materials due to the surface crystallization or ox-

idation, presence of defects or other fabrication issues. The rest of the steps of the computational procedure except for the step (3) stay unchanged.

The proposed procedure is applicable for an arbitrary composition of the stack and arbitrary thicknesses of the layers, provided only that the bulk nonlinearity can be neglected. Under these circumstances the proposed computational method yields analytically strict results for the amplitudes of the emitted SHW. However, when the stack consists of ultrathin layers, the obtained equations can be largely simplified and even explicit expressions for the SHW amplitude can be derived. Therefore the next section is devoted to this interesting specific case of the multilayer stack.

2.5 Multilayer structure of ultrathin layers

In this section I aim to derive approximate analytical expressions for the emitted SHW, when the multilayer structure is composed of ultrathin layers. More specifically, the following inequality is assumed to hold for each layer in the stack:

$$k_m^{\text{SH}} h_m \ll 1, \quad (2.34)$$

i.e. the SH wavelength inside the materials is much smaller than thickness of the respective layer. If the condition Eq. (2.34) is fulfilled, Eq. (2.1) can also be expected to be fulfilled, since both of them put the upper limit on the thicknesses of the layers.

When the condition Eq. (2.34) holds for each layer in the stack, the entries of the transfer matrix Eq. (2.13) for a TE-polarized SHW can be expanded up to the first-order terms of the small parameters $k_{z,m}^{\text{SH}} h_m$:

$$\hat{M}_m^{\text{TE}} \approx \begin{bmatrix} 1 & i Z'_m k_{m,z}^{\text{SH}} h_m \\ \frac{i}{Z'_m} k_{m,z}^{\text{SH}} h_m & 1 \end{bmatrix},$$

and for the transfer matrix of the whole stack $\hat{M}_\Sigma^{\text{TE}}$ the following approximate expression can be obtained:

$$\hat{M}_\Sigma^{\text{TE}} = \prod_{m=1}^N \hat{M}_m \approx \begin{bmatrix} 1 & i \sum_{m=1}^N Z'_m k_{m,z}^{\text{SH}} h_m \\ i \sum_{m=1}^N \frac{1}{Z'_m} k_{m,z}^{\text{SH}} h_m & 1 \end{bmatrix}. \quad (2.35)$$

The source vector $\hat{S}_\Sigma^{\text{TE}}$ for the TE-polarized SHW is also expanded as:

$$\hat{S}_\Sigma^{\text{TE}} = \sum_{m=0}^N \hat{S}_{m(m+1)}^{\text{TE}} + \tilde{S}_\Sigma^{\text{TE}}$$

where the last term is given as:

$$\tilde{S}_\Sigma^{\text{TE}} = \begin{bmatrix} i \sum_{m=0}^{N-1} S_{m(m+1),2}^{\text{TE}} \cdot \sum_{j=m+1}^N Z'_j k_{j,z}^{\text{SH}} h_j \\ i \sum_{m=0}^{N-1} S_{m(m+1),1}^{\text{TE}} \cdot \sum_{j=m+1}^N \frac{k_{j,z}^{\text{SH}} h_j}{Z'_j} \end{bmatrix}. \quad (2.36)$$

Now Eq. (2.19) for the amplitude of the electric field in the outgoing SHW can be expanded into the series over small parameters $k_{z,m}^{\text{SH}} h_m$. Introducing A_0^+ as the limiting value of A^+ when all $k_{z,m}^{\text{SH}} h_m \rightarrow 0$, one reduces Eq. (2.19) to:

$$A^+ = A_0^+ - \tilde{S}_{\Sigma,1}^{\text{TE}} \cdot \frac{Z'_0}{Z'_0 + Z'_{\text{sub}}} + \tilde{S}_{\Sigma,2}^{\text{TE}} \cdot \frac{Z'_0 Z'_{\text{sub}}}{Z'_0 + Z'_{\text{sub}}} + \frac{\tilde{S}_{\Sigma,2}^{\text{TE}} Z'_0 Z'_{\text{sub}} - \tilde{S}_{\Sigma,1}^{\text{TE}} Z'_0}{(Z'_0 + Z'_{\text{sub}})^2} \times (M_{12} + M_{21} Z'_0 Z'_{\text{sub}}). \quad (2.37)$$

The same treatment for a TM-polarized SHW under the same condition Eq. (2.34) yields for the entries of the transfer matrix Eq. (2.21) up to the first-order terms of small parameters $k_{z,m}^{\text{SH}} h_m$:

$$\hat{M}_m^{\text{TM}} \approx \begin{bmatrix} 1 & -i Z''_m k_{m,z}^{\text{SH}} h_m \\ -\frac{i}{Z''_m} k_{m,z}^{\text{SH}} h_m & 1 \end{bmatrix},$$

and the transfer matrix of the whole stack is obtained as:

$$\hat{M}_{\Sigma}^{\text{TM}} = \prod_{m=1}^N \hat{M}_m \approx \begin{bmatrix} 1 & -i \sum_{m=1}^N Z''_m k_{m,z}^{\text{SH}} h_m \\ -i \sum_{m=1}^N \frac{1}{Z''_m} k_{m,z}^{\text{SH}} h_m & 1 \end{bmatrix}. \quad (2.38)$$

Now, the source vector $\hat{S}_{\Sigma}^{\text{TM}}$ can be represented in the following form:

$$\hat{S}_{\Sigma}^{\text{TM}} = \sum_{m=0}^N \hat{S}_{m(m+1)}^{\text{TM}} + \tilde{S}_{\Sigma}^{\text{TM}}$$

with the last term:

$$\tilde{S}_{\Sigma}^{\text{TM}} = - \begin{bmatrix} i \sum_{m=0}^{N-1} \hat{S}_{m(m+1),2}^{\text{TM}} \cdot \sum_{j=m+1}^N Z''_m k_{m,z}^{\text{SH}} h_m \\ i \sum_{m=0}^{N-1} \hat{S}_{m(m+1),1}^{\text{TM}} \cdot \sum_{j=m+1}^N \frac{k_{m,z}^{\text{SH}} h_m}{Z''_m} \end{bmatrix}. \quad (2.39)$$

The expression Eq. (2.26) for the SHW amplitude in the upper space is again to be expanded into a series over the small parameters $k_{z,m}^{\text{SH}} h_m$. Introducing A_0^+ as the limiting value of A^+ when all $k_{z,m}^{\text{SH}} h_m \rightarrow 0$, one gets using Eq. (2.26):

$$A^+ = A_0^+ - \tilde{S}_{\Sigma,1}^{\text{TM}} \cdot \frac{Z''_0}{Z''_0 + Z''_{\text{sub}}} - \tilde{S}_{\Sigma,2}^{\text{TM}} \cdot \frac{Z''_0 Z''_{\text{sub}}}{Z''_0 + Z''_{\text{sub}}} + \frac{\tilde{S}_{\Sigma,2}^{\text{TM}} Z''_0 Z''_{\text{sub}} + \tilde{S}_{\Sigma,1}^{\text{TM}} Z''_0}{(Z''_0 + Z''_{\text{sub}})^2} \cdot (M_{12} + M_{21} Z''_0 Z''_{\text{sub}}). \quad (2.40)$$

To apply the expressions for the emitted SHW Eq. (2.37) and (2.40) one has to get a corresponding expansion of the source terms, and thus the electric field components at the fundamental frequency at each interface. For a TE-polarized illuminating plane wave at the fundamental frequency Eq. (2.28)

for the electric field just below the m -th layer can be simplified using Eq. (2.35) to :

$$E_{\text{TE},\parallel} \Big|_{m+1,+} = E_I \left[1 + R_{\text{TE}} + i \frac{1 - R_{\text{TE}}}{Z'_{0,\text{FW}}} \cdot \sum_{j=1}^m Z'_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j \right]. \quad (2.41)$$

For a TM-polarized incident plane wave at the fundamental frequency Eqs. (2.29) and (2.38) give for the electric field just below the m -th layer:

$$E_{\text{TM},\parallel} \Big|_{m+1,+} = E_{I,\parallel} \left[1 + R_{\text{TM}} + i \frac{1 - R_{\text{TM}}}{Z''_{0,\text{FW}}} \cdot \sum_{j=1}^m Z''_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j \right]. \quad (2.42)$$

The magnetic field component $H_{\text{TM},y}$ inside the stack can be found from Eq. (2.29):

$$H_{\text{TM},y} \Big|_{m+1,+} = E_{I,x} \left[-\frac{1 - R_{\text{TM}}}{Z''_{0,\text{FW}}} - i(1 + R_{\text{TM}}) \cdot \sum_{j=1}^m \frac{k_{j,z}^{\text{FW}} h_j}{Z''_{j,\text{FW}}} \right],$$

which gives for the normal component $E_{\text{TM},\perp}$ according to Eq. (2.30):

$$E_{\text{TM},\perp} \Big|_{m+1,+} = -\frac{k_{m+1,x}^{\text{FW}}}{k_{m+1}^{\text{FW}}} \cdot Z_{m+1,\text{FW}} \cdot E_{I,x} \cdot \left[-\frac{1 - R_{\text{TM}}}{Z''_{0,\text{FW}}} - i(1 + R_{\text{TM}}) \cdot \sum_{j=1}^m \frac{k_{j,z}^{\text{FW}} h_j}{Z''_{j,\text{FW}}} \right]. \quad (2.43)$$

In the Eqs. (2.41)-(2.43) the amplitude reflection coefficients for the stack R_{TE} and R_{TM} are still used, which also need to be expanded into a series of small parameters $k_{z,m}^{\text{SH}} h_m$. From Eq. (2.32) one can obtain for R_{TE} :

$$\begin{aligned} R_{\text{TE}} &= \frac{Z'_{\text{sub},\text{FW}} - Z'_{0,\text{FW}}}{Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}}} + \frac{M_{21} Z'_{0,\text{FW}} Z'_{\text{sub},\text{FW}} - M_{12}}{Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}}} + \frac{Z'_{\text{sub},\text{FW}} - Z'_{0,\text{FW}}}{(Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}})^2} \\ &\quad \times (M_{21} Z'_{0,\text{FW}} Z'_{\text{sub},\text{FW}} + M_{12}) \\ &= R_{\text{TE}}^0 + \frac{2i Z_{\text{sub},\text{FW}}'^2 Z'_{0,\text{FW}}}{(Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}})^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z'_{j,\text{FW}}} - \frac{2i Z'_{0,\text{FW}}}{(Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}})^2} \cdot \sum_{j=1}^N Z'_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j, \end{aligned} \quad (2.44)$$

where one introduces

$$R_{\text{TE}}^0 = \frac{Z'_{\text{sub},\text{FW}} - Z'_{0,\text{FW}}}{Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}}}$$

as the amplitude reflection coefficient at the interface between air and substrate, i.e. in the absence of the stack. Eq. (2.44) allows expanding Eq. (2.41) up to first-order small terms as follows:

$$\begin{aligned} E_{\text{TE},\parallel} \Big|_{m+1,+} &= E_I \left[1 + R_{\text{TE}}^0 + \frac{2i Z_{\text{sub},\text{FW}}'^2 Z'_{0,\text{FW}}}{(Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}})^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z'_{j,\text{FW}}} \right. \\ &\quad \left. - \frac{2i Z'_{0,\text{FW}}}{(Z'_{\text{sub},\text{FW}} + Z'_{0,\text{FW}})^2} \cdot \sum_{j=1}^N Z'_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j + i \frac{1 - R_{\text{TE}}^0}{Z'_{0,\text{FW}}} \cdot \sum_{j=1}^m Z'_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j \right]. \end{aligned} \quad (2.45)$$

In a similar way from Eq. (2.33) one can find for the amplitude reflection coefficient R_{TM} :

$$\begin{aligned}
R_{\text{TM}} &= \frac{Z''_{\text{sub,FW}} - Z''_{0,\text{FW}}}{Z''_{\text{sub,FW}} + Z''_{0,\text{FW}}} + \frac{-M_{21} Z''_{0,\text{FW}} Z''_{\text{sub,FW}} + M_{12}}{Z''_{\text{sub,FW}} + Z''_{0,\text{FW}}} - \frac{Z''_{\text{sub,FW}} - Z''_{0,\text{FW}}}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \\
&\quad \times (M_{21} Z''_{0,\text{FW}} Z''_{\text{sub,FW}} + M_{12}) \\
&= R_{\text{TM}}^0 + \frac{2i Z''_{\text{sub,FW}}{}^2 Z''_{0,\text{FW}}}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z''_{j,\text{FW}}} - \frac{2i Z''_{0,\text{FW}}}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \cdot \sum_{j=1}^N Z''_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j, \quad (2.46)
\end{aligned}$$

where

$$R_{\text{TM}}^0 = \frac{Z''_{\text{sub,FW}} - Z''_{0,\text{FW}}}{Z''_{\text{sub,FW}} + Z''_{0,\text{FW}}}$$

is the amplitude reflection coefficient between air and substrate in the absence of the stack. Using Eq. (2.46), Eqs. (2.42) and (2.43) can be expanded up to first-order small terms as follows:

$$\begin{aligned}
E_{\text{TM},\parallel} \Big|_{m+1,+} &= E_{\text{I},\parallel} \left[1 + R_{\text{TM}}^0 + \frac{2i Z''_{\text{sub,FW}}{}^2 Z''_{0,\text{FW}}}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \right. \\
&\times \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z''_{j,\text{FW}}} - \frac{2i Z''_{0,\text{FW}}}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \cdot \sum_{j=1}^N Z''_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j \\
&\quad \left. + i \frac{1 - R_{\text{TM}}^0}{Z''_{0,\text{FW}}} \cdot \sum_{j=1}^m Z''_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j \right], \\
E_{\text{TM},\perp} \Big|_{m+1,+} &= -\frac{k_{m+1,x}^{\text{FW}}}{k_{m+1}^{\text{FW}}} \cdot Z_{m+1,\text{FW}} \cdot E_{\text{I},\parallel} \left[-\frac{1 - R_{\text{TM}}^0}{Z''_{0,\text{FW}}} \right. \\
&+ \frac{2i Z''_{\text{sub,FW}}{}^2}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z''_{j,\text{FW}}} - \frac{2i}{(Z''_{\text{sub,FW}} + Z''_{0,\text{FW}})^2} \cdot \sum_{j=1}^N Z''_{j,\text{FW}} k_{j,z}^{\text{FW}} h_j \\
&\quad \left. - i(1 + R_{\text{TM}}^0) \cdot \sum_{j=1}^m \frac{k_{j,z}^{\text{FW}} h_j}{Z''_{j,\text{FW}}} \right]. \quad (2.47)
\end{aligned}$$

The obtained expressions can be used to find the source terms and eventually to express the emitted SHW Eq. (2.37) and (2.40) through the thicknesses and refractive indices of the layers and the amplitude reflection coefficients in the absence of the stack. Keeping only the first terms of this expansion allows representing the SHW amplitude through the effective surface nonlinear tensor of the whole stack. In the next section I proceed with explicit expressions for this effective tensor and discuss the applicability limits for such approximation.

2.6 Effective surface nonlinear tensor of the stack

If one keeps only zeroth-order terms in the field expansions from the previous section, the results attain the same form as if the SHW is emitted by a single nonlinear interface between the air and the

substrate. Let us thus introduce the surface nonlinear tensor of such an interface, which effectively describes the SSHG from the whole multilayer stack. Summarizing the equations from the previous section and the analytical solution for a SSHG from a single flat interface [89, 90], one obtains the following expressions for the effective surface nonlinear tensor:

$$\chi_{\Sigma,ijk}^{(2)} = \sum_{m=0}^N \chi_{m(m+1),ijk}^{(2)} \cdot \left(\frac{\varepsilon_{\text{up,SH}}}{\varepsilon_{m,\text{SH}}} \right)^{n_2} \cdot \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \right)^{n_1}, \quad (2.48)$$

where $\chi_{m(m+1)}^{(2)}$ is the surface nonlinear tensor of the interface between m -th and $m+1$ -th layer in the stack and ε_{up} is the relative dielectric permittivity of the upper medium. As already mentioned, this surface nonlinear tensor is to be attributed to the interface between the upper medium and the substrate. It is important to clarify the origin of both factors arising in Eq. (2.48). The first factor on the right-hand side, i.e. the ratio of the relative permittivities at the second-harmonic frequency, appears, since the sheet of the surface nonlinear polarization at the interface between m -th and $m+1$ -th layer is assumed to be just above the interface, i.e. inside m -th layer with the relative permittivity ε_m . The value of the index of power n_2 is correspondingly given as:

$$\begin{aligned} n_2 &= 1, & \text{when } i = z, \\ n_2 &= 0, & \text{when } i = x, y. \end{aligned}$$

The second factor on the right-hand side in Eq. (2.48), i.e. the ratio of the relative permittivities at the fundamental frequency, results from the jump of the normal component of the electric field at each interface inside the stack, as compared to a single interface between the substrate and the upper medium, and appears only if the normal component of the electric field E_z is used for the calculation of the surface nonlinear polarization P^{NL} . The value of the index of power n_1 is therefore provided by:

$$\begin{aligned} n_1 &= 2, & \text{when } j = z, k = z, \\ n_1 &= 1, & \text{when } j = z, k \neq z \text{ or } j \neq z, k = z, \\ n_1 &= 0, & \text{when } j \neq z, k \neq z. \end{aligned}$$

In the important case of the isotropic surfaces inside the stack the expressions for the surface nonlinear tensor Eq. (2.48) can be written as:

$$\begin{aligned} \chi_{\Sigma,\perp\perp\perp}^{(2)} &= \sum_{m=0}^N \chi_{m(m+1),\perp\perp\perp}^{(2)} \cdot \frac{\varepsilon_{\text{up,SH}}}{\varepsilon_{m,\text{SH}}} \cdot \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \right)^2, \\ \chi_{\Sigma,\perp\parallel\parallel}^{(2)} &= \sum_{m=0}^N \chi_{m(m+1),\perp\parallel\parallel}^{(2)} \cdot \frac{\varepsilon_{\text{up,SH}}}{\varepsilon_{m,\text{SH}}}, \\ \chi_{\Sigma,\parallel\perp\parallel}^{(2)} &= \sum_{m=0}^N \chi_{m(m+1),\parallel\perp\parallel}^{(2)} \cdot \frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}}. \end{aligned}$$

Let us now take a closer look at the effective surface tensor for different polarizations of the FW. For a TE-polarized incident plane wave at fundamental frequency one gets $E_{\perp}^{(\omega)} = 0$ inside the stack and only the term $\sim \chi_{\perp||}^{(2)}$ in Eq. (2.27) is non-zero at each interface. The surface nonlinear polarization thus possesses only normal component $P_{S,\perp}^{\text{NL}}$, therefore one obtains $\Delta \vec{H}_{||} = 0$ and $S_{mm+1,2}^{\text{TM}} = 0$ in Eq. (2.39). Following the reasoning in Eqs. (2.10)-(2.11), it means that only a TM-polarized SHW is emitted in this case from the stack. The expression for the effective surface nonlinear tensor can be extended with respect to Eq. (2.48) to include also first-order terms as follows:

$$\chi_{\Sigma,\perp||}^{(2)} = \sum_{m=0}^N \chi_{\perp||,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} + \delta \chi_{\Sigma,\perp||}^{(2)} + O\left[(k_{j,z}^{\text{SH}} h_j)^2\right].$$

The explicit expressions for the first-order small terms $\delta \chi_{\Sigma,\perp||}^{(2)}$ derived using the field expansions from the previous section, are quite cumbersome and put to Appendix B.

Next, let us consider similarly TM-polarized incident plane wave at the fundamental frequency. Here all terms in the expression for the surface nonlinear polarization in Eq. (2.27) are non-zero. Eqs. (2.2) and (2.27) show that in this case $\vec{P}_{S,||}^{\text{NL}} \parallel \vec{E}_{||}^{(\omega)}$, and hence $\Delta \vec{H}_{||} \perp \vec{E}_{||}^{(\omega)}$ and $\Delta \vec{H}_{||} \perp \Delta \vec{E}_{||}$. According to Eqs. (2.10)-(2.11) one finds again, that only TM-polarized SHW is emitted. The expressions for the effective surface nonlinear tensor with the inclusion of the first-order terms are given as:

$$\begin{aligned} \chi_{\Sigma,\perp\perp\perp}^{(2)} &= \sum_{m=0}^N \chi_{\perp\perp\perp,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} \cdot \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}}\right)^2 + \delta \chi_{\Sigma,\perp\perp\perp}^{(2)} + O\left[(k_{j,z}^{\text{SH}} h_j)^2\right], \\ \chi_{\Sigma,||\perp||}^{(2)} &= \sum_{m=0}^N \chi_{||\perp||,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} + \delta \chi_{\Sigma,||\perp||}^{(2)} + O\left[(k_{j,z}^{\text{SH}} h_j)^2\right], \\ \chi_{\Sigma,\perp||}^{(2)} &= \sum_{m=0}^N \chi_{\perp||,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} + \delta \chi_{\Sigma,\perp||}^{(2)} + O\left[(k_{j,z}^{\text{SH}} h_j)^2\right]. \end{aligned}$$

Again the explicit expressions for the first-order small terms $\delta \chi_{\Sigma,\perp||}^{(2)}$ derived using the field expansions from Eqs. (2.40) and (2.47) and presented in Appendix B.

The derived expressions for the first-order correction terms of the effective surface nonlinear tensor contain only the imaginary part, even if the absorption in the ultrathin layers is neglected. The contribution to the real part of $\chi_{\text{surf},\Sigma}^{(2)}$ can be thus given by the second-order correction terms. In order to derive the second-order correction terms, one has to keep correspondingly up to the second-order small terms in all expansions of the fields inside the stack in the previous section. These second-order terms, however, turn out to be much more cumbersome and are not provided here. It should be also noted that Eqs. (B.1)-(B.2) were derived for pure TE- or TM-polarized incident plane wave and are not directly applicable for other illuminating fields. In particular, in the case of the plane-wave excitation with the mixed polarization, TE- and TM-polarized field components at the fundamental frequency will be mixed through the terms $\chi_{||\perp||,m(m+1)}^{(2)}$, resulting in the generation of both TE- and TM-polarized SHWs. With Eqs. (2.37) and (2.40) it would be possible to generalize the expressions for the effective surface nonlinear tensor for such incident FW as well.

An important issue is related to the applicability of the derived expansions for the effective surface nonlinearity of the stack $\chi_{\text{surf},\Sigma}^{(2)}$. In order to estimate the limits of validity of the simplified Eq. (2.48), one has to find the ratio between the zeroth-order and first-order correction terms in the expansion for Eq. (2.48). Using the first-order correction terms in Eqs. (B.1), (B.2) the condition of smallness of this ratio can be reduced to the smallness of the following quantities:

$$\begin{aligned}\xi_{\text{TE}} &= \frac{1}{Z'_{0,\text{FW/SH}} + Z'_{\text{sub,FW/SH}}} \cdot \sum_{m=1}^N Z'_{m,\text{FW/SH}} k_{m,z}^{\text{FW/SH}} h_m \ll 1; \\ \xi_{\text{TM}} &= \frac{1}{Z''_{0,\text{FW/SH}} + Z''_{\text{sub,FW/SH}}} \cdot \sum_{m=1}^N Z''_{m,\text{FW/SH}} k_{m,z}^{\text{FW/SH}} h_m \ll 1,\end{aligned}\quad (2.49)$$

and

$$\begin{aligned}\zeta_{\text{TE}} &= \frac{Z'_{0,\text{FW/SH}} Z'_{\text{sub,FW/SH}}}{Z'_{0,\text{FW/SH}} + Z'_{\text{sub,FW/SH}}} \cdot \sum_{m=1}^N \frac{k_{m,z}^{\text{FW/SH}} h_m}{Z'_{m,\text{FW/SH}}} \ll 1; \\ \zeta_{\text{TM}} &= \frac{Z''_{0,\text{FW/SH}} Z''_{\text{sub,FW/SH}}}{Z''_{0,\text{FW/SH}} + Z''_{\text{sub,FW/SH}}} \cdot \sum_{m=1}^N \frac{k_{m,z}^{\text{FW/SH}} h_m}{Z''_{m,\text{FW/SH}}} \ll 1.\end{aligned}\quad (2.50)$$

Given that Eqs. (2.49), (2.50) are fulfilled, the simplest form of the effective nonlinear tensor Eq. (2.48) can be well justified. As long as the conditions Eqs. (2.49), (2.50) cease to be valid, the first-order correction terms in Eqs. (B.1), (B.2) must be also accounted.

2.7 Simulation results and discussion

In this section I aim to test the proposed nonlinear TMM with some exemplary multilayer structure. I take a 10-nm stack composed of five ultrathin layers made of Si, SiO₂ and TiO₂, which are placed on the top of a glass substrate (BK7). The thicknesses of all layers together with their materials are listed in Table 2.1. The fundamental wavelength of the illuminating plane wave is chosen 1 μm , so that the second-harmonic wavelength is 500 nm. The relative permittivities of all materials both at fundamental and second-harmonic frequencies are taken from Refs. [91, 92, 93].

Layer, №	1	2	3	4	5	Substrate
Material	Si	SiO ₂	TiO ₂	Si	SiO ₂	BK7
Thickness, nm	1	3	2	1	3	-

Table 2.1: The composition of the considered multilayer stack.

The intensity of the incident FW was taken to 1.33 GW/cm², meaning that the electric field strength in the FW is $E^{\text{FW}} = 10^8$ V/m. The values of the entries of the surface nonlinear tensor for the top interface between silicon and air were experimentally measured in [94]. Since experimental data for the frequency dispersion of the surface nonlinear susceptibilities of Si/Air interface is not available, the Miller rule [2] is applied to calculate the surface nonlinear tensor at the fundamental

frequency. Miller rule states the following relation to hold:

$$\frac{\chi^{(2)}(2\omega)}{\chi^{(1)}(2\omega) \cdot \left[\chi^{(1)}(\omega)\right]^2} = \text{const}, \quad (2.51)$$

where $\chi^{(1)}(\omega)$ is the usual linear susceptibility of the material. Using Eq. (2.51) one can calculate the surface nonlinear tensor for Si/Air interface at the frequency of interest, given that its value at least at one frequency is known. Substituting the experimental values for the pump wavelength 800 nm from [94], one obtains the following values of the tensor components: $\chi_{\perp\perp\perp}^{(2)} \approx 7 \cdot 10^{-18} \text{ m}^2/\text{V}$, $\chi_{\perp\parallel\parallel}^{(2)} \approx 3.7 \cdot 10^{-19} \text{ m}^2/\text{V}$, $\chi_{\parallel\perp\parallel}^{(2)} \approx 3.7 \cdot 10^{-19} \text{ m}^2/\text{V}$.

For most interfaces between different material pairs the experimental data for the respective surface nonlinearities is lacking. Therefore it seems to be not possible in general to find experimental data for the values of the surface nonlinear tensor for different adjacent media in the stack. To overcome this problem an assumption is taken that the strength of the surface nonlinearity is simply linearly proportional to the ratio of the relative permittivities of adjacent materials. This is just a rough approximation, which is inspired by the fact that the surface nonlinearity originates from the discontinuity of the material properties across the interface. As far as I know, there are no theoretical models developed so far, which could provide a way to calculate the strength of the surface nonlinearity for an arbitrary pair of adjacent dielectric media. With the described approximation, the experimental values for the Si/Air interface above are used to get the estimated values of the entries of the surface nonlinear tensor for each interface inside the considered stack. It is important that the values of the surface nonlinear tensors are not relevant for testing the performance of the numerical method itself. Moreover, approximate values of the surface nonlinearities are still expected to yield correct order-of-magnitude amplitude of the emitted SHW from the stack. Along with that, one has to take care of the correct direction of the vector of the surface nonlinear polarization at each interface. It is assumed here that the vector of the surface nonlinear polarization is always directed from the medium with stronger bulk second-order nonlinearity into the medium with weaker second-order nonlinearity. This means that, for example, the normal component of the surface nonlinear polarization constantly changes its sign as moving across the stack and thus the surface sources of the SHG can interfere both constructively and destructively.

Let us start with simulation of the SSHG from the exemplary stack in Table 2.1 and the comparison of the simulation results with the analytically obtained SSHG from a single interface with the effective nonlinear tensor Eq. (2.48). Both plots are shown in Fig. 2.2 for both TE- and TM-polarized incident plane wave at the fundamental frequency. The angle of incidence of the FW is varied. The agreement between both curves in Fig. 2.2 is very good, with noticeable deviations only close to the maxima of the plots. One can also see that the maximal SHW amplitudes for both polarizations of the FW differ slightly.

It is interesting next to estimate the quantities $\xi_{\text{TE/TM}}$, $\zeta_{\text{TE/TM}}$ from Eqs. (2.49)-(2.50) in order to check the validity of the zeroth-order term for the effective nonlinear tensor Eq. (2.48). Both quantities are plotted in Fig. 2.3 vs. the angle of incidence both for the fundamental and SH frequency. One can see that the requirement of the smallness of both ξ and ζ can be well enough satisfied within a

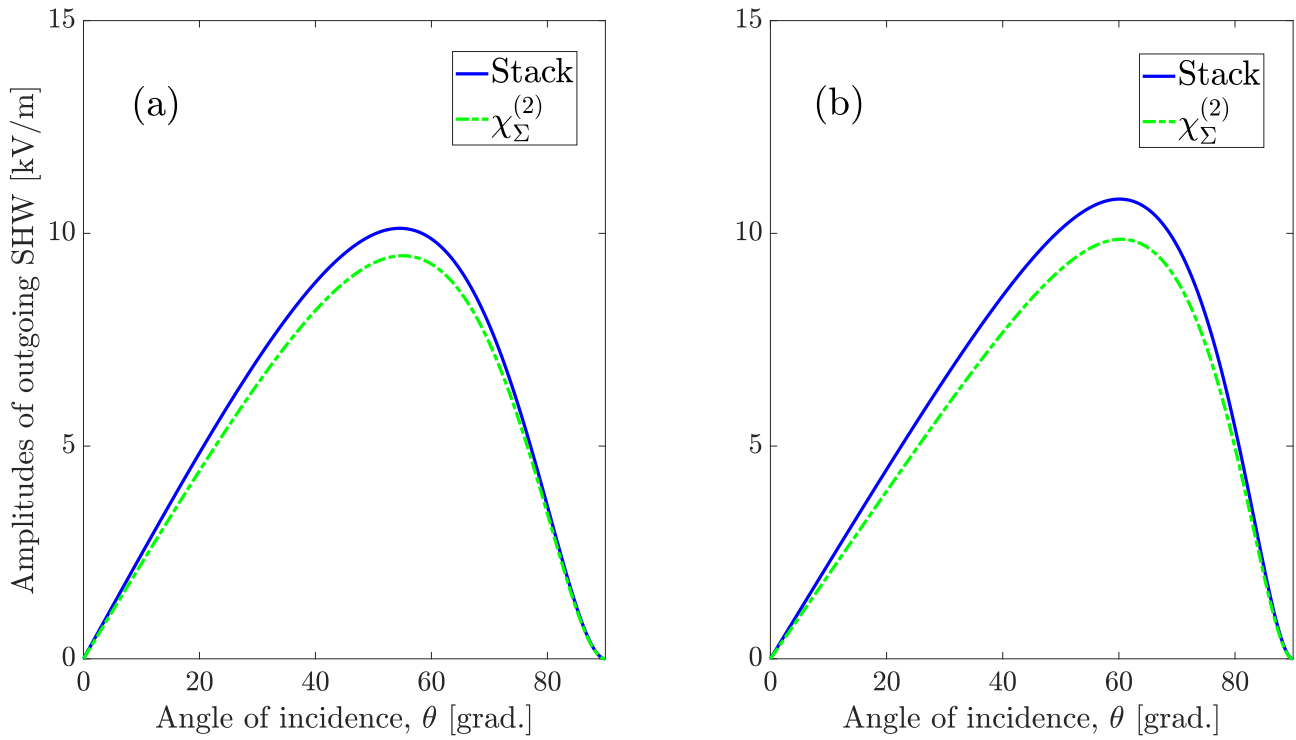


Figure 2.2: Amplitude of the upwards outgoing SHW from the considered stack for: (a) TE-polarized; (b) TM-polarized incident wave at fundamental frequency, obtained using proposed TMM from the stack (blue solid line) and using effective surface nonlinear tensor Eq. (2.48) for a single interface between the substrate and the air (green dashed line).

certain range of the values of the incidence angle close to the normal incidence. For close to grazing incidence parameters ξ_{TM} and ζ_{TE} at SH frequency significantly grow and cannot be assumed much smaller than 1 anymore. In this case some terms in the first-order corrections to Eq. (2.48) are not negligible and must be taken into account. The values of the discrepancies can be reduced further by selecting materials with smaller refractive index contrast and/or reducing the thicknesses. The extra simulations were performed and it was checked that with decreasing the thicknesses of all layers in the stack the blue curves in Fig. 2.2 displaying the SSHG from the considered multilayer stack get closer to and eventually fuse together with the green curves corresponding to the SSHG from a single interface possessing the effective surface nonlinearity $\chi_{\text{surf},\Sigma}^{(2)}$ from Eq. (2.48). The reduction of the layer thicknesses could be, however, quite challenging and is limited by the fabrication techniques.

Lastly, it would be interesting to compare the amplitude of the SHW emitted from the considered exemplary stack in Table 2.1 with the one from a bulk layer of comparable thickness made of non-centrosymmetric material with strong bulk nonlinearity. Thereby one can estimate the efficiency of the multilayer structures for potential applications in nonlinear photonic devices. For comparison, a GaAs layer placed on the top of the same glass substrate was selected. The components of the vector of the bulk nonlinear polarization in GaAs are expressed as follows:

$$P_i^{\text{NL}} = \varepsilon_0 \chi_{ijk}^{(2)} E_j E_k,$$

with indices i, j, k , corresponding to the crystal symmetry axes. Due to the zinc-blende crystalline structure of bulk GaAs, the entries of its bulk second-order nonlinear susceptibility tensor obey the

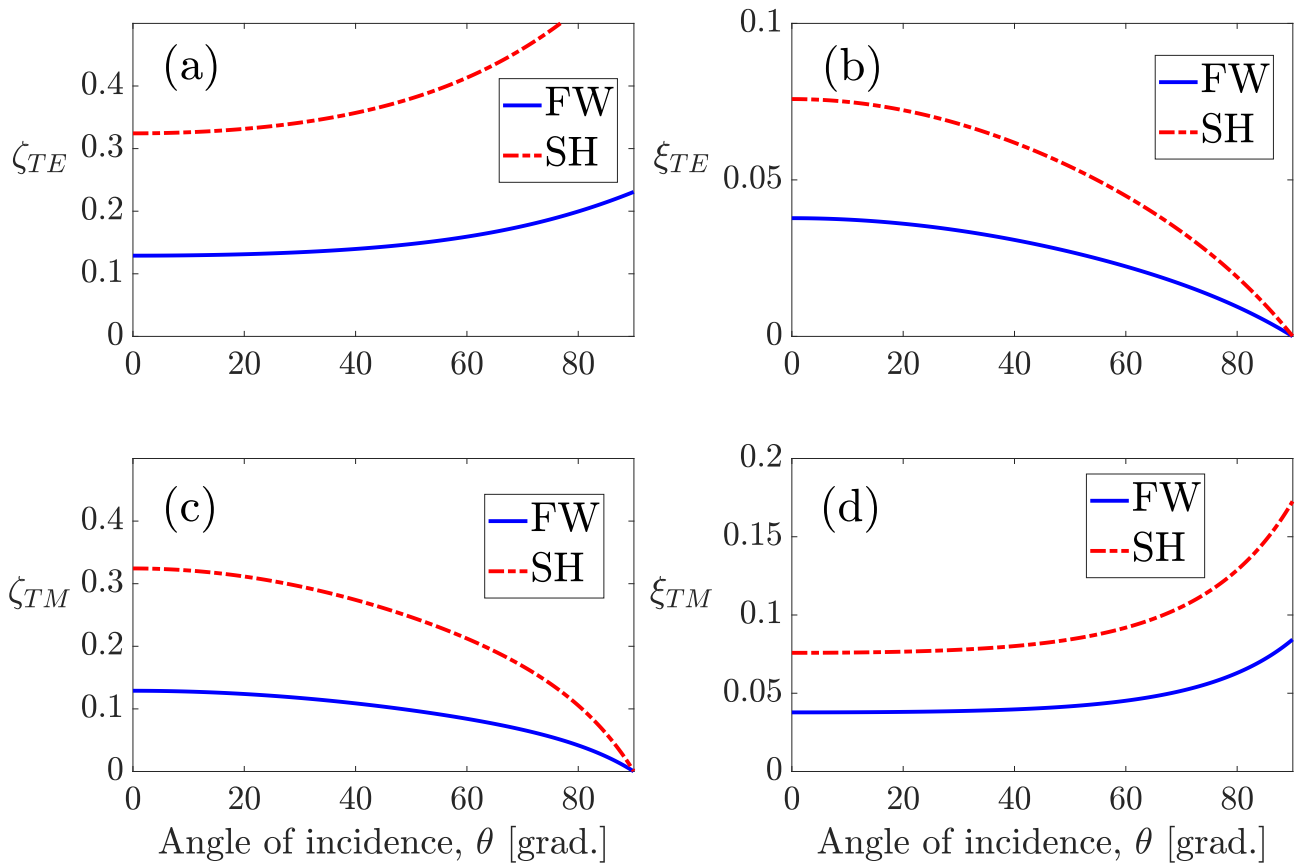


Figure 2.3: Values of the discrepancies $\xi_{TE/TM}$ and $\zeta_{TE/TM}$ from Eqs. (2.49)-(2.50) both for fundamental (FW) and second-harmonic (SH) frequencies.

following symmetry relations:

$$\chi_{ijk}^{(2)} \neq 0 \text{ if } i \neq j \neq k,$$

i.e. only entries with all indices different are non-zero. Moreover, all non-zero tensor elements are equal to each other [41]. The respective value in GaAs was measured as $\chi_{ijk}^{(2)} = 100$ pm/V [95]. The top plane of the considered GaAs layer is assumed to coincide with the (100) crystal face for simplicity. As the illuminating field at the fundamental frequency TE- and TM-polarized plane waves are taken. TM-polarized plane wave was taken with the electric field vector in the crystal xz -plane and for the TE-polarized plane wave the angle 45° between the electric field vector and both x - and y - crystal axis is assumed. The simulation results for GaAs layer are shown in Fig. 2.4 together with the data for the exemplary stack from Fig. 2.2. The thickness of the GaAs layer was varied and finally chosen 70 nm in order to get stronger amplitude of the upwards propagating SHW for TM-polarized incident plane wave. For a thinner GaAs layer, e.g., close to 10 nm for the stack in Table 2.1, the amplitude of the emitted SHW was even weaker. One can see that even though the GaAs layer possesses strong bulk nonlinearity and has much larger thickness, the efficiency of the SHG from the stack is still much higher. It should be recalled here that just approximate values for the surface nonlinear tensors inside the stack were used. However, the simulated data is expected to be of the correct order of magnitude, so that the comparison between the conversion efficiencies from the multilayer structure and GaAs layer still holds. These findings let to expect the considered multilayer

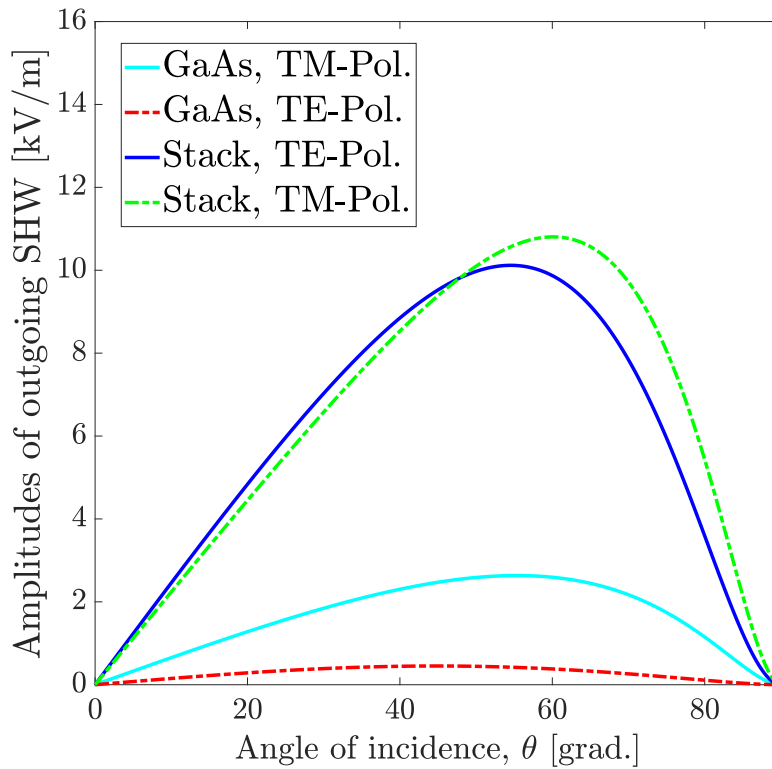


Figure 2.4: Amplitude of the upwards outgoing SHW from a GaAs layer of thickness 70 nm, located on the same substrate (BK7). The top plane coincide with (100)-plane of the crystal. TM-polarized incident wave is polarized in xz -plane; in the case of TE-polarized incident wave the electric field vector forms angles 45° both with x - and y -axis.

structures as promising candidates for applications in nonlinear photonic devices. The conversion efficiency can be also further boosted, for instance, through the nano-patterning of the stacks as it has been demonstrated for all-dielectric metasurfaces [44].

2.8 Summary

This chapter has been devoted to SHG from multilayer structures arising due to the surface nonlinearities at multiple interfaces between adjacent layers. With such composition of many layers, the surface sources of the nonlinear polarization interfere, which can result in the strong enhancement of SHG from multilayer structures. I did not account for the bulk nonlinearity in the layers and fully concentrated on the surface nonlinearities only. Such assumption is well justified, when the multilayer structure is made of centrosymmetric semiconductors, dielectrics or metals. The layers of noncentrosymmetric semiconductors can be also treated in such a way, provided that the thickness is small enough, that the bulk nonlinearity can be safely neglected.

For the efficient modeling of the SSHG from multilayer structures there have been no convenient numerical methods so far. Therefore a transfer matrix method was developed well suited for the simulations of the SH field from an arbitrary one-dimensional layered stack of any possible composition. The proposed TMM inherently accounts for such effects, as multiple reflections of the field both at fundamental frequency and the SH frequency. This TMM was tested for an exemplary multilayer structure and the simulation results have demonstrated high enough conversion efficiency of

the SSHG. To be more specific, the SHW amplitude from the stack was compared with the SHW amplitude from a GaAs layer of comparable or even larger thickness. The calculation showed the SHW amplitude emitted from the exemplary multilayer structure is several times stronger than SHW amplitude from the GaAs layer. This means at least one order of magnitude larger energy conversion efficiency for the multilayer stack, despite neglecting the bulk nonlinearities. The proposed multilayer structures can be thus potentially interesting for applications in nonlinear nanophotonics, for example, for the frequency doubling in nanophotonic devices. Besides that, the nano-patterning of the stack into a dielectric metasurface can allow to further boost the efficiency of SHG thanks to the excitation of Mie-type or other resonances. Finally, it is worth noting that the developed TMM can be potentially applied to examine nonlinear effects in some novel classes of layered materials, such as multilayered transition metal dichalcogenide monolayer structures (TMDCs) [96] or hyperbolic metamaterials [97], which are actively studied in recent years and can be also interesting for nonlinear optical applications [98, 99].

An important particular case that was addressed is the multilayer structure composed of ultrathin layers (of the order of tens of nm or smaller). This specific case is interesting for several reasons. First, for such ultrathin layers even in noncentrosymmetric materials the contribution of the bulk nonlinearity is negligibly small and the SHG response is fully driven by the surface nonlinearities. Second, small thicknesses of the layers lead to small phase shifts between the sources of the nonlinear polarization at the neighbouring interfaces and thus facilitate the constructive interference of the nonlinear surface sources. Third, the analytical derivation can be greatly simplified in this limit and even explicit approximate equations can be derived for the emitted SHW. The analytical expressions of proposed TMM were reduced in the limit of ultrathin layers and it was shown that in the zeroth-order approximation the SSHG from an ultrathin stack coincides with the SSHG from a single interface between the upper medium and the substrate. For this replacement interface the effective surface nonlinear tensor was derived describing the nonlinear optical properties of the whole stack. Moreover, the first-order correction terms for the effective surface nonlinear tensor were explicitly obtained. Performed simulations have shown that for an exemplary ultrathin stack the effective surface nonlinear tensor can indeed describe the nonlinear response with high enough accuracy.

The proposed TMM is only applicable to one-dimensional multilayer structures. Therefore it cannot be used in a variety of important applications, like the SSHG from plasmonic or dielectric nanostructures or metasurfaces composed of them. Hence, other numerical methods are needed for efficient modeling of SSHG from arbitrary-shape structures. In the following chapter I develop such method based on the finite-element formalism.

Chapter 3

Adaptive finite-element method for modeling of surface SHG from nanostructures

As discussed in the introduction, different nonlinear mechanisms are primarily responsible for SHG from different materials depending on the medium symmetry. In noncentrosymmetric semiconductor materials the inversion symmetry is naturally broken, leading to the strong bulk nonlinearity. The surface nonlinearity can be often disregarded, even though sometimes the surface contribution is still able to noticeably effect the total SH response and needs to be accounted for [43, 44, 45]. The surface nonlinearity, on the contrary, is typically dominating in centrosymmetric media (Si, Ge, TiO₂, metals and others) [39], because the inversion symmetry prohibits the bulk second-order nonlinearity in the dipole approximation.

Bulk and surface nonlinearities need different numerical methods to be used for their efficient modeling. The vectorial finite-difference time-domain method (FDTD) with uniform rectangular mesh is well suited for the bulk SHG [100], but is hardly applicable for the surface SHG from arbitrary-shaped nanoparticles. There are a number of numerical methods developed so far for the modeling of the SSHG in nanophotonics. Among these are the boundary element methods (BEM) [61], volume integral formulations [59, 60] and different versions of the surface integral method [62, 63, 64, 65]. All these methods have certain advantages for SSHG simulations, but also share some downsides. The limitations, to mention just a few, include the complexity of incorporating the bulk nonlinearity in the medium in addition to the surface one, relatively slow convergence rates compared to FEM, the limited flexibility in handling the surface discretization for complex geometries and often the difficulty to achieve acceptable trade-off between high computational accuracy and reasonable computational costs.

Besides all the above stated methods, finite-element methods are also widely used for SHG simulations from different nanostructures, but mainly for the case of the bulk nonlinearity only. FEM are especially popular in the area of nanophotonics, since their nonuniform discretization meshes (see Fig. 3.1,a) allow easy handling of the complex geometries. It is also possible to use FEM for SSHG simulations, if one assumes the finite thickness of the surface layer. In this case the surface nonlinearity is modeled as the bulk one, but existing only within the ultrathin surface layer. This approach, however, harms the FEM performance due to the appearance of the largely different length scales,

since the light wavelength and the typical dimensions of the nanostructures are orders of magnitude larger than the efficient thickness of the surface layer. This results in the strongly irregular computational mesh and very poor efficiency of FEM simulations, since the performance of the FEMs greatly benefits from the regularity of the computational mesh.

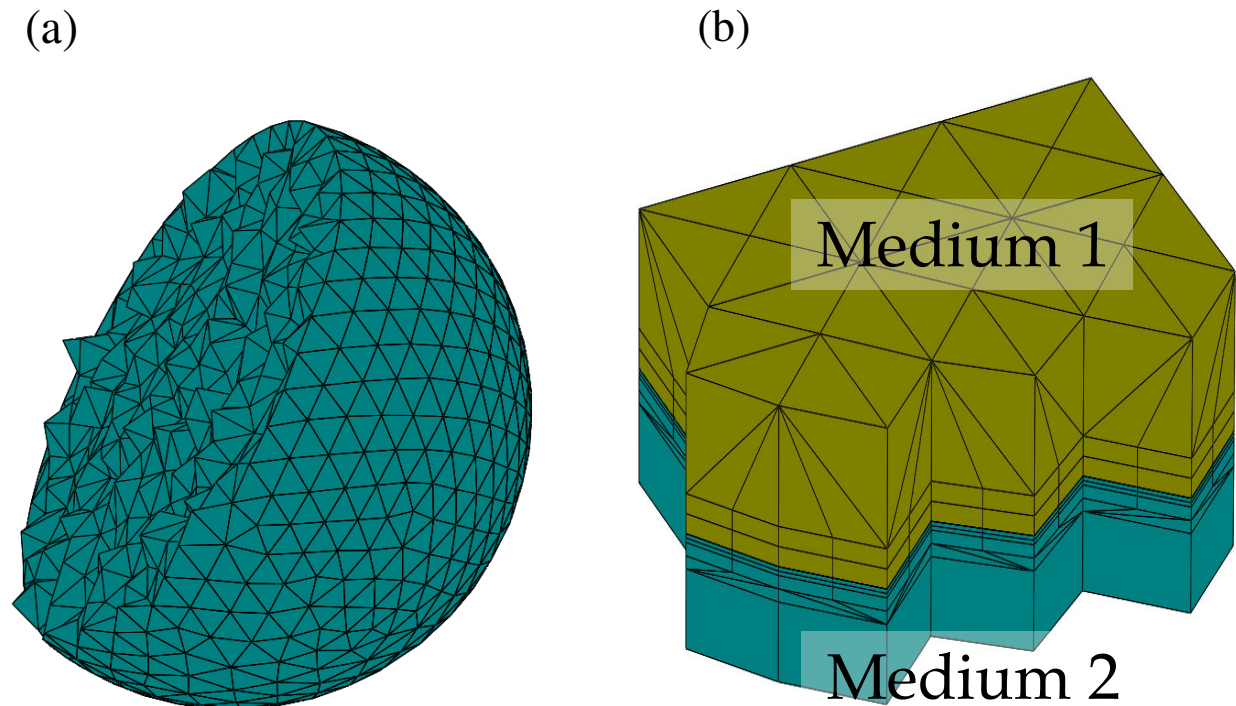


Figure 3.1: (a) An example cut of the discretization grid for FEM simulation of a three-dimensional spherical nanostructure; (b) an example of the refinement of the finite-element grid in the vicinity of the interface between two media.

Here a finite-element implementation of the surface nonlinearity is developed, where the surface nonlinear polarization is treated as a true surface source, embedded in the interface between two adjacent media. This method enables the SSHG simulation from arbitrary-shaped nanostructures, either isolated or arranged into metasurfaces. Also, both all-dielectric and plasmonic nanostructures can be considered. Proposed implementation allows taking advantage of the major benefits of FEM in general, such as the fast exponential convergence rate, the simplicity of handling complex geometries, the possibility of the local mesh refinement just in a single subdomain of the computational domain, for instance close to the interface (an example is shown in Fig. 3.1,b), the feasibility of *hp*-adaptive strategies for reducing the computational costs, the direct addressing of each interface between adjacent media and thus the possibility to locally tune the boundary conditions etc. [58]. Along with that, direct FEM implementation of the surface nonlinear polarization as strictly two-dimensional surface source embedded in the corresponding interface lets us to overcome the problem arising for the

finite-thickness surface layer, namely the undesirable grid refinement in the vicinity of the interface.

This chapter is organized as follows. Section 3.1 provides the basic introduction into the finite-element methods. In the following Sections 3.2–3.4 the detailed description of the proposed method is provided. In particular, in Section 3.2 a general scheme of the proposed FEM is given. Section 3.3 is devoted to the mathematical representation of the surface nonlinear polarization and calculation of its spatial derivatives. The approach of introducing the generalized boundary conditions into the FEM code is described in Section 3.4. In Section 3.5 the simulation results for test problem are presented and the issues related to the uniform and adaptive refinement strategies are discussed. Section 3.6 outlines the extension of the proposed method to the modeling of both linear and nonlinear optical properties of 2D materials, e.g. graphene or transition metal dichalcogenide monolayers. Finally, in Section 3.7 the key results are summarized and concluding remarks are provided.

3.1 Overview of the finite-element methods

The finite-element methods rely on the so-called weak formulation of the Maxwell's equations. For the computational domain $\Omega \subset R^3$ possessing the boundary $\partial\Omega$, weak formulation is given as [58]:

$$\begin{aligned} \int_{\Omega} [(\mu^{-1}\vec{\nabla} \times \vec{E}) \cdot \vec{\nabla} \times \vec{\phi}^* - \frac{\omega^2}{c^2}(\varepsilon\vec{E}) \cdot \vec{\phi}^*] dV + \int_{\partial\Omega} \vec{n} \times (\mu^{-1}\vec{\nabla} \times \vec{E}) \cdot \vec{\phi}_T^* dS \\ = \int_{\Omega} \vec{F} \cdot \vec{\phi}^* dV, \quad \forall \vec{\phi} \in V, \end{aligned} \quad (3.1)$$

where \vec{E} is the sought electric field, \vec{F} is the source term, vector-valued test functions $\vec{\phi}$ belong to the space $V = H(\text{curl}, \Omega)$ and $\vec{\phi}_T = (\vec{n} \times \vec{\phi}) \times \vec{n}$ with the surface normal vector \vec{n} . In particular, for modeling SHG at frequency 2ω one has to use the source term:

$$\vec{F}(2\omega, \vec{r}) = \frac{4\omega^2}{c^2} \vec{P}^{\text{NL}}(2\omega, \vec{r}),$$

with the vector of the second-order nonlinear polarization \vec{P}^{NL} .

Outside the computational domain the homogeneous exterior domain $R^3 \setminus \Omega$ with relative dielectric permittivity ε_{ext} and permeability μ_{ext} is assumed, as sketched in Fig. 3.2. The incident field \vec{E}_{inc} propagates from the exterior domain $R^3 \setminus \Omega$ and enters the interior domain Ω across its boundary $\partial\Omega$. Interaction of the incident field with the scatterers inside the computational domain causes the scattered field \vec{E}_{scat} . The scattered field \vec{E}_{scat} propagates outwards from the scatterers into the exterior domain. Therefore the total field represents the sum of two terms, namely the incident field \vec{E}_{inc} and the scattered field \vec{E}_{scat} :

$$\vec{E} = \vec{E}_{\text{inc}} + \vec{E}_{\text{scat}}.$$

If one assumes an arbitrary boundary of the different materials inside the computational domain, the tangential components of the electric and magnetic field have to satisfy the well-known continuity boundary conditions coming from the Maxwell's equations [2]. Specifically, if the limiting values of

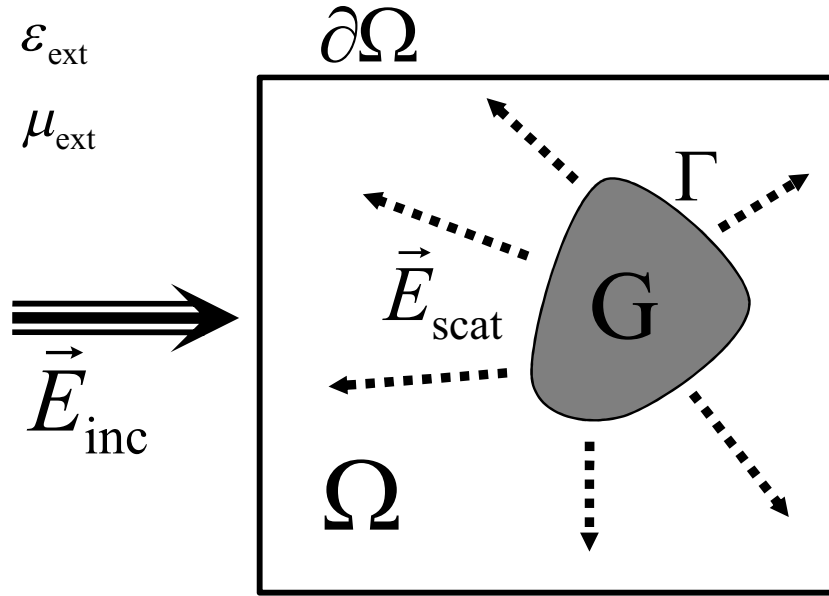


Figure 3.2: Geometry for the scattering problem with a single scatterer G inside the computational domain Ω .

the electric field as the boundary is approached from different sides are denoted as \vec{E}_+ and \vec{E}_- , the following equalities have to be fulfilled:

$$\begin{aligned}\vec{E}_+ \times \vec{n} &= \vec{E}_- \times \vec{n}, \\ (\mu_+^{-1} \vec{\nabla} \times \vec{E}_+) \times \vec{n} &= (\mu_-^{-1} \vec{\nabla} \times \vec{E}_-) \times \vec{n},\end{aligned}$$

where \vec{n} is the unit vector to the boundary at any specific point. In a similar way, at the boundary $\partial\Omega$ between the internal and external domains the boundary conditions can be written as:

$$\begin{aligned}(\vec{E}_{\text{inc}} + \vec{E}_{\text{scat}}) \times \vec{n} &= \vec{E} \times \vec{n}, \\ (\mu_{\text{ext}}^{-1} \vec{\nabla} \times (\vec{E}_{\text{inc}} + \vec{E}_{\text{scat}})) \times \vec{n} &= (\mu^{-1} \vec{\nabla} \times \vec{E}) \times \vec{n}.\end{aligned}$$

Here for convenience the total field is written in the internal domain and the superposition of the scattered and incident fields is written in the external domain. The incident field is prescribed, while for the scattered field in the external domain one has to impose additional conditions to assure it to be strictly outgoing. For this purpose the Silver-Müller radiation condition, also known as the "boundary condition at infinity", is often applied [58]:

$$\lim_{r \rightarrow \infty} r \left(\vec{\nabla} \times \vec{E}_{\text{scat}}(\vec{r}) \times \vec{r}_0 - i \frac{\omega \sqrt{\epsilon_{\text{ext}} \mu_{\text{ext}}}}{c} \vec{E}_{\text{scat}}(\vec{r}) \right) = 0, \quad (3.2)$$

with the coordinate vector \vec{r} , its norm r and unit vector $\vec{r}_0 = \vec{r}/r$. The Silver-Müller radiation condition Eq. (3.2) is imposed at the infinitely distant boundary. Since in realistic simulation problem the computational domain Ω is always finite, the radiation condition Eq. (3.2) must be imposed on a

surface located at the finite distance far enough from the scatterer. With this assumption, the boundary condition Eq. (3.2) on $\partial\Omega$ can be turned to the following form:

$$(\vec{\nabla} \times \vec{E}) \times \vec{n} - i\frac{\omega}{c}\vec{E} = (\vec{\nabla} \times \vec{E}_{\text{inc}}) \times \vec{n} - i\frac{\omega}{c}\vec{E}_{\text{inc}} , \quad (3.3)$$

where \vec{n} is the outward unit normal vector [58]. The boundary condition in the form Eq. (3.3) allows reducing the infinite computational domain of the scattering problem to the finite one. The solution of this reduced problem can serve as an approximation of the exact solution to a high accuracy, provided that the boundary $\partial\Omega$ is located far enough from the scatterer. The expression Eq. (3.3) gives just one possible way to treat the boundaries of the finite computational domain. Another widely used option is perfectly-matched layer (PML). PML is an extra layer surrounding the computational domain, which absorbs (in an ideal case completely) the outgoing radiation suppressing any back-reflection at the boundary of the computation domain [101]. There are a number of different realizations of PML, which are designed in order to exhibit the best possible absorption of the outgoing radiation incident on the boundary at any angle. It has been also proven that PML is a rigorous method of treating the boundaries of a finite computational domain, meaning that any numerical errors introduced due to the PML discretization would vanish with the PML refinement [102].

In FEM instead of the exact solution of the scattering problem one aims to find the approximate solution which belongs to a certain finite-dimensional subspace V_h of the space V . In order to facilitate the treatment of the functions in V_h , the usage of polynomial subspaces is especially suitable, since the integration and the differentiation are easily performed with the polynomial functions. In particular, the edge elements of Nedelec are often used as the polynomial subspaces [103]. The initial scattering problem then reduces to the calculation of such polynomial function $\vec{E}_h \in V_h$, which obeys the weak form of Maxwell's equations Eq. (3.1) for any $\vec{\phi} \in V_h$. Let us assume that a set of polynomial functions $\{\vec{\phi}_1, \vec{\phi}_2, \dots, \vec{\phi}_{N_h}\}$ is chosen, so that they form a basis in the subspace V_h . Given that, the approximate solution \vec{E}_h of the scattering problem is sought as their linear combination with weight factors α_i :

$$\vec{E}_h = \sum_{i=1}^{N_h} \alpha_i \vec{\phi}_i.$$

Substituting this expansion into Eq. (3.1), one gets the system of linear equations for unknowns α_i :

$$A \cdot \vec{\alpha} = B,$$

where the matrix and the source vector are given as:

$$A_{ij} = \int_{\Omega} [(\mu^{-1}\vec{\nabla} \times \vec{\phi}_i) \cdot \vec{\nabla} \times \vec{\phi}_j^* - \frac{\omega^2}{c^2}(\varepsilon\vec{\phi}_i) \cdot \vec{\phi}_j^*] dV + \int_{\partial\Omega} \vec{n} \times (\mu^{-1}\vec{\nabla} \times \vec{\phi}_i) \cdot \vec{\phi}_{j,T}^* dS,$$

$$B_{ij} = \int_{\Omega} \vec{F} \cdot \vec{\phi}_i^* dV.$$

The most convenient choice of the ansatz functions $\vec{\phi}_i$ is to take the polynomials with the support restricted just to several patches of the computational domain, so that outside the respective patches

each $\vec{\phi}_i$ is strictly zero. With such basic functions one gets the sparse matrix A with $O(N_h)$ nonzero entries for $N_h \times N_h$ matrix. Finding the inverse matrix A^{-1} can be then largely simplified. Some special algorithms, developed for dealing with sparse matrices, enable close to linear scaling of the computational costs with N_h for the computation of the inverse matrix.

FEMs are also known for their advantageous convergence properties. In the absence of field singularities in the computational domain, FEMs exhibit exponential convergence to the exact solution. To be more specific, for smooth enough interfaces between different media in the scattering problem and regular meshes (i.e. the tetrahedra or prisms of the computational domain decomposition do not flatten out as the mesh is refined), the error of the approximation satisfies the following estimate [58]:

$$\begin{aligned} & \sqrt{\int_{\Omega} |\vec{E} - \vec{E}_h|^2 dV} + \sqrt{\int_{\Omega} |\vec{\nabla} \times \vec{E} - \vec{\nabla} \times \vec{E}_h|^2 dV} \\ & \leq Ch^p \left(\sqrt{\sum_{|m|_1 \leq p} \int_{\Omega} \left| \frac{\partial^m \vec{E}}{\partial \vec{x}^m} \right| dV} + \sqrt{\sum_{|m|_1 \leq p} \int_{\Omega} \left| \frac{\partial^m \vec{\nabla} \times \vec{E}}{\partial \vec{x}^m} \right| dV} \right), \end{aligned} \quad (3.4)$$

where $|\vec{m}|_1 = |m_1| + |m_2| + |m_3|$ for the three-dimensional problem, p is the degree of the ansatz polynomials, h is the maximum dimension of the subdomains of the mesh (tetrahedra or prisms) and C is a certain constant which is independent on h and \vec{E} . According to Eq. (3.4), the convergence rate $O(h^p)$ is to be obtained for the solution of a scattering problem by means of FEM.

3.2 Treatment of the surface SHG

The concept of the surface nonlinearity only applies in the interfacial layer with the thickness of a few atomic layers, where the breaking of the medium symmetry by the interfaces influence the optical properties. The surface nonlinearity is described by the vector of the surface nonlinear polarization, which is only present within the surface region. In order to simulate the SSHG phenomena one needs an appropriate mathematical treatment of the surface nonlinear polarization.

The most natural way is to consider the surface layer of very small but finite thickness. In this case the surface nonlinear polarization can be dealt with in the same manner as the bulk one, but the respective surface nonlinear susceptibility tensor is nonzero only inside the thin surface layer. It was shown that the exact value of the layer thickness H has only minor effect on the simulation results if H is much smaller than the SH wavelength [36]. The exact solution of the scattering problem then corresponds to the limit $H \rightarrow 0$ with $\chi_{\text{surf}}^{(2)} H = \text{const}$, where $\chi_{\text{surf}}^{(2)}$ is the surface nonlinear tensor. However, in finite-element simulations it seems to be exceedingly inefficient to discretize very thin surface layer. Indeed, the grid far from the interface is sparse enough with the mesh size several times smaller than the wavelength, and much finer mesh within the finite-thickness surface layer would lead to unnecessary grid refinement close to the interface and eventually degrade the performance of FEM.

Alternatively, the surface layer can be regarded as an exactly two-dimensional source of the nonlinear polarization [36]. The implementation of a two-dimensional delta-surface source embedded in the interface between two subdomains with different materials seems also suitable with FEM. The

main difference from usual FEM simulations in this case are the modified boundary conditions for the electric and magnetic fields. The presence of the surface nonlinear polarization results in jumps of the electric and magnetic field components across the interface, as given by Eq. (2.2) in the previous chapter.

For the surface nonlinear polarization the expression Eq. (2.27) from the previous chapter is used. Eq. (2.27) contains only those terms, which are nonzero for the isotropic surface. Other nonzero terms can arise on the surface of a noncentrosymmetric semiconductors, with their number and values determined by the crystal symmetry and the orientation of the surface.

In the treatment below the undepleted pump approximation (UPA) is applied, introduced in the introduction chapter. UPA implies that the inverse action of the SHW on the pump wave can be safely neglected. With UPA the system of equations Eqs. (1.6)-(1.7) gets uncoupled and one has to separately solve the linear Maxwell's equations for two field components - homogeneous equations for the fundamental (pump) field and inhomogeneous equations with the surface nonlinear polarization as the source term for the second-harmonic field. One can thus formulate the following three-step procedure to run the simulations of SSHG:

- 1) solve the scattering problem at the fundamental frequency with FEM, using the incident (pump) field as the source;
- 2) calculate the vector of the surface nonlinear polarization using Eq. (2.27), in order to use it as the source for the emitted field at the second-harmonic frequency;
- 3) solve the scattering problem at the second-harmonic frequency with FEM, using the surface nonlinear polarization calculated in the step (2) as the source term in the weak formulation of Maxwell's equations Eq. (3.1).

3.3 Mathematical formalism of the surface nonlinear polarization

The expression for the surface nonlinear polarization Eq. (2.27) requires the knowledge of the normal and tangent vectors at every point on the surface. In order to directly implement Eq. (2.27) into the programming code a more convenient representation of all terms in Eq. (2.27) would be desirable. One can first simplify Eq. (2.27) from the first term and note that it can be put in the form:

$$P_{S,j}^{\text{NL}} = \varepsilon_0 \chi_{\perp\perp\perp}^{(2)}(\vec{E}\vec{n}) (\vec{E}\vec{n}) \cdot n_j = \varepsilon_0 \vec{E}^T \cdot (\vec{n} \cdot \chi_{\perp\perp\perp}^{(2)} \cdot \vec{n}^T) \cdot \vec{E} \cdot n_j.$$

At this point a new second-rank tensor $\tilde{\chi}_{\perp\perp\perp}^{(2)}$ should be introduced according to:

$$\tilde{\chi}_{\perp\perp\perp}^{(2)} = \vec{n} \cdot \chi_{\perp\perp\perp}^{(2)} \cdot \vec{n}^T.$$

The tangential component of the electric field can be calculated from the normal component as:

$$\vec{E}_{\parallel} = \vec{E} - \vec{E}_{\perp} = (\hat{I} - \vec{n}\vec{n}^T)\vec{E},$$

with the identity matrix \hat{I} . Now one can rewrite the second term in Eq. (2.27) as follows:

$$\begin{aligned} P_{S,j}^{\text{NL}} &= \varepsilon_0 \chi_{\perp||}^{(2)} (\vec{E}_{||}^T \vec{E}_{||}) \cdot n_j = \varepsilon_0 \vec{E}^T \cdot \left((\hat{I} - \vec{n}\vec{n}^T)^T \cdot \chi_{\perp||}^{(2)} \cdot (\hat{I} - \vec{n}\vec{n}^T) \right) \cdot \vec{E} \cdot n_j \\ &= \varepsilon_0 \vec{E}^T \cdot \chi_{\perp||}^{(2)} \cdot (\hat{I} - \vec{n}\vec{n}^T) \cdot \vec{E} \cdot n_j, \end{aligned}$$

and again a new second-rank tensor $\tilde{\chi}_{\perp||}^{(2)}$ is to be defined:

$$\tilde{\chi}_{\perp||}^{(2)} = \chi_{\perp||}^{(2)} \cdot (\hat{I} - \vec{n}\vec{n}^T).$$

Finally, the last term in Eq. (2.27) can be written as:

$$\begin{aligned} P_{S,j}^{\text{NL}} &= \varepsilon_0 \chi_{||\perp||}^{(2)} (\vec{E}\vec{n}) \cdot E_{||,j} = \varepsilon_0 \vec{E}^T \cdot \vec{n} \cdot \chi_{||\perp||}^{(2)} \cdot \left\{ (\hat{I} - \vec{n}\vec{n}^T) \vec{E} \right\}_j \\ &= \varepsilon_0 \vec{E}^T \cdot \left(\vec{n} \cdot \chi_{||\perp||}^{(2)} \cdot (\hat{I} - \vec{n}\vec{n}^T)_{(j,:)} \right) \cdot \vec{E}, \end{aligned}$$

where $(\hat{I} - \vec{n}\vec{n}^T)_{(j,:)}$ stands for the j -th row of the matrix $\hat{I} - \vec{n}\vec{n}^T$, $j = 1, 2, 3$. As before, one should introduce a new second-rank tensor $\tilde{\chi}_{||\perp||,j}^{(2)}$, which depends also on the index j as the parameter:

$$\tilde{\chi}_{||\perp||,j}^{(2)} = \vec{n} \cdot \chi_{||\perp||}^{(2)} \cdot (\hat{I} - \vec{n}\vec{n}^T)_{(j,:)}$$

Bringing together the introduced tensors, one finally obtains the following convenient form for the surface nonlinear polarization:

$$P_{S,j}^{\text{NL}} = \varepsilon_0 \vec{E}^T \cdot \tilde{\chi}_{\text{surf},j}^{(2)} \cdot \vec{E}. \quad (3.5)$$

with

$$\tilde{\chi}_{\text{surf},j}^{(2)} = \tilde{\chi}_{\perp\perp\perp}^{(2)} \cdot n_j + \tilde{\chi}_{\perp||}^{(2)} \cdot n_j + \tilde{\chi}_{||\perp||,j}^{(2)}. \quad (3.6)$$

In the representation with Eq. (3.5)-(3.6) one only needs to know the normal vector at each patch of the surface discretization. In FEM the triangulation of the surface allows direct addressing of each patch of the surface. Hence, the unit normal \vec{n} for each patch is also easily accessible. As long as the normal vector \vec{n} is known, one is able to readily calculate the nonlinear polarization at every point of the surface using Eq. (3.5)-(3.6). It is also worth noting that both the normal vector \vec{n} and the electric field vector \vec{E} can be expressed in the global coordinates on the whole surface.

The generalized boundary conditions Eq. (2.2) demand the calculation of the spatial derivatives of the surface nonlinear polarization in every specific patch of the surface mesh. Therefore one has to introduce first the local coordinate system $x'y'z'$ in every patch and then make a coordinate transformation from the global coordinates to the local ones. The spatial derivatives in both coordinate system can be connected through the rotation matrix R of this transformation. Namely, the Jacobian

in rotated (local) coordinate system can be expressed as:

$$J_{x'y'z'} = R^{-1} \cdot J_{xyz} \cdot R,$$

where x, y, z and x', y', z' denote the initial global and the rotated local coordinate axis respectively and R is the corresponding rotation matrix. In the considered case one gets some uncertainty in choosing the local coordinate system, since only one axis direction provided by the patch normal \vec{n} is fixed. The rotation in the plane of the patch can be taken arbitrary. In general this may cease to be correct for anisotropic interfaces.

However, the usage of this general expression for calculation of the right-hand side of Eq. (2.2) appears to be excessive. Indeed, for an arbitrary scalar field, e.g. the normal component of the surface nonlinear polarization, the following equality holds:

$$\vec{\nabla}_{||} = \vec{\nabla}_{xyz} - \vec{n} \frac{\partial}{\partial \vec{n}} = \vec{\nabla}_{xyz} - \vec{n} \cdot (\vec{\nabla}_{xyz} \cdot \vec{n}),$$

where $\vec{\nabla}_{xyz}$ denotes the gradient taken in the global coordinate system. In such a manner the in-plane gradient gets expressed through the usual gradient of the scalar field. But the gradient of any scalar field is an invariant vector field, meaning that it does not depend on the choice of the coordinate system. So one can for convenience calculate it always in the global coordinates x, y, z regardless of the specific patch on the interface. Finally, one can write down the right-hand side of Eq. (2.2) in the following form:

$$\vec{\nabla}_{||} P_{S,\perp}^{NL} = \vec{\nabla}_{||} (\vec{P}_S^{NL} \cdot \vec{n}) = \vec{\nabla}_{xyz} (\vec{P}_S^{NL} \cdot \vec{n}) - \vec{n} \cdot (\vec{n} \cdot \vec{\nabla}_{xyz} (\vec{P}_S^{NL} \cdot \vec{n})). \quad (3.7)$$

Eq. (3.7) can be now readily implemented into the code. At every specific patch of the surface mesh Eqs. (3.5)-(3.6) are to be used to find the surface nonlinear polarization and then Eq. (3.7) to get the jump of the electric field. In such a way one does not need to make a transformation to the local coordinates at all, since only normal vector \vec{n} and the components of the fundamental field, expressed in the global coordinates, are needed. It should be also stated that calculation of the spatial derivatives in FEM is especially suitable, because the fields are approximated with the polynomial ansatz functions.

3.4 Implementation of generalized boundary conditions

The next important issue is related with the numerical implementation of the generalized boundary conditions Eq. (2.2) in the presence of the surface nonlinear polarization, containing the jumps of the tangential field components. Typically the approximate solution for the electric field in FEM is sought using polynomial subspaces called edge elements of Nedelec [103]. An important property of these ansatz functions is that they inherently assure the continuity of the tangential components of electric and magnetic field across each interface in the computational domain. Therefore the implementation of the generalized boundary conditions Eq. (2.2) with discontinuous tangential components would

require significant changes in the basic algorithms of FEM.

However, here a special mathematical trick is applied that allows getting over the discontinuity of the fields and treat them as well with edge elements of Nedelec. Let us assume that the source of the surface nonlinear polarization is located on the surface of an arbitrary scatterer G , e.g. an all-dielectric nanoparticle, inside the computational domain Ω , possessing the boundary $\Gamma = \partial G$. Now one can write the weak formulation of Maxwell's equations Eq. (3.1) separately for domains G and Ω/G and sum them up, keeping in mind the discontinuity of the tangential component of the magnetic field across the interface Γ , what gives:

$$\begin{aligned} \int_{\Omega} [(\mu^{-1} \vec{\nabla} \times \vec{E}) \cdot \vec{\nabla} \times \vec{\phi}^* - \frac{\omega^2}{c^2} (\varepsilon \vec{E}) \cdot \vec{\phi}^*] dV - 2i\omega \int_{\Gamma} \vec{n} \times \Delta \vec{H}_{\parallel} \cdot \vec{\phi}_T^* dS \\ + \int_{\partial\Omega} \vec{n} \times (\mu^{-1} \vec{\nabla} \times \vec{E}) \cdot \vec{\phi}_T^* dS = 0 \end{aligned} \quad (3.8)$$

for $\forall \vec{\phi} \in V$. Next, an auxiliary vector field \vec{Y} is defined with support only in the vicinity of Γ , which is tangentially continuous across the interface Γ together with its spatial derivatives and obeys the equalities:

$$\begin{aligned} \vec{Y}|_{\Gamma} &= \Delta \vec{E}_{\parallel} \\ \vec{\nabla} \times \vec{Y}|_{\Gamma} &= -2i\mu\omega \cdot \Delta \vec{H}_{\parallel}. \end{aligned} \quad (3.9)$$

The sought-for electric field \vec{E} can be represented in the following way:

$$\vec{E} = \underbrace{(\vec{E} + \vec{Y})}_{=\vec{E}_Y} - \vec{Y} = \vec{E}_Y - \vec{Y}. \quad (3.10)$$

Now it is convenient to introduce instead of the electric field \vec{E} another vector field \vec{E}_C as follows:

$$\vec{E}_C = \begin{cases} \vec{E}, & \vec{r} \in G; \\ \vec{E}_Y, & \vec{r} \notin G. \end{cases} \quad (3.11)$$

One can see that new electric field \vec{E}_C is designed to have continuous tangential components across Γ . Moreover, the magnetic field \vec{H}_C which is related with the electric field \vec{E}_C by the Maxwell's equation (at the SH frequency):

$$\vec{\nabla} \times \vec{E} = -2i\mu\omega \vec{H}$$

also turns out to possess continuous tangential components across Γ .

Using Eq. (3.11) the weak formulation of the electromagnetic scattering problem Eq. (3.1) can be

rewritten for the introduced electric field \vec{E}_C :

$$\begin{aligned} & \int_{\Omega} [(\mu^{-1}\vec{\nabla} \times \vec{E}_C) \cdot \vec{\nabla} \times \vec{\phi}^* - \frac{\omega^2}{c^2}(\varepsilon\vec{E}_C) \cdot \vec{\phi}^*]dV + \int_{\partial\Omega} \vec{n} \times (\mu^{-1}\vec{\nabla} \times \vec{E}_C) \cdot \vec{\phi}_T^* dS \\ & = 2i\omega \int_{\Gamma} \vec{n} \times \Delta \vec{H}_{\parallel} \cdot \vec{\phi}_T^* dS + \int_{\Omega/G} [(\mu^{-1}\vec{\nabla} \times \vec{Y}) \cdot \vec{\nabla} \times \vec{\phi}^* - \frac{\omega^2}{c^2}(\varepsilon\vec{Y}) \cdot \vec{\phi}^*]dV \\ & \quad + \int_{\partial\Omega} \vec{n} \times (\mu^{-1}\vec{\nabla} \times \vec{Y}) \cdot \vec{\phi}_T^* dS, \end{aligned} \quad (3.12)$$

for $\forall \vec{\phi} \in V$. The surface integral in the left-hand side of Eq. (3.12) taken over the outer boundary $\partial\Omega$ can be neglected, if a PML is used, which absorbs all outgoing SH radiation. Another surface integral in the right-hand side in Eq. (3.12) turns to zero because the auxiliary vector field \vec{Y} is by design nonzero only in the vicinity of Γ . Eventually the weak formulation Eq. (3.12) attains the final form:

$$\begin{aligned} & \int_{\Omega} [(\mu^{-1}\vec{\nabla} \times \vec{E}_C) \cdot \vec{\nabla} \times \vec{\phi}^* - \frac{\omega^2}{c^2}(\varepsilon\vec{E}_C) \cdot \vec{\phi}^*]dV - 2i\omega \int_{\Gamma} \vec{n} \times \Delta \vec{H}_{\parallel} \cdot \vec{\phi}_T^* dS \\ & = \int_{\Omega/G} [(\mu^{-1}\vec{\nabla} \times \vec{Y}) \cdot \vec{\nabla} \times \vec{\phi}^* - \frac{\omega^2}{c^2}(\varepsilon\vec{Y}) \cdot \vec{\phi}^*]dV, \quad \forall \vec{\phi} \in V. \end{aligned} \quad (3.13)$$

Direct implementation of Eq. (3.13) demands the explicit functional form of the auxiliary vector field \vec{Y} . Indeed, the function \vec{Y} by definition has to fulfill the expressions Eq. (3.9), but no other conditions are imposed. Therefore one has certain freedom in constructing the exact shape of the function \vec{Y} . It seems especially suitable to consider the limit, when the compact of \vec{Y} gets arbitrary close to Γ , i.e. \vec{Y} is nonzero only in the close vicinity of Γ . The equalities Eq. (3.9) still stay valid upon taking such limit. In this limit the last two integrals in Eq. (3.13) turn to the surface integrals over the boundary Γ . Therefore upon the numerical integration of the last two integrals in Eq. (3.13) the evaluation points with nonzero values of the function \vec{Y} and its spatial derivatives are only located on the surface Γ , with the respective values provided by Eq. (3.9). Finally, when the solution of Eq. (3.13) for the new vector field \vec{E}_C is obtained, one can easily find the electric field \vec{E} using the relations Eq. (3.10)-(3.11).

3.5 Extension to 2D materials

The presented method looks well suited for modeling optical response of the two-dimensional (2D) materials. 2D materials consist of a single layer of atoms or several layers which are held together by the weak van der Waals forces. The most well known of them is graphene, a monolayer of carbon atoms arranged in a 2D honeycomb lattice, which is actively studied since its discovery in 2004 [104] and is currently gaining increasing interest in electronics. Other important examples include transition metal dichalcogenide monolayers (TMDCs) with the chemical formula MX_2 , where M stands for a transition metal from group IV, V and VI, and X is a chalcogen (e.g. sulfur, selenium or tellurium), for instance MoS_2 , MoSe_2 , MoTe_2 , WS_2 and WSe_2 . TMDCs have layered structure and their band gaps changes from indirect to direct depending on the number of layers what can find applications in

optoelectronics [105, 106]. Due to the inherent lack of symmetry of the monolayers, TMDCs usually exhibit a strong second-order nonlinear response and their nonlinear properties are therefore actively investigated [107, 108, 109, 110].

The nonlinear properties of 2D materials can be directly simulated by means of the proposed finite-element method. Indeed, the presence of a few-atom-thick or single-atom-thick sheet of a 2D material on the top of a substrate is completely analogous to the single-atom-thick surface layer of any nonlinear material with symmetry properties different from the bulk. Thus, the nonlinear susceptibility tensor of the 2D material must be simply used in this case in the expression for the surface nonlinear polarization.

Moreover the proposed method can be extended to describe also the linear optical properties of 2D materials. Here the sheet of 2D material can be treated in the Maxwell's equations through the generalized boundary conditions similar to Eq. (2.2):

$$\begin{aligned}\Delta\vec{E}_{\parallel} &= -\frac{1}{\varepsilon_0\varepsilon'}\vec{\nabla}_{\parallel}P_{S,\perp}^L, \\ \Delta\vec{H}_{\parallel} &= -2i\omega\vec{P}_S^L\times\vec{r}_{\perp},\end{aligned}\quad (3.14)$$

where instead of the nonlinear polarization now usual linear polarization of a sheet P_S^L is written.

The calculation of the linear optical response of a sheet would follow the four-step procedure:

- 1) Solution of the linear scattering problem without contribution of a sheet of 2D material;
- 2) Calculation of the linear polarization of a sheet;
- 3) Solution of the linear scattering problem at the same pump frequency as in step (1), but using the linear polarization calculated in step (2) as the source and applying the generalized boundary conditions Eq. (3.14);
- 4) Summing the fields obtained in the step (1) and in the step (3).

Such simulations of the linear properties of 2D materials are not exact ones, since, similar to UPA for nonlinear simulations, the reverse action of the sheet on the incident field is neglected. This approximation is justified when the total polarization of the sheet P_S^L is negligibly small as compared to the total polarization of other materials inside the computational domain. Such condition must be fulfilled for atomic-thick layers on a substrate.

3.6 Test results for benchmark problem

I aim now to illustrate the performance of the developed FEM for a test problem. There are only few benchmark problems for SSHG that allow for the exact analytical solution under UPA and can be thus used for testing the simulation accuracy. Those include the illumination of a semi-infinite slab by a plane-wave [89, 90] and illumination of a sphere by a plane-wave [111]. The testing with a sphere can be simplified, if one assumes the illumination by a single vector spherical harmonic (VSH) as FW. Indeed, exact solution for a plane-wave illumination relies first on the expansion of a plane-wave into VSHs at fundamental frequency and then on the expansion of the electric field discontinuity due to the surface nonlinear polarization from Eq. (2.2) into VSHs at SH frequency. For testing of the

numerical method it would be therefore convenient to consider just one VSH as the pump FW, since it greatly simplifies the analytical solution. The derived exact solution for several low-order VSHs is presented in Appendix C.

FEMs are mathematically proven to exhibit the exponential convergence as h^p , where h is the sidelength of the computational mesh and p is the finite-element degree, see Eq. (3.4). The convergence of the approximate finite-element solution $E_{\text{approx}}(h, p)$ is investigated using the parameter of the relative error ε , defined as:

$$\varepsilon = \left| \frac{E_{\text{exact}} - E_{\text{approx}}(h, p)}{E_{\text{exact}}} \right|.$$

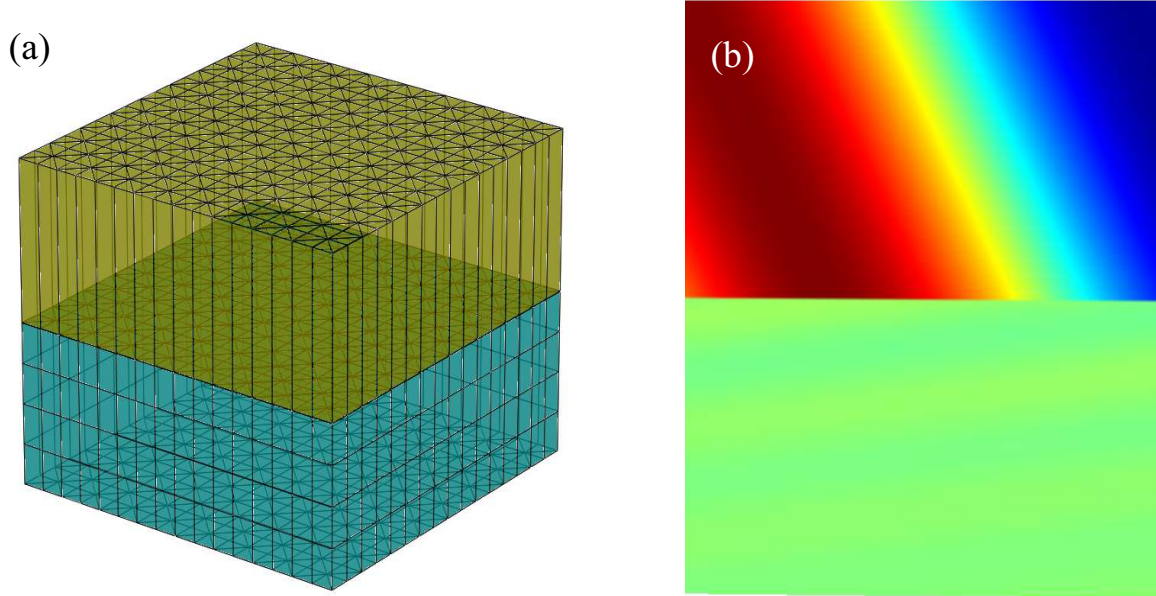


Figure 3.3: Test problem of a plane-wave illumination of a semi-infinite silicon slab; (a) an exemplary uniform grid; (b) an exemplary cross-section of the real part of the emitted SH field.

As a starting point the case of the plane-wave illumination of a slab of crystalline silicon is taken, see Fig. 3.3. The dielectric permittivities of the crystalline silicon at both frequencies are taken from the literature [112]. Different polarizations of the FW as well as different values of the angle of incidence are considered. Besides that, an interesting issue in the simulations would be the role of the local mesh refinement in the vicinity of the interface. To address this question the simulations were also run for the computational mesh, which is a-priori selectively pre-refined close to the Si/Air interface. Specifically, all the prisms of the computational mesh bordering the interface were divided each 4 identical prisms with twice shorter sides in the plane of the interface. At the same time all

other prisms in the mesh were not changed.

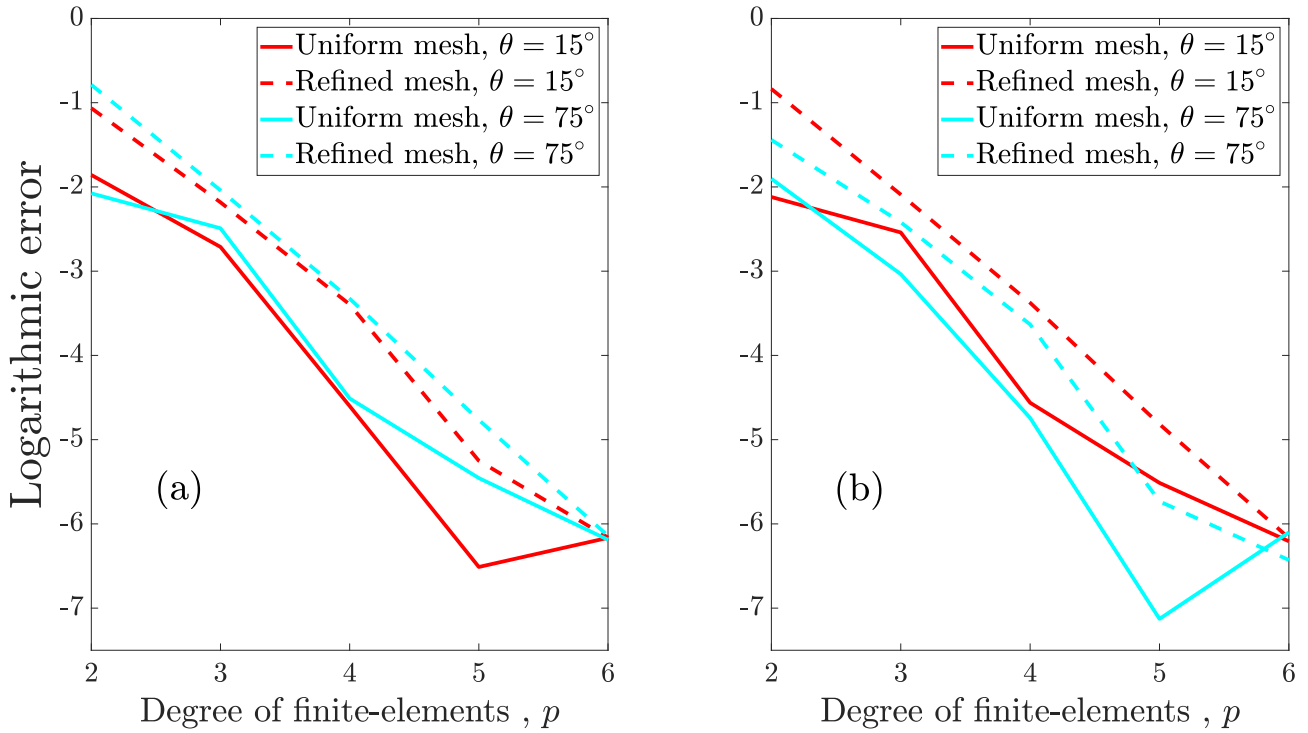


Figure 3.4: Relative error of the simulations for a plane-wave illumination of a semi-infinite silicon slab; uniform and refined meshes are considered for two angles of incidence: $\theta = 15^\circ$ and $\theta = 75^\circ$: (a) TE-polarized FW; (b) TM-polarized FW.

The simulation results are summarized in Fig. 3.4. Here the convergence plots are presented in the logarithmic scale for both polarization of FW, for both uniform mesh and the pre-refined one next to the interface and for two angles of incidence θ : 15° and 75° . The exponential convergence is well seen for both cases. The convergence rate does not appear to noticeably depend neither on the FW polarization, nor on the angle of incidence. Along with that the selectively pre-refined mesh in all cases yields definitely worse performance compared to the uniform one. This finding seems to show that for simple geometries introducing of an additional irregularity of the mesh just degrades the overall FEM performance. However, FEMs are known to be well suitable for complex geometries with multiple length scales involved. Therefore the pre-refinement of the mesh should still be able to boost the computational accuracy for more complex arrangements of the nanoparticles.

As multiple comparative studies demonstrate [113, 114, 115], FEMs are in general relatively time- and memory-efficient, if compared, for example, with the time-domain methods and/or surface-integral methods. One can especially expect that simulations of SSHG with the FEM would profit in terms of the computational costs, when complex geometries with multiple length scales are involved in the calculations. Besides that, other advantages of FEMs can also make a lot difference and favour the usage of FEMs. Apart from the local mesh refinement and huge flexibility in handling the mesh, one can, for example, readily add the bulk nonlinear polarization into the simulated problem through the right-hand side term in Eq. (3.1).

Finally, it should be mentioned that FEMs currently evolve towards the so-called *hp*-adaptivity. This property implies that a specially designed algorithm analyses the scattering problem before the

actual FEM-simulation, and selects for each subdomain of the computational mesh simultaneously both the local mesh sidelength h and local finite-element degree p for polynomial ansatz functions. As the results h and p are not constant within the computational domain, but vary between different subdomains in such a way that the accuracy of the final solution is achieved with the possibly small computational costs. In contrast to the uniform strategy with constant values of h and p , adaptive strategy is expected to lead to significant improvement, when the field varies over strongly different length scales in different part of the computational domain. Such situation typically arises in nanophotonics, especially when simulating resonant nanostructures or metasurfaces. Availability of the elaborated fully hp -adaptive strategies of FEM in the nearest future can be expected to increase the performance of SSHG modeling with the proposed method.

3.7 Summary

A fully finite-element numerical method was developed for efficient modeling of SSHG. This method treats the surface nonlinear polarization as the delta-surface source embedded into the respective interfaces between two media. The influence of this surface source is implemented through the generalized boundary conditions for the tangential components of the electric and magnetic field, imposed at the boundaries of the nonlinear media. The proposed method allows modeling of SSHG from arbitrary-shaped all-dielectric and plasmonic isolated nanoparticles, as well as from the ones periodically-arranged into the metasurfaces.

The convergence tests were performed with an analytically solvable benchmark problem, which confirmed the exponential convergence of the finite-element solution. With the FEM it is possible to locally refine the mesh, for example, close to the interface and/or sharp edges of the geometry, what can be barely done with uniform meshes. Also FEM allows to easily include the bulk nonlinearity in the material into the calculations, what would be difficult to do with BIMs. FEMs are generally known to exhibit profitable performance in terms of the computational costs compared to time-domain methods or BIMs for computational problems with complex geometries and multiple length scales. The performance of the developed FEM can be further improved by means of hp -adaptive strategies. Once fully available, they would yield better performance in terms of the simulation accuracy and the computational costs. Also it is important to state that the developed FEM implementation for modeling SSHG will be included into one of the nearest versions of the commercial software tool for nanophotonic simulations JCMsuite [116, 117].

Next, the developed numerical methods should be applied to the simulations of SSHG for some relevant problems, where the corresponding simulations are needed. In the next chapter one of such problems related to the interplay of the bulk and surface SHG from a slab of a noncentrosymmetric material is investigated by means of the FEM proposed in the current chapter.

Chapter 4

Signatures of surface nonlinearity in far-field patterns of SHG from noncentrosymmetric semiconductors

The interplay of the bulk and surface second-order nonlinearities is proved to be especially complicated in nanophotonics, because the surface-to-volume ratio gets largely increased in the nanostructures as compared to the bulk crystals. In centrosymmetric optical materials the surface second-order nonlinearity prevails in the SH response, but even in centrosymmetric semiconductors the bulk contribution arising from the higher-order terms in the multipole expansion of the nonlinear polarization Eq. (1.3) can sometimes become comparable with the surface contribution, e.g. due to the excitation of optical resonances [39]. That is why the separation and comparison of both nonlinearities have been studied. For example, it was theoretically demonstrated that the surface and bulk contributions to SHG in isotropic materials can be unambiguously identified by their polarization signatures [37, 38]. In such a way the bulk- and surface-induced nonlinear responses were experimentally separated and compared by means of a two-beam technique in thin gold films [38] and in poled polymer films [37]. Though the surface response was shown to play the major role, the bulk nonlinearity under certain conditions was still found to provide an appreciable contribution.

Noncentrosymmetric semiconductor materials are widely used in nonlinear nanophotonic applications because of their strong second-order nonlinearities [30]. Together with the typically low absorption of semiconductors it allows for increasing the energy conversion efficiency of SHG by orders of magnitude compared to the plasmonic structures [118, 119]. In contrast to centrosymmetric semiconductors, for noncentrosymmetric ones the surface SHG is often assumed negligible as compared to the bulk SHG [120]. Some recent studies demonstrate, however, that this assumption may become invalid. In particular, comparable surface and bulk contributions to the SHG signal due to the strong interband resonances were observed at semi-insulating and N^+ -doped GaAs-oxide interfaces [121]. In Refs. [43, 122] the surface nonlinear response was found to dominate in the SHG signal from GaP nanopillars for small enough nanostructures with diameters below 200 nm. In Ref. [66] it was experimentally shown that the surface SHG prevails in the nonlinear response from 50-nm and 100-nm-thick GaAs layers. The strong contribution of the surface resonances to the spectrum of the second-order

nonlinear susceptibility of GaAs in a wide frequency range was experimentally measured in Ref. [46]. In Ref. [44] it was found that resonantly enhanced SHG from GaAs dielectric metasurfaces cannot be adequately described with only bulk nonlinearity, because the surface nonlinearity gives a noticeable contribution. The strong surface-bulk interference was recently observed in the sum-frequency generation from a GaAs crystal with incident visible and infrared pulses [47]. A significant surface contribution to SHG was also measured in a number of experiments at the interface between GaAs and aqueous electrolyte [123, 124, 125]. Finally, the interplay between comparable surface and bulk nonlinear contributions was experimentally detected in thin van der Waals crystals [126]. It should be noted that the presence of the static electric fields within the surface depletion region of a doped GaAs was found to largely enhance the surface-like SHG response from the near-surface region due to the third-order nonlinearity, which significantly contributes to the measured SHG and can even dominate in highly-doped samples [72, 73]. The listed findings point out that the surface nonlinearity in noncentrosymmetric semiconductors cannot generally be neglected. Moreover, it can even prevail in the total SHG nonlinear signal, at least for the nanoscale structures and ultrathin films with large surface-to-bulk ratio and/or for the pump frequencies close to the surface resonances. Hence, the experimental evaluation of the surface optical nonlinearity is essential for different applications related with the frequency conversion in the nanophotonic devices.

Only a few works so far have addressed the question of the efficient separation of the surface- and bulk-driven contributions to the SHG in noncentrosymmetric semiconductors and the quantitative evaluation of the ratio of the surface and bulk nonlinearities. The feasibility of the polarization-resolved analysis for the identification of surface nonlinearities in noncentrosymmetric semiconductors based on the distinct symmetry properties of the surface and the bulk nonlinear tensors of a crystal was first proposed by Stehlin *et al.* [127]. They showed theoretically and confirmed experimentally, that for a certain polarization of the illuminating FW the total second-order nonlinear response of a GaAs slab can be entirely determined by the surface nonlinearity. Following a similar idea, Hollering [128] and Takebayashi *et al.* [129] experimentally separated and compared the surface and bulk nonlinear optical responses in GaAs crystals. The angular dependence of the surface SHG from noncentrosymmetric cubic crystals as the crystal is rotated was theoretically examined in Refs. [130, 131] for varying ratios of the surface and bulk nonlinearities and the influence of the surface orientation was studied. In Ref. [87] the surface nonlinearity was experimentally estimated in GaP nanopillars using some simplified assumptions related to the relative contributions of both nonlinearities to the total SHG. The proposed evaluation scheme, however, neglected the varying normal directions of the curved nanopillar sidewalls. Despite all listed works, comprehensive studies devoted to the separation of the bulk and surface SHG in noncentrosymmetric semiconductors as well as to the quantitative evaluation of the ratio of the surface and bulk nonlinearities are still lacking. Specifically, despite some qualitative estimations of the surface nonlinearity from experimental rotation-angle dependencies of the SHG intensity [128, 129], no attempts have been made to optimize the sensitivity of these approaches and to find their limits of applicability in terms of the reliably detectable values of the elements of the surface nonlinear tensor. Furthermore, no efforts have been put into extending this detectable range with alternative illumination geometries. The better understanding of the interplay

and mutual effect of the surface and bulk nonlinearities in semiconductor nanostructures is of definite interest for improving the design of nonlinear nanoscale components.

This chapter is concerned therefore with the possibility of the identification and the quantitative comparison of the bulk and surface contributions to SHG in noncentrosymmetric semiconductors. Proposed approach is based on finding the specific signatures in the polarization-resolved far-field SHG patterns attributed to the bulk and surface optical nonlinearities, respectively. A semi-infinite slab of a noncentrosymmetric semiconductor possessing a single flat top interface is considered to get rid of the interference of the multiple sources of the surface nonlinear polarization. I start with the plane-wave illumination and theoretically analyse the characteristic features in the dependences of the bulk- and surface-induced SH responses on the illumination geometry. This analysis should allow optimizing the illumination parameters in order to maximize the sensitivity of this setup in terms of the detectable signal of the surface SHG. Then this approach is extended to the case of other pump field, namely a linearly-polarized tightly-focused Gaussian beam (TFGB).

The chapter is organized as follows. In Section 4.1 the problem statement together with the basic idea of proposed approach is described. Moreover, the analytical results for the bulk and surface SHG under plane-wave illumination are presented here. In Section 4.2 the far-field SHG pattern under plane-wave illumination is analysed and the optimal illumination parameters for the detection of the surface-driven SH field are found as well as the estimates of the detectable threshold for the ratio of both nonlinearities. Next, the illumination by a linearly-polarized TFGB is considered. The issues related with the representation of the TFGB and its implementation as the input field for numerical simulations are addressed in Section 4.3. In Section 4.4 the signatures of the surface- and bulk-driven SH signals in the simulated polarization-resolved far-field SHG patterns are found and it is shown, how the quantitative evaluation of the surface nonlinearity can be performed. Finally, Section 4.5 provides the discussion of the obtained results together with some concluding remarks.

4.1 Model and analytical solution

The setup that is considered represents a flat unstructured interface between an upper medium (air) and a semi-infinite slab made of a noncentrosymmetric semiconductor, as shown in Fig. 4.1. The slab is illuminated from the top side by a FW at the fundamental frequency. In particular case of the plane-wave illumination, which is studied below, the incident plane wave possesses the wavevector \vec{k}^{FW} and its angle of incidence is denoted as θ . Let now introduce two coordinate systems that are suitable for further calculations. The first coordinate system is determined by the illuminating FW. The direction of the surface normal is fixed as z -axis and for the plane wave as FW the plane of incidence is taken, formed by the wavevector \vec{k}^{FW} and the surface normal, as the xz -plane of this coordinate system (see Fig. 4.1). Therefore, the x -axis of the FW coordinate system is directed along the tangential component of the wavevector of the FW \vec{k}^{FW} , and y -axis is correspondingly directed orthogonal to it. For a linearly-polarized FW it is convenient to decompose the electric field into TE- and TM-polarized components. It is also suitable to introduce the auxiliary polarization angle α , defined as the angle between the full electric field vector and its TM-component. With such notation

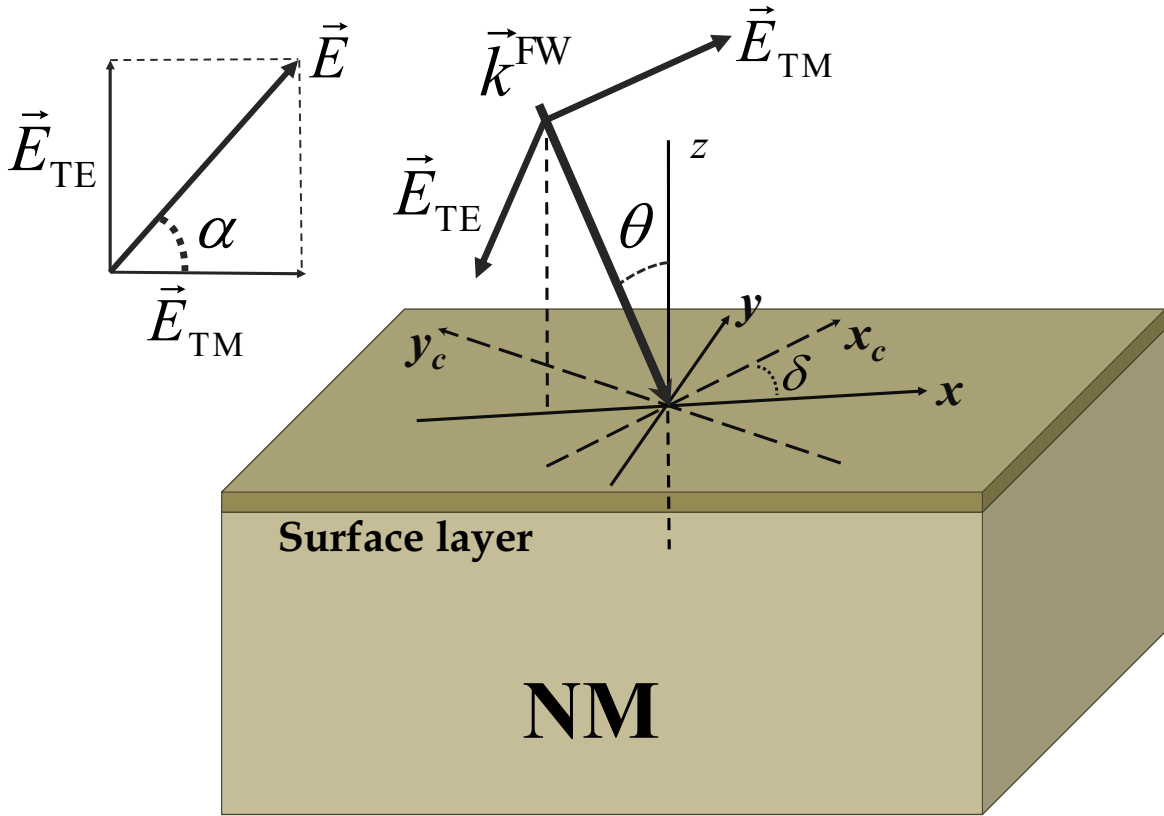


Figure 4.1: Schematic picture of the considered slab of a nonlinear medium (NM), illuminated by a plane wave at the fundamental frequency with the wavevector \vec{k}^{FW} . The inset at the upper left illustrates the polarization components of the fundamental plane wave \vec{E}_{TE} and \vec{E}_{TM} .

one can express the field components of different polarization in the FW coordinate system as:

$$\begin{aligned}
 E_x^{(\omega)} &= E_{TM} \cos \theta = E \cos \alpha \cos \theta, \\
 E_y^{(\omega)} &= E_{TE} = E \sin \alpha, \\
 E_z^{(\omega)} &= E_{TM} \sin \theta = E \cos \alpha \sin \theta.
 \end{aligned} \tag{4.1}$$

Next, let introduce the coordinate system of the crystal (x_c, y_c, z_c) , i.e. coinciding with the crystal symmetry axis. For definiteness the z -axis of the nonlinear crystal is assumed oriented normal to the interface, i.e. $z_c = z$. In the plane of the interface the rotation between the crystal coordinate axes x_c, y_c and the axes x, y of the FW coordinate system is described by the angle δ , as shown in Fig. 4.1. Now one can express the components of the electric field in the crystal coordinate system through the field components in the FW coordinate system Eq. (4.1):

$$\begin{aligned}
 E_{x_c}^{(\omega)} &= E_x^{(\omega)} \cos \delta + E_y^{(\omega)} \sin \delta, \\
 E_{y_c}^{(\omega)} &= -E_x^{(\omega)} \sin \delta + E_y^{(\omega)} \cos \delta, \\
 E_{z_c}^{(\omega)} &= E_z^{(\omega)}.
 \end{aligned} \tag{4.2}$$

Here two parts of the nonlinear polarization of the medium, namely bulk and surface ones, should

be accounted for. Therefore, the second-order nonlinear susceptibility $\chi_{ijk}^{(2)}(\vec{r})$ can be written as the sum of the bulk and surface contributions, i.e., $\chi_{ijk}^{(2)} = \chi_{ijk}^{\text{bulk}} + \chi_{ijk}^{\text{surf}}$, where the surface term is non-zero only inside the interfacial surface region and the bulk term is non-zero in the rest of the medium volume. The overall nonlinear polarization of the medium has to be written in the crystal coordinate system and is given as:

$$\begin{aligned} P_i^{\text{NL}}(2\omega, \vec{r}) &= P_{\text{bulk},i}^{\text{NL}}(2\omega, \vec{r}) + P_{\text{surf},i}^{\text{NL}}(2\omega, \vec{r}) = \varepsilon_0 \chi_{ijk}^{(2)}(\vec{r}) E_j^{(\omega)}(\vec{r}) E_k^{(\omega)}(\vec{r}) \\ &= \varepsilon_0 \chi_{ijk}^{\text{bulk}}(\vec{r}) E_j^{(\omega)}(\vec{r}) E_k^{(\omega)}(\vec{r}) + \varepsilon_0 \chi_{ijk}^{\text{surf}}(\vec{r}) E_j^{(\omega)}(\vec{r}) E_k^{(\omega)}(\vec{r}), \end{aligned} \quad (4.3)$$

where $i, j, k \in x_c, y_c, z_c$. It is more suitable for calculations to express the medium nonlinear polarization also in the FW coordinate system. Performing the respective transformation of the coordinate system, one obtains for the components of the nonlinear polarization:

$$\begin{aligned} P_x^{\text{NL}} &= P_{x_c}^{\text{NL}} \cos \delta - P_{y_c}^{\text{NL}} \sin \delta, \\ P_y^{\text{NL}} &= P_{x_c}^{\text{NL}} \sin \delta + P_{y_c}^{\text{NL}} \cos \delta, \\ P_z^{\text{NL}} &= P_{z_c}^{\text{NL}}. \end{aligned} \quad (4.4)$$

In the treatment below a III-V semiconductor is taken as the nonlinear material, since this class of semiconductors is especially widely used in photonics. All III-V semiconductors possess a zincblende crystal lattice with the $\bar{4}3m$ symmetry group. The symmetry properties of the crystal determine a certain symmetry of the tensor of the nonlinear susceptibility. In the bulk nonlinear tensor in III-V semiconductors only elements with all indices different are non-zero, i.e. $\chi_{ijk}^{(2)} \neq 0$ if $i \neq j \neq k$, and all of them are equal: $\chi_{x_c y_c z_c}^{(2)} = \chi_{y_c z_c x_c}^{(2)} = \chi_{z_c x_c y_c}^{(2)} = \chi_{x_c z_c y_c}^{(2)} = \chi_{y_c x_c z_c}^{(2)} = \chi_{z_c y_c x_c}^{(2)} = \chi_{\text{bulk}}^{(2)}$. Furthermore, the bulk nonlinear tensor is homogeneous in the whole volume of the nonlinear material. As a specific semiconductor $\text{Al}_x\text{Ga}_{1-x}\text{As}$ is used in the simulations. AlGaAs is one of the most promising material platforms for nanophotonic applications, since it exhibits a number of beneficial optical properties, including strong second-order nonlinearity, direct band gap, operation without two-photon absorption at $1.55 \mu\text{m}$ for a molar fraction of Al of $x \geq 0.18$ and the mature nanostructuring technology [31, 40]. SHG from the semiconductor metasurfaces made of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ is actively studied over last years [44, 31, 40, 132, 42, 133, 41, 134]. The molar fraction of Al is fixed to $x = 0.18$ and the medium dispersion is taken from the experimental work [135]. The following values of the complex relative permittivity were obtained at the fundamental and the SH frequencies: $\varepsilon_\omega = 11.46$ and $\varepsilon_{2\omega} = 20.44 + 3.68i$. However, the specific choice of the Al molar fraction x influences noticeably only the linear properties of the material. Therefore the simulation results for the differentiation of the bulk and surface nonlinearities are expected to hold at least qualitatively for different values of the Al content x as well as for other III-V semiconductors.

The surface nonlinearity arises due to symmetry breaking at the surface of the slab, which is only felt by the few closest atomic layers. The consistent treatment of this thin surface layer is crucial for any theoretical study dealing with the surface nonlinearity. As was discussed in the chapter 2, for modeling of the surface nonlinearity thin interfacial layer can be either considered as possessing very

small but finite thickness (as shown in Fig. 4.1) or as an exactly two-dimensional surface source [36]. In the case of a deeply subwavelength thickness of the surface layer both approaches have to yield the same results. Specifically, one has to take a limit for the thickness H of the surface layer, when $H \rightarrow 0$ with $\chi_{ijk}^{\text{surf}} H = \text{const}$. The simulation results for the SSHG in this limiting case should then coincide with the simulation results with the two-dimensional surface source and the value of the constant product $\chi_{ijk}^{\text{surf}} H$ represents the surface nonlinear tensor of the truly two-dimensional surface source, which is used, for example, in Eq. (2.27). Besides that, the sheet of the surface nonlinear polarization can be optionally placed on any side of the interface provided that the corresponding rescaling of the surface nonlinear tensor χ^{surf} is done [38]. I hold onto the reasoning in the chapter 3 and keep using the standard approach, when the nonlinear polarization sheet is located above the interface. Along with that an exactly two-dimensional source is assumed again, i.e. it is supposed in Eq. (4.3):

$$\chi_{ijk}^{\text{surf}}(\vec{r}) = \chi_{ijk}^{2D} \cdot \delta(z).$$

It is worth mentioning that the dimension of the elements of the surface second-order susceptibility tensor should be multiplied by the length dimension with respect to the bulk tensor.

Since the symmetry of the bulk crystal is always broken at the interface, the surface layer possesses different symmetry properties with respect to the bulk of the crystal. In the considered setup in Fig. 4.1, one finds the $mm2$ symmetry group of the surface layer. The tensor of the surface nonlinear susceptibility then has the following non-zero components: $\chi_{z_c z_c z_c}^{(2)}$, $\chi_{z_c x_c x_c}^{(2)}$, $\chi_{z_c y_c y_c}^{(2)}$, $\chi_{x_c z_c x_c}^{(2)}$, $\chi_{x_c x_c z_c}^{(2)}$, $\chi_{y_c y_c z_c}^{(2)}$ and $\chi_{y_c z_c y_c}^{(2)}$, which have to be accounted for [2]. Here a different notation is used for the subscripts than before in order to emphasize that the tensor entries are tied to the crystal coordinate system. However, some available experimental data for the surfaces of III-V semiconductors demonstrate the possibility to simplify the treatment of the surface nonlinearity. In particular, in papers [44, 43, 122, 87] the tensor component $\chi_{z_c z_c z_c}^{(2)}$ was found to give the major contribution to the overall surface nonlinear response. One can also come up with a simple physical explanation for this fact. Indeed, only normal component of the fundamental field $E_z^{(\omega)}$ undergoes a jump at the interface and also the material properties at the interface have a discontinuity only along normal z -axis. Therefore the element $\chi_{z_c z_c z_c}^{(2)}$ of the surface nonlinear tensor has to play the prevailing role, which allows to limit the treatment of the surface nonlinearity to this term only, while neglecting all other non-zero tensor entries.

A few important remarks have to be made now regarding the radiation pattern of the SHG. SHW produced by the surface nonlinear polarization in the setup under consideration is roughly equally emitted upwards into the the air and downwards into the substrate. This is, however, not the case for the SHG in the volume of the slab. Here, the phase mismatch between the pump FW and the generated SHW comes into play. FW in the considered setup propagates downwards into the substrate, hence SHW due to the bulk nonlinearity is generated more efficiently in the same direction. For the upwards outgoing SHW (in the reflection mode) the phase mismatch between FW and SHW is very large because of the opposite propagation directions. Therefore the upwards outgoing SHW seems to be appropriate for the analysis of the interplay between the surface and bulk nonlinearities, since the contribution of the bulk nonlinearity to this SHW is much smaller than to the downwards outgoing

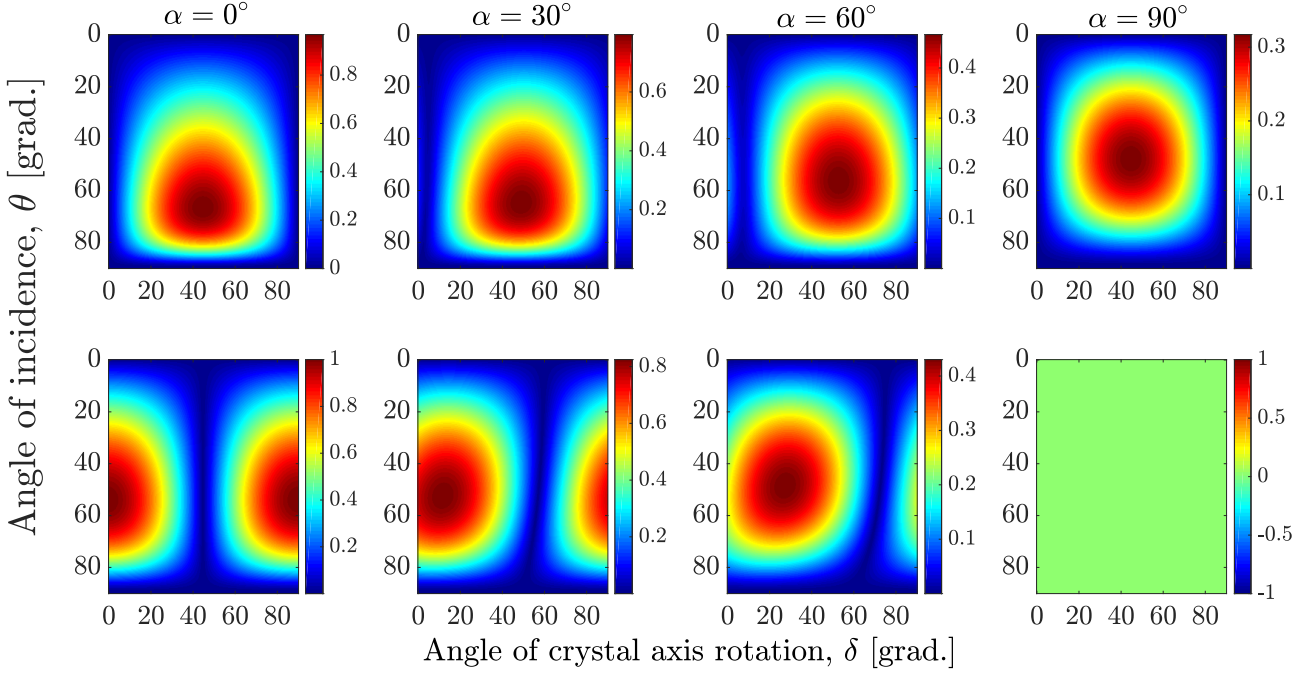


Figure 4.2: Diagram illustrating the dependence of the normalized amplitude of the SHW generated by the bulk second-order nonlinearity only in Eq. (4.5) on the angle of incidence θ and on the angle of the crystal axis rotation δ . The different columns correspond to the different values of the polarization angle α . Top row: normalized amplitude of the TM-polarized component of the SHW. Bottom row: normalized amplitude of the TE-polarized component of the SHW. All subplots are normalized to the maximal value among all, which is achieved in the leftmost subplot in the bottom row.

SHW, and the surface contribution must be much better pronounced against the bulk one.

The analytical expressions for the SHW emitted from a nonlinear slab under illumination by a plane wave were derived in Ref. [89]. If one considers both nonlinear sources separately, the TE- and TM-polarized components of the upwards outgoing SHW are given as:

$$\begin{aligned}
 E_{\text{TE,surf}}^{(2\omega)} &= 0, \\
 E_{\text{TM,surf}}^{(2\omega)} &= \frac{4i(k^{\text{FW}})^2 \varepsilon_{2\omega} P_{\text{surf},z}^{\text{NL}} \sin \theta}{W_0 \varepsilon_{2\omega} + W_2}, \\
 E_{\text{TE,bulk}}^{(2\omega)} &= -\frac{4(k^{\text{FW}})^2 P_{\text{bulk},y}^{\text{NL}}}{\varepsilon_0 (W_0 + W_2) (W_2 + 2W_1)}, \\
 E_{\text{TM,bulk}}^{(2\omega)} &= -\frac{4(k^{\text{FW}})^2 \left(P_{\text{bulk},x}^{\text{NL}} \sqrt{\varepsilon_{2\omega} - \sin^2 \theta} + P_{\text{bulk},z}^{\text{NL}} \sin \theta \right)}{\sqrt{\varepsilon_{2\omega}} \varepsilon_0 (W_0 \varepsilon_{2\omega} + W_2) (W_2 + 2W_1)}, \tag{4.5}
 \end{aligned}$$

with

$$W_0 = 2k^{\text{FW}} \cos \theta, \quad W_1 = k^{\text{FW}} \sqrt{\varepsilon_{2\omega} - \sin^2 \theta}, \quad W_2 = 2k^{\text{FW}} \sqrt{\varepsilon_{2\omega} - \sin^2 \theta}. \tag{4.6}$$

Eqs. (4.5)-(4.6) allow to calculate the emitted SHW from an arbitrary nonlinear medium under illumination at an arbitrary angle of incidence.

In Figs. 4.2 and 4.3 the SHW amplitudes Eq. (4.5) emitted from the AlGaAs slab are plotted, when only bulk and surface nonlinearities are taken into account respectively. The wavelength of the FW was fixed to $\lambda = 1 \mu\text{m}$, so that the SH wavelength is 500 nm. Fig. 4.2 shows the amplitude of

the upwards outgoing SHW generated solely by the bulk nonlinear polarization, i.e. upon assuming the surface nonlinear tensor equal to zero. Each column in Fig. 4.2 corresponds to a specific value of the polarization angle α , so that $\alpha = 0^\circ$ corresponds to the pure TM-polarization of the FW and $\alpha = 90^\circ$ corresponds to the pure TE-polarized FW. Similarly the top and bottom rows correspond to the TM- and TE-polarized component of the emitted SHW. Each subplot in Fig. 4.2 shows the SHW amplitude vs. the angle of incidence θ of the FW and the angle δ between the crystal axis and the plane of incidence (see Fig. 4.1). All figures are normalized to the maximal value of the SHW amplitude among both field polarizations and all values of α in order to illustrate their relative amplitudes. Such normalization is deliberately done, since the absolute values of the SHW amplitude are linearly proportional to the strength of the bulk nonlinear tensor of the specific chosen material and here only the dependence of SHW on the illumination parameters is to be analysed. The SHW amplitude appears to be a harmonic function of the rotational angle δ with a period of 180° , what may be easily found from Eqs. (4.2)-(4.4). At the same time for the absolute value of the SHW amplitude the period amounts to 90° , that is why in all subplots only the range from $\delta = 0^\circ$ to $\delta = 90^\circ$ is displayed. Remarkably, the SHW amplitude in Fig. 4.2 vanishes in both polarizations of SHW, when $\theta = 0^\circ$ and $\theta = 90^\circ$. The case $\theta = 0^\circ$ means the normal incidence of the FW, and the plane SHW in the upper domain cannot be excited, because the normally incident FW has the electric field parallel to the surface and consequently only a normal component of the bulk nonlinear polarization P_z^{NL} is induced through the $\chi_{zcx_cy_c}^{(2)}$ element of the bulk nonlinear tensor. This normal component of the nonlinear polarization can only generate SH fields polarized in the same z -direction, which cannot propagate upwards. In the case of the glancing incidence $\theta = 90^\circ$, both TE- and TM-polarized FWs get fully reflected from the interface, meaning that the pump field at the fundamental frequency does not penetrate into the nonlinear medium and cannot induce a bulk nonlinear polarization. One can also see some other points in the diagrams in Fig. 4.2, where the emitted SHW amplitude turns to zero. Their location on the diagram can be analysed with Eq. (4.5). So, for example, the SHW in TE-polarization (bottom row in Fig. 4.2) vanishes for those combinations of angles θ and δ , which give a zero value for the component P_y^{NL} in Eq. (4.4). Since Fresnel reflection coefficients for both polarizations of the FW are not equal, for $\alpha > 0$ this condition is fulfilled for varying values of the angle of incidence θ , when plotted vs. the angle δ .

Similar plots for the SHW emitted solely by the surface nonlinear polarization are given in Fig. 4.3. Since for the employed form of the surface nonlinear susceptibility tensor a TE-polarized SHW cannot be emitted, only the amplitude of the TM-polarized upwards outgoing SHW is displayed. One can see that all subplots in Fig. 4.3 coincide up to a scaling factor driven by the polarization angle α . This result appears, because the surface nonlinear polarization is determined only by the TM-polarized component of the FW. Moreover, no dependence on the crystal rotation angle δ arises. It should be noted that other weaker components of the surface nonlinear tensor besides $\chi_{zzz}^{2\text{D}}$ would in general introduce a dependence of the surface-driven SHW on the angle δ as well. The dependence of the surface-only SHW on the angle of incidence θ exhibits a single pronounced maximum. This angular dependence is governed by the joint action of two factors: the Fresnel reflection of the FW and the dependence of the surface nonlinear polarization on the z -component of the electric field of

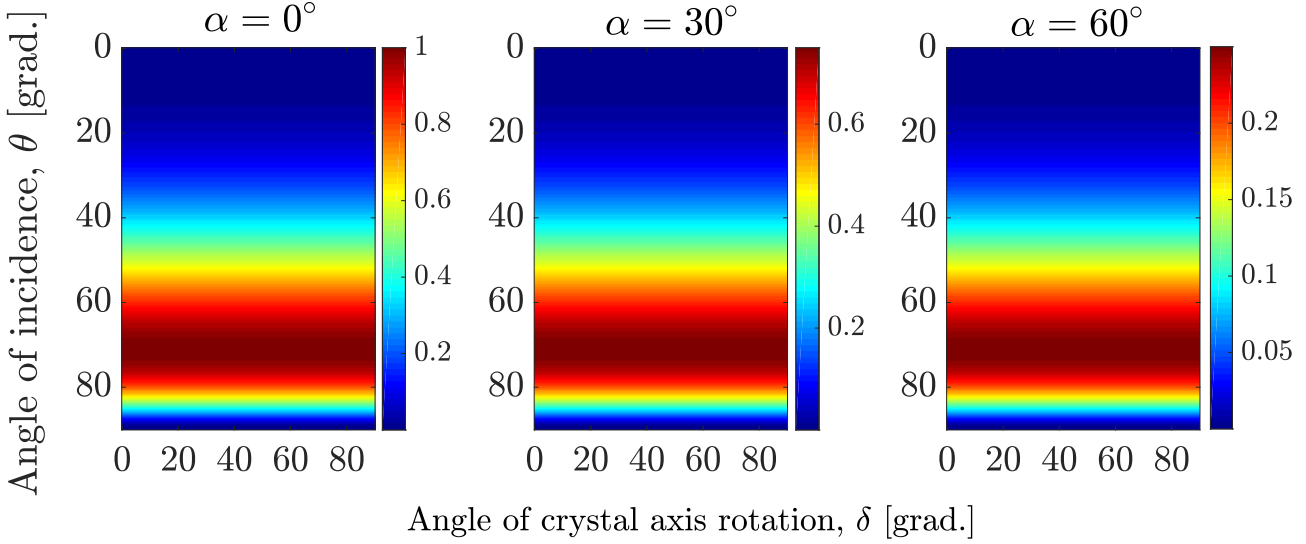


Figure 4.3: Diagram illustrating the dependence of the normalized amplitude of the TM-polarized SHW generated by the surface nonlinearity only in Eq. (4.5) on the angle of incidence θ and on the angle of the crystal axis rotation δ . The different columns correspond to the different values of the polarization angle α . All values are normalized to the maximal field amplitude for $\alpha = 0$.

the FW. For the normal incidence $\theta = 0^\circ$ the normal field component E_z turns to zero and the surface nonlinear polarization can not be induced through the term $\chi_{z_c z_c z_c}^{(2)}$. For the glancing incidence $\theta = 90^\circ$ the FW is fully reflected and also cannot induce a nonlinear polarization in the medium, exactly as for the bulk-induced SHW. For all intermediate values of θ the non-zero surface nonlinear polarization appears and thus the SHW can be emitted.

4.2 SHG patterns for plane-wave illumination

Up to now the SH response provided solely by the surface and bulk nonlinearities was studied, and now their interference should be examined in order to find the optimal illumination conditions, which would allow the most efficient detection of the surface-driven SH signal from a slab of the non-centrosymmetric semiconductor. Let introduce the following dimensionless ratio η of the bulk and surface nonlinearities of the medium:

$$\eta = \frac{\chi_{zzz}^{2D}}{\chi_{\text{bulk}}^{(2)} \lambda}, \quad (4.7)$$

where the pump wavelength λ is added in the denominator. It is assumed in the following that the bulk nonlinearity of a noncentrosymmetric semiconductor is known from the experimental studies, so that one has to determine the value of η to obtain the strength of the surface nonlinearity. It is worth to mention again that the bulk nonlinear tensor has all non-zero components equal to $\chi_{\text{bulk}}^{(2)}$, while for the surface nonlinear tensor only the term χ_{zzz}^{2D} is kept. It should be emphasized that the bulk and surface nonlinear polarizations appear as separate terms in the right-hand side of the system of the linear Maxwell's equations. Therefore, the overall SHG radiation pattern can be obtained just as the linear superposition of bulk-induced and surface-induced complex second-harmonic fields. Since the

evaluation of the ratio of the surface and bulk nonlinear coefficients η is of main interest, the distinct features in the total far-field SHG radiation pattern are to be analysed to allow for the unambiguous separation of the surface contribution and for quantification of the desired ratio from experimental far-field measurements.

In order to find suitable measurement conditions for finding the value of η , Figs. 4.2 and 4.3 should be compared for the specific features of the surface nonlinearity in the SHG radiation pattern. First of all, one can see that the surface nonlinearity only results in the SHW emission for the TM-polarized FW. Next, the main qualitative difference between the bulk- and surface-driven SHWs in Figs. 4.2 and 4.3 is their distinctive dependences on the angle δ of the crystal rotation. Namely, the surface-induced SHW remains constant when varying δ , whereas the bulk-induced SHW exhibits the harmonic dependence. It is worth noticing that taking into account other smaller components of the surface nonlinear tensor would also introduce a weak dependence of the surface signal on δ . Therefore the measurement of the emitted SHW vs. the crystal rotation angle δ with fixed incidence angle θ under the illumination by the TM-polarized plane wave appears to be the best suited experimental layout for the detection of the surface contribution. Figs. 4.2 and 4.3 show that the strongest SHG signal from both nonlinear sources is found for the values of the incidence angle in the range $\theta \sim 50^\circ - 70^\circ$, while for the normal incidence $\theta \approx 0^\circ$ as well as for the glancing incidence $\theta \approx 90^\circ$ both bulk- and surface-induced SHWs vanish. Now one can proceed to determine the optimal values of the angles θ and α , which allow maximizing the surface contribution to the measured SHG radiation pattern.

Summarizing the above, one can conclude that the polarization-resolved measurement of the SHG amplitude, while rotating the crystal sample around the z -axis (and hence changing the δ angle) should enable to quantitatively compare the surface and bulk contributions. The analytical dependence of the total SHG intensity on the crystal rotation angle δ is given as:

$$I_{\text{SHG}}(\delta, \alpha, \theta) \sim \left| A_{\text{surf}}(\alpha, \theta) + A_{\text{bulk}}(\alpha, \theta) \sin \left(2[\delta - \delta_0] \right) \right|^2, \quad (4.8)$$

where the dependence on the crystal rotation angle δ only is explicitly stated. The terms $A_{\text{bulk}}(\alpha, \theta)$ and $A_{\text{surf}}(\alpha, \theta)$ that are introduced here represent the complex amplitudes of the δ -dependence of the fields from Eq. (4.5). The value $\delta_0(\theta, \alpha)$ represents the angular shift of the maximum of A_{bulk} . The dependence of the bulk-induced SH field on 2δ directly follows from Eqs. (4.2)-(4.4). Since $E_{\text{TM,surf}}^{(2\omega)}$ in Eq. (4.5) does not depend on the crystal rotation angle δ , one finds:

$$A_{\text{surf}}(\alpha, \theta) = E_{\text{TM,surf}}^{(2\omega)}(\alpha, \theta) = \frac{4i(k^{\text{FW}})^2 \varepsilon_{2\omega} P_{\text{surf},z}^{\text{NL}}(\alpha, \theta) \cdot \sin \theta}{W_0 \varepsilon_{2\omega} + W_2}. \quad (4.9)$$

At the same time $A_{\text{bulk}}(\alpha, \varphi)$ depends on the angle δ and from Eqs. (4.2)-(4.6) one can derive the

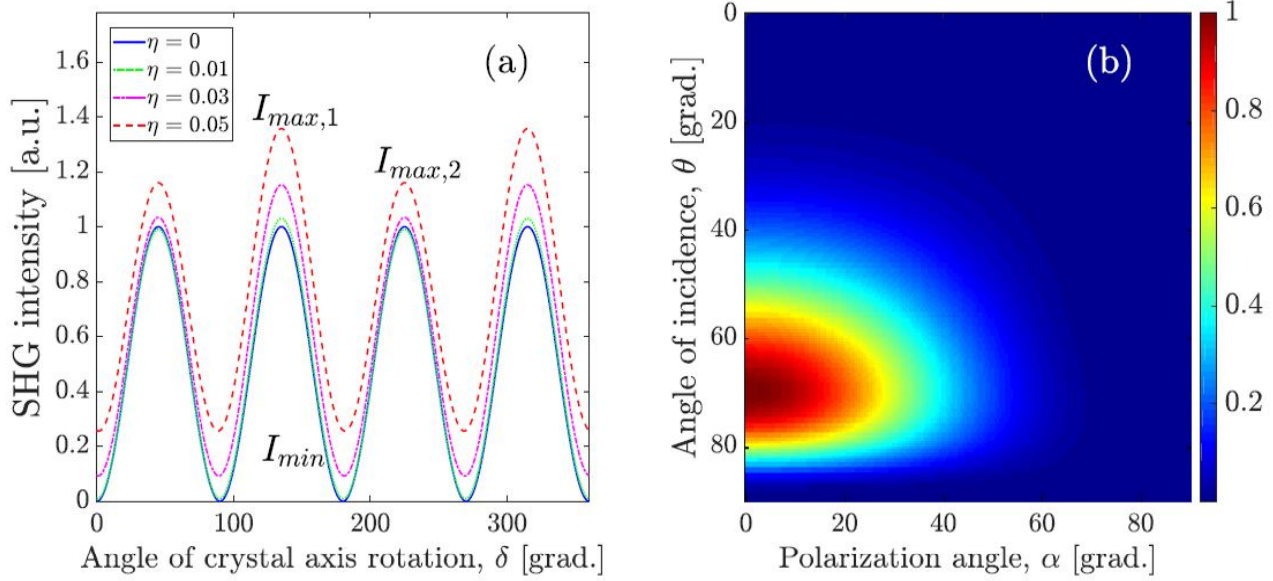


Figure 4.4: (a) Dependence of the SHW intensity I_{SHG} on the rotation-angle δ for the different η and the parameter values $\theta = 60^\circ$ and $\alpha = 0^\circ$, i.e. the pure TM-polarization of the incident plane wave. (b) Normalized values of the difference of the maxima shown in (a), defined by Eq. (4.13), in dependence on the polarization angle α and the angle of incidence θ .

following expressions:

$$A_{\text{bulk}}(\alpha, \theta) = -\frac{4\chi_{\text{bulk}}^{(2)}(k^{\text{FW}})^2 \sqrt{\Pi_1^2 + \Pi_2^2}}{\varepsilon_0(W_0\varepsilon_{2\omega} + W_2)(W_2 + 2W_1)},$$

$$\tan(2\delta_0) = -\frac{\Pi_1}{\Pi_2}, \quad (4.10)$$

with

$$\Pi_1 = \frac{\varepsilon_0}{\sqrt{\varepsilon_{2\omega}}} \left(E_x E_y \sin \theta + E_y E_z \sqrt{\varepsilon_{2\omega} - \sin^2 \theta} \right),$$

$$\Pi_2 = \frac{\varepsilon_0}{\sqrt{\varepsilon_{2\omega}}} \left(\frac{E_x^2 - E_y^2}{2} \sin \theta - E_x E_z \sqrt{\varepsilon_{2\omega} - \sin^2 \theta} \right).$$

In order to find the ratio of the bulk and surface nonlinearities η , one has to measure first the SHG intensity as the function of the angle δ and then fit Eq. (4.8) to the experimental data. The treatment above as well as the diagrams in Figs. 4.2 and 4.3 can be considered as the generalization of the earlier studies [128, 129, 130, 131], where only the rotational-angle dependence of SHG response (angle δ in the notation here) was examined. One can thus improve the performance of measurements in terms of the detectable contribution of the surface nonlinearity by considering also the dependence on the angle of incidence θ and the polarization angle α . Indeed, from the comparison of the dependences of the bulk- and surface-driven SHWs in TM-polarization in Figs. 4.2-4.3 on the angle of incidence θ , one can see that for certain values of the angles θ and α the surface contribution can be expected to be especially strongly pronounced against the background of the bulk-driven SH signal. Particularly, close to the maximum of the surface-induced SHW in Fig. 4.3, the bulk-driven SH signal can be much

weaker than its maximal value. Therefore, the dependence of the total SHG response on the angle δ should be more sensitive to the surface contribution for these values of the angles θ and α and weaker surface nonlinearities can be detected.

Fig. 4.4,a demonstrates the calculated dependence of the total SHG intensity on the crystal rotation angle δ for the illumination by the TM-polarized FW (i.e. $\alpha = 0^\circ$) with the angle of incidence $\theta = 60^\circ$ and for different ratios of the bulk and surface nonlinearities η . One can see alternating maxima and minima of the intensity every 90° , but due to the interference of the bulk and surface contributions with a certain phase shift between them the neighboring maxima have different intensities. Let denote them as $I_{\max,1}$ for the larger one and $I_{\max,2}$ for the smaller one. To achieve the highest sensitivity to the contribution of the surface nonlinearity, the difference between these alternating maxima of Eq. (4.8),

$$T = |I_{\max,1} - I_{\max,2}|, \quad (4.11)$$

has to be maximized. Assuming that the surface contribution is much weaker than the bulk one, i.e. $|A_{\text{surf}}|/|A_{\text{bulk}}| \ll 1$, for both maxima $I_{\max,1}$ and $I_{\max,2}$ one directly finds from Eq. (4.8) up to a nonrelevant constant factor:

$$\begin{aligned} I_{\max,1} &= |A_{\text{surf}}(\alpha, \theta)|^2 + |A_{\text{bulk}}(\alpha, \theta)|^2 + 2|A_{\text{bulk}}(\alpha, \theta)|^2 \cdot \left| \text{Re} \left(\frac{A_{\text{surf}}(\alpha, \theta)}{A_{\text{bulk}}(\alpha, \theta)} \right) \right|; \\ I_{\max,2} &= |A_{\text{surf}}(\alpha, \theta)|^2 + |A_{\text{bulk}}(\alpha, \theta)|^2 - 2|A_{\text{bulk}}(\alpha, \theta)|^2 \cdot \left| \text{Re} \left(\frac{A_{\text{surf}}(\alpha, \theta)}{A_{\text{bulk}}(\alpha, \theta)} \right) \right|, \end{aligned} \quad (4.12)$$

where the last terms arise due to the phase shift between the bulk- and surface-induced SH fields, which is in turn caused by their specific dependences on the complex relative permittivities of AlGaAs both at the fundamental and SH frequencies. Finally, one gets for the difference of both maxima in Eq. (4.11):

$$T = 4|A_{\text{bulk}}(\alpha, \theta)|^2 \cdot \left| \text{Re} \left(\frac{A_{\text{surf}}(\alpha, \theta)}{A_{\text{bulk}}(\alpha, \theta)} \right) \right|. \quad (4.13)$$

The normalized difference of the SHG maxima T from Eq. (4.13) is plotted in Fig. 4.4,b vs. the polarization angle α and the angle of incidence θ . One can clearly get the optimal values of the angles θ and α , which give the maximum of T :

$$\begin{aligned} \theta_{\text{optim}} &\approx 70^\circ, \\ \alpha_{\text{optim}} &= 0^\circ. \end{aligned} \quad (4.14)$$

The optimal polarization angle α_{optim} corresponds to the TM-polarized FW as has been discussed before. From Fig. 4.4,b one concludes that the target function T describing the detectability of the surface contribution in the total SHG signal can be largely increased with the optimized values of the angles Eq. (4.14) as compared to an arbitrary set of their values. It is important to emphasize, that since the function T is linearly proportional to A_{surf} according to Eq. (4.13), the parameters Eq. (4.14) maximize the target function Eq. (4.13) regardless of the value of $\chi_{\text{surf}}^{(2)}$. In other words, no matter how weak is the surface nonlinearity in a specific noncentrosymmetric semiconductor, the function Eq. (4.13) would anyway have its maximum for the parameter values given by Eq. (4.14).

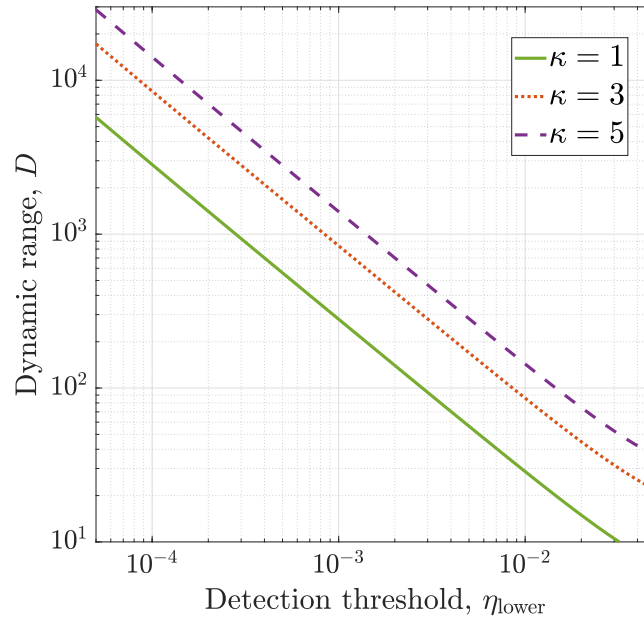


Figure 4.5: Minimal dynamic range of a measurement system D needed to detect the surface contribution to the SHW intensity for the different ratios η of the surface and bulk nonlinearities in the medium.

The values of the illumination angles Eq. (4.14) provide the optimal conditions for the experimental identification of the weak surface contribution to the overall SH response. It is then interesting to estimate quantitatively the range of the detectable values of the surface nonlinearity. It should be noticed first that in an experimental setup one has to ensure the difference of the alternating maxima in Eq. (4.13) to be measured to exceed the noise level in the setup. Only with this requirement fulfilled one can reliably separate the actual surface SH signal from the technical noise. This condition can be expressed in mathematical terms through the dynamic range of the detection system D , which is defined as the ratio of the saturation signal level to the noise level and represents a characteristic of the experimental setup itself. In order to reliably identify the difference of the alternating maxima in the δ -dependence of the SHG intensity, the following requirement has to be fulfilled:

$$D > \frac{I_{\max, 1}}{|I_{\max, 1} - I_{\max, 2}|}. \quad (4.15)$$

In this inequality it is implicitly assumed that the larger maximum in the measured SHG signal $I_{\max, 1}$ can be increased up to the saturation level of the experimental setup by increasing the pump (FW) intensity. To be on the safe side, one can relax the condition Eq. (4.15) and assume that the SHG signal to be measured should be larger than the noise level by an extra factor $\kappa > 1$. In this case one arrives at the following condition

$$\frac{I_{\max, 1}}{|I_{\max, 1} - I_{\max, 2}|} = \frac{D}{\kappa}, \quad (4.16)$$

with the parameter $\kappa > 1$, which actually describes the accuracy of the evaluation of the surface signal against the background of the noise in the camera. Eq. (4.16) implicitly contains the minimum measurable ratio of the nonlinearities η_{lower} .

The lower detectable limit η_{lower} as obtained from Eq. (4.16) for the optimized parameter values Eq. (4.14) and several exemplary values of κ is shown in Fig. 4.5 vs. the dynamic range of the detector

D . The smallest value of η_{lower} , which can be potentially measured with the proposed experimental scheme, corresponds to $\kappa = 1$. With larger values of κ and thus the higher accuracy of the experimental results, sensitivity of the measurements degrades in terms of the detectable range of the strength of the surface nonlinearity. For example, with the camera possessing the dynamic range $D = 100$, the lower limit η_{lower} can be estimated to be around:

$$\eta_{\text{lower}} \approx 0.01 ,$$

if assuming that $\kappa = 3$ is sufficient to reliably identify the surface SHG signal in the experimental curve. One can also find with Eq. (4.16) that for the fixed parameters provided by Eq. (4.14) the detection threshold η_{lower} is almost inversely proportional to the dynamic range D and linearly proportional to the accuracy parameter κ .

The results presented in this section describe an optimized way of measuring the surface nonlinearity. The case of the plane-wave illumination of a single flat medium interface is the only one, which allows for the explicit analytical treatment. However, the optimized plane-wave illumination with large angle of incidence as in Eq. (4.14) may be hard to achieve experimentally. Thus, in the following a focused FW beam will be considered, which experimentally is readily available and would also yield high conversion efficiency. More realistic finite-sized beams can in principle also be expanded into multiple plane-wave components, but the nonlinearity in the problem makes the theoretical analysis increasingly cumbersome. However, the results of this section can provide the analytical framework for a qualitative explanation of the far-field signatures of the surface nonlinearity for other types of illuminating beams at the fundamental frequency. In the next sections these findings are extended to more complex focused pump beams and it is examined, how the detectable range of η values changes for such pump sources.

4.3 Representation of tightly-focused Gaussian beam

Following the reasoning in the previous section, I proceed with the illumination of the same slab made of a noncentrosymmetric semiconductor by a focused beam at the fundamental frequency. Specifically, a normally-incident linearly-polarized tightly-focused Gaussian beam (TFGB) was chosen as the pump source. The motivation for such choice is governed by several reasons. Firstly, TFGB seems to be the most widely studied focused beam and multiple analytical approximations for its field have been derived so far beyond the paraxial approximation. Secondly, the results of the previous section show that large angles of incidence θ are needed for the efficient detection of the surface SH signal. However, using of the obliquely incident focused beam would greatly complicate the detection of the reflected SH signal in the experimental setup. With the normally-incident focused beam, on the other hand, the detection can be directly done with the same objective, which is used for the focusing of the incident beam. Therefore the incident beam has to be tightly focused to assure that the beam incorporates field components with large incidence angles. Thirdly, the major difference between the surface- and bulk-induced SHWs, as it was found above, is their angular dependence on the angle δ between the crystal axis and the plane of incidence of the illuminating wave. The plane-wave Fourier

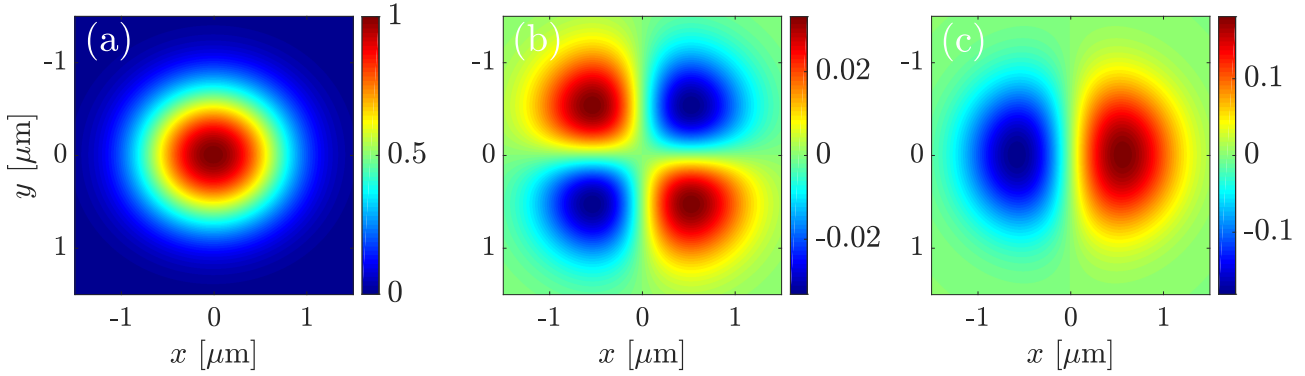


Figure 4.6: Near-field images of the TFGB used in the simulations on the transverse plane through the focus: (a) x -component; (b) y -component; (c) z -component of the electric field. Beam waist is equal to $w_0 = 750$ nm. The beam is linearly-polarized along x -axis. All field components are normalized to the field value at the beam center.

components of a normally-incident TFGB cover the whole range of values of the angle δ . Thus one does not have to do the angular scan over the angle δ , instead the illumination by a normally-incident TFGB would allow performing the whole quantitative comparison of both nonlinearities by the polarization-resolved analysis of the measured far-field SHG radiation pattern.

The interplay of the bulk and surface SHG under TFGB illumination has been already considered in different materials, such as the metal nano-objects [136], glass slides [137, 138], silicon nanowires [139] or isolated centrosymmetric nanospheres [140, 141]. It was shown that the analysis of the SHG radiation patterns allows for the separation of the surface and bulk SH signals [140, 138], second-harmonic fields stemming from the different entries of the surface nonlinear susceptibility tensor [140, 141, 139] or different contributing terms in the bulk quadrupolar tensor [140, 141, 137]. Importantly, all these studies were concerned with SHG from centrosymmetric materials. To the best of my knowledge, SHG under TFGB illumination has not been considered and analysed so far in non-centrosymmetric semiconductors. Because of this fact and the experimental data, demonstrating the possible significant contribution of the surface nonlinearity in nanostructures made of noncentrosymmetric semiconductors, here the characteristic signatures of the bulk and surface nonlinearities in the far-field SHG radiation patterns from $\text{Al}_x\text{Ga}_{1-x}\text{As}$ slabs are examined under the TFGB illumination.

Since the analytical solution for the SHG induced by a pump TFGB is not available even for a single flat interface, one has to perform numerical simulations of the SHG under the UPA to figure out the characteristic signatures of the surface nonlinearity in the SHG radiation pattern. For the numerical simulations the FEM developed in chapter 3 for modeling the SSHG is used. This method was implemented into the commercial finite-element software for optical simulations JCMsuite [116, 117]. JCMsuite was also used for modeling of the purely bulk SHG from the considered slab, representing the vector of the bulk nonlinear polarization as the spatially distributed source for the SH field. TFGB was introduced as the source of the input field at the fundamental frequency.

As long as the exact analytical solution for the fields of a linearly-polarized TFGB is lacking, it is worth to address in more detail the numerical representation of an illuminating TFGB. When a Gaussian beam is focused to the subwavelength waist radius, the standard paraxial approximation appears to be not valid anymore, because it cannot accurately describe the longitudinal component of

the electric field. Indeed, a standard expression for a linearly-polarized Gaussian beam is derived from the scalar Helmholtz equation, so that the Gaussian beam is assumed to possess only one non-zero component of the electric field. When a linearly-polarized Gaussian beam gets tightly focused, this assumption, i.e. the paraxial approximation, is not applicable and the longitudinal field component becomes significant and cannot be neglected.

Beyond the paraxial approximation the fields of a Gaussian beam can be described using expansions into a series over a small parameter ε , also known as Lax series, with:

$$\varepsilon = \frac{w_0}{z_R} = \frac{2}{kw_0} < 1, \quad (4.17)$$

where w_0 is the beam waist and $z_R = kw_0^2/2$ is the Rayleigh length [142, 143, 144, 145, 146, 147]. It is interesting to note that in the limit of the paraxial approximation the parameter ε Eq. (4.17) coincides with the tangent of the beam divergence angle. As it was found in Ref. [146], the electric field of a linearly-polarized Gaussian beam beyond the paraxial approximation can be obtained from a linearly-polarized vector potential of the form:

$$\vec{A}(x, y, z, t) = \vec{e}_x A_0 \Psi(x, y, z) e^{i(\omega t - kz)},$$

with a constant amplitude A_0 . Substituting this expression into the wave equation for the vector potential one arrives to:

$$\Delta \Psi - 2ik \frac{\partial \Psi}{\partial z} = 0. \quad (4.18)$$

For small values of the parameter ε Eq. (4.17), it can be used as an expansion parameter for the sought-for function Ψ :

$$\Psi = \sum_{n=0}^{+\infty} \varepsilon^{2n} \Psi_{2n}. \quad (4.19)$$

The series expansion for the function Ψ now has to be inserted into Eq. (4.18) and the terms of the same power of ε^{2n} have to be grouped together, which results in the following relations:

$$\begin{aligned} \nabla_{\perp}^2 \Psi_0 - 4i \frac{\partial \Psi_0}{\partial \zeta} &= 0, \quad \text{if } n = 0, \\ \nabla_{\perp}^2 \Psi_{2n} - 4i \frac{\partial \Psi_{2n}}{\partial \zeta} + \frac{\partial \Psi_{2n-2}}{\partial \zeta^2} &= 0, \quad \text{if } n \geq 1, \end{aligned} \quad (4.20)$$

with the rescaled coordinates:

$$\xi = \frac{x}{w_0}, \quad \nu = \frac{y}{w_0}, \quad \zeta = \frac{z}{z_R}, \quad \rho^2 = \xi^2 + \nu^2. \quad (4.21)$$

The first equation in Eq. (4.20) for Ψ_0 provides the standard solution for a Gaussian beam in the paraxial approximation. The first terms of the expansion Eq. (4.19) were obtained as follows [142,

143, 144]:

$$\begin{aligned}\Psi_0 &= f e^{-f\rho^2}, \\ \Psi_2 &= \left(\frac{f^2}{2} - \frac{f^4\rho^4}{4}\right) e^{-f\rho^2},\end{aligned}\tag{4.22}$$

where it is denoted:

$$f = \frac{i}{\zeta + i}.$$

The standard paraxial approximation corresponds to the limit $\varepsilon \rightarrow 0$. With tighter focusing of the Gaussian beam the beam waist w_0 becomes smaller and ε from Eq. (4.17) increases respectively. For larger ε one cannot neglect the correction higher-order terms in the expansion Eq. (4.19) to assure the high enough accuracy of the simulations. For very tight focusing as $\varepsilon \rightarrow 1$ the convergence of the Lax series is in doubt, but for moderate levels of focusing the truncated Lax series was shown to yield a reasonable analytical approximation for the fields of the TFGB [148]. In the simulations a TFGB with the beam waist $w_0 = 750$ nm was taken, which gives the value of the parameter Eq. (4.17):

$$\varepsilon \approx 0.42.$$

It is also assumed that the beam waist is located exactly in the top interface of the considered slab of AlGaAs. The number of the expansion terms in Eq. (4.19) is selected to meet the accuracy of the finite-element simulations. The expected relative accuracy of the simulations in the software was taken as 0.01, which determines the settings of the PML. According to the accuracy estimates for the Lax series obtained in [144], one correspondingly keeps the higher-order corrected terms for a linearly-polarized TFGB up to ε^3 . The respective expressions for the components of the electric field in the transverse plane through the focus are [146]:

$$\begin{aligned}E_x &= E_0 e^{-\rho^2} \left(1 + \varepsilon^2 \left[\xi^2 - \frac{\rho^4}{4}\right]\right), \\ E_y &= E_0 e^{-\rho^2} \cdot \varepsilon^2 \xi \nu, \\ E_z &= E_0 e^{-\rho^2} \xi \left(\varepsilon + \varepsilon^3 \left[-\frac{1}{2} + \rho^2 - \frac{\rho^4}{4}\right]\right).\end{aligned}\tag{4.23}$$

The near-field images of the TFGB used in the simulations on the transverse plane through the focus were calculated with Eq. (4.23) and are plotted in Fig. 4.6. It is convenient to expand the electric field in the transverse plane through the focus into the two-dimensional Fourier series, then the electric field can be written as the superposition of the plane-wave components as:

$$\vec{E}(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \vec{F}(k_x, k_y) \cdot e^{i(k_x x + k_y y + k_z z)} dk_x dk_y,\tag{4.24}$$

where

$$\begin{aligned}
 F_x &= \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E_x(x, y, z) e^{-i(k_x x + k_y y)} dx dy, \\
 F_y &= \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E_y(x, y, z) e^{-i(k_x x + k_y y)} dx dy, \\
 F_z &= -\frac{F_x k_x + F_y k_y}{k_z}, \\
 k_z &= \sqrt{(k^{\text{FW}})^2 - k_x^2 - k_y^2}.
 \end{aligned}$$

Using Eq. (4.23) one readily finds:

$$\begin{aligned}
 F_x(k_x, k_y) &= \frac{E_0 w_0^2}{4\pi} e^{-\frac{w_0^2(k_x^2 + k_y^2)}{4}} \left(1 + \varepsilon^2 \left[\frac{2 - k_x^2 w_0^2}{4} - \frac{3}{16} \left(1 - w_0^2(k_x^2 + k_y^2) + \frac{w_0^4(k_x^4 + k_y^4)}{12} \right) \right. \right. \\
 &\quad \left. \left. - \frac{3}{32} (2 - k_x^2 w_0^2)(2 - k_y^2 w_0^2) + \frac{k_x k_y w_0^2}{16} (12 - w_0^2(k_x^2 + k_y^2)) \right] \right), \\
 F_y(k_x, k_y) &= -\frac{1}{16\pi} E_0 \varepsilon^2 k_x k_y w_0^4 e^{-\frac{w_0^2(k_x^2 + k_y^2)}{4}}.
 \end{aligned} \tag{4.25}$$

For the numerical simulations the TFGB is to be represented as the sum of its multiple plane-wave components, i.e. through the discrete form of the integral Eq. (4.24):

$$\vec{E}(x, y, z) \approx \sum_{l_1=-K_x}^{l_1=+K_x} \sum_{l_2=-K_y}^{l_2=+K_y} \vec{F}(k_x(l_1), k_y(l_2)) \cdot e^{i(k_x(l_1)x + k_y(l_2)y + k_z(l_1, l_2)z)} \Delta k_x \Delta k_y, \tag{4.26}$$

where indices l_1, l_2 numerate the discretization points along k_x - and k_y -axis. It is reasonable to turn this equation into the cylindrical coordinates as:

$$\vec{E}(x, y, z) \approx \sum_{l_r=0}^{l_r=K_r} \sum_{l_\varphi=-\pi}^{l_\varphi=+\pi} \vec{F}(k_x(l_r, l_\varphi), k_y(l_r, l_\varphi)) \cdot e^{i(k_x(l_r, l_\varphi)x + k_y(l_r, l_\varphi)y + k_z(l_r, l_\varphi)z)} k_r \Delta k_r \Delta \varphi, \tag{4.27}$$

with

$$\begin{aligned}
 k_x(l_r, l_\varphi) &= k_r(l_r) \cos \varphi(l_\varphi), \\
 k_y(l_r, l_\varphi) &= k_r(l_r) \sin \varphi(l_\varphi).
 \end{aligned}$$

The upper integration boundary for k_r is taken as:

$$K_r = \frac{6}{w_0},$$

i.e. three times the width of the Gaussian profile in the Fourier transform of a Gaussian beam in the paraxial limit. Eqs. (4.25) and (4.27) together provide the required representation of an incident linearly-polarized TFGB for the simulations.

4.4 Far-field SHG patterns under TFGB illumination

When the slab is illuminated by the linearly-polarized normally-incident TFGB, the only varying parameter is the angle between the direction of the beam polarization and the crystal axis. Here the notation similar to the case of the plane-wave illumination shown in Fig. 4.1 is applied. Namely, the angle δ is defined between the polarization direction of the TFGB and the crystal x_c -axis. In order to find out the specific features of the bulk- and surface-driven SHG in the total radiation pattern the simulations are first performed separately, when only the bulk or surface nonlinearity acts.

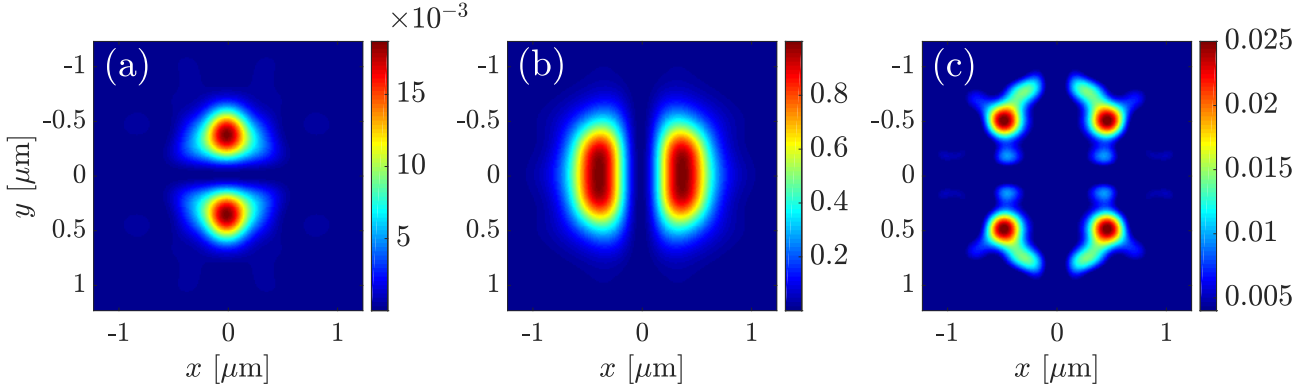


Figure 4.7: Near-field SHG patterns induced by the bulk nonlinearity only, as calculated in the horizontal plane $z = 0.7 \mu\text{m}$ above the surface for $\delta = 0^\circ$: the intensity distributions of (a) x -component; (b) y -component; (c) z -component of the electric field. All plots are normalized to the largest value among all, which is achieved in the subplot (b).

Firstly the case of the bulk nonlinearity only is considered. Fig. 4.7 demonstrates the near-field SHG pattern obtained in the horizontal plane $z = 0.7 \mu\text{m}$ above the interface for the value of the rotation angle $\delta = 0^\circ$. It seems more appropriate, however, to analyse the far-field pattern as if measured in the focal plane of a collection lens or objective. The respective far-field SHG patterns for different values of the rotation angle δ are brought together in Fig. 4.8. The top row shows the intensity of the SH field polarized along the x_c -direction, while the bottom row shows the intensity of the SH field polarized along the y_c -direction. When plotting Fig. 4.8, the presence of the longitudinal field component was accounted for, as illustrated in Fig. 4.7. Therefore subplots in Fig. 4.8 do not represent just the Fourier transforms of the x - and y -components of the electric field from Fig. 4.7, but are calculated as:

$$F_x(k_x, k_y) = \frac{k_z}{k_{\text{FW}}} FT [E_x(x, y)] - \frac{k_x}{k_{\text{FW}}} FT [E_z(x, y)],$$

$$F_y(k_x, k_y) = \frac{k_z}{k_{\text{FW}}} FT [E_y(x, y)] - \frac{k_y}{k_{\text{FW}}} FT [E_z(x, y)].$$

The $x_c y_c$ -coordinate system is fixed to the crystal axes in the interface plane. Since the illuminating TFGB is linearly-polarized under an angle δ with respect to the crystal x_c -axis, its polarization direction is anticlockwise rotating from left to right in Fig. 4.8. All subplots in Fig. 4.8 are normalized to the largest value among all, which is achieved twice in the leftmost subplot in the bottom row and in the rightmost subplot in the top row.

Fig. 4.8 shows the largest SHG intensity in y_c -polarization for the TFGB polarized along x_c -

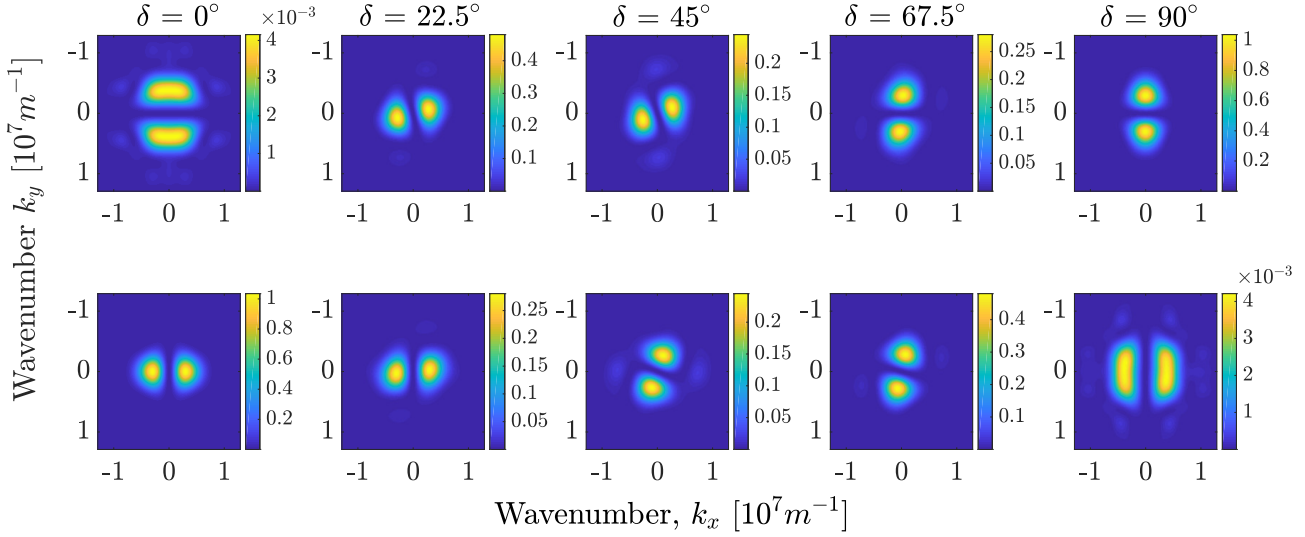


Figure 4.8: Polarization-resolved far-field SHG intensity distributions induced by the bulk nonlinearity only for different values of the angle δ ; top row: x_c -polarization, bottom row: y_c -polarization. The FW wavelength is $\lambda = 1 \mu\text{m}$, thus the SH wavelength is 500 nm. All subplots are normalized to the maximal value among all, which is achieved both in the rightmost subplot in the top row and in the leftmost subplot in the bottom row.

axis as well as in x_c -polarization for the TFGB polarized along y_c -axis. In both cases, when the pump TFGB is polarized along one of the crystal axes, i.e. for $\delta = 0^\circ$ and $\delta = 90^\circ$, the amplitudes of the strongest peaks in the intensity patterns for both polarizations differ by a factor of around 250. The radiation patterns for all values of the angle δ have two-lobed form in both polarizations. The appearance of such two-lobed far-field pattern results from the symmetry properties of the bulk nonlinear susceptibility tensor of III-V semiconductors. Indeed, when illuminating TFGB is, for instance, linearly-polarized along the crystal x_c -axis, i.e. $\delta = 0^\circ$, the largest components of the electric field are the x_c - and z_c -components (see Fig. 4.6), what results in a dominant y_c -component of the bulk nonlinear polarization $\vec{P}_{\text{bulk}}^{\text{NL}}$. Moreover, from the images of the TFGB in Fig. 4.6 one can see that the product of x_c - and z_c - electric field components consists of two domains shifted along x_c -axis with respect to each other. Since these domains of the bulk nonlinear polarization have the primary direction along y_c -axis, one can expect the radiation pattern to be qualitatively similar to the interference radiation pattern of two classical dipoles. The corresponding radiation pattern from two dipoles exhibits a two-lobed y_c -polarized far-field pattern with lobes aligned along the k_x -direction. At the same time the x_c -polarized SH field turns out to be much weaker, because the component of the bulk nonlinear polarization $P_{\text{bulk},x_c}^{\text{NL}}$ is much weaker than $P_{\text{bulk},y_c}^{\text{NL}}$. As the angle δ is increased, the dominant direction of the bulk nonlinear polarization rotates, but the two-lobed shape of the far-field pattern is still preserved.

In a similar way the case of the surface nonlinearity only is considered next. Respective far-field SHG intensity patterns are plotted in Fig. 4.9. As before the top row corresponds to the intensity of the SH field polarized along the x_c -direction, while the bottom row displays the intensity of the SH field polarized along the y_c -direction. All subplots are scaled again to the maximum value among all, which is achieved twice in the leftmost subplot in the top row and in the rightmost subplot in the bottom row. The symmetry properties of the surface-driven far-field patterns appear to be significantly different as compared to the case of the bulk-induced SH field. Particularly, two domains of the

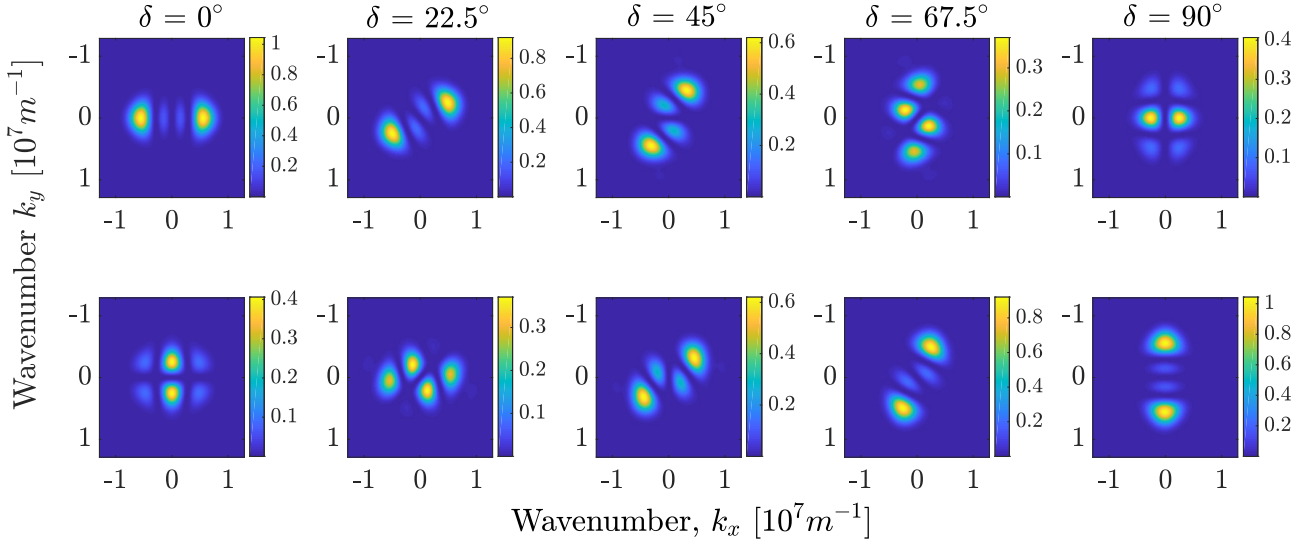


Figure 4.9: Polarization-resolved surface-induced far-field SHG intensity distributions for different values of the rotation angle δ . Top row: x_c -polarization, bottom row: y_c -polarization. The FW wavelength is $\lambda = 1 \mu\text{m}$, thus the SH wavelength is 500 nm. All subplots are normalized to the maximal value among all, which is achieved both in the leftmost subplot in the top row and in the rightmost subplot in the bottom row.

normal component of the surface nonlinear polarization $P_{\text{surf},z_c}^{\text{NL}}$ prevail due to the dominating $\chi_{zzz}^{(2)}$ element of the surface nonlinear susceptibility tensor. One gets two domains of $P_{\text{surf},z_c}^{\text{NL}}$, since z -component of the illuminating TFGB vanishes at the beam symmetry axis and exhibits two maxima symmetrically located along the polarization direction, see Fig. 4.6. The interference of two domains of the surface nonlinear polarization directed along the surface normal leads to the multi-lobed SHG radiation patterns with the comparable far-field amplitudes in both polarizations and for all values of the rotation angle δ .

The SH field emitted, when both the bulk and surface nonlinearities are included in the simulations, is simply obtained as the linear superposition of the bulk- and surface-induced fields shown in Figs. 4.8 and 4.9. It should be noted that this statement is valid for the complex amplitudes of the electric field, but not for the intensities of both SH fields. If one assumes the bulk nonlinearity known, this superposition can be considered as the sum of two fields, where the second one has a weight factor proportional to the unknown ratio of the surface and bulk nonlinear coefficients Eq. (4.7). It is interesting here to find the illumination conditions that allow for the best possible spatial isolation of the far-field lobes coming from the bulk and surface sources of the nonlinear polarization. Such separation would let to reliably identify and analyse the surface-induced and bulk-induced features in the total far-field SHG intensity distribution. From the comparative analysis of Figs. 4.8 and 4.9 one can conclude that the bulk- and surface-induced lobes are maximally separated for the polarization angles $\delta = 0^\circ$ and $\delta = 90^\circ$, i.e. when the TFGB is linearly-polarized along one of the crystal axes. It is also worth noting that thanks to the symmetry of the problem the far-field SHG patterns for these two values of δ can be obtained from each other by rotating for 90° and exchanging both polarizations. In particular, when $\delta = 0^\circ$ the surface-induced and bulk-induced peaks in the x_c -polarized far-field SHG pattern are symmetrically shifted from the beam axis along the k_x - and k_y -axis, respectively, what minimizes their mutual overlap. The overlap of the surface-induced and bulk-induced lobes in the y_c -polarized far-field SHG pattern is much larger, but the contributions are still distinguishable, because

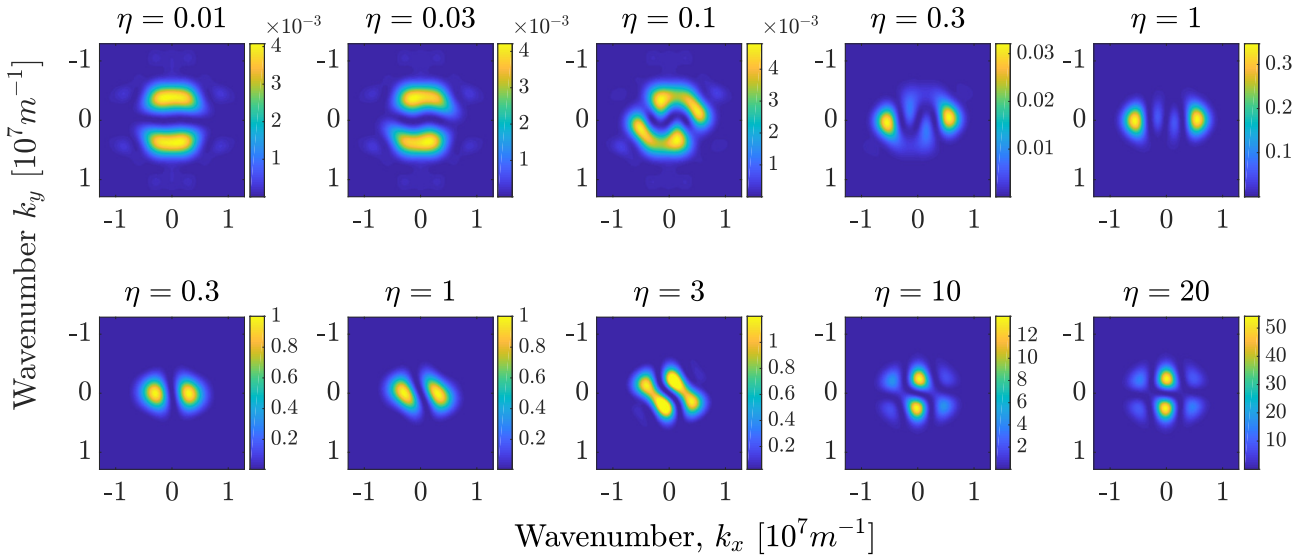


Figure 4.10: Polarization-resolved far-field SHG intensity distributions for different values of the parameter η with $\delta = 0^\circ$. Top row: x_c -polarization, bottom row: y_c -polarization. The FW wavelength is $\lambda = 1 \mu\text{m}$, thus the SH wavelength is 500 nm.

the bulk-induced far-field possesses its main lobes on the k_x -axis and vanishes on the k_y -axis, while the surface-induced far-field possesses its main lobes on the k_y -axis and vanishes on the k_x -axis.

In the following the treatment is limited to the specific case of $\delta = 0^\circ$, when the illuminating TFGB is polarized along x_c -axis. Fig. 4.10 shows the total far-field SHG intensity patterns for different values of the parameter η from Eq. (4.7). Again, in the top row the x_c -polarized SH field is plotted and in the bottom row the y_c -polarized SH field. All plots are scaled to the largest amplitude in the bulk-induced far-field pattern. Different values of η for both polarizations were used to illustrate the transition from the pure-bulk to the pure-surface SHG patterns in x_c - and y_c -polarization, since the maxima of the bulk-induced far-field patterns in the two polarizations in Fig. 4.8 were found to be largely different. Fig. 4.10 demonstrates that in each polarization the obtained far-field pattern closely follows the pure-bulk pattern for small values of η and undergoes the gradual transition towards the pure-surface far-field pattern with the increase of η . The intermediate range of η values, where the specific features of both contributions can be identified, is of main interest. Within this range of η values each specific far-field intensity distribution with certain relative amplitudes of the bulk-induced and surface-induced lobes is unambiguously connected to a certain value of η . The boundaries of the relevant range of η values are governed by the condition of the pure-bulk/pure-surface characteristic lobes being still safely distinguishable against the background of the far-field pattern from another nonlinear source.

In particular, in x_c -polarization the surface-induced lobes are located on the k_x -axis in the far-field, where the bulk-induced SH field vanishes. In contrast to this, the bulk-induced lobes are the strongest on the k_y -axis, where the surface-induced SH field vanishes. This fact enables measurement of the surface-to-bulk ratio from the x_c -polarized SHG radiation pattern. Indeed, the ratio of the maxima along k_x - and k_y -axis in the far-field intensity distribution is linearly proportional to η . In Fig. 4.11 the cross-sections of the far-field SHG intensity pattern in x_c -polarization for $\eta = 0.01$ from the leftmost subplot in the top row in Fig. 4.10 are shown. Fig. 4.11,a gives the SH intensity along the k_x -axis

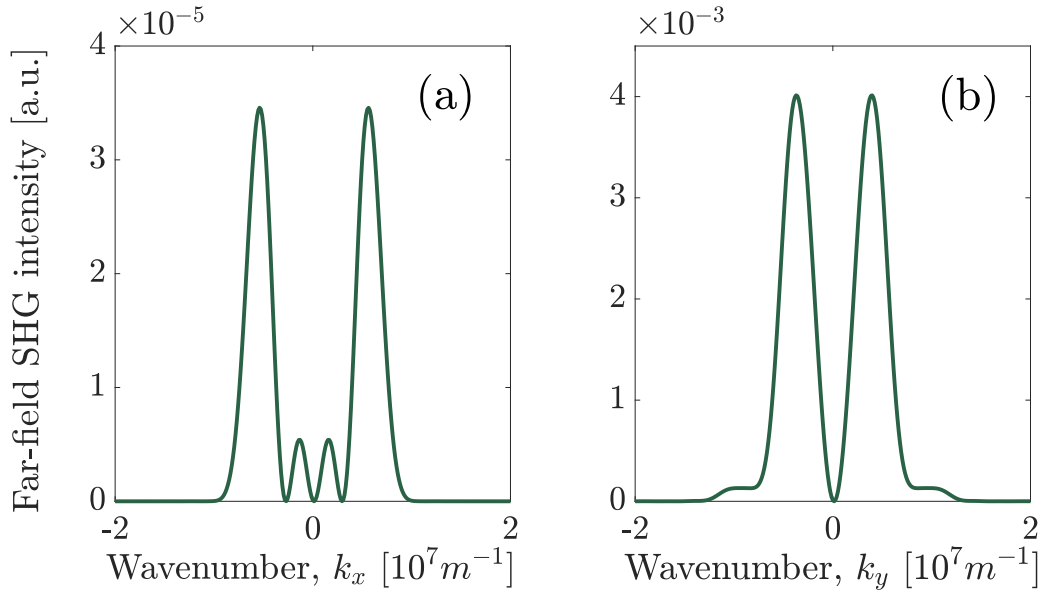


Figure 4.11: Cross-sections through the far-field intensity pattern in x_c -polarization for $\eta = 0.01$ in Fig. 4.10: (a) SH intensity along the k_x -axis for $k_y = 0$ and (b) along the k_y -axis for $k_x = 0$.

for $k_y = 0$ and Fig. 4.11,b displays the SH intensity along the k_y -axis for $k_x = 0$. As the value of η gets increased, the bulk-induced SH field in Fig. 4.11,b remains unchanged, while the surface-induced SH field in Fig. 4.11,a grows linearly in amplitude without changing its profile. Similar to the plane-wave illumination in the previous section, now the estimated boundaries of the range of the measurable values of η are to be calculated. These boundaries are determined by the ratio of the amplitudes of the lobes in the SHG intensity patterns induced solely by the bulk and surface nonlinear polarizations. Let us denote the respective maxima along the k_x - and k_y -axes in the far-field pattern as I_1^{\max} for the stronger one and I_2^{\max} for the weaker one. Typically, one expects in noncentrosymmetric semiconductors the surface nonlinearity to be weaker, so that I_2^{\max} corresponds to the surface-induced SH field. Now one can state the following condition to quantify the detectable range of the surface nonlinearities:

$$\frac{I_1^{\max}}{I_2^{\max}} = \frac{D}{\kappa}, \quad (4.28)$$

where D is again the dynamic range of the measurement system and κ is a constant. Eq. (4.28) is to be separately evaluated for the far-field SHG intensity patterns measured in the x_c - and y_c -polarizations, what would give accordingly the upper and the lower boundaries of the detectable range of η .

The values of η obtained from the relation Eq. (4.28) for $\kappa = 1$ are plotted in Fig. 4.12 for both polarizations. The curves in Fig. 4.12 were calculated numerically using the simulated far-field SHG intensity patterns for the bulk- and surface-induced SH fields shown in Fig. 4.8-4.11. For each polarization the lower limit η_{lower} implies that I_2^{\max} corresponds to the surface-driven lobe, and the upper limit η_{upper} implies that I_2^{\max} corresponds to the bulk-driven lobe. If, for example, the dynamic range $D = 100$ is fixed, then for the lower detectable limit η_{lower} in x_c -polarization one gets:

$$\eta_{\text{lower}} \approx 0.01.$$

Calculating in exactly the same way the upper limit for η , one obtains the following range, where the

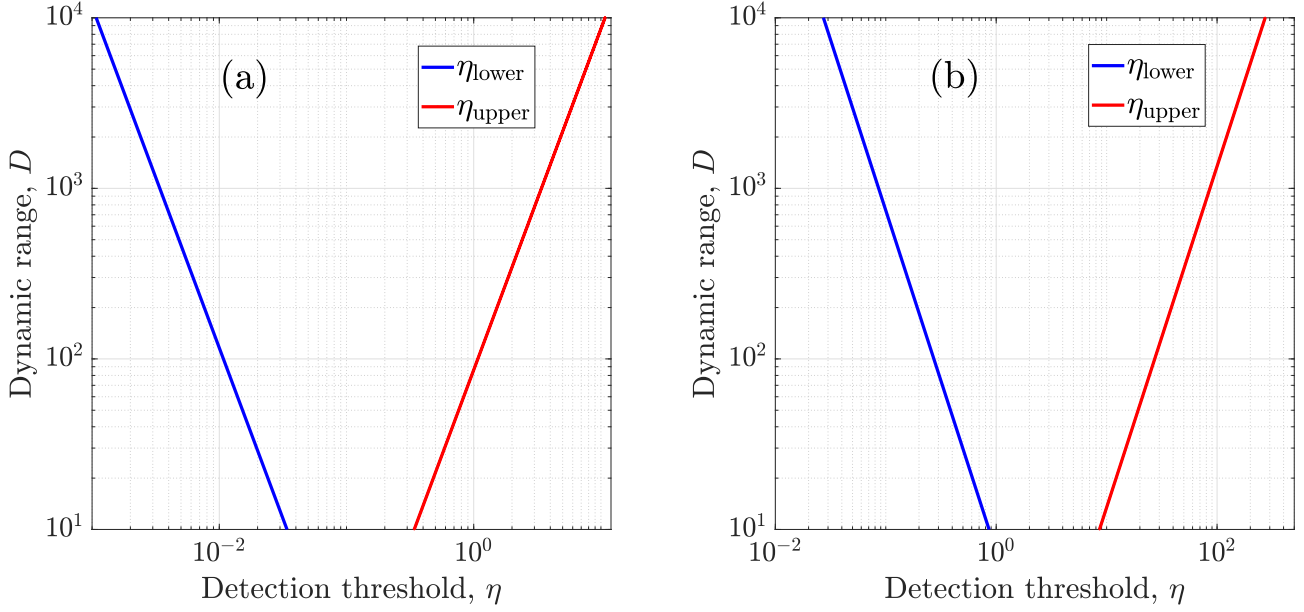


Figure 4.12: The diagrams, showing the upper and lower detection thresholds of η Eq. (4.28) vs. the dynamic range of the camera for $\kappa = 1$ for x_c -polarization (a) and y_c -polarization (b).

surface nonlinearity can be quantified from the x_c -polarized far-field SHG pattern:

$$0.01 \lesssim \eta \lesssim 1 .$$

Performing analogous calculations for the y_c -polarized far-field distribution, one can find the following range of the detectable η values:

$$0.1 \lesssim \eta \lesssim 10 .$$

Remarkably, the detectable ranges of η values largely differ in both polarizations and just slightly overlap. According to Fig. 4.12 this conclusion holds for all considered values of the dynamic range D . Hence, the far-field signatures of the surface nonlinearity have to be predominantly manifested in one polarization only depending on the value of η in a specific semiconductor. For $D = 100$ one can estimate the detectable range of the surface-to-bulk ratios, as expressed through the parameter η , from the polarization-resolved measurements to cover:

$$0.01 \lesssim \eta \lesssim 10 , \tag{4.29}$$

i.e. more than 3 orders of magnitude of η values. The specific upper and lower boundaries of the detectable values of η range would depend on the specific material, namely on the refractive indices of the slab both at the fundamental and the SH frequencies. However, this range is expected to always span more than 3 orders of magnitude. It is notable that the range Eq. (4.29) closely matches the range found in the previous section for the polarization-resolved rotation-angle scan under the plane-wave illumination. The main advantage of the TFGB illumination can be attributed to its relative simplicity, because one does not need to rotate the crystal in order to measure the angular dependence of the emitted SHW. Instead a single measurement of the far-field SHG radiation pattern in different

polarizations without any further movements in the setup would suffice to quantify the ratio of the surface and bulk nonlinearities.

It seems reasonable to compare the obtained range Eq. (4.29) of the detectable ratios of the nonlinearities with values obtained in the literature. Values of the bulk nonlinear coefficients $\chi_{\text{bulk}}^{(2)}$ in noncentrosymmetric semiconductor materials are usually of the order of 10^{-12} m/V, though for III-V semiconductors it achieves $\sim 10^{-10}$ m/V [46]. The strength of the surface nonlinearity χ_{zzz}^{2D} in different solids varies quite strongly [149, 129, 94, 87, 38, 39], mainly attaining values in the range $\sim 10^{-20} - 10^{-17}$ m²/V. For the parameter η from Eq. (4.7) this would give the range $0.01 \lesssim \eta \lesssim 10$ for $\chi_{\text{bulk}}^{(2)} \sim 10^{-12}$ m/V and respectively smaller values for larger $\chi_{\text{bulk}}^{(2)}$. Along with that, experimental measurements in [87] for III-V semiconductor, namely GaP nanopillars, yielded in current notation $\eta \sim 0.1$, what allows to expect η in other III-V semiconductors to also fall within the same limits. One can therefore see that the expected values of η in different noncentrosymmetric media correlate well with the detectable range in Eq. (4.29).

It would be interesting to discuss the ways to extend the detectable range given by Eq. (4.29). The most natural approach is to get larger longitudinal component of the electric field, provided by more complex illuminating beams. Stronger longitudinal field would increase the surface contribution to the total far-field SHG pattern, because the surface nonlinear polarization scales quadratically with the E_z -component of the illuminating beam due to the dominating $\chi_{zzz}^{(2)}$ tensor element, while the bulk nonlinear polarization scales linearly, since all entries of the bulk nonlinear tensor with more than one z -index are zero. The longitudinal component of the electric field can be enlarged by the oblique incidence of the TFGB, but the measurement of the SH radiation with such arrangement would become more complicated. Alternatively, a number of more complex beams have been shown to possess the strong longitudinal electric field, including optical needles [150, 151], spatially phase-shaped beams [152], radially-polarized tightly-focused beams [153, 154, 155], polarization vortices [156], higher-order cylindrical beams [157] or photonic nanojets [158]. With these alternative illuminating beams the boundaries of the detectable surface nonlinearity can be stretched even further.

4.5 Summary

In nanostructures and ultrathin films made of noncentrosymmetric semiconductors the largely increased surface-to-volume ratio makes it important to account for the surface nonlinearity, despite the strong bulk second-order nonlinearity. Besides, the surface signal can be additionally gained in certain wavelength ranges, for instance, in the vicinity of surface resonances or for the SH frequencies above the fundamental bandgap. Therefore the separation and the quantitative comparison of the surface and bulk contributions to the SHG appears to be the relevant problem for the efficient frequency conversion.

The findings demonstrate that the different symmetry properties of the bulk and surface nonlinear tensors in noncentrosymmetric semiconductors result in specific signatures in the polarization-resolved far-field SHG pattern. Under the plane-wave illumination of a slab of a noncentrosymmetric semiconductor the differences arise in the angular dependences of the emitted SHW on the angle of

incidence and the polarization angle. It is shown that the surface and bulk contributions can be unambiguously separated and quantitatively compared from the polarization-resolved angular scans of SHG within a wide range of surface-to-bulk ratios, which spans over at least three orders of magnitude.

Using a normally-incident linearly-polarized TFGB as the pump field, the distinct features in the far-field SHG intensity patterns were also obtained. The polarization-resolved analysis of the SHG radiation patterns thus enables to reliably identify, separate and compare the strength of both nonlinearities. With the setup with an illuminating TFGB the surface optical nonlinearity can be evaluated in a wide range of its strengths from just a couple of the polarization-resolved far-field SHG radiation patterns. The applicability of this approach could be further boosted by devising special configurations of the illumination field allowing to selectively increase the surface response in the total SHG field.

Obtained findings are directly applicable to other III-V semiconductor materials that belong to the $\bar{4}3m$ crystal symmetry group, e.g., GaAs, GaP, InSb, InAs, AlAs, or InP. Besides that, the extension to other noncentrosymmetric media is also straightforward and can be done with due regard for the specific type of the bulk crystal symmetry. One can believe these results to provide an appreciable contribution to gaining the SHG conversion efficiency from the nanoscale semiconductor structures.

In the next chapter, I aim to numerically investigate another interesting computational problem, namely the SSHG from multilayer structures of different possible compositions. The TMM developed earlier in chapter 2 is to be used for the simulations. The main issue to be solved is finding the optimal geometry of the multilayer structures for the efficient SSHG.

Chapter 5

Optimization of surface SHG from multilayer structures

The enhancement of the SHG from nanooptic components can be achieved in several ways, but mainly based on the excitation of resonant modes in the nanostructures. However, one can also think of using the surface SHG to achieve efficient frequency conversion, especially when multiple surfaces are brought together. Then the constructive interference of the surface-driven second-harmonic waves from multiple interfaces can serve as an efficient way of boosting the SHG conversion efficiency from the nanoscale optical elements, especially made of centrosymmetric materials. The most efficient enhancement of SHG should take place when the spatial separation between interfaces is much smaller than the wavelength. Then the phase shift between the SHWs originating from the sources of surface nonlinear polarization can be neglected and the SHWs interfere in-phase. As the distance between the interfaces is increased, new effects come into play. The phase shifts now become important and the interference of the emitted SHWs starts to strongly depend on the separation between layers. Still certain resonances can also appear for non-negligible separation between interfaces when the constructive interference of a number of sources is reached.

In Chapter 2, a transfer-matrix-based method was developed for the efficient modeling of SSHG from arbitrary multilayer structures. Moreover, with an exemplary stack the possibility to obtain strong enough SH signal was demonstrated. It was even found that an exemplary structure consisting of 5 layers with total thickness of 10 nm yields stronger SH response than a GaAs layer with bulk nonlinearity and larger thickness. In this chapter I aim to investigate in more detail the SHG enhancement in multilayer structures of different compositions. In particular, periodic stacks of nonlinear materials with several layers in each period can exhibit interesting behaviour. It is worth noting that existing deposition techniques nowadays, like atomic layer deposition or molecular beam epitaxy, make it possible to grow layered structures of largely variable composition and required layer thicknesses with high precision down to just a few nanometers [69, 70, 71].

The enhancement of SHG in layered structures has been already studied before. However, as far as I know only bulk nonlinearities have been considered in this respect. It was shown, that the phase-matching condition plays an important role in this case. Besides that, SHG is additionally enhanced if the fundamental frequency and SH frequency are resonant with the band-edge states in the finite

periodic structure [159, 160, 161, 162, 163]. Combining the field enhancement at the fundamental frequency and SH frequency results in a scaling of the SHG conversion efficiency as $\sim N^6$ [161] with the number of periods N . If additionally the PMC is fulfilled, scaling up to $\sim N^8$ can be achieved [164]. Moreover, $\sim N^8$ scaling of SHG was even obtained in properly designed Bragg multilayers without PMC [165]. An additional way for enhancing SHG can be provided by the optimal modulation of the nonlinear optical coefficient, e.g., through modulated polarization directions in a PhC consisting of ferroelectric domains [166, 167]. Along with that, the strong enhancement of the SHG conversion was also predicted in planar microcavities with periodic dielectric mirrors, when double resonances for the pump and harmonic fields, as well as phase matching, are achieved [168, 169].

In this chapter the treatment is limited to the surface second-order nonlinearity only. For centrosymmetric optical materials the bulk nonlinearity can be safely neglected in nanophotonic application. As stated in Chapter 2, this reasoning can be also applied to layers made of noncentrosymmetric semiconductors provided that the thicknesses of the layers are small enough (below ~ 100 nm). In the following, centrosymmetric semiconductor materials will be mainly used for simulations. Since only media with bulk nonlinearity can support the propagation of inhomogeneous waves at the second-harmonic frequency and thus the efficient energy transfer [2], the standard PMC plays an essential role in such media only. Given that the bulk nonlinearity is not taken into account in this chapter, the PMC ceases to play a significant role in the considered layout. Therefore, the scaling laws found for layered media with bulk nonlinearities are not fulfilled anymore. With the surface nonlinearities the outcome of the interference of the multiple interfaces would rely on the strengths of the surface nonlinear polarization at different interfaces as well as on the spatial separation between them.

This chapter is organized as follows. In Section 5.1 several exemplary nonperiodic stacks are considered and the SHG enhancement is searched for with varying the layer thicknesses and the incidence angle. Section 5.2 deals with a periodic stack consisting of 2 layers in each period and the optimization of the SHG vs. the number of periods and thicknesses of both layers in the period is performed. The generalization to the periodic stack with 3 layers in each period is done in Section 5.3. The discussion of the main findings is given in the concluding Section 5.4.

5.1 Nonperiodic stacks

The expected constructive interference of the surface SHWs from multiple interfaces should in general arise regardless of the composition of the stack. The total emitted field would depend on both the strengths of the surface nonlinearities and the linear transmission properties of the stack both at the fundamental and the second-harmonic frequency. The periodicity of the structure can also bring new properties due to the presence of the photonic band gaps and corresponding transmission dips. Therefore the dependence of the SSHG from the multilayer structures deserves detailed investigation.

In order to clarify the role of the top interfaces on the total response let us begin with the simplest case of a single flat layer on a substrate. A silicon layer located on the top of a glass (BK7) substrate is selected. Such choice of materials provides high contrast of the refractive indices between the silicon layer and both the air above and the substrate below. For simulations, the transfer-matrix

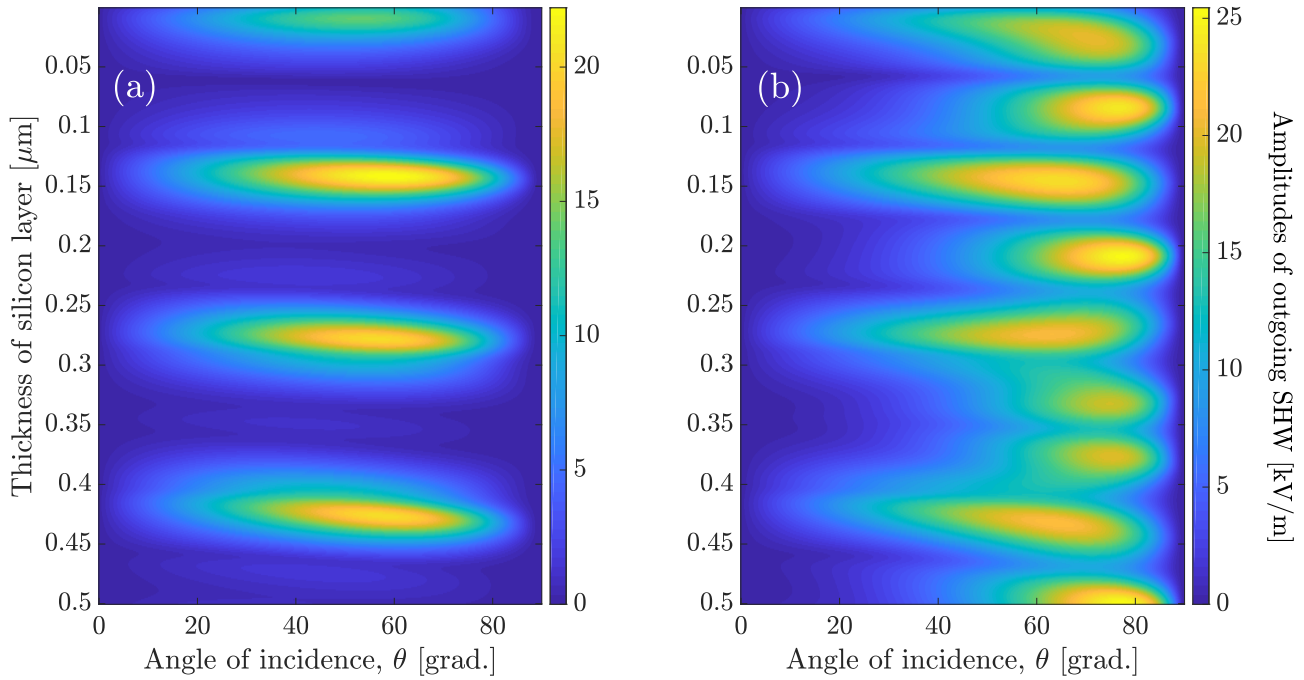


Figure 5.1: Amplitude of the upwards outgoing SHW from a silicon layer on a glass (BK7) substrate [in kV/m] for: (a) TE-polarized; (b) TM-polarized incident wave at fundamental frequency vs. the angle of incidence θ and the thicknesses of the layer. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

method described in Chapter 2 is used. The glass substrate is again assumed semi-infinite and the thickness of the silicon layer is varied. The fundamental wavelength of the illuminating plane wave is $1 \mu\text{m}$, so that the second-harmonic wavelength is 500 nm . The intensity of the incident FW was again taken to 1.33 GW/cm^2 , so that the electric field strength in the FW is $E^{\text{FW}} = 10^8$ V/m. The simulation results are presented in Fig. 5.1 for TE- and TM-polarized incident plane wave at the fundamental frequency. Here, the amplitude of the upwards outgoing plane wave at the second-harmonic frequency is plotted vs. the angle of incidence and the thickness of the silicon layer. For arbitrary thickness one gets a single maximum in the angular dependence with zero field at $\theta = 0^\circ$ and $\theta = 90^\circ$. Regarding the layer thickness one can intuitively expect the appearance of alternating maxima and minima because of the varying phase shift leading to constructive or destructive interference between two sources of the surface nonlinear polarization. Similar dependence is indeed obtained along the vertical axis in Fig. 5.1. The differences between the dependences for TM- and TE-polarized incident waves at the fundamental frequency seem to be related to the fact that for TE-polarization only one term in the expression for the surface nonlinear polarization Eq. (2.27) given by the tensor element $\chi_{\perp||\parallel}^{(2)}$ contributes to SSHG. On the other hand, for TM-polarization, all three terms in Eq. (2.27) are contributing, leading to a more complex interference pattern.

As the next step, let us move to a more complex multilayer example. Namely, the multilayer stack from section 2.7 is taken, the corresponding parameters for convenience are provided again in Table 5.1. Now one tries to numerically optimize the SHG efficiency by changing the thicknesses of the layers. In order to do it, it is assumed that the thicknesses of all layers from Table 5.1 are multiplied by an additional factor α , which is a fixed constant for all layers. The amplitude of the upwards outgoing SHW from the considered stack vs. the angle of incidence θ and the extra factor α is plotted in Fig. 5.2. The case of $\alpha = 1$ in Fig. 5.2 then matches the plots in Fig. 2.2. Within the

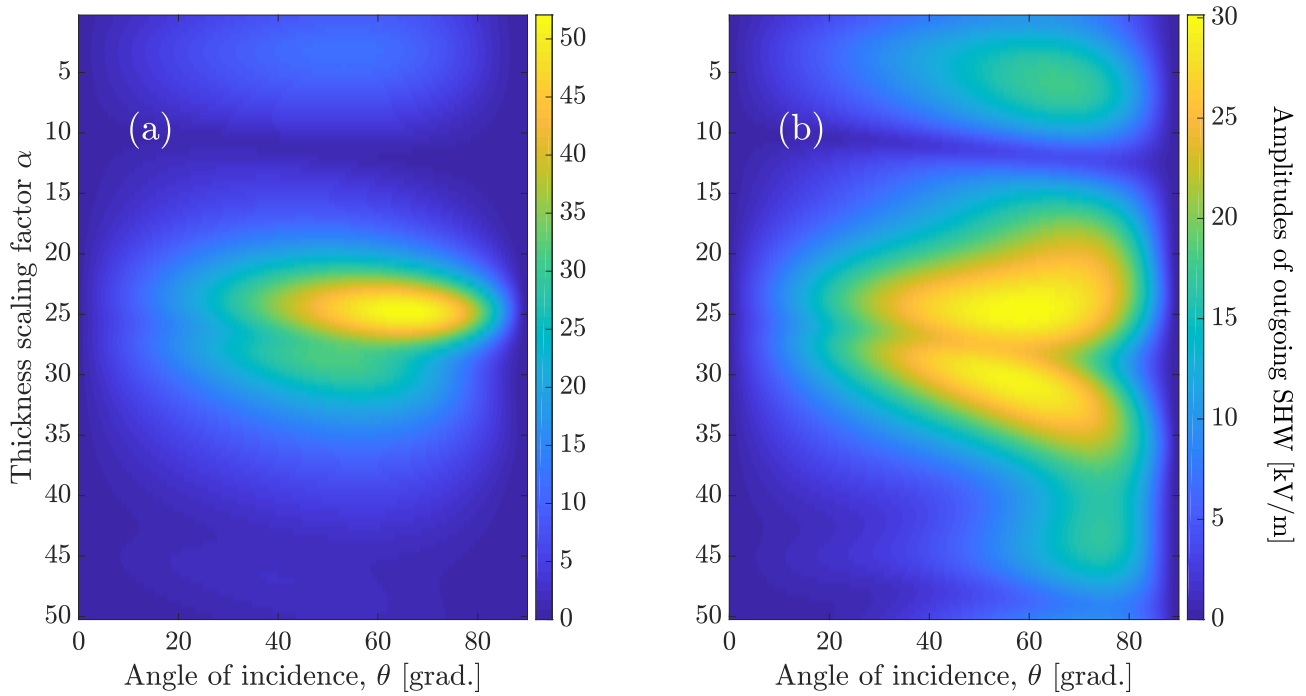


Figure 5.2: Amplitude of the upwards outgoing SHW from the considered stack from Table 5.1 [in kV/m] for: (a) TE-polarized; (b) TM-polarized incident wave at the fundamental frequency (respective wavelength is $1 \mu m$) vs. the angle of incidence θ and the scaling factor α of the thicknesses of the layers in the stack. The amplitude of the FW (pump) $E^{FW} = 10^8$ V/m.

considered range of the values of α one gets a few pronounced maxima here. One can notice that the achievable strength of the SH field is larger than the one from a single silicon layer in Fig. 5.1 due to the larger number of interfaces.

Layer, №	1	2	3	4	5	Substrate
Material	Si	SiO ₂	TiO ₂	Si	SiO ₂	BK7
Thickness, nm	1	3	2	1	3	-

Table 5.1: The composition of the considered nonperiodic multilayer stack.

It would be important to relate the location of the maxima in Fig. 5.2 with the linear optical properties of the stack both at the fundamental and second-harmonic frequencies. Therefore in Figs. 5.3 and 5.4 the reflectivities of the stack for both TE- and TM-polarized incident waves at both frequencies are plotted. It should be noted that Fig. 5.3,b shows the reflectivity for TM-polarization at the SH frequency, since in the considered example for an incident TE-polarized FW one gets only TM-polarized SHW. Even though the maxima in Fig. 5.2 and minima in Figs. 5.3 and 5.4 are slightly shifted with respect to each other, one can still notice the correlation between the SHG from the considered stack and the reflection properties of the stack. Therefore, one can conclude that the appearance of the maxima in Fig. 5.2 is largely related to the linear reflection/transmission of the stack both at the fundamental and at the SH frequencies. Specifically, one can see that the maxima of the emitted SHW for both polarizations of the incident wave in Fig. 5.2 are located near the positions, where the stack has dips in reflection at both frequencies simultaneously. Certain discrepancies between the locations of the maxima in the diagrams result from the complex interplay of the surface

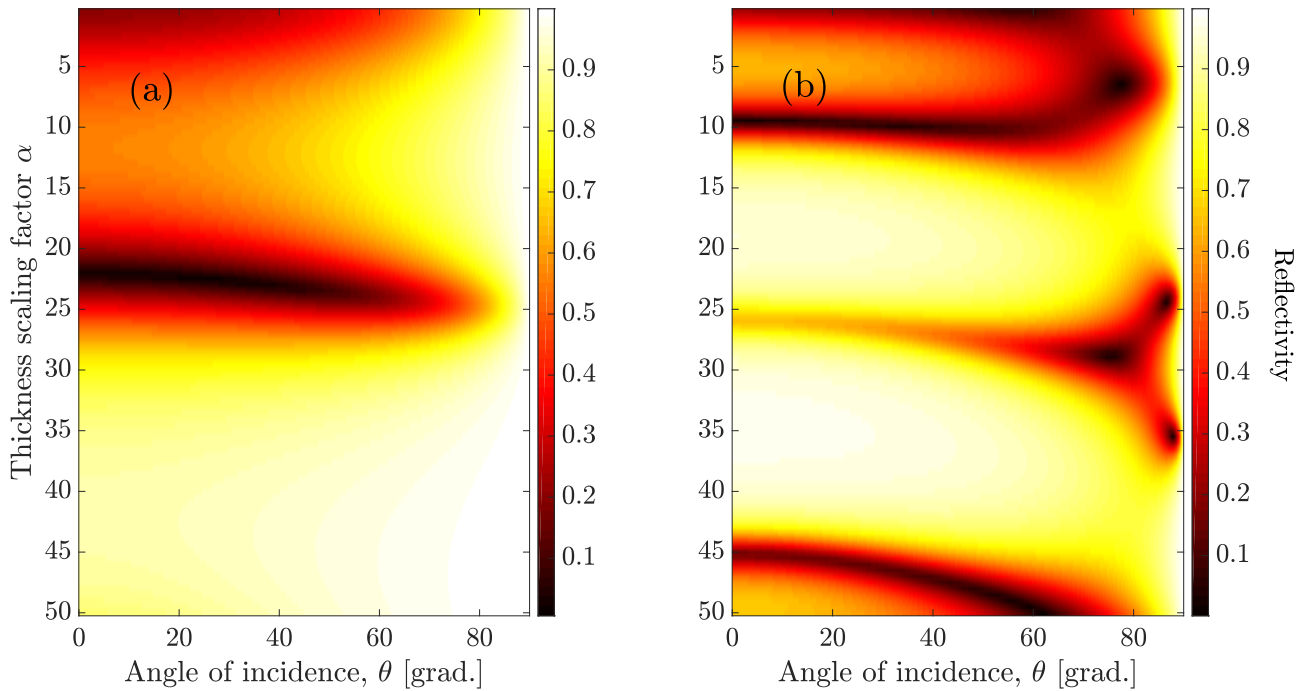


Figure 5.3: Reflectivity of the multilayer stack from Table 5.1 for: (a) TE-polarization at the fundamental frequency (respective wavelength is $1 \mu\text{m}$); (b) TM-polarization at the SH frequency (respective wavelength is 500 nm) vs. the angle of incidence θ and the scaling factor α of the thicknesses of the layers in the stack.

nonlinear polarizations at different interfaces inside the stack, which is not directly related with the maxima in the reflection/transmission.

5.2 Periodic stack of $\text{SiO}_2/\text{TiO}_2$ layer pairs

In contrast to the stacks of an arbitrary composition, the periodic ones are known to exhibit multiple resonances and can be thus expected to exhibit better enhancement of the surface SHG as well. Let us start with considering periodic stacks consisting of two layers in each period. As the material pair $\text{SiO}_2/\text{TiO}_2$ is selected in order to provide high enough refractive index contrast. Specifically at the pump wavelength $1 \mu\text{m}$ the refractive index of TiO_2 is 2.49, while the refractive index of SiO_2 is 1.45. Also in contrast to silicon, which absorbs strongly at the wavelengths below 400 nm , SiO_2 and TiO_2 are transparent in the whole optical range. Besides that, SiO_2 and TiO_2 are widely used materials in photonic technologies. The stack is assumed to be placed on the top of a SiO_2 substrate for convenience in fabrication. In the following, the SSHG from a multilayer periodic stack of alternating $\text{SiO}_2/\text{TiO}_2$ layers is investigated by varying the number of layers and their thicknesses. Therefore both the layers of equal thickness in each period and different ratios of the thicknesses of SiO_2 - and TiO_2 -layers are considered.

Firstly, the case of equal thicknesses of both layers is taken. Fig. 5.5 shows the amplitude of the upwards outgoing SHW from a stack of 2 periods vs. the angle of incidence and the thickness of the layers for TE- and TM-polarizations of the pump plane wave (FW). One can see that the dependence on the angle of incidence exhibits a single well-pronounced maximum in both polarizations. It is interesting to note that the angular position of this maximum does not noticeably change when varying

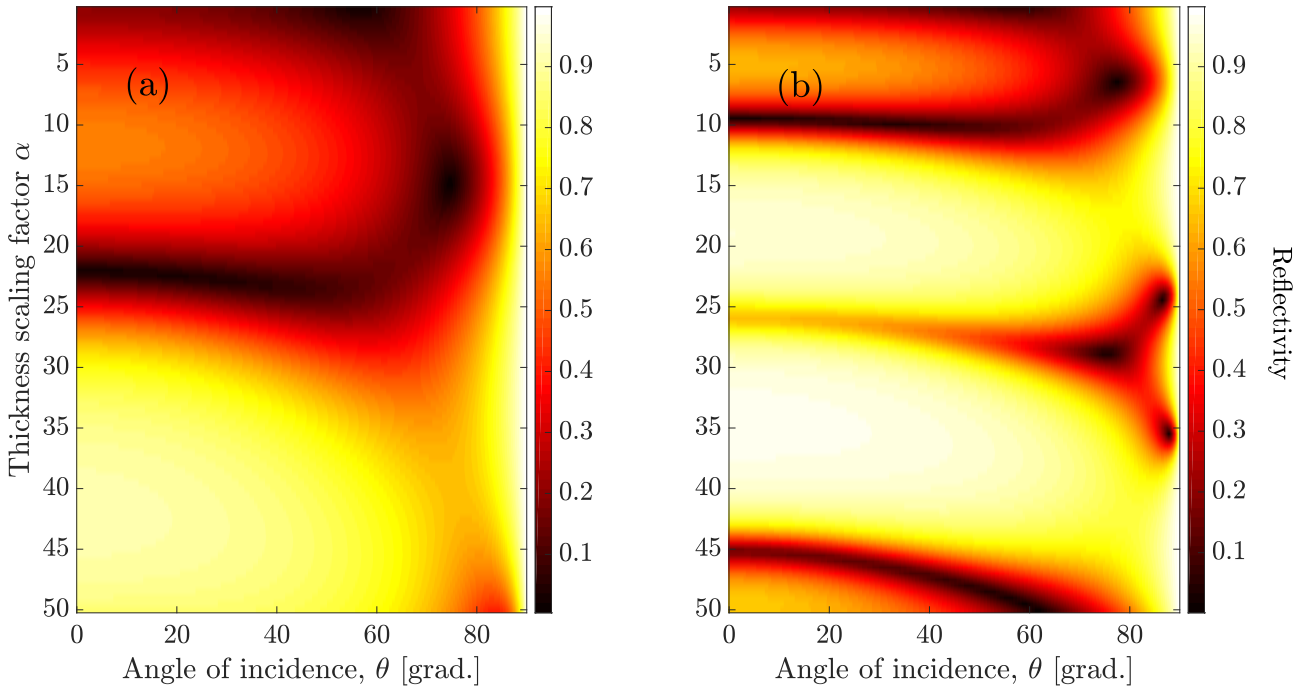


Figure 5.4: Reflectivity of the multilayer stack from Table 5.1 for TM-polarization at the: (a) fundamental frequency (respective wavelength is $1 \mu\text{m}$); (b) SH frequency (respective wavelength is 500 nm) vs. the angle of incidence θ and the scaling factor α of the thicknesses of the layers in the stack.

the thickness of the layers in wide limits. The dependence of the SHW on the thicknesses turns out to be much more complicated. One can observe here a non-monotonous function with alternating maxima and minima. Importantly, the maximal SHW amplitude for TM-polarized FW is almost one order of magnitude larger than for TE-polarized (i.e. the ratio of the corresponding intensities is around 100). The largest maximum in Fig. 5.5,b gives the SHW amplitude $E^{\text{SHG}} \approx 150 \text{ kV/m}$, and with the amplitude of the FW $E^{\text{FW}} = 10^8 \text{ V/m}$ (i.e. the pump intensity is 1.33 GW/cm^2) one obtains the energy conversion efficiency CE_{SHG} :

$$CE_{\text{SHG}} = \left| \frac{E^{\text{SHG}}}{E^{\text{FW}}} \right|^2 \approx 2.25 \cdot 10^{-6}.$$

Such values exceed by several orders of magnitude the typical values of the conversion efficiency in plasmonic and in all-dielectric nanostructures made of centrosymmetric semiconductors.

As it was shown in the previous section, the dependence of the SHW on the thicknesses can be to a large extent connected to the usual resonances of the multilayer structure at both the fundamental and the second-harmonic frequencies. The radiated field, however, results not only from the transmission properties of the stack at both frequencies, but also from the relative values and signs of the surface nonlinear polarization sources at each interface. The latter depends crucially on the nonlinear optical properties of the respective material pair. Specifically when the period of the stack consists of just two layers, one has the same pair of adjacent media at every interface inside the stack, but with alternating relative location of both materials. As it was stated in Chapter 2, in the simulations the vector of the surface nonlinear polarization is assumed to be always directed from the medium with weaker second-order nonlinearity into the medium with stronger second-order nonlinearity. Therefore in the case of

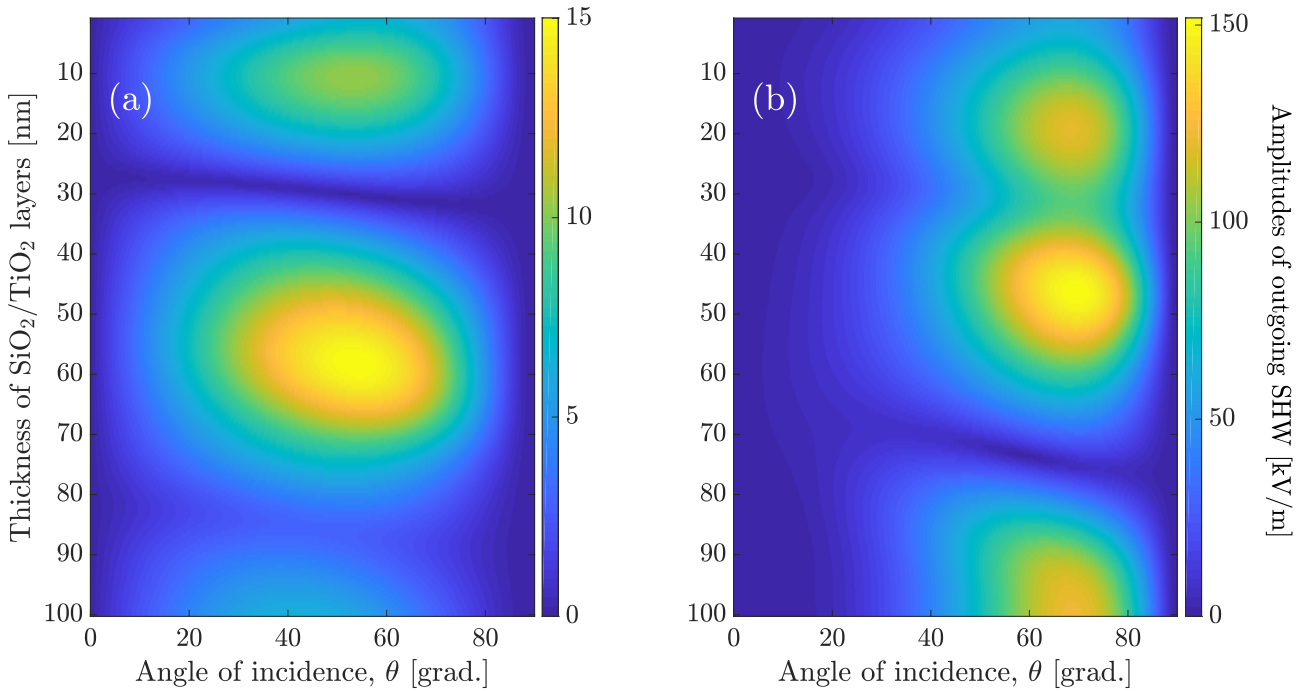


Figure 5.5: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 2 periods of SiO₂/TiO₂ layer pairs, $h_1 = h_2$: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

a material pair one can conclude that the surface nonlinear polarization at neighbouring interfaces has opposite directions. At the same time, the relative values of the components of the vector of the surface nonlinear polarization stay in general unknown and would vary for different number of periods in the stack.

Let us increase now the number of periods in order to figure out their role on the simulation findings. Fig. 5.6 demonstrates the amplitude of the upwards outgoing SHW from the stacks of 5 periods vs. the angle of incidence and the thickness of the layers. One sees that the dependence on the angle of incidence does not change as compared to the case of 2 periods in Fig. 5.5. At the same time the dependence of SHW on the thickness yields now more maxima and minima. From comparison of Figs. 5.5 and 5.6 one can infer that with increasing the number of periods the maxima split into several ones with corresponding minima in between. The maximal amplitudes for TM-polarization increases just slightly when going from 2 periods to 5 periods. For TE-polarization significant increase of the maximal SHW amplitude is found when going from 2 periods to 5 periods and negligible further growth for larger number of periods. Increasing the number of layers even further leads to further splitting and shrinking of maxima in the diagram due to interference effects. This is illustrated by the blue curves in Fig. 5.7, which show the maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) from the stack of 10 periods with equal layer thicknesses.

It is important to address the question of the experimental verification of the simulation results. Furthermore, it would be useful for applications to be able to determine the components of the surface nonlinear tensor from the measurements of the emitted SHW from the stack. From Figs. 5.5, 5.6 and 5.7 one can see that experimental measurements of SSHG from periodic multilayer structures would be most suitable for the number of layers in the range $N \approx 5 - 10$. For smaller number of

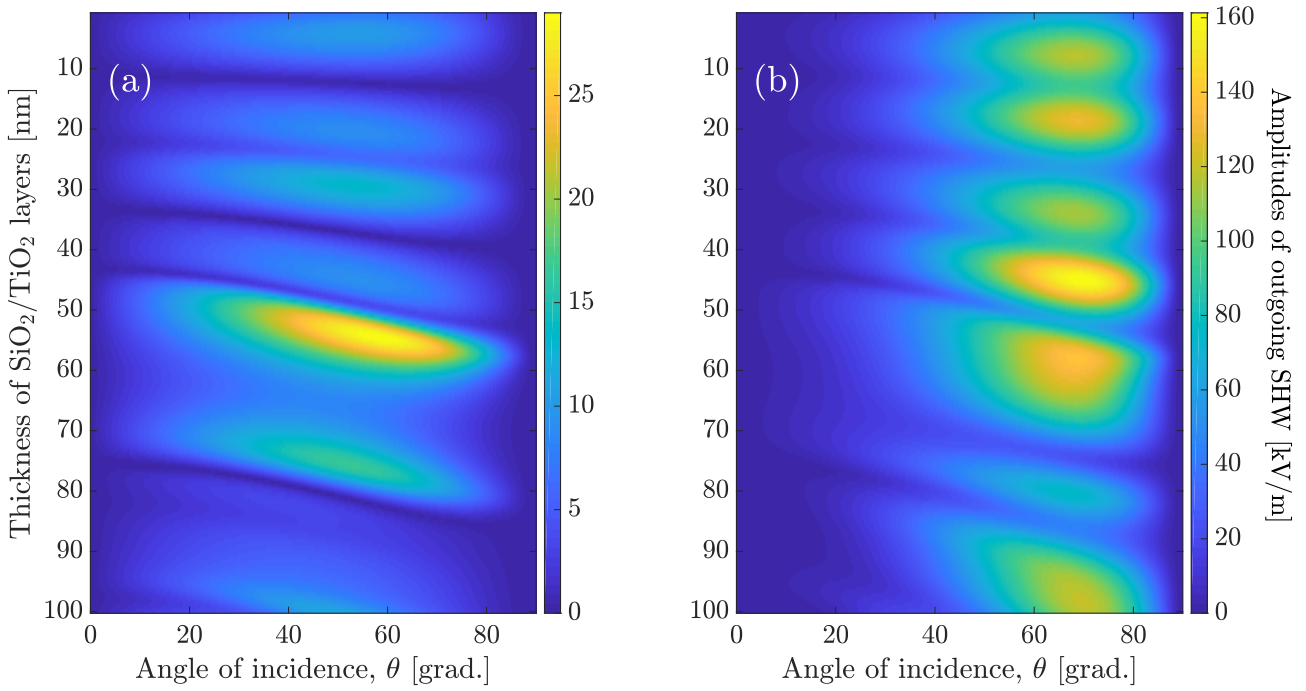


Figure 5.6: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 5 periods of SiO₂/TiO₂ layer pairs, $h_1 = h_2$: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

layers the maximal SHW amplitude can be significantly smaller. On the other hand, for larger N the emitted SHW becomes much more sensitive to inaccuracies of the thickness measurements (due to the large number of maxima in Fig. 5.7) with no significant increase in the SHW amplitude due to the interference effects. Therefore the optimal trade-off is achieved at medium values of the number of layers.

In order to obtain the nonlinear tensor from the measured SH field, one would need to perform a fit of simulated data to the experimental ones. If the period of the multilayer structure consists of 2 layers, one gets in general 6 unknowns, namely the tensor entries $\chi_{\perp\perp\perp}^{(2)}, \chi_{\perp\parallel\parallel}^{(2)}, \chi_{\parallel\perp\parallel}^{(2)}$ for material interfaces SiO₂/TiO₂ and SiO₂/Air. It is assumed here that the interfaces are isotropic, since all media are centrosymmetric. If one measures, for instance, the dependence of the outgoing SH field on the angle of incidence for several fabricated structures with different thicknesses and/or number of layers, the numerical fit must be obtained to the measured curves with 6 unknown parameters. The problem can be greatly simplified, if the surface nonlinear tensor for one of the material pair, e.g. SiO₂/Air, is already known from the literature. In this case the fit to the experimental data with just 3 unknown tensor entries for SiO₂/TiO₂ interface has to be performed.

Secondly, I proceed with the case of different thicknesses of both layers. The thickness of the top layer is fixed now and the diagram for 10 periods is calculated for varying thickness of the bottom TiO₂ layer. Fig. 5.8 corresponds to the fixed thicknesses $h_1 = 2$ nm. Similar figures were also obtained for other values, e.g. $h_1 = 1$ nm and $h_1 = 3$ nm (not shown), and showed no noticeable differences. Hence for such small thicknesses the value h_1 has just minor influence on the SHW and the locations of maxima on the diagram. Next, a fixed ratio between the thicknesses h_1 and h_2 is examined. In Fig. 5.7 the plots for the several values of this ratio are put together, namely $h_2 = 2h_1$, $h_2 = h_1$ and $h_1 = 2h_2$. I have checked that the angular dependence of the SHW amplitude in all these cases

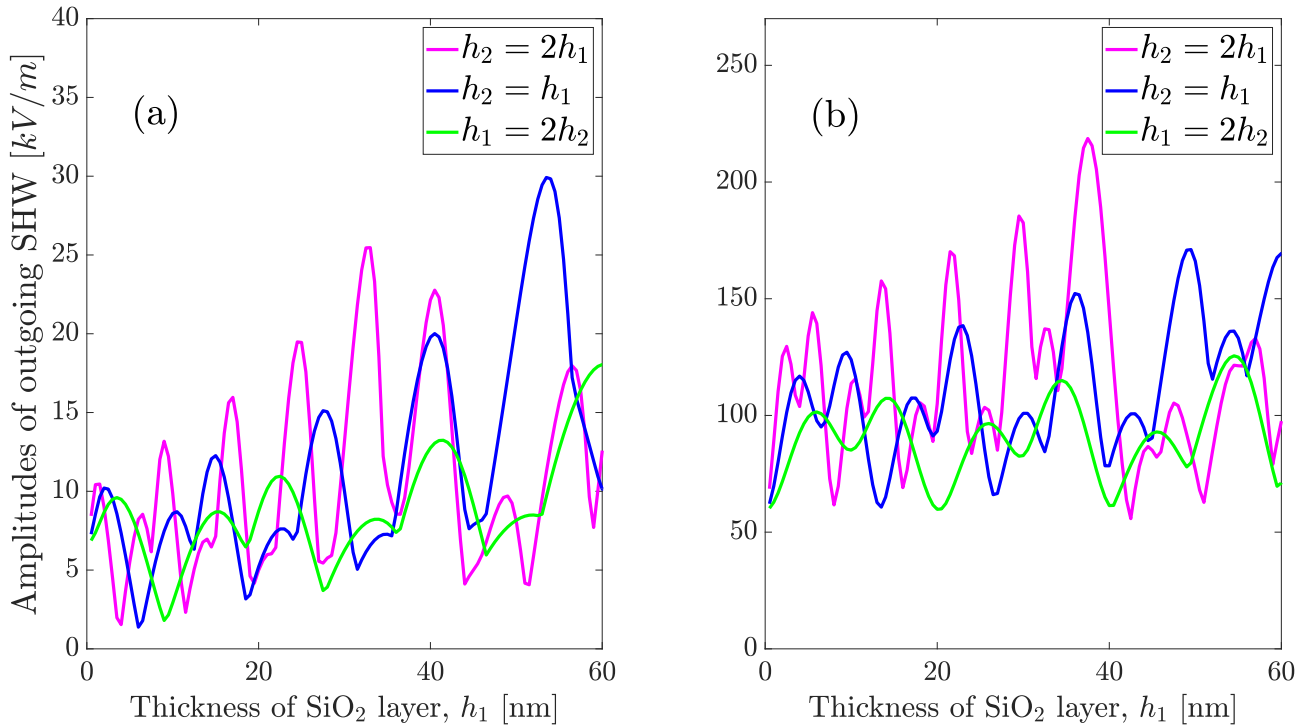


Figure 5.7: Maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) from the stack of 10 periods of SiO₂/TiO₂ layer pairs for different ratios of the thicknesses: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

exhibits a single maximum and therefore plotted the maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) vs. the thickness of the first layer h_1 . When $h_2 = 2h_1$ a certain increase in the maximal SHW amplitude for TM-polarization with respect to the case of equal thicknesses of both layers is found here. In the case $h_1 = 2h_2$ one does not observe any improvement both for TM-polarized FW and for TE-polarized FW as compared to the equal layer thicknesses.

Having checked that for different ratios of thicknesses one gets a single isolated maximum in the angular dependence, whose location changes insignificantly with the thicknesses, I now aim to generalize the findings for the different thicknesses of both layers by plotting a two-dimensional diagram of the maximal SH amplitude among all values of the angle of incidence vs. the thicknesses h_1 and h_2 . This diagram is shown in Fig. 5.9. One sees a number of "fringes", corresponding to the maxima of the SH field. For both polarizations of the FW one gets a pronounced strongest peak on the diagram, with much stronger one for TM-polarized FW in Fig. 5.9,b. This largest maximum in Fig. 5.9,b amounts to the SHW amplitude $E^{\text{SHG}} \approx 240$ kV/m, what gives for the best value of the energy conversion efficiency CE_{SHG} :

$$CE_{\text{SHG}} = \left| \frac{E^{\text{SHG}}}{E^{\text{FW}}} \right|^2 \approx 6 \cdot 10^{-6}.$$

Another interesting issue is the role of the layer ordering in the periodic stack. Specifically, one can consider a similar SiO₂/TiO₂ multilayer structure with a TiO₂ layer on the top. Fig. 5.10 shows the corresponding amplitude of the outgoing SHW from such stack composed of 9.5 periods, i.e. a TiO₂ layer plus 9 periods of SiO₂/TiO₂ layer pairs, placed on the top of the SiO₂ substrate. The

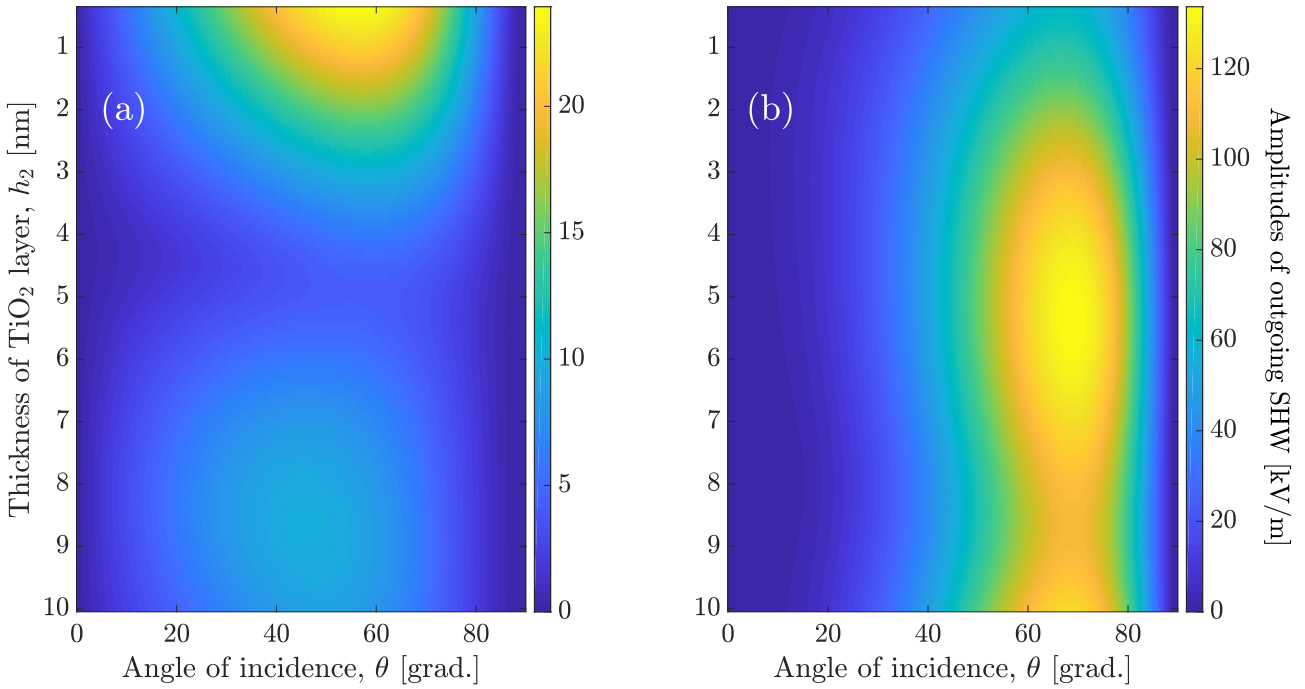


Figure 5.8: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 10 periods of SiO₂/TiO₂ layer pairs, $h_1 = 2$ nm (SiO₂): (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

thicknesses of both layers were taken equal. One can see again the maximal SHW amplitude for TM-polarized FW to be several times larger than for TE-polarized FW. Comparison with Fig. 5.7 yields smaller maximal SHW amplitude for the inverted case, therefore the usual arrangement of SiO₂/TiO₂ layers in Fig. 5.7 turns out to be more advantageous in this respect.

Finally, the interesting point would be the scaling of the second-harmonic radiation with the number of layers in the stack. This should clarify the effect of the increasing the number of interfaces as well as the effect of the top interface on the emitted SHW. The stack containing N periods of SiO₂/TiO₂ and located on the same SiO₂ substrate is taken. In Fig. 5.11,a the maximal amplitude of the outgoing SHW over the angle of incidence is depicted as the function of the number of periods N , when $h_1 = h_2 = 1$ nm. For ultrathin layers one can naturally expect that the emitted field would exhibit close to linear dependence on N , at least for moderate values of N until the total thickness of the stack becomes large enough. When the thickness of the whole stack gets comparable to the SH wavelength in the materials of the layer, more complex nonmonotonic behaviour is to be expected due to the interference of multiple reflected waves inside the stack both at the fundamental frequency and at the second-harmonic frequency. Fig. 5.11,a indeed confirms the monotonic increasing dependence for $N \lesssim 20$ for TE-polarized pump FW and until almost $N \lesssim 40$ for TM-polarized FW as well as the nonmonotonic functional dependence for larger N . Similar behaviour is observed in Fig. 5.11,b, where the same curves are depicted for different thicknesses $h_1 = 2$ nm, $h_2 = 1$ nm. The values of the SHW amplitude change slightly, but the linear dependence on N now breaks for smaller number of periods thanks to larger thickness of a single period of the stack. Hence, one can conclude that for ultrathin layers the outgoing SHW scales roughly linearly with the number of layers, where the constant term of this linear function corresponds to the contribution of the top SiO₂/Air interface.

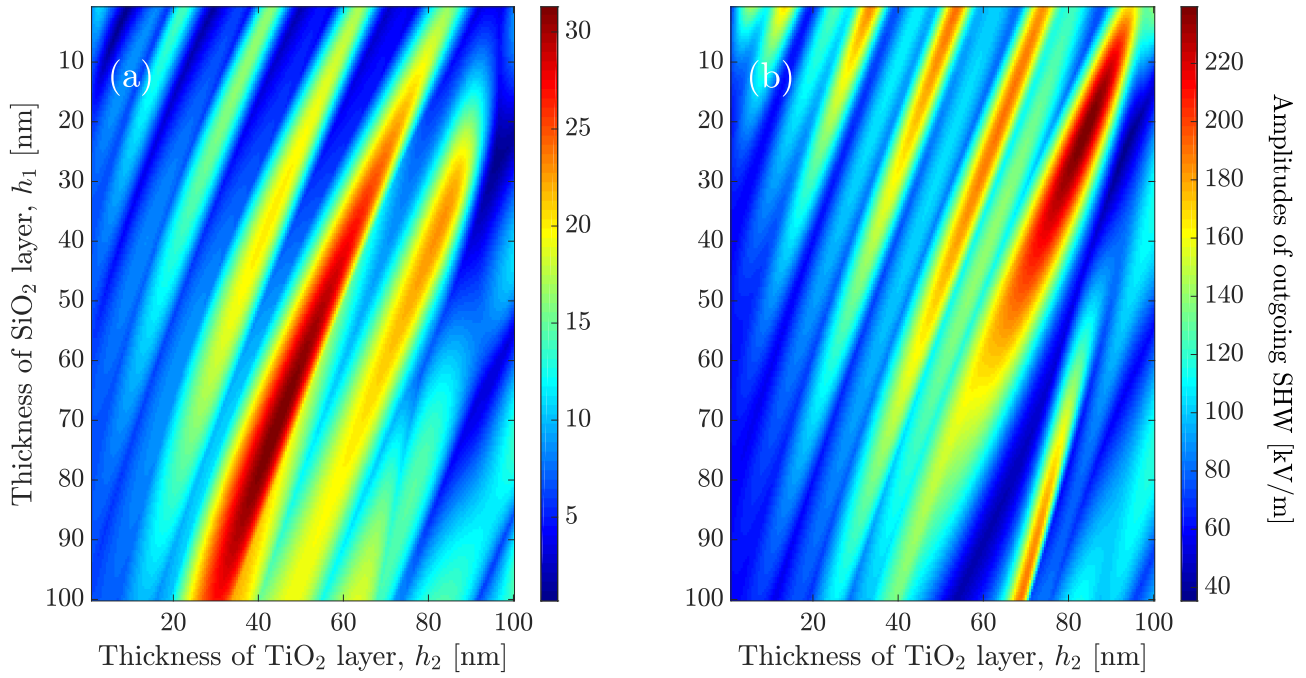


Figure 5.9: Maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) from the periodic stack of 10 periods of SiO₂/TiO₂ layer pairs [in kV/m] vs. the thicknesses h_1 , h_2 : (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

5.3 Periodic stack of SiO₂/Al₂O₃/TiO₂ layer triples

I proceed with considering periodic stacks of SiO₂/Al₂O₃/TiO₂ layer triples. A layer of Al₂O₃ was added here compared to the previous section, since its refractive index (1.67 at the pump wavelength 1 μm) falls in between the refractive indices of SiO₂ and TiO₂. Introducing an extra layer into each period would give an extra source of the surface nonlinear polarization interfering with the contributions of the rest interfaces. I will follow the treatment in the previous section here and consider both the layers of equal thickness in each period and different ratios between the thicknesses of the layers in each period.

Fig. 5.12 shows the amplitude of the upwards outgoing SHW from the stack of 2 periods vs. the angle of incidence and the thickness of the layers, when $h_1 = h_2 = h_3$. Similar to the periodic stack in the previous section, one gets a single maximum in the angular dependence and multiple maxima over the thickness of the layers. The angle of the maximal SH field for TM-polarized pump FW gets shifted to much smaller values. The obtained field amplitude in the maxima are significantly smaller than the ones obtained for SiO₂/TiO₂ stack with 2 periods at TM-polarization. Also the field amplitudes turn out to be comparable in their values for both polarizations of the FW.

The respective results for the increased number of periods $N = 5$ are depicted in Fig. 5.13. The field amplitude at the resonances gets almost doubled at both polarizations, however the positions of the maxima on the diagram change slightly. It should be noticed again that maxima split into several narrower ones when taking larger number of periods. The further increase of the number of periods leads to very slow growth of the maximal amplitudes of the diagram together with the ongoing splitting and shrinking of the maxima. It should be stated that for the selected 3 layers in the period the achieved SH field strengths for both polarizations of the pump plane wave have close values, so

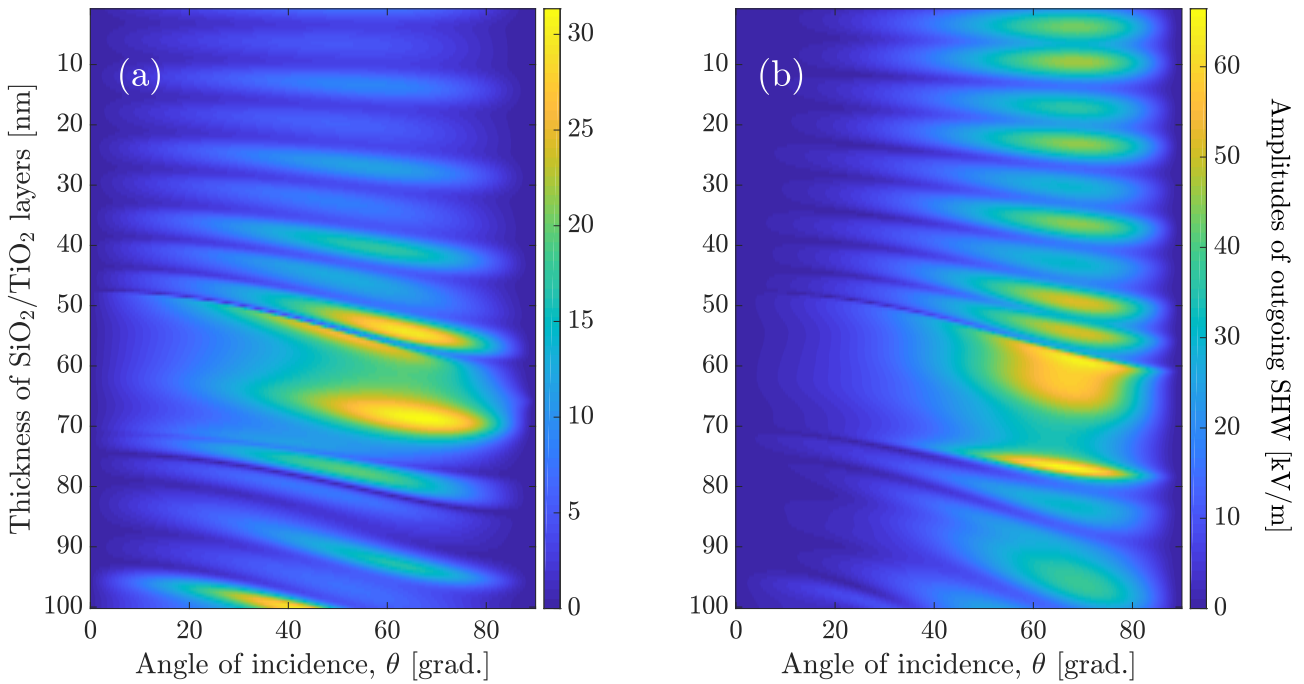


Figure 5.10: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 9.5 periods of SiO₂/TiO₂ layer pairs, i.e. with a TiO₂ layer on the top, $h_1 = h_2$: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

that no definite preferable polarization is found.

Fig. 5.14 shows the amplitude of the outgoing SHW from the stack of 10 periods for the fixed ratio of the thicknesses of SiO₂ layer h_1 , Al₂O₃ layer h_2 and TiO₂ layer h_3 , namely $h_1 = h_2 = 2h_3$. One sees that the obtained values of the SH field strength here are smaller than ones for the equal thicknesses above. Also one gets a noticeable difference in the SHW amplitudes for both polarizations of FW, and the maximal SH field for TM-polarized FW becomes much stronger. As before for the periodic stack of 2 layers in each period, I proceed to generalize the results to the case of different thicknesses of the layers by plotting a two-dimensional diagram of the maximal SH amplitude among all values of the angle of incidence vs. the thicknesses. Since there are three layers in every period now, the thickness of the SiO₂ layer h_1 is fixed and the diagram is plotted vs. the thicknesses h_2 and h_3 . An exemplary diagram is shown in Fig. 5.15 for the value $h_1 = 20$ nm. Similarly to the case of SiO₂/TiO₂ layers, one gets a picture of "fringes", since the generated SHW exhibits oscillatory-type dependence on each thickness separately. Again for both polarizations of the FW a definite largest maximum is seen, which is significantly stronger for TM-polarized FW in Fig. 5.15,b. Still the maximal SHW amplitude in Fig. 5.15,b turns out to be much smaller than the one for periodic SiO₂/TiO₂ layers in Fig. 5.9,b. The diagrams like in Fig. 5.15 were calculated for many other values of the thickness h_1 . In all cases the obtained diagrams look similar to Fig. 5.15 and the maximal achieved SHW amplitude does not noticeably alter when varying h_1 . As an example, also the diagram for the fixed value $h_1 = 60$ nm is added in Fig. 5.16. As one can easily see, the "fringes" in the diagrams get shifted with increasing h_1 , but the SHW amplitudes in the maxima remain almost unchanged.

In total, introducing the third layer of Al₂O₃ does not bring any benefits in terms of the conversion efficiency. The highest achieved SHW amplitude for SiO₂/TiO₂ turns out to be larger for the same

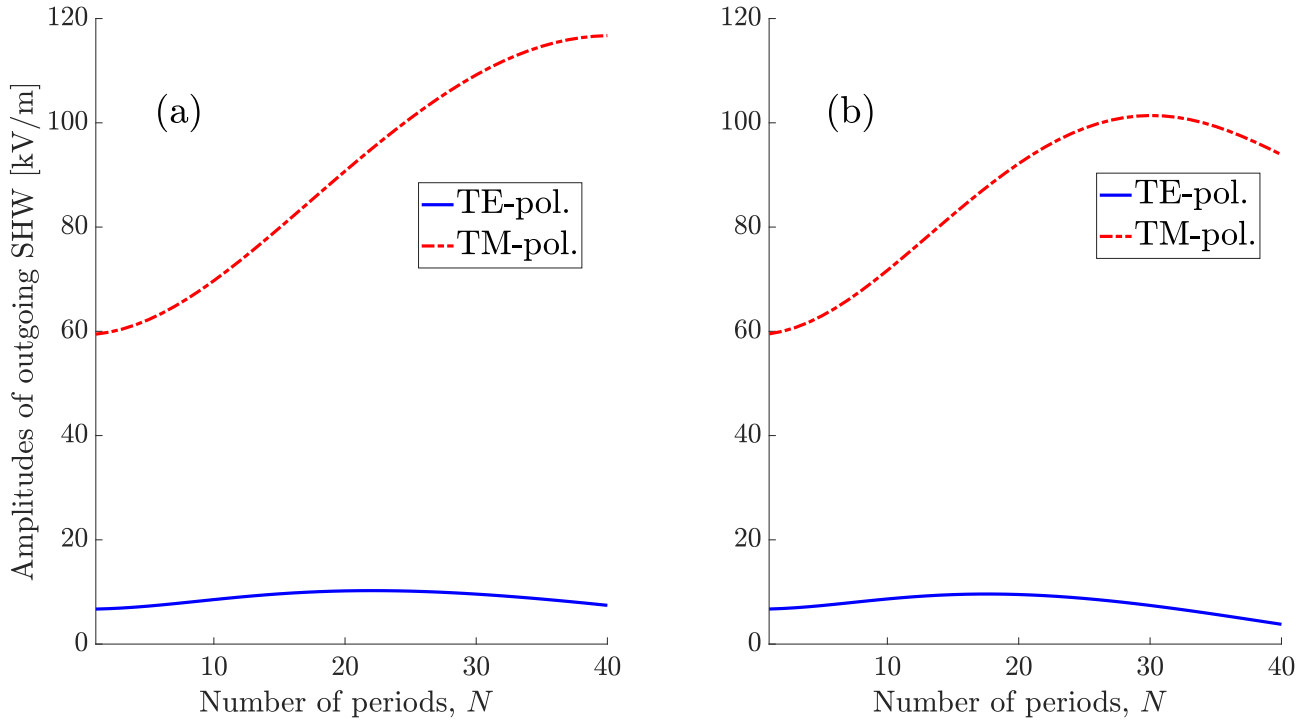


Figure 5.11: Maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) from the SiO₂/TiO₂ periodic stack [in kV/m] vs. the number of layer pairs: (a) $h_1 = 1$ nm (SiO₂), $h_2 = 1$ nm (TiO₂); (b) $h_1 = 2$ nm (SiO₂), $h_2 = 1$ nm (TiO₂). The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

N , despite the proportionally smaller number of interfaces inside the whole structure. This increased number of interfaces is confronted with the smaller contrast of the refractive indices between the intermediate Al₂O₃ layer and other layers. The simulations show that the latter effect prevails resulting in reduced SHW amplitude.

5.4 Summary

In this chapter the possibility of the enhancement of the emitted surface-driven SHW from multilayer structures of different compositions was studied. I have searched through several arrangement of the layers of Si, SiO₂ and TiO₂ layers, specifically two-layer periodic structures made of SiO₂/TiO₂ layers and three-layer periodic structures made of SiO₂/Al₂O₃/TiO₂ layers. The varying of the layer thicknesses, the number of the periods as well as the incidence angle were considered to find the strongest amplitude of the SSHG. The performed simulations provide therefore the optimized parameters for the highest achievable energy conversion efficiency into the SHW for the considered compositions of the multilayer structures.

In all cases the dependence on the angle of incidence was found with a single maximum at oblique incidence. The dependence on the thicknesses of the layers appears typically complex with multiple maxima as the layer thicknesses are varied. The dependences of the emitted SHW on both the angle of incidence and on the layer thicknesses exhibit obvious correlation with the linear reflection/transmission of the stack both at the fundamental and at the SH frequency, namely the maxima of the emitted SHW closely follow the minima of the reflection of the stack at both frequencies. The

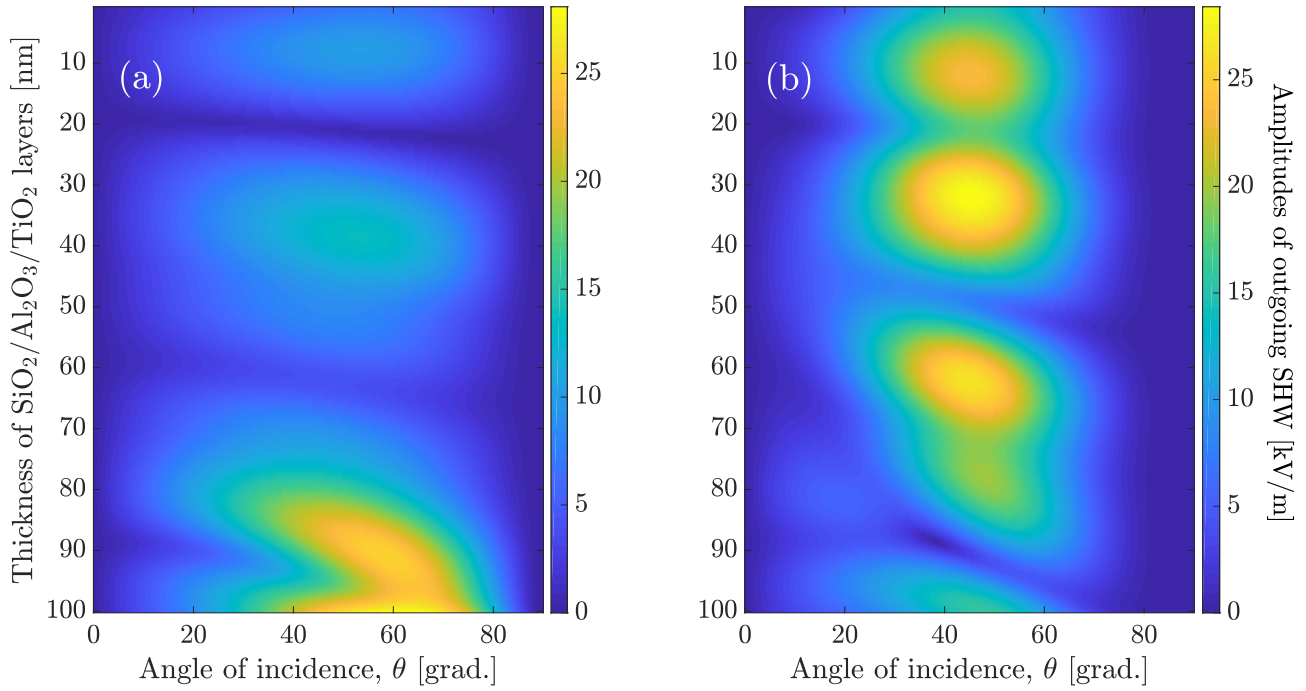


Figure 5.12: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 2 periods of $\text{SiO}_2/\text{Al}_2\text{O}_3/\text{TiO}_2$ layer triples, $h_1 = h_2 = h_3$: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

specific dependence on the thicknesses of all layers in the period allows the best optimization of the SH signal. In the case of the periodic $\text{SiO}_2/\text{TiO}_2$ stack optimal thicknesses were found that give the strongest SH field for a fixed number of layers. At the same time the dependence on the number of layers exhibits several conflicting features. On one hand, increasing the number of layers results in the growth of the maximal amplitude of the SHW. This dependence is close to linear for small number of periods, but around $N \sim 5$ one usually gets a saturation and the further increasing of the number of periods N leads to the increasingly slow growth of the maximal SH field strength. On the other hand, for larger N the maxima over the thicknesses of the layer split into several narrower maxima. As the result for large enough number of layers, for instance several tens of periods, the maxima and minima over the layer thicknesses are located too densely, so that the separation between the neighbouring maxima amounts to around 1 nm or even smaller. Given that the fabrication techniques possess some limited precision, it seems challenging to assure the required accuracy of the thicknesses of the grown layers. Hence, the periodic stack with the large number of periods can be considered to be hardly suitable for the experimental investigations.

The scaling law of the SHG with the number of periods N can also represent an interesting question for studies. An important issue here is related with the presence of a substrate. The power scaling as $\sim N^8$ has been theoretically found for the finite periodic multilayer structures with the bulk nonlinearity. However, the presence of the substrate breaks the symmetry together with the corresponding scaling law. In the simulations for the ultrathin layers close to linear scaling in the field amplitude (i.e. quadratic scaling in the intensity) was shown for $\text{SiO}_2/\text{TiO}_2$ layers, see Fig. 5.11. The faster scaling can be expected to be barely possible for a stack on a substrate, since the presence of a substrate breaks the periodicity of the structure and leads to more complicated interference effects. Although

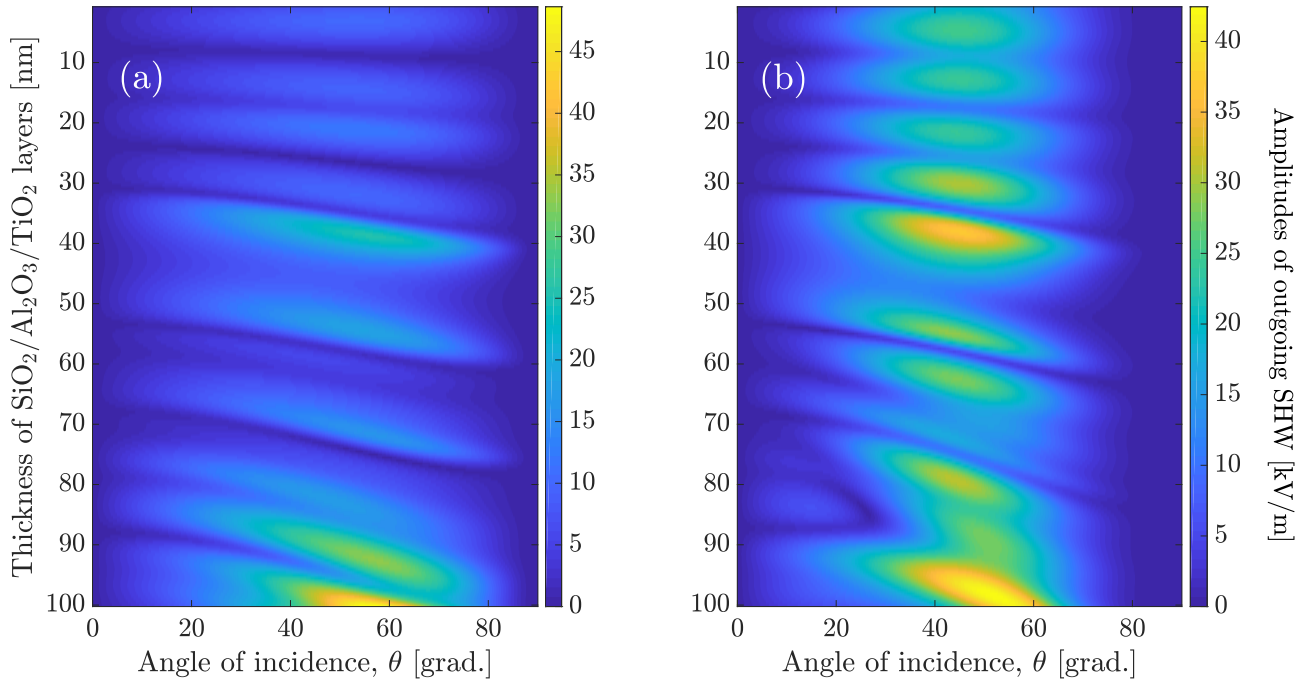


Figure 5.13: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 5 periods of $\text{SiO}_2/\text{Al}_2\text{O}_3/\text{TiO}_2$ layer triples, $h_1 = h_2 = h_3$: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

the stack is normally grown on the top of a substrate, it could be also of interest to search for the scaling law faster than $\sim N^2$ also in the multilayer structures with the surface nonlinearities in the absence of a substrate. This issue is outside the scope of the current work.

Among the considered compositions $\text{SiO}_2/\text{TiO}_2$ periodic multilayer structure was found to yield the highest SH field strength. The largest obtained energy conversion efficiency amounts to almost 10^{-5} for the intensity of the incident FW of $1.33 \text{ GW}/\text{cm}^2$, which is comparable to the values achieved in resonant GaAs nanostructures with similar intensities [31, 40]. The efficient surface-driven SHG in such stacks seemingly comes from the high contrast in the refractive indices between the constituting materials, what results in the strong surface nonlinearity at their interface. These findings thus demonstrate the potential of the multilayer stacks made of centrosymmetric materials as promising optical components for nonlinear nanophotonic applications.

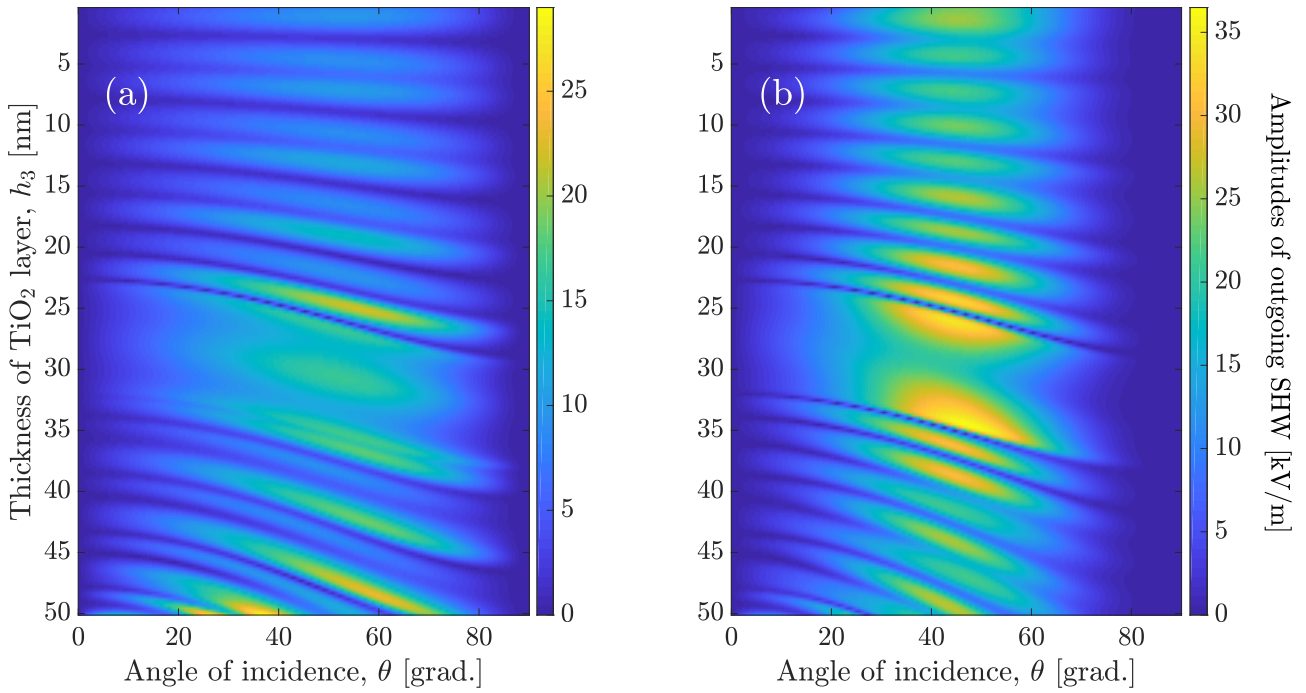


Figure 5.14: Amplitude of the upwards outgoing SHW [in kV/m] from the stack of 10 periods of $\text{SiO}_2/\text{Al}_2\text{O}_3/\text{TiO}_2$ layer triples, $h_1 = h_2 = 2h_3$: (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

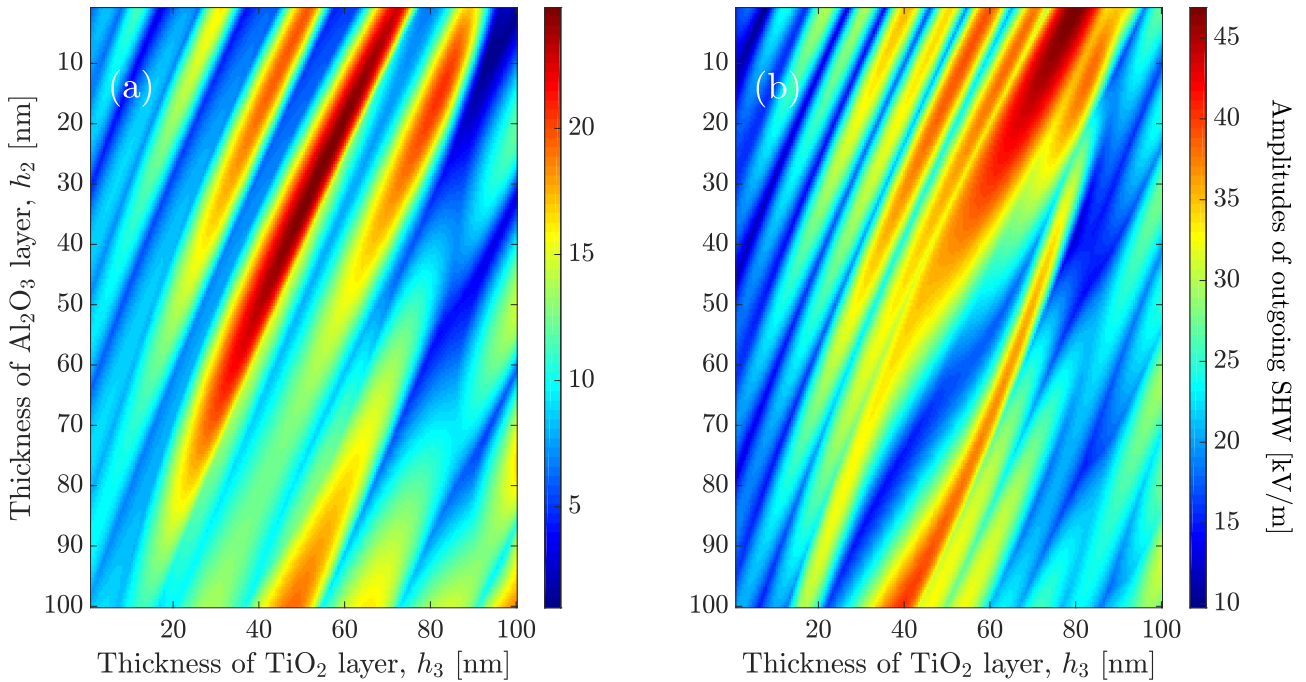


Figure 5.15: Maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) from the periodic stack of 10 periods of $\text{SiO}_2/\text{Al}_2\text{O}_3/\text{TiO}_2$ layer triples [in kV/m] for fixed $h_1 = 20$ nm vs. the thicknesses h_2, h_3 : (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

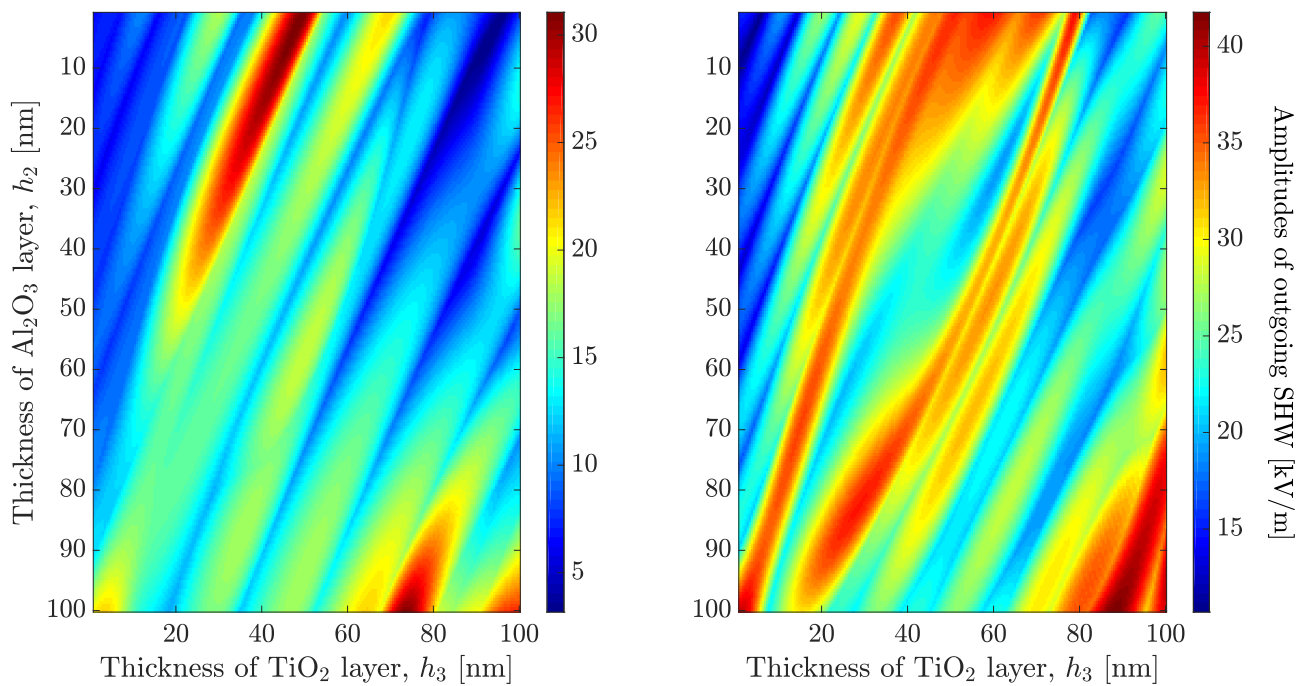


Figure 5.16: Maximal amplitude of the upwards outgoing SHW (among all values of the angle of incidence) from the periodic stack of 10 periods of $\text{SiO}_2/\text{Al}_2\text{O}_3/\text{TiO}_2$ layer triples [in kV/m] for fixed $h_1 = 60$ nm vs. the thicknesses h_2, h_3 : (a) TE-polarized FW; (b) TM-polarized FW. The amplitude of the FW (pump) $E^{\text{FW}} = 10^8$ V/m.

Conclusion and outlook

The research efforts in this thesis have been primarily focused on the implementation of efficient numerical methods for the modeling of surface second-harmonic generation in nanophotonic applications. Therefore the research was performed towards the main goals stated in the introduction section. The first case considered was just a one-dimensional stack of multiple flat layers of nonlinear materials. This layout is expected to be promising for boosting the SSHG. However, the discretization of the whole computational domain for such structure is unreasonable. For simulation of the SSHG from such multilayer structures an analytical framework was formulated based on the transfer matrix method (TMM). TMM allows treating the field propagation through an arbitrary layered structure, including the effects of the multiple reflections. The developed TMM approach enables the analytical calculation of the SSHG from the multilayer structure of an arbitrary composition within UPA and with only surface nonlinearity considered. Moreover, it was shown that in the case of the ultrathin layers in the multilayer structure the analytical results can be greatly simplified. As a result, it is possible to derive the tensor of the effective surface nonlinear susceptibility. The derived effective tensor acts as the surface nonlinear tensor of the whole stack and provides the correct expression for the generated SHW in the limit of vanishing thicknesses of all layers. The first-order correction terms were obtained over the small parameters, namely, the normalized thicknesses, which allow estimating the applicability limits of the effective surface nonlinear tensor.

The developed numerical method was tested with an exemplary multilayer structure. In particular, the applicability of the effective surface tensor for the ultrathin layers was demonstrated. Importantly, it was shown that in the multilayer structure the surface nonlinear responses of single interfaces can efficiently sum up, which leads to the strongly enhanced conversion efficiency of the SSHG. The simulation results were compared with a GaAs layer of similar thickness and it was found that fully surface-driven SHG in the multilayer stack turns out to yield much better performance than a bulk-driven SHG from a layer of a noncentrosymmetric semiconductor.

TMM can only be applied to modeling of layered one-dimensional structures. For more complex geometries other numerical methods are needed. Here, an alternative finite-element implementation of the surface second-order nonlinear polarization has been developed, which allows efficient simulations of SSHG from both all-dielectric and plasmonic nanoparticles and metasurfaces. The implementation was done in the collaboration with JCMwave GmbH and was based on the realisation of FEM in the JCMsuite solver. For this purpose, an analytical framework was developed for the implementation of the surface nonlinearity as a surface-localized 2D source, which introduces the jump of the tangential components of the electric and magnetic field.

The developed FEM was then applied to several test problems, which also allow for an analytical solution. The performed tests confirmed the expected exponential convergence rate of the numerical solution with the polynomial degree of the ansatz functions. The implemented method provides the possibility to treat the geometrical parameters of the discretization mesh in a flexible way. In particular, the mesh can be easily refined close to the interface between two media, nearby the corners of the nanostructure or in any other places, where large gradients of the electric field can be expected to occur. Furthermore, a bulk nonlinearity can be incorporated in a simple way into the simulation without any extra costs.

FEMs are well known to exhibit advantageous performance with respect to many other numerical simulation techniques. FEMs provide the possibility to not only adapt the mesh settings over a wide range, but also to vary the polynomial degree of the ansatz functions and the PML settings. Selecting coarser mesh and PML discretization together with the smaller polynomial degree would strongly reduce the computation time and resources needed at the cost of the accuracy of the simulations. Therefore one usually is able to find an optimal trade-off between the required computational costs and the achievable numerical accuracy. Finally, it should be noted that FEMs enable the *hp*-adaptive strategy, when the discretization grid is adaptively refined in every part of the computation domain and at the same time the polynomial degree is also accordingly adjusted in every part of the computation domain. This approach provides the optimal performance of FEM for the required accuracy of the solution. Such *hp*-adaptive strategies are currently actively studied for FEM and could potentially represent a powerful tool to boost the performance of any FEM-based simulation. In total, it is expected that the developed FEM-based implementation of the surface nonlinearity can be successfully applied for modeling SSHG in different problems in nonlinear nanophotonics.

The developed numerical methods have made it possible to proceed with considering several specific problems related to the modeling of SSHG. The first of these problems was related to the reliable detection of the surface nonlinear signal from noncentrosymmetric semiconductor materials. This problem has arisen, since some recent experiments demonstrated the interplay of both bulk and surface contribution in nanoparticles and metasurfaces made of noncentrosymmetric semiconductors. At the same time, the question of the separation of both contribution has not been theoretically studied well enough. Therefore this issue required more detailed numerical investigation to determine the anticipated signatures of the surface and bulk contributions to the total nonlinear response. A detailed numerical analysis was performed for the illumination of a semi-infinite slab of a noncentrosymmetric semiconductor with two experimentally feasible illumination sources: plane-wave and tightly-focused Gaussian beam. It was demonstrated that in III-V semiconductors the surface and bulk optical nonlinearities exhibit specific features in the polarization-resolved far-field pattern, which can be used to not only identify and separate both contributions, but also for the direct comparison of the strengths of both nonlinearities. Finally, a numerical measure was introduced for the ratio of the surface and bulk nonlinearities and it was connected to the amplitude parameters of the observed characteristic signatures in the far-field. With these findings the applicability limits of this approach were estimated, namely, the strength of the surface nonlinearity, which can be measured. The proposed setup both with a plane wave and with a tightly-focused Gaussian beam allows quantifying the strength

of the surface optical nonlinearity in a wide range from several polarization-resolved measurements of the far-field radiation pattern. Moreover, the obtained estimates show that the expected values of the surface nonlinearity of the widely-used nonlinear materials should be covered with the presented approach.

Finally, the enhancement of surface-driven SHG from the multilayer structures was investigated using the developed transfer-matrix-based method. Both periodic and nonperiodic compositions of the stack were considered with varying number of layers and the dependence of the SHG response on the parameters of the layers was studied. In particular, the nonperiodic compositions of Si, SiO₂ and TiO₂ layers were studied together with the periodic structures made of SiO₂/TiO₂ layer pairs and periodic structures made of SiO₂/Al₂O₃/TiO₂ layer triples. The dependence on the incidence angle in all cases yields a single maximum. At the same time the dependence on the thicknesses of the layers in general shows a complex behaviour with multiple maxima. It was found that the SiO₂/TiO₂ periodic multilayer structure provides the highest efficiency of SHG among the considered compositions. The relatively efficient SSHG in this case seems to be related with the high refractive index contrast between the neighbouring layers in the stack. The performed simulation have thus confirmed that multilayer structures made of centrosymmetric materials with high refractive index contrast between the layers enable boosting the conversion efficiency into SHG for the applications in the nanoscale optical devices.

In conclusion, it would be worth to outline the prospects for the future research and the development of the findings of the thesis. Specifically the further elaboration of the proposed numerical methods for modeling of SSHG could be of interest. The developed finite-element method for SSHG simulations can be further improved through the application of *hp*-adaptive algorithms. This topic currently represents the mainstream research direction in the area of the finite-element numerical methods. Applying *hp*-adaptive strategies allows significant optimization of the required computational resources for the fixed accuracy of the simulations as compared to the standard FEM. Such improvement is obtained thanks to the efficient algorithm of the apriori estimates, which adaptively selects the varying mesh sidelength and varying polynomial degree of the ansatz functions within the computational domain to assure the best possible performance of the method with lowest possible costs. The usage of a *hp*-adaptive strategy for modeling of SSHG would be particularly promising, since the simulations of SSHG are inherently very sensitive to the treatment of the geometry and the calculated field close to the interfaces and much less sensitive to their treatment in the rest of the computational domain.

The analysis of SSHG from multilayer structures can be also further evolved. First of all, the contribution of the bulk nonlinearity should be accounted for. This would imply the modification of the source matrices in the transfer-matrix method to include the bulk nonlinear polarization. Although in ultrathin layers of several nm or tens of nm the surface nonlinearity should prevail, this may cease to be true for the thicker layers (hundreds of nm). In such intermediate case the interplay of the bulk and surface nonlinear contributions can become important. Hence, the correct modeling of SHG phenomena would rely on the elaborated TMM. The possibility to gain high SHG conversion efficiency in the intermediate-thickness stacks can be an interesting problem to address. Moreover,

the possibility of the power scaling law for the emitted SHW in the multilayer structures with the surface nonlinearities in the absence of a substrate deserves more detailed consideration. A power scaling up to $\sim N^8$ has been theoretically found for finite periodic multilayer structures with bulk nonlinearity, when the fundamental frequency and the SH frequency are resonant with the band-edge states in the finite periodic structure and the PMC is fulfilled simultaneously. It would be therefore of certain interest to examine the possible enhancement of the SSHG, when similar resonances are excited in the finite periodic multilayer structures with the surface nonlinearities.

Deutschsprachige Zusammenfassung

In meiner Promotion habe ich die nichtlineare Frequenzkonversion, insbesondere die Erzeugung der zweiten Harmonischen einer einfallenden Welle, durch Oberflächennichtlinearitäten an Grenzflächen dielektrischer Materialien untersucht. Dabei habe ich für die Berechnung der zweiten Harmonischen optimierte numerische Verfahren entwickelt und diese auf verschiedene Strukturen angewendet um deren Eigenschaften für die Frequenzkonversion zu bestimmen.

Im Speziellen habe ich ein Transfer-Matrix-Verfahren für die Berechnungen von mehrschichtigen Strukturen beliebiger Zusammensetzung entwickelt, welches die Erzeugung der zweiten Harmonische an deren Grenzflächen berechnet. Darauf aufbauend habe ich analytische Gleichungen für den effektiven nichtlinearen Tensor der Oberflächennichtlinearität zweiter Ordnung abgeleitet, welcher die nichtlineare Wechselwirkung der Gesamtstruktur beschreibt. Ich habe das entwickelte Verfahren mit einigen exemplarischen Strukturen getestet, um seine Leistung zu überprüfen. Dabei konnte ich zeigen, dass mehrschichtige Strukturen mit ultradünnen Schichten die zweite Harmonische mit hoher Effizienz erzeugen können, obwohl nur die relativ schwache Oberflächennichtlinearität genutzt wird.

Danach habe ich ein numerisches Verfahren für die Berechnungen der Erzeugung der zweiten Harmonische in Nanostrukturen entwickelt. Das entwickelte Verfahren basiert auf der Finite-Elemente-Methode und ermöglicht Berechnungen unter Berücksichtigung sowohl von Oberflächen- als auch Volumennichtlinearitäten. Der Vorteil der Finite-Elemente-Methode ist dabei, dass das numerische Gitter zur Berechnung umfassend an das zu analysierende Problem angepasst werden kann. Das entwickelte Verfahren wurde mit einigen Beispielproblemen getestet, die exponentielle Konvergenz der berechneten Lösung konnte dabei nachgewiesen werden.

Ich habe das entwickelte Finite-Elemente-Verfahren angewandt, um die Signaturen von interferierenden Oberflächen- und Volumennichtlinearitäten in nicht-zentrosymmetrischen Halbleitern voneinander abzugrenzen. Dabei habe ich entdeckt, dass beide Arten der Nichtlinearität wesentlich unterschiedliche räumliche Verteilungen der ins Fernfeld emittierten zweiten Harmonischen aufweisen, wenn eine Grenzflächen eines nicht-zentrosymmetrischen Halbleiters, z.B. AlGaAs, mit einer ebenen Welle oder mit einem Gausstrahl angeregt wird. Ich konnte zeigen, dass es für beide Anregungen einen breiten Parameterbereich von mehr als drei Größenordnungen der Stärke der Oberflächennichtlinearität gibt, in dem beide Nichtlinearitäten experimentell sicher voneinander unterschieden werden können. Es ist erwartbar, dass diese Ergebnisse nicht nur für AlGaAs, sondern auch für andere III-V-Verbindungshalbleiter nutzbar sind.

Schlussendlich habe ich mithilfe des entwickelten Transfer-Matrix-Verfahrens mehrschichtige Strukturen mit Oberflächennichtlinearitäten für die Erzeugung der zweiten Harmonische optimiert.

Insbesondere habe ich periodische Strukturen aus $\text{SiO}_2/\text{TiO}_2$ und $\text{SiO}_2/\text{Al}_2\text{O}_3/\text{TiO}_2$ Schichten erforscht. Davon hat ein periodischer Schichtstapel aus abwechselnden $\text{SiO}_2/\text{TiO}_2$ Schichten die höchste Konversionseffizienz. Basierend auf meinen Ergebnissen können periodischen Mehrschichtsysteme nun für die nichtlineare Frequenzkonversion in verschiedenen Anwendungen verwendet werden.

List of own contributions

Published papers:

1. A. Pakhomov, F.J.F. Löchner, L. Zschiedrich, S. Saravi, M. Hammerschmidt, S. Burger, T. Pertsch, F. Setzpfandt, "Far-field polarization signatures of surface optical nonlinearity in noncentrosymmetric semiconductors", *Sci. Rep.* **10**, 10545 (2020).
2. A. Pakhomov, M. Hammerschmidt, S. Burger, T. Pertsch, F. Setzpfandt, "Modeling of surface-induced second-harmonic generation from multilayer structures by transfer matrix method", *Opt. Express* **29**, 9098-9122 (2021).
3. A. Pakhomov, L. Zschiedrich, S. Burger, F. Setzpfandt, T. Pertsch, M. Hammerschmidt, "Adaptive finite-element method for modeling of surface second-harmonic generation from arbitrary-shaped nanostructures", *to be submitted*.

Conferences:

1. A. Pakhomov, L. Zschiedrich, S. Burger, F. Setzpfandt, T. Pertsch, "HP-adaptive FEM modeling of bulk and surface second-harmonic generation from $\text{Al}_x\text{Ga}_{1-x}\text{As}$ nanostructures", 36 th EU-PROMETA School "Metamaterials and nanophotonic structures in applications", Karlsruhe Institute of technology, Karlsruhe, Germany, 28 May - 1 June (2018);
2. A. Pakhomov, L. Zschiedrich, S. Saravi, S. Burger, F. Setzpfandt, T. Pertsch, "Polarization-resolved quantification of bulk and surface contributions to second-harmonic generation in AlGaAs", European Optical Society Biennial Meeting 2018, Technical University of Delft, Delft, Netherlands, October 8-12 (2018);
3. A. Pakhomov, L. Zschiedrich, M. Hammerschmidt, S. Burger, S. Saravi, F. Setzpfandt, T. Pertsch, "Optimization of illumination geometry for evaluation of surface nonlinearity in zinc-blende semiconductors", 12-th Annual Meeting Photonic Devices 2019, Zuse Institute Berlin, February 14-15 (2019).

Workshops & seminars:

1. Seminar talk, 20.06.2017, Jena, Abbe Center of Photonics, Nanooptics group seminar, "Modeling of spatio-temporal dynamics in transversely extended resonant optical systems";
2. Seminar talk, 10.07.2017, Berlin, Zuse Institute Berlin, Computational Nanooptics group seminar, "Comparison of hp-adaptive finite-element algorithms for benchmark 2D problems";

3. Seminar talk, 11.09.2017, Berlin, Zuse Institute Berlin, Computational Nanooptics group seminar, "Numerical simulations of the resonantly-enhanced SHG from AlGaAs semiconductor nanodisks";
4. Seminar talk, 15.03.2018, Jena, Abbe Center of Photonics, Nanooptics group seminar, "Numerical modeling of surface second-harmonic generation from AlGaAs nanostructures using hp-adaptive finite-element method";
5. Seminar talk, 05.04.2018, Berlin, Zuse Institute Berlin, Computational Nanooptics group seminar, "Numerical modeling of surface second-harmonic generation from AlGaAs nanostructures";
6. Seminar talk, 07.05.2019, Jena, Abbe Center of Photonics, Nanooptics group seminar, "Finite-element methods for simulation of surface nonlinearities";
7. Seminar talk, 23.06.2020, Jena, Abbe Center of Photonics, Nanooptics group seminar, "Surface second-harmonic generation from ultrathin multilayer structures".

Bibliography

- [1] Y. R. Shen, *The principles of nonlinear optics* (John Wiley & Sons, 1984).
- [2] R. W. Boyd, *Nonlinear Optics* (Academic Press, 2008).
- [3] T. H. Maiman, “Stimulated optical radiation in ruby,” *Nature* **187**, 493–494 (1960).
- [4] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, “Generation of optical harmonics,” *Phys. Rev. Lett.* **7**, 118–119 (1961).
- [5] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, “Interactions between light waves in a nonlinear dielectric,” *Phys. Rev.* **127**, 1918–1939 (1962).
- [6] F. S. Pavone and P. J. Campagnola, *Second Harmonic Generation Imaging* (CRC Press, 2016).
- [7] P. J. Campagnola, H. A. Clark, W. A. Mohler, A. Lewis, and L. M. Loew, “Second harmonic imaging microscopy of living cells,” *Journal of Biomedical Optics* **6**, 277 – 286 (2001).
- [8] P. J. Campagnola and L. M. Loew, “Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms,” *Nature Biotechnology* **21**, 1356—1360 (2003).
- [9] M. Han, G. Giese, and J. F. Bille, “Second harmonic generation imaging of collagen fibrils in cornea and sclera,” *Opt. Express* **13**, 5791–5797 (2005).
- [10] R. Paschotta, P. Kürz, R. Henking, S. Schiller, and J. Mlynek, “82% efficient continuous-wave frequency doubling of 1.06 μm with a monolithic MgO:LiNbO₃ resonator,” *Opt. Lett.* **19**, 1325–1327 (1994).
- [11] P. Zeller and P. Peuser, “Efficient, multiwatt, continuous-wave laser operation on the ${}^4F_{3/2} - {}^{-4}I_{9/2}$ transitions of Nd:YVO₄ and Nd:YAG,” *Opt. Lett.* **25**, 34–36 (2000).
- [12] Q. H. Xue, Q. Zheng, Y. K. Bu, F. Q. Jia, and L. S. Qian, “High-power efficient diode-pumped Nd:YVO₄/LiB₃O₅ 457 nm blue laser with 4.6 W of output power,” *Opt. Lett.* **31**, 1070–1072 (2006).
- [13] L. Schneider and W. Peukert, “Second harmonic generation spectroscopy as a method for in situ and online characterization of particle surface properties,” *Particle & Particle Systems Characterization* **23**, 351–359 (2006).

- [14] Y. Wang, J. Xiao, S. Yang, Y. Wang, and X. Zhang, “Second harmonic generation spectroscopy on two-dimensional materials,” *Opt. Mater. Express* **9**, 1136–1149 (2019).
- [15] T. F. Heinz, M. M. T. Loy, and W. A. Thompson, “Study of Si(111) surfaces by optical second-harmonic generation: Reconstruction and surface phase transformation,” *Phys. Rev. Lett.* **54**, 63–66 (1985).
- [16] Y. R. Shen, “Surface nonlinear optics,” *J. Opt. Soc. Am. B* **28**, A56–A66 (2011).
- [17] S. K. Kurtz and T. T. Perry, “A powder technique for the evaluation of nonlinear optical materials,” *Journal of Applied Physics* **39**, 3798–3813 (1968).
- [18] R. Trebino, K. W. DeLong, D. N. Fittinghoff, J. N. Sweetser, M. A. Krumbügel, B. A. Richman, and D. J. Kane, “Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating,” *Review of Scientific Instruments* **68**, 3277–3295 (1997).
- [19] M. Fiebig, D. Fröhlich, B. B. Krichevstov, and R. V. Pisarev, “Second harmonic generation and magnetic-dipole-electric-dipole interference in antiferromagnetic Cr_2O_3 ,” *Phys. Rev. Lett.* **73**, 2127–2130 (1994).
- [20] M. Wegener, *Extreme nonlinear optics* (Springer, 2005).
- [21] K. D. Moll, D. Homoelle, A. L. Gaeta, and R. W. Boyd, “Conical harmonic generation in isotropic materials,” *Phys. Rev. Lett.* **88**, 153901 (2002).
- [22] W. Lehr and R. von Baltz, “Quantum theory of radiation-pressure induced electrical currents: I. Free electrons,” *Z. Physik B - Condensed Matter* **51**, 25–28 (1983).
- [23] N. Bloembergen and A. J. Sievers, “Nonlinear optical properties of periodic laminar structures,” *Applied Physics Letters* **17**, 483–486 (1970).
- [24] J. Webjorn, S. Siala, D. W. Nam, R. G. Waarts, and R. J. Lang, “Visible laser sources based on frequency doubling in nonlinear waveguides,” *IEEE Journal of Quantum Electronics* **33**, 1673–1686 (1997).
- [25] M. Kauranen and A. V. Zayats, “Nonlinear plasmonics,” *Nature Photonics* **6**, 737–748 (2012).
- [26] J. Butet, P.-F. Brevet, and O. J. F. Martin, “Optical second harmonic generation in plasmonic nanostructures: From fundamental principles to advanced applications,” *ACS Nano* **9**, 10545–10562 (2015).
- [27] D. Smirnova and Y. S. Kivshar, “Multipolar nonlinear nanophotonics,” *Optica* **3**, 1241–1255 (2016).
- [28] Y. S. Kivshar and A. E. Miroshnichenko, “Meta-optics with Mie resonances,” *Opt. Photon. News* **28**, 24–31 (2017).

- [29] S. S. Kruk and Y. S. Kivshar, “Functional meta-optics and nanophotonics governed by Mie resonances,” *ACS Photonics* **4**, 2638–2649 (2017).
- [30] A. I. Kuznetsov, A. E. Miroschnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk’yanchuk, “Optically resonant dielectric nanostructures,” *Science* **354**, aag2472 (2016).
- [31] L. Carletti, A. Locatelli, O. Stepanenko, G. Leo, and C. De Angelis, “Enhanced second-harmonic generation from magnetic resonance in AlGaAs nanoantennas,” *Opt. Expr.* **23**, 26544–26550 (2015).
- [32] L. Carletti, K. Koshelev, C. De Angelis, and Y. Kivshar, “Giant nonlinear response at the nanoscale driven by bound states in the continuum,” *Phys. Rev. Lett.* **121**, 033903 (2018).
- [33] K.-C. Shen, Y.-T. Huang, T. L. Chung, M. L. Tseng, W.-Y. Tsai, G. Sun, and D. P. Tsai, “Giant efficiency of visible second-harmonic light by an all-dielectric multiple-quantum-well metasurface,” *Phys. Rev. Applied* **12**, 064056 (2019).
- [34] P. Guyot-Sionnest, W. Chen, and Y. R. Shen, “General considerations on optical second-harmonic generation from surfaces and interfaces,” *Phys. Rev. B* **33**, 8254–8263 (1986).
- [35] P. Guyot-Sionnest and Y. R. Shen, “Local and nonlocal surface nonlinearities for surface optical second-harmonic generation,” *Phys. Rev. B* **35**, 4420–4426 (1987).
- [36] P.-F. Brevet, *Surface Second Harmonic Generation* (PPUR, 1997).
- [37] S. Cattaneo and M. Kauranen, “Polarization-based identification of bulk contributions in surface nonlinear optics,” *Phys. Rev. B* **72**, 033412 (2005).
- [38] F. X. Wang, F. J. Rodriguez, W. M. Albers, R. Ahorinta, J. E. Sipe, and M. Kauranen, “Surface and bulk contributions to the second-order nonlinear optical response of a gold film,” *Phys. Rev. B* **80**, 233402 (2009).
- [39] D. Timbrell, J. W. You, Y. S. Kivshar, and N. C. Panoiu, “A comparative analysis of surface and bulk contributions to second-harmonic generation in centrosymmetric nanoparticles,” *Sci. Rep.* **8**, 3586 (2018).
- [40] V. F. Gili, L. Carletti, A. Locatelli, D. Rocco, M. Finazzi, L. Ghirardini, I. Favero, C. Gomez, A. Lemaitre, M. Celebrano, C. De Angelis, and G. Leo, “Monolithic AlGaAs second-harmonic nanoantennas,” *Opt. Expr.* **24**, 15965–15971 (2016).
- [41] L. Carletti, D. Rocco, A. Locatelli, C. De Angelis, V. F. Gili, M. Ravaro, I. Favero, G. Leo, M. Finazzi, L. Ghirardini, M. Celebrano, G. Marino, and A. V. Zayats, “Controlling second-harmonic generation at the nanoscale with monolithic AlGaAs-on-AlOx antennas,” *Nanotechn.* **28**, 114005 (2017).

- [42] L. Ghirardini, L. Carletti, V. Gili, G. Pellegrini, L. Duo, M. Finazzi, D. Rocco, A. Locatelli, C. De Angelis, I. Favero, M. Ravaro, G. Leo, A. Lemaitre, and M. Celebrano, “Polarization properties of second-harmonic generation in AlGaAs optical nanoantennas,” *Opt. Lett.* **42**, 559–562 (2017).
- [43] R. Sanatinia, M. Swillo, and S. Anand, “Surface second-harmonic generation from vertical GaP nanopillars,” *Nano Lett.* **12**, 820–826 (2012).
- [44] S. Liu, M. B. Sinclair, S. Saravi, G. A. Keeler, Y. Yang, J. Reno, G. M. Peake, F. Setzpfandt, I. Staude, T. Pertsch, and I. Brener, “Resonantly enhanced second-harmonic generation using III-V semiconductor all-dielectric metasurfaces,” *Nano Lett.* **16**, 5426–5432 (2016).
- [45] L. Rodríguez-Suné, J. Trull, M. Scalora, R. Vilaseca, and C. Cojocar, “Harmonic generation in the opaque region of GaAs: the role of the surface and magnetic nonlinearities,” *Opt. Express* **27**, 26120–26130 (2019).
- [46] S. Bergfeld and W. Daum, “Second-harmonic generation in GaAs: Experiment versus theoretical predictions of $\chi_{xyz}^{(2)}$,” *Phys. Rev. Lett.* **90**, 036801 (2003).
- [47] Z. Zhang, J. Kim, R. Khoury, M. Saghayezhian, L. H. Haber, and E. W. Plummer, “Surface sum frequency generation spectroscopy on non-centrosymmetric crystal GaAs(001),” *Surf. Sci.* **664**, 21–28 (2017).
- [48] K. Yee, “Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media,” *IEEE Transactions on Antennas and Propagation* **14**, 302–307 (1966).
- [49] K. L. Shlager and J. B. Schneider, “A selective survey of the finite-difference time-domain literature,” *IEEE Antennas and Propagation Magazine* **37**, 39–57 (1995).
- [50] P. M. Goorjian and A. Taflove, “Direct time integration of Maxwell’s equations in nonlinear dispersive media for propagation and scattering of femtosecond electromagnetic solitons,” *Opt. Lett.* **17**, 180–182 (1992).
- [51] R. W. Ziolkowski and J. B. Judkins, “Nonlinear finite-difference time-domain modeling of linear and nonlinear corrugated waveguides,” *J. Opt. Soc. Am. B* **11**, 1565–1575 (1994).
- [52] R. W. Ziolkowski, J. M. Arnold, and D. M. Gogny, “Ultrafast pulse interactions with two-level atoms,” *Phys. Rev. A* **52**, 3082–3094 (1995).
- [53] S. A. Basinger and D. J. Brady, “Finite-difference time-domain modeling of dispersive nonlinear Fabry-Perot cavities,” *J. Opt. Soc. Am. B* **11**, 1504–1511 (1994).
- [54] S. Radic and N. George, “Ultrafast pulse propagation in periodic optical media: a generalized finite-difference time-domain approach,” *Opt. Lett.* **19**, 1064–1066 (1994).

- [55] S. Zivanovic and K. Mei, "A subgridding method for the time-domain finite-difference method to solve Maxwell's equations," *IEEE Transactions on Microwave Theory and Techniques* **39**, 471–479 (1991).
- [56] M. Okoniewski, E. Okoniewska, and M. Stuchly, "Three-dimensional subgridding algorithm for FDTD," *IEEE Transactions on Antennas and Propagation* **45**, 422–429 (1997).
- [57] A. Zakharian, M. Brio, C. Dineen, and J. Moloney, "Second-order accurate FDTD space and time grid refinement method in three space dimensions," *IEEE Photonics Technology Letters* **18**, 1237–1239 (2006).
- [58] P. Monk, *Finite-element methods for Maxwell's equations* (Clarendon Press, 2003).
- [59] A. Benedetti, M. Centini, C. Sibilia, and M. Bertolotti, "Engineering the second harmonic generation pattern from coupled gold nanowires," *J. Opt. Soc. Am. B* **27**, 408–416 (2010).
- [60] A. Benedetti, M. Centini, M. Bertolotti, and C. Sibilia, "Second harmonic generation from 3D nanoantennas: on the surface and bulk contributions by far-field pattern analysis," *Opt. Expr.* **19**, 26752–26767 (2011).
- [61] J. Mäkitalo, S. Suuriniemi, and M. Kauranen, "Boundary element method for surface nonlinear optics of nanoparticles," *Opt. Expr.* **19**, 23386–23399 (2011).
- [62] A. M. Kern and O. J. F. Martin, "Surface integral formulation for 3D simulations of plasmonic and high permittivity nanostructures," *J. Opt. Soc. Am. B* **26**, 732–740 (2009).
- [63] C. Forestiere, A. Capretti, and G. Miano, "Surface integral method for second harmonic generation in metal nanoparticles including both local-surface and nonlocal-bulk sources," *J. Opt. Soc. Am. B* **30**, 2355–2364 (2013).
- [64] J. Butet, B. Gallinet, K. Thyagarajan, and O. J. F. Martin, "Second-harmonic generation from periodic arrays of arbitrary shape plasmonic nanostructures: a surface integral approach," *J. Opt. Soc. Am. B* **30**, 2970–2979 (2013).
- [65] L. Zhang, S. Tao, Z. Fan, and R. Chen, "Efficient method for evaluation of second-harmonic generation by surface integral equation," *Opt. Expr.* **25**, 28010–28021 (2017).
- [66] Q. T. Vu, H. Haug, O. D. Mücke, T. Tritschler, M. Wegener, G. Khitrova, and H. M. Gibbs, "Light-induced gaps in semiconductor band-to-band transitions," *Phys. Rev. Lett.* **92**, 217403 (2004).
- [67] S. V. Makarov, M. I. Petrov, U. Zywiets, V. Milichko, D. Zuev, N. Lopanitsyna, A. Kuksin, I. Mukhin, G. Zograf, E. Ubyivovk, D. A. Smirnova, S. Starikov, B. N. Chichkov, and Y. S. Kivshar, "Efficient second-harmonic generation in nanocrystalline silicon nanoparticles," *Nano Letters* **17**, 3047–3053 (2017).

- [68] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University Press, 1999).
- [69] J. Reiner, A. Kolpak, Y. Segal, K. Garrity, S. Ismail-Beigi, C. Ahn, and F. J. Walker, “Crystalline oxides on silicon,” *Advanced Materials* **22**, 2919–2938 (2010).
- [70] S. M. George, “Atomic layer deposition: An overview,” *Chemical Reviews* **110**, 111–131 (2010).
- [71] U. W. Pohl, *Epitaxy of Semiconductors: Introduction to Physical Principles* (Springer Science & Business Media, 2013).
- [72] J. Qi, M. S. Yeganeh, I. Koltover, A. G. Yodh, and W. M. Theis, “Depletion-electric-field-induced changes in second-harmonic generation from GaAs,” *Phys. Rev. Lett.* **71**, 633–636 (1993).
- [73] T. A. Germer, K. W. Kołasiński, J. C. Stephenson, and L. J. Richter, “Depletion-electric-field-induced second-harmonic generation near oxidized GaAs(001) surfaces,” *Phys. Rev. B* **55**, 10694–10706 (1997).
- [74] V. Fomenko, E. P. Gusev, and E. Borguet, “Optical second harmonic generation studies of ultrathin high- k dielectric stacks,” *Journal of Applied Physics* **97**, 083711 (2005).
- [75] D. S. Bethune, “Optical harmonic generation and mixing in multilayer media: analysis using optical transfer matrix techniques,” *J. Opt. Soc. Am. B* **6**, 910–916 (1989).
- [76] Y. Jeong and B. Lee, “Matrix analysis for layered quasi-phase-matched media considering multiple reflection and pump wave depletion,” *IEEE Journal of Quantum Electronics* **35**, 162–178 (1999).
- [77] S. Enoch and H. Akhouayri, “Second-harmonic generation in multilayered devices: theoretical tools,” *J. Opt. Soc. Am. B* **15**, 1030–1041 (1998).
- [78] M. Cherchi, “Exact analytic expressions for electromagnetic propagation and optical nonlinear generation in finite one-dimensional periodic multilayers,” *Phys. Rev. E* **69**, 066602 (2004).
- [79] J.-J. Li, Z.-Y. Li, and D.-Z. Zhang, “Second harmonic generation in one-dimensional nonlinear photonic crystals solved by the transfer matrix method,” *Phys. Rev. E* **75**, 056606 (2007).
- [80] M. F. Saleh, L. D. Negro, and B. E. A. Saleh, “Second-order parametric interactions in 1-D photonic-crystal microcavity structures,” *Opt. Express* **16**, 5261–5276 (2008).
- [81] A. Rose, S. Larouche, D. Huang, E. Poutrina, and D. R. Smith, “Nonlinear parameter retrieval from three- and four-wave mixing in metamaterials,” *Phys. Rev. E* **82**, 036608 (2010).
- [82] X. Liu, A. Rose, E. Poutrina, C. Ciraci, S. Larouche, and D. R. Smith, “Surfaces, films, and multilayers for compact nonlinear plasmonics,” *J. Opt. Soc. Am. B* **30**, 2999–3010 (2013).

- [83] M. C. Larciprete, A. Belardini, M. G. Cappeddu, D. de Ceglia, M. Centini, E. Fazio, C. Sibilìa, M. J. Bloemer, and M. Scalora, “Second-harmonic generation from metallodielectric multilayer photonic-band-gap structures,” *Phys. Rev. A* **77**, 013809 (2008).
- [84] M. Scalora, M. A. Vincenti, D. de Ceglia, V. Roppo, M. Centini, N. Akozbek, and M. J. Bloemer, “Second- and third-harmonic generation in metal-based structures,” *Phys. Rev. A* **82**, 043828 (2010).
- [85] M. Scalora, M. A. Vincenti, D. de Ceglia, N. Akozbek, V. Roppo, M. J. Bloemer, and J. W. Haus, “Dynamical model of harmonic generation in centrosymmetric semiconductors at visible and UV wavelengths,” *Phys. Rev. A* **85**, 053809 (2012).
- [86] M. Scalora, M. A. Vincenti, D. de Ceglia, N. Akozbek, M. J. Bloemer, C. De Angelis, J. W. Haus, R. Vilaseca, J. Trull, and C. Cojocar, “Harmonic generation from metal-oxide and metal-metal boundaries,” *Phys. Rev. A* **98**, 023837 (2018).
- [87] R. Sanatinia, S. Anand, and M. Swillo, “Experimental quantification of surface optical nonlinearity in GaP nanopillar waveguides,” *Opt. Expr.* **23**, 756–764 (2015).
- [88] T. F. Heinz, “Nonlinear optics of surfaces and adsorbates,” Ph.D. thesis, University of California (1984).
- [89] J. E. Sipe, D. J. Moss, and H. M. van Driel, “Phenomenological theory of optical second- and third-harmonic from cubic centrosymmetric crystals,” *Phys. Rev. B* **35**, 1129–1141 (1987).
- [90] V. Mizrahi and J. E. Sipe, “Phenomenological treatment of surface second-harmonic generation,” *J. Opt. Soc. Am. B* **5**, 660–667 (1988).
- [91] I. H. Malitson, “Interspecimen comparison of the refractive index of fused silica,” *J. Opt. Soc. Am.* **55**, 1205–1209 (1965).
- [92] J. R. DeVore, “Refractive indices of rutile and sphalerite,” *J. Opt. Soc. Am.* **41**, 416–419 (1951).
- [93] M. A. Green, “Self-consistent optical parameters of intrinsic silicon at 300K including temperature coefficients,” *Solar Energy Materials and Solar Cells* **92**, 1305 – 1310 (2008).
- [94] M. Falasconi, L. Andreani, A. Malvezzi, M. Patrini, V. Mulloni, and L. Pavesi, “Bulk and surface contributions to second-order susceptibility in crystalline and porous silicon by second-harmonic generation,” *Surface Science* **481**, 105 – 112 (2001).
- [95] M. Ohashi, T. Kondo, R. Ito, S. Fukatsu, Y. Shiraki, K. Kumata, and S. S. Kano, “Determination of quadratic nonlinear optical coefficient of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ system by the method of reflected second harmonics,” *Journal of Applied Physics* **74**, 596–601 (1993).
- [96] L. Kou, T. Frauenheim, and C. Chen, “Nanoscale multilayer transition-metal dichalcogenide heterostructures: Band gap modulation by interfacial strain and spontaneous polarization,” *The Journal of Physical Chemistry Letters* **4**, 1730–1736 (2013).

- [97] A. Poddubny, I. Iorsh, P. Belov, and Y. Kivshar, “Hyperbolic metamaterials,” *Nature Photonics* **7**, 948–957 (2013).
- [98] G. Marino, P. Segovia, A. Krasavin, P. Ginzburg, N. Olivier, G. Wurtz, and A. Zayats, “Second-harmonic generation from hyperbolic plasmonic nanorod metamaterial slab,” *Laser Photonics Reviews* **12**, 1700189 (2018).
- [99] I. Kolmychek, I. Malysheva, V. Novikov, A. Leontiev, K. Napol'skii, and T. Murzina, “Phase-matched optical second harmonic generation in a hyperbolic metamaterial based on silver nanorods,” *Physical Review B* **102**, 241405 (2020).
- [100] M. A. Alsunaidi, H. M. Al-Mudhaffar, and H. M. Masoudi, “Vectorial FDTD technique for the analysis of optical second-harmonic generation,” *IEEE Photon. Tech. Lett.* **21**, 310–312 (2009).
- [101] J.-P. Berenger, “A perfectly matched layer for the absorption of electromagnetic waves,” *J. Comput. Phys.* **114**, 185–200 (1994).
- [102] L. Zschiedrich, R. Klose, A. Schädle, and F. Schmidt, “A new finite element realization of the perfectly matched layer method for Helmholtz scattering problems on polygonal domains in 2D,” *J. Comput. Appl. Math.* **188**, 12–32 (2006).
- [103] J. C. Nedelec, “Mixed finite elements in \mathbb{R}^3 ,” *Numer. Math.* **35**, 315–341 (1980).
- [104] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, “Electric field effect in atomically thin carbon films,” *Science* **306**, 666–669 (2004).
- [105] Y. P. Venkata Subbaiah, K. J. Saji, and A. Tiwari, “Atomically thin MoS_2 : A versatile non-graphene 2D material,” *Advanced Functional Materials* **26**, 2046–2069 (2016).
- [106] K. Mak and J. Shan, “Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides,” *Nature Photon.* **10**, 216–226 (2016).
- [107] L. M. Malard, T. V. Alencar, A. P. M. Barboza, K. F. Mak, and A. M. de Paula, “Observation of intense second harmonic generation from MoS_2 atomic crystals,” *Phys. Rev. B* **87**, 201401 (2013).
- [108] G. Wang, X. Marie, I. Gerber, T. Amand, D. Lagarde, L. Bouet, M. Vidal, A. Balocchi, and B. Urbaszek, “Giant enhancement of the optical second-harmonic emission of WSe_2 monolayers by laser excitation at exciton resonances,” *Phys. Rev. Lett.* **114**, 097403 (2015).
- [109] T. Wang and S. Zhang, “Large enhancement of second harmonic generation from transition-metal dichalcogenide monolayer on grating near bound states in the continuum,” *Opt. Express* **26**, 322–337 (2018).

- [110] F. J. F. Löchner, R. Mupparapu, M. Steinert, A. George, Z. Tang, A. Turchanin, T. Pertsch, I. Staude, and F. Setzpfandt, “Controlling second-harmonic diffraction by nano-patterning MoS₂ monolayers,” *Opt. Express* **27**, 35475–35484 (2019).
- [111] Y. Pavlyukh and W. Hübner, “Nonlinear Mie scattering from spherical particles,” *Phys. Rev. B* **70**, 245434 (2004).
- [112] C. Schinke, P. Christian Peest, J. Schmidt, R. Brendel, K. Bothe, M. R. Vogt, I. Kröger, S. Winter, A. Schirmacher, S. Lim, H. T. Nguyen, and D. MacDonald, “Uncertainty analysis for the coefficient of band-to-band absorption of crystalline silicon,” *AIP Advances* **5**, 067168 (2015).
- [113] J. R. de Lasson, L. H. Frandsen, P. Gutsche, S. Burger, O. S. Kim, O. Breinbjerg, A. Ivinskaya, F. Wang, O. Sigmund, T. Häyrynen, A. V. Lavrinenko, J. Mørk, and N. Gregersen, “Benchmarking five numerical simulation techniques for computing resonance wavelengths and quality factors in photonic crystal membrane line defect cavities,” *Opt. Express* **26**, 11366–11392 (2018).
- [114] P. Lalanne, W. Yan, A. Gras, C. Sauvan, J.-P. Hugonin, M. Besbes, G. Demésy, M. D. Truong, B. Gralak, F. Zolla, A. Nicolet, F. Binkowski, L. Zschiedrich, S. Burger, J. Zimmerling, R. Remis, P. Urbach, H. T. Liu, and T. Weiss, “Quasinormal mode solvers for resonators with dispersive materials,” *J. Opt. Soc. Am. A* **36**, 686–704 (2019).
- [115] N. Anttu, H. Mäntynen, T. Sadi, A. Matikainen, J. Turunen, and H. Lipsanen, “Comparison of absorption simulation in semiconductor nanowire and nanocone arrays with the Fourier modal method, the finite element method, and the finite-difference time-domain method,” *Nano Express* **1**, 030034 (2020).
- [116] See official JCMSuite documentation at: <https://www.docs.jcmwave.com/JCMSuite/html/>.
- [117] J. Pomplun, S. Burger, L. Zschiedrich, and F. Schmidt, “Adaptive finite element method for simulation of optical nano structures,” *Phys. Stat. Sol.* **244**, 3419 – 3434 (2007).
- [118] M. Celebrano, X. Wu, M. Baselli, S. Grossmann, P. Biagioni, A. Locatelli, C. De Angelis, G. Cerullo, R. Osellame, B. Hecht, L. Duo, F. Ciccacci, and M. Finazzi, “Mode matching in multiresonant plasmonic nanoantennas for enhanced second harmonic generation,” *Nat. Nanotechnol.* **10**, 412–417 (2015).
- [119] R. Chandrasekar, N. K. Emani, A. Lagutchev, V. M. Shalaev, C. Ciraci, D. R. Smith, and A. V. Kildishev, “Second harmonic generation with plasmonic metasurfaces: direct comparison of electric and magnetic resonances,” *Opt. Mat. Expr.* **5**, 2682–2691 (2015).
- [120] Y. R. Shen, “Surface contribution versus bulk contribution in surface nonlinear optical spectroscopy,” *Appl. Phys. B* **68**, 295–300 (1999).
- [121] S. Janz and Z. H. Lu, “Interband resonances in the optical second-harmonic response of the (001) GaAs–oxide interface,” *J. Opt. Soc. Am. B* **14**, 1647–1650 (1997).

- [122] R. Sanatinia, S. Anand, and M. Swillo, "Modal engineering of second-harmonic generation in single GaP nanopillars," *Nano Lett.* **14**, 5376–5381 (2014).
- [123] V. Lazarescu, M. F. Lazarescu, H. Jones, and W. Schmickler, "Second harmonic generation at the GaAs(111)-BIsolution interface," *J. Electroanal. Chem.* **567**, 257–261 (2004).
- [124] V. Lazarescu, R. Scurtu, M.-F. Lazarescu, E. Santos, H. Jones, and W. Schmickler, "Field effects and surface states in second harmonic generation at n-GaAs(*hkl*) electrodes," *Electrochim. Acta* **50**, 4830–4836 (2005).
- [125] R. Scurtu, N. I. Ionescu, M. Lazarescu, and V. Lazarescu, "Surface states- and field-effects at p- and n-doped GaAs(111)A/solution interface," *Phys. Chem. Chem. Phys.* **11**, 1765–1770 (2009).
- [126] S. Deckoff-Jones, J. Zhang, C. E. Petoukhoff, M. K. Man, S. Lei, R. Vajtai, P. M. Ajayan, D. Talbayev, J. Madéo, and K. M. Dani, "Observing the interplay between surface and bulk optical nonlinearities in thin van der Waals crystals," *Sci. Rep.* **6**, 22620 (2016).
- [127] T. Stehlin, M. Feller, P. Guyot-Sionnest, and Y. R. Shen, "Optical second-harmonic generation as a surface probe for noncentrosymmetric media," *Opt. Lett.* **13**, 389–391 (1988).
- [128] R. W. Hollering, "Bulk and surface second-harmonic generation in noncentrosymmetric semiconductors," *Opt. Commun.* **90**, 147–150 (1992).
- [129] M. Takebayashi, G. Mizutani, and S. Ushioda, "Azimuthal angle dependence of optical second harmonic intensity from a vicinal GaAs(001) wafer," *Opt. Commun.* **133**, 116–122 (1997).
- [130] C. Yamada and T. Kimura, "Anisotropy in second-harmonic generation from reconstructed surfaces of GaAs," *Phys. Rev. Lett.* **70**, 2344–2347 (1993).
- [131] C. Yamada and T. Kimura, "Rotational symmetry of the surface second-harmonic generation of zinc-blende-type crystals," *Phys. Rev. B* **49**, 14372–14381 (1994).
- [132] R. Camacho-Morales, M. Rahmani, S. Kruk, L. Wang, L. Xu, D. A. Smirnova, A. S. Solntsev, A. Miroshnichenko, H. H. Tan, F. Karouta, S. Naureen, K. Vora, L. Carletti, C. De Angelis, C. Jagadish, Y. S. Kivshar, and D. N. Neshev, "Nonlinear generation of vector beams from AlGaAs nanoantennas," *Nano Lett.* **16**, 7191–7197 (2016).
- [133] L. Carletti, A. Locatelli, D. Neshev, and C. De Angelis, "Shaping the radiation pattern of second harmonic generation from AlGaAs dielectric nanoantennas," *ACS Photonics* **3**, 1500–1507 (2016).
- [134] F. J. F. Löchner, A. N. Fedotova, S. Liu, G. A. Keeler, G. M. Peake, S. Saravi, M. R. Shcherbakov, S. Burger, A. A. Fedyanin, I. Brener, T. Pertsch, F. Setzpfandt, and I. Staude, "Polarization-dependent second harmonic diffraction from resonant GaAs metasurfaces," *ACS Photonics* **5**, 1786–1793 (2018).

- [135] S. Gehrsitz, F. K. Reinhart, C. Gourgon, N. Herres, A. Vonlanthen, and H. Sigg, “The refractive index of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ below the band gap: Accurate determination and empirical modeling,” *J. Appl. Phys.* **87**, 7825–7837 (2000).
- [136] G. Bautista, M. J. Huttunen, J. Makitalo, J. M. Kontio, J. Simonen, and M. Kauranen, “Second-harmonic generation imaging of metal nano-objects with cylindrical vector beams,” *Nano Lett.* **12**, 3207–3212 (2012).
- [137] X. Wang, S. Fardad, S. Das, A. Salandrino, and R. Hui, “Direct observation of bulk second-harmonic generation inside a glass slide with tightly focused optical fields,” *Phys. Rev. B* **93**, 161109(R) (2016).
- [138] X. Wang, S. Shen, J. Sun, F. Fan, and S. Chang, “Surface and bulk second-harmonic responses from a glass slide using tightly focused radially polarized light,” *Opt. Lett.* **41**, 5652–5655 (2016).
- [139] P. R. Wiecha, A. Arbouet, C. Girard, T. Baron, and V. Pailard, “Origin of second-harmonic generation from individual silicon nanowires,” *Phys. Rev. B* **93**, 125421 (2016).
- [140] B. Huo, X. Wang, S. Chang, M. Zeng, and G. Zhao, “Second harmonic generation of individual centrosymmetric sphere excited by a tightly focused beam,” *J. Opt. Soc. Am. B* **28**, 2702–2711 (2011).
- [141] B. Huo, X. Wang, S. Chang, and M. Zeng, “Second harmonic generation of a single centrosymmetric nanosphere illuminated by tightly focused cylindrical vector beams,” *J. Opt. Soc. Am. B* **29**, 1631–1640 (2012).
- [142] M. Lax, W. H. Louisell, and W. B. McKnight, “From Maxwell to paraxial wave optics,” *Phys. Rev. A* **11**, 1365–1370 (1975).
- [143] L. W. Davis, “Theory of electromagnetic beams,” *Phys. Rev. A* **19**, 1177–1179 (1979).
- [144] J. P. Barton and D. R. Alexander, “Fifth-order corrected electromagnetic field components for a fundamental Gaussian beam,” *J. Appl. Phys.* **66**, 2800–2802 (1989).
- [145] G. Wang and J. F. Webb, “Calculation of electromagnetic field components for a fundamental Gaussian beam,” *Phys. Rev. E* **72**, 046501 (2005).
- [146] Y. I. Salamin, “Fields of a Gaussian beam beyond the paraxial approximation,” *Appl. Phys. B* **86**, 319–326 (2007).
- [147] Y. I. Salamin, “Fields of a focused linearly polarized Gaussian beam: truncated series versus the complex-source-point spherical-wave representation,” *Opt. Lett.* **34**, 683–685 (2009).
- [148] H. Luo, S. Liu, Z. Lin, and C. T. Chan, “Method for accurate description of a radially polarized Gaussian laser beam beyond the paraxial approximation,” *Opt. Lett.* **32**, 1692–1694 (2007).

- [149] V. Gavrilenko and F. Rebenrost, “Nonlinear optical susceptibility of the surfaces of silicon and diamond,” *Surface Science* **331-333**, 1355 – 1360 (1995).
- [150] H. Dehez, A. April, and M. Piche, “Needles of longitudinally polarized light: guidelines for minimum spot size and tunable axial extent,” *Opt. Expr.* **20**, 14891–14905 (2012).
- [151] L. Turquet, X. Zang, J.-P. Kakko, H. Lipsanen, G. Bautista, and M. Kauranen, “Demonstration of longitudinally polarized optical needles,” *Opt. Expr.* **26**, 27572–27584 (2018).
- [152] L. Turquet, J.-P. Kakko, X. Zang, L. Naskali, L. Karvonen, H. Jiang, T. Huutio, E. Kauppinen, H. Lipsanen, M. Kauranen, and G. Bautista, “Tailorable second-harmonic generation from an individual nanowire using spatially phase-shaped beams,” *Laser & Photon. Rev.* **11**, 1600175 (2016).
- [153] R. Dorn, S. Quabis, and G. Leuchs, “Sharper focus for a radially polarized light beam,” *Phys. Rev. Lett.* **91**, 233901 (2003).
- [154] Y. I. Salamin, “Fields of a radially polarized Gaussian laser beam beyond the paraxial approximation,” *Opt. Lett.* **31**, 2619–2621 (2006).
- [155] S. Yan and B. Yao, “Accurate description of a radially polarized Gaussian beam,” *Phys. Rev. A* **77**, 023827 (2008).
- [156] D. P. Biss and T. G. Brown, “Polarization-vortex-driven second-harmonic generation,” *Opt. Lett.* **28**, 923–925 (2003).
- [157] S. N. Khonina, S. V. Alferov, and S. V. Karpeev, “Strengthening the longitudinal component of the sharply focused electric field by means of higher-order laser beams,” *Opt. Lett.* **38**, 3223–3226 (2013).
- [158] M. Wu, R. Chen, J. Ling, Z. Chen, X. Chen, R. Ji, and M. Hong, “Creation of a longitudinally polarized photonic nanojet via an engineered microsphere,” *Opt. Lett.* **42**, 1444–1447 (2017).
- [159] M. Centini, C. Sibilia, M. Scalora, G. D’Aguanno, M. Bertolotti, M. J. Bloemer, C. M. Bowden, and I. Nefedov, “Dispersive properties of finite, one-dimensional photonic band gap structures: Applications to nonlinear quadratic interactions,” *Phys. Rev. E* **60**, 4891–4898 (1999).
- [160] G. D’Aguanno, M. Centini, C. Sibilia, M. Bertolotti, M. Scalora, M. J. Bloemer, and C. M. Bowden, “Enhancement of $\chi^{(2)}$ cascading processes in one-dimensional photonic bandgap structures,” *Opt. Lett.* **24**, 1663–1665 (1999).
- [161] C. D. Angelis, F. Gringoli, M. Midrio, D. Modotto, J. S. Aitchison, and G. F. Nalesso, “Conversion efficiency for second-harmonic generation in photonic crystals,” *J. Opt. Soc. Am. B* **18**, 348–351 (2001).

- [162] G. D'Aguanno, M. Centini, M. Scalora, C. Sibilia, Y. Dumeige, P. Vidakovic, J. A. Levenson, M. J. Bloemer, C. M. Bowden, J. W. Haus, and M. Bertolotti, "Photonic band edge effects in finite structures and applications to $\chi^{(2)}$ interactions," *Phys. Rev. E* **64**, 016609 (2001).
- [163] Y. Dumeige, I. Sagnes, P. Monnier, P. Vidakovic, I. Abram, C. Mériadec, and A. Levenson, "Phase-matched frequency doubling at photonic band edges: Efficiency scaling as the fifth power of the length," *Phys. Rev. Lett.* **89**, 043901 (2002).
- [164] T. Ochiai and K. Sakoda, "Scaling law of enhanced second-harmonic generation in finite Bragg stacks," *Opt. Express* **13**, 9094–9114 (2005).
- [165] M. Liscidini, A. Locatelli, L. C. Andreani, and C. De Angelis, "Maximum-exponent scaling behavior of optical second-harmonic generation in finite multilayer photonic crystals," *Phys. Rev. Lett.* **99**, 053907 (2007).
- [166] L.-M. Zhao, B.-Y. Gu, and Y.-S. Zhou, "A way for enhancing second harmonic generation in one-dimensional nonlinear photonic crystals," *Optics Communications* **281**, 2954 – 2958 (2008).
- [167] L.-M. Zhao, C. Li, Y.-S. Zhou, and F.-H. Wang, "Multiple wavelength second-harmonic generation in one-dimensional nonlinear photonic crystals," *J. Opt. Soc. Am. B* **25**, 2010–2014 (2008).
- [168] M. Liscidini and L. C. Andreani, "Highly efficient second-harmonic generation in doubly resonant planar microcavities," *Applied Physics Letters* **85**, 1883–1885 (2004).
- [169] M. Liscidini and L. C. Andreani, "Second-harmonic generation in doubly resonant microcavities with periodic dielectric mirrors," *Phys. Rev. E* **73**, 016613 (2006).

Appendix A

Generalized boundary conditions for an anisotropic medium

Generalized boundary conditions Eq. (2.2) were derived upon assuming the medium is isotropic. Here I aim to extend Eq. (2.2) for the case of surfaces of optically anisotropic crystals. In an anisotropic crystal the electric displacement vector \vec{D} can be written in the following form:

$$D_j(\vec{r}, 2\omega) = \varepsilon_0 \varepsilon_{jk}^{(2\omega)}(\vec{r}) E_k(\vec{r}, 2\omega) + P_{S,j}^{\text{NL}}(x', y') \cdot \delta(z'), \quad (\text{A.1})$$

where x', y', z' are assumed to be the local coordinate system at certain point on the interface and z' -axis is pointing along the surface normal (see Fig. A.1). The first term in Eq. (A.1) represents usual linear displacement in the anisotropic medium with dielectric permittivity tensor ε_{jk} , while the second term corresponds to the surface nonlinear polarization.

Following Heinz [88], one can state that the continuity of the normal component of the electric displacement vector D_{\perp} implies:

$$\int_{0^-}^{0^+} D_{\perp} dz' = 0,$$

what gives:

$$\varepsilon_0 \varepsilon_{jk}^{(2\omega)} n_j \int_{0^-}^{0^+} E_k dz' = -P_{S,z'}^{\text{NL}} \quad (\text{A.2})$$

From the continuity of the tangential component of the electric field across the interface it follows that:

$$\int_{0^-}^{0^+} \vec{E}_{\parallel} dz' = \int_{0^-}^{0^+} \left(\vec{E} - (\vec{E}\vec{n})\vec{n} \right) dz' = \vec{0}.$$

This vectorial expression is equivalent to the set of three scalar equations:

$$K \cdot \vec{u} = \vec{0}, \quad (\text{A.3})$$

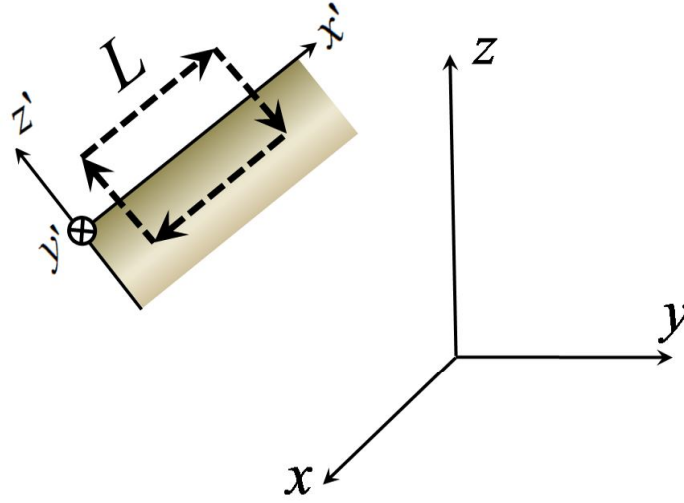


Figure A.1: Integration contour L at certain point of the surface of the nonlinear slab; local coordinate system $x'y'z'$ for the selected point and global coordinate system xyz are depicted.

with

$$K = \begin{bmatrix} (1-n_x^2) & -n_x n_y & -n_x n_z \\ -n_x n_y & (1-n_y^2) & -n_y n_z \\ -n_x n_z & -n_y n_z & (1-n_z^2) \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix},$$

where:

$$s_x = \int_{0^-}^{0^+} E_x dz', \quad s_y = \int_{0^-}^{0^+} E_y dz', \quad s_z = \int_{0^-}^{0^+} E_z dz'. \quad (\text{A.4})$$

It is worth noting, that the set of equations Eq. (A.3) is overdetermined. Thus one can now bring together Eq. (A.2) and, for example, two first equations in Eq. (A.3) to get the closed set of equations for unknowns s_x , s_y and s_z :

$$M \cdot \vec{u} = \vec{v}, \quad (\text{A.5})$$

where

$$M = \begin{bmatrix} \varepsilon_0 \varepsilon_{jx}^{(2\omega)} n_j & \varepsilon_0 \varepsilon_{jy}^{(2\omega)} n_j & \varepsilon_0 \varepsilon_{jz}^{(2\omega)} n_j \\ (1-n_x^2) & -n_x n_y & -n_x n_z \\ -n_x n_y & (1-n_y^2) & -n_y n_z \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} -P_{S,z'}^{\text{NL}} \\ 0 \\ 0 \end{bmatrix}. \quad (\text{A.6})$$

If one denotes Δ^M the determinant of the matrix M : $\Delta^M = \det M$, and $\Delta_{i,j}^M$ the minors of the matrix M , then the solution of Eq. (A.6) according to the Cramer's rule yields:

$$s_j = (-1)^j \cdot \frac{\Delta_{1,j}^M}{\Delta^M} \cdot P_{S,z'}^{\text{NL}}. \quad (\text{A.7})$$

I introduce now the jumps of the electric field components across the interface:

$$\Delta E_x = E_x(z' = 0^+) - E_x(z' = 0^-),$$

$$\Delta E_y = E_y(z' = 0^+) - E_y(z' = 0^-).$$

If now the contour integral over the contour L is considered, one finds:

$$\Delta E_{x'} = \frac{\partial}{\partial x'} \int_{0^-}^{0^+} \vec{E} \vec{n} dz' = \frac{\partial}{\partial x'} s_j n_j, \quad \Delta E_{y'} = \frac{\partial}{\partial y'} \int_{0^-}^{0^+} \vec{E} \vec{n} dz' = \frac{\partial}{\partial y'} s_j n_j.$$

Using Eq. (A.6) and Eq. (A.7) the above equations can be rewritten as:

$$\Delta \vec{E}_{\parallel} = (-1)^j \cdot \frac{n_j \Delta_{1,j}^M}{\Delta^M} \cdot \vec{\nabla}_{\parallel} P_{S,z'}^{\text{NL}}. \quad (\text{A.8})$$

From Eq. (A.6) one can directly find:

$$\begin{aligned} \Delta_{1,1}^M &= n_x n_z; & \Delta_{1,2}^M &= -n_y n_z; & \Delta_{1,3}^M &= n_z^2; \\ \Delta^M &= \varepsilon_0 n_z (\varepsilon_{jx}^{(2\omega)} n_x n_j + \varepsilon_{jy}^{(2\omega)} n_y n_j + \varepsilon_{jz}^{(2\omega)} n_z n_j). \end{aligned}$$

Now Eq. (A.8) simplifies to:

$$\Delta \vec{E}_{\parallel} = - \frac{1}{\varepsilon_0 (\varepsilon_{jx}^{(2\omega)} n_x n_j + \varepsilon_{jy}^{(2\omega)} n_y n_j + \varepsilon_{jz}^{(2\omega)} n_z n_j)} \cdot \vec{\nabla}_{\parallel} P_{S,z'}^{\text{NL}}. \quad (\text{A.9})$$

Specifically, if the global coordinate system x, y, z coincides with the crystal axis, one gets:

$$\Delta \vec{E}_{\parallel} = - \frac{1}{\varepsilon_0 (\varepsilon_{xx}^{(2\omega)} n_x^2 + \varepsilon_{yy}^{(2\omega)} n_y^2 + \varepsilon_{zz}^{(2\omega)} n_z^2)} \cdot \vec{\nabla}_{\parallel} P_{S,z'}^{\text{NL}}. \quad (\text{A.10})$$

Finally, in the case of optically isotropic medium, i.e. $\varepsilon_{xx}^{(2\omega)} = \varepsilon_{yy}^{(2\omega)} = \varepsilon_{zz}^{(2\omega)}$, one arrives to the first equation in Eq. (2.2).

It is important to notice here that expression Eq. (A.9) is readily generalized for the case of the interface between two anisotropic media. Indeed, in the derivation above the integration was performed over the ultrathin surface layer which is fully located inside the anisotropic medium. This reasoning is quite natural for an interface between an anisotropic medium and vacuum. On the other hand, when an interface between two adjacent anisotropic media is considered, the choice of the medium where the surface layer is located is in principle arbitrary. Then the actual jump of the electric field across the interface would be still given by expression Eq. (A.9) but one has to take into account that the surface nonlinear tensor must be rescaled accordingly with the choice of the medium. It is also worth noting here that in the case of two adjacent anisotropic media the simple representation Eq. (2.27) is not valid and the full third-rank tensor of surface second-order nonlinearity must be taken.

The rescaling of the surface nonlinear tensor depends in general on two issues. Firstly, one has to fix on which side from the interface the pump field for the calculation of the surface nonlinear

polarization is evaluated. Specifically, the usual boundary conditions for the fundamental (pump) wave give the following equalities connecting the field on both sides from the interface:

$$\begin{aligned} E_x^{(\omega)}(z' = 0^+) &= E_x^{(\omega)}(z' = 0^-), \\ E_y^{(\omega)}(z' = 0^+) &= E_y^{(\omega)}(z' = 0^-), \\ \varepsilon_{jk}^{(\omega),+} n_j E_k^{(\omega)}(z' = 0^+) &= \varepsilon_{jk}^{(\omega),-} n_j E_k^{(\omega)}(z' = 0^-), \end{aligned} \quad (\text{A.11})$$

where $\varepsilon_{jk}^{(\omega),+}$ and $\varepsilon_{jk}^{(\omega),-}$ stand for the dielectric permittivity tensors of both media at the pump frequency and the last expression in Eq. (A.11) yields:

$$\begin{aligned} E_z^{(\omega)}(z' = 0^+) &= \frac{\varepsilon_{jz}^{(\omega),-} n_j}{\varepsilon_{jz}^{(\omega),+} n_j} E_z^{(\omega)}(z' = 0^-) + \frac{n_j (\varepsilon_{jx}^{(\omega),-} - \varepsilon_{jx}^{(\omega),+})}{\varepsilon_{jz}^{(\omega),+} n_j} E_x^{(\omega)}(z' = 0^-) \\ &+ \frac{n_j (\varepsilon_{jy}^{(\omega),-} - \varepsilon_{jy}^{(\omega),+})}{\varepsilon_{jz}^{(\omega),+} n_j} E_y^{(\omega)}(z' = 0^-). \end{aligned} \quad (\text{A.12})$$

If one wants to use the field on another side from the interface, the surface nonlinear tensor should be changed respectively according to Eqs. (A.11)-(A.12) so that the final expression for the surface nonlinear polarization stays the same. Secondly, the sheet of nonlinear polarization can be placed on any side from the interface. From the above Eq. (A.9) it follows that the components $\chi_{\perp jk}^{(2)}$ of the surface nonlinear tensor should be multiplied by the factor:

$$\kappa = \frac{\varepsilon_{jx}^{(2\omega),-} n_x n_j + \varepsilon_{jy}^{(2\omega),-} n_y n_j + \varepsilon_{jz}^{(2\omega),-} n_z n_j}{\varepsilon_{jx}^{(2\omega),+} n_x n_j + \varepsilon_{jy}^{(2\omega),+} n_y n_j + \varepsilon_{jz}^{(2\omega),+} n_z n_j},$$

if one wants to move the polarization sheet from $z' = 0^-$ to $z' = 0^+$ and vice versa.

Appendix B

First-order correction terms for the effective surface nonlinear tensor $\chi_{\Sigma,ijk}^{(2)}$

For TE-polarized incident wave at fundamental frequency combining Eqs. (2.40) and (2.45) and keeping zeroth- and first-order small terms one finds:

$$\begin{aligned}
 \chi_{\Sigma,\perp\parallel\parallel}^{(2)} = & \sum_{m=0}^N \chi_{\perp\parallel\parallel,m(m+1)}^{(2)} \cdot \frac{\epsilon_{up}}{\epsilon_m} + \sum_{m=0}^N \chi_{\perp\parallel\parallel,m(m+1)}^{(2)} \cdot \frac{\epsilon_{up}}{\epsilon_m} \cdot \\
 & \left[\frac{4iZ_{sub,FW}'^2 Z_{0,FW}'}{(1+R_{TE}^0)(Z_{sub,FW}' + Z_{0,FW}')^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{FW} h_j}{Z_{j,FW}'} - \frac{4iZ_{0,FW}'}{(1+R_{TE}^0)(Z_{sub,FW}' + Z_{0,FW}')^2} \cdot \right. \\
 & \left. \sum_{j=1}^N Z_{j,FW}' k_{j,z}^{FW} h_j + 2i \frac{1-R_{TE}^0}{(1+R_{TE}^0)Z_{0,FW}'} \cdot \sum_{j=1}^m Z_{j,FW}' k_{j,z}^{FW} h_j \right] - \\
 & iZ_{sub}'' \cdot \sum_{m=0}^{N-1} \chi_{\perp\parallel\parallel,m(m+1)}^{(2)} \cdot \frac{\epsilon_{up}}{\epsilon_m} \cdot \sum_{j=m+1}^N \frac{k_{j,z}^{SH} h_j}{Z_j''} + \\
 & \frac{i}{Z_0'' + Z_{sub}''} \cdot \sum_{m=0}^N \chi_{\perp\parallel\parallel,m(m+1)}^{(2)} \cdot \frac{\epsilon_{up}}{\epsilon_m} \cdot \sum_{m=0}^N k_{m,z}^{SH} h_m \left(Z_m'' + \frac{Z_0'' Z_{sub}''}{Z_m''} \right). \tag{B.1}
 \end{aligned}$$

Here it is assumed for simplicity, that all parameters without "FW" subscript are taken at the second-harmonic frequency. The first term in Eq. (B.1) represents the zeroth-order effective nonlinear tensor from Eq. (2.48), while the rest terms give the first-order correction.

For TM-polarized incident wave at the fundamental frequency combining Eqs. (2.40) and (2.47)

and keeping only zeroth-order and first-order small terms one obtains:

$$\begin{aligned}
\chi_{\Sigma,\perp\perp\perp}^{(2)} = & \sum_{m=0}^N \chi_{\perp\perp\perp,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} \cdot \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \right)^2 + \sum_{m=0}^N \chi_{\perp\perp\perp,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} \cdot \\
& \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \right)^2 \cdot \left[- \frac{4iZ_{\text{sub}}''^2 Z_0''}{\varepsilon_m (1 - R_{\text{TM}}^0) (Z_{\text{sub}}'' + Z_0'')^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{SH}} h_j}{Z_j''} + \right. \\
& \frac{4iZ_0''}{\varepsilon_m (1 - R_{\text{TM}}^0) (Z_{\text{sub}}'' + Z_0'')^2} \cdot \sum_{j=1}^N Z_j'' k_{j,z}^{\text{SH}} h_j + \frac{2iZ_0'' (1 + R_{\text{TM}}^0)}{\varepsilon_m (1 - R_{\text{TM}}^0)} \cdot \sum_{j=1}^m \frac{k_{j,z}^{\text{SH}} h_j}{Z_j''} \left. \right] - \\
& iZ_{\text{sub}}'' \cdot \sum_{m=0}^{N-1} \chi_{\perp\perp\perp,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} \cdot \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \right)^2 \cdot \sum_{j=m+1}^N \frac{k_{j,z}^{\text{SH}} h_j}{Z_j''} + \\
& \frac{i}{Z_0'' + Z_{\text{sub}}''} \cdot \sum_{m=0}^N \chi_{\perp\perp\perp,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{up}}}{\varepsilon_m} \cdot \left(\frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \right)^2 \cdot \sum_{m=0}^N k_{m,z}^{\text{SH}} h_m \left(Z_m'' + \frac{Z_0'' Z_{\text{sub}}''}{Z_m''} \right);
\end{aligned}$$

$$\begin{aligned}
\chi_{\Sigma,\parallel\perp\parallel}^{(2)} = & \sum_{m=0}^N \chi_{\parallel\perp\parallel,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} + \sum_{m=0}^N \chi_{\parallel\perp\parallel,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \cdot \\
& \left[- \frac{2iZ_{\text{sub,FW}}''^2 Z_{0,\text{FW}}''}{\varepsilon_m (1 - R_{\text{TM}}^0) (Z_{\text{sub,FW}}'' + Z_{0,\text{FW}}'')^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z_{j,\text{FW}}''} + \right. \\
& \frac{2iZ_{0,\text{FW}}''}{\varepsilon_m (1 - R_{\text{TM}}^0) (Z_{\text{sub,FW}}'' + Z_{0,\text{FW}}'')^2} \cdot \sum_{j=1}^N Z_{j,\text{FW}}'' k_{j,z}^{\text{FW}} h_j + \\
& \frac{iZ_{0,\text{FW}}'' (1 + R_{\text{TM}}^0)}{\varepsilon_m (1 - R_{\text{TM}}^0)} \cdot \sum_{j=1}^m \frac{k_{j,z}^{\text{FW}} h_j}{Z_{j,\text{FW}}''} + \frac{2iZ_{\text{sub,FW}}''^2 Z_{0,\text{FW}}''}{(1 + R_{\text{TM}}^0) (Z_{\text{sub,FW}}'' + Z_{0,\text{FW}}'')^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z_{j,\text{FW}}''} - \\
& \frac{2iZ_{0,\text{FW}}''}{(1 + R_{\text{TM}}^0) (Z_{\text{sub,FW}}'' + Z_{0,\text{FW}}'')^2} \cdot \sum_{j=1}^N Z_{j,\text{FW}}'' k_{j,z}^{\text{FW}} h_j + \\
& \left. i \frac{1 - R_{\text{TM}}^0}{(1 + R_{\text{TM}}^0) Z_{0,\text{FW}}''} \cdot \sum_{j=1}^m Z_{j,\text{FW}}'' k_{j,z}^{\text{FW}} h_j \right] - \\
& \frac{i}{Z_{\text{sub}}''} \cdot \sum_{m=0}^{N-1} \chi_{\parallel\perp\parallel,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \cdot \sum_{j=m+1}^N Z_j'' k_{j,z}^{\text{SH}} h_j + \\
& \frac{i}{Z_0'' + Z_{\text{sub}}''} \cdot \sum_{m=0}^N \chi_{\parallel\perp\parallel,m(m+1)}^{(2)} \cdot \frac{\varepsilon_{\text{sub,FW}}}{\varepsilon_{m,\text{FW}}} \cdot \sum_{m=0}^N k_{m,z}^{\text{SH}} h_m \left(Z_m'' + \frac{Z_0'' Z_{\text{sub}}''}{Z_m''} \right);
\end{aligned}$$

$$\begin{aligned}
\chi_{\Sigma,\perp||}^{(2)} &= \sum_{m=0}^N \chi_{\perp||}^{(2),m(m+1)} \cdot \frac{\epsilon_{\text{up}}}{\epsilon_m} + \sum_{m=0}^N \chi_{\perp||}^{(2),m(m+1)} \cdot \frac{\epsilon_{\text{up}}}{\epsilon_m} \cdot \\
&\left[\frac{4iZ_{\text{sub},\text{FW}}''^2 Z_{0,\text{FW}}''}{(1+R_{\text{TM}}^0)(Z_{\text{sub},\text{FW}}'' + Z_{0,\text{FW}}'')^2} \cdot \sum_{j=1}^N \frac{k_{j,z}^{\text{FW}} h_j}{Z_{j,\text{FW}}''} - \frac{4iZ_{0,\text{FW}}''}{(1+R_{\text{TM}}^0)(Z_{\text{sub},\text{FW}}'' + Z_{0,\text{FW}}'')^2} \cdot \right. \\
&\quad \left. \sum_{j=1}^N Z_{j,\text{FW}}'' k_{j,z}^{\text{FW}} h_j + 2i \frac{1-R_{\text{TM}}^0}{(1+R_{\text{TM}}^0)Z_{0,\text{FW}}''} \cdot \sum_{j=1}^m Z_{j,\text{FW}}'' k_{j,z}^{\text{FW}} h_j \right] - \\
&\quad iZ_{\text{sub}}'' \cdot \sum_{m=0}^{N-1} \chi_{\perp||}^{(2),m(m+1)} \cdot \frac{\epsilon_{\text{up}}}{\epsilon_m} \cdot \sum_{j=m+1}^N \frac{k_{j,z}^{\text{SH}} h_j}{Z_j''} + \\
&\quad \frac{i}{Z_0'' + Z_{\text{sub}}''} \cdot \sum_{m=0}^N \chi_{\perp||}^{(2),m(m+1)} \cdot \frac{\epsilon_{\text{up}}}{\epsilon_m} \cdot \sum_{m=0}^N k_{m,z}^{\text{SH}} h_m \left(Z_m'' + \frac{Z_0'' Z_{\text{sub}}''}{Z_m''} \right).
\end{aligned} \tag{B.2}$$

Appendix C

Surface SHG from a dielectric sphere under VSH illumination

I assume a dielectric sphere made of a nonlinear material and illuminated by the FW, which represents a vector spherical harmonic (VSH). The material is assumed to be centrosymmetric, so that the bulk nonlinearity can be neglected and only surface nonlinearity should be taken into account. I start here with the case $\chi_{\perp||}^{(2)} \neq 0$, and all other components of the surface nonlinear susceptibility tensor equal to zero. Two different vector spherical harmonics M_{11} and N_{11} are considered below. Solution for any other VSH could be found in a similar way.

Incident FW vector spherical harmonic N_{11} is given as:

$$N_{11}^{FW} = \sqrt{\frac{3}{16\pi}} \frac{1}{kr} \frac{\partial(r \cdot j_1(kr))}{\partial r} e^{i\varphi} (i \cdot \vec{e}_\varphi + \cos\theta \cdot \vec{e}_\theta),$$

where k is the wavenumber in the outer medium (vacuum or air). Nonrelevant normal component of the electric field, which is also nonzero, was skipped for clarity. The field inside the sphere is given by the same function, but with the wavenumber k replaced by the wavenumber k_1 inside the sphere with the permittivity ε :

$$k_1 = k\sqrt{\varepsilon}.$$

The amplitude of the incident field is assumed equal to unity, then the field inside the sphere of radius R :

$$E^{FW} = a^{FW} N_{11}^{FW},$$

where a^{FW} is the amplitude factor. Surface nonlinear polarization is given as:

$$\vec{P}^{NL} = \chi_{\perp||}^{(2)} (E_{||}^{FW})^2 \cdot \vec{e}_r = \chi_{\perp||}^{(2)} (a^{FW} N_{11}^{FW})^2 \cdot \vec{e}_r, \quad (\text{C.1})$$

and the jump of the SH electric field at the interface according to Eq. (2.2):

$$\Delta E_{\parallel} = -\nabla_{\parallel} P_{\perp}^{NL} = -\frac{3\chi_{\perp\parallel\parallel}^{(2)}}{16\pi} \left(\frac{a^{FW}}{k_1^{FW} R} \frac{\partial(r \cdot j_1(k_1 r))}{\partial r} \Big|_{r=R} \right)^2 \cdot \nabla_{\parallel} (e^{2i\varphi} \cdot (-1 + \cos^2 \theta)) =$$

$$\frac{3\chi_{\perp\parallel\parallel}^{(2)}}{16\pi R} \left(\frac{a^{FW}}{k_1^{FW} R} \frac{\partial(r \cdot j_1(k_1^{FW} r))}{\partial r} \Big|_{r=R} \right)^2 \cdot 2 \sin \theta \cdot e^{2i\varphi} \cdot (i \cdot \vec{e}_{\varphi} + \cos \theta \cdot \vec{e}_{\theta}).$$

As the next step, the jump of the electric field has to be expanded into the VSHs. Each term in the expansion would generate the same VSH at the SH frequency. Since the VSHs form an orthonormal function system on the sphere surface, the generalized boundary conditions Eq. (2.2) have to be written separately for each VSH term. In particular, for the considered VSH N_{11} one can see that the field at the SH frequency contains only one vector spherical harmonic, namely:

$$N_{22}^{SH} = \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{1}{k_1^{SH} R} \frac{\partial \left(r \cdot \begin{Bmatrix} j_2(k_1^{SH} r) \\ h_2(k_1^{SH} r) \end{Bmatrix} \right)}{\partial r} \cdot \sin \theta \cdot e^{2i\varphi} (i \cdot \vec{e}_{\varphi} + \cos \theta \cdot \vec{e}_{\theta}).$$

Here spherical Bessel function $j_2(k_1 r)$ describes the field inside the sphere and spherical Hankel function $h_2(kr)$ – the field outside the sphere.

If corresponding amplitude coefficients inside and outside (scattered field) the sphere of radius R are a_{inside}^{SH} and a_{scat}^{SH} respectively, the boundary conditions yield:

$$a_{scat}^{SH} \cdot k^{SH} \cdot h_2(k^{SH} R) = a_{inside}^{SH} \cdot k_1^{SH} \cdot j_2(k_1^{SH} R)$$

$$\frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{a_{scat}^{SH}}{k^{SH} R} \frac{\partial(r h_2(k^{SH} r))}{\partial r} \Big|_{r=R} - \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{a_{inside}^{SH}}{k_1^{SH} R} \frac{\partial(r j_2(k_1^{SH} r))}{\partial r} \Big|_{r=R} =$$

$$\frac{3\chi_{\perp\parallel\parallel}^{(2)}}{8\pi R} \left(\frac{a^{FW}}{k_1^{FW} R} \frac{\partial(r \cdot j_1(k_1^{FW} r))}{\partial r} \Big|_{r=R} \right)^2,$$

what gives for the sought amplitude a_{scat}^{SH} :

$$a_{scat}^{SH} = \frac{3\chi_{\perp\parallel\parallel}^{(2)}}{2R\sqrt{5}\pi} \left(\frac{a^{FW}}{k_1^{FW} R} \frac{\partial(r \cdot j_1(k_1^{FW} r))}{\partial r} \Big|_{r=R} \right)^2 \cdot \frac{j_2(k_1^{SH} R)}{\frac{j_2(k_1^{SH} R)}{k_1^{SH} R} \frac{\partial(r h_2(k^{SH} r))}{\partial r} \Big|_{r=R} - \frac{k^{SH} \cdot h_2(k^{SH} R)}{(k_1^{SH})^2 R} \frac{\partial(r j_2(k_1^{SH} r))}{\partial r} \Big|_{r=R}}.$$

I proceed with the analogous analysis for the incident vector spherical harmonic M_{11} :

$$M_{11}^{FW} = \sqrt{\frac{3}{16\pi}} j_1(k_1 r) e^{i\varphi} (i \cdot \vec{e}_{\theta} - \cos \theta \cdot \vec{e}_{\varphi}).$$

Given that the amplitude of the incident VSH is unity, the field inside the sphere is:

$$E^{FW} = a^{FW} M_{11}^{FW},$$

and the expression for the surface nonlinear polarization is analogous to Eq. (C.1):

$$\vec{P}^{NL} = \chi_{\perp\parallel\parallel}^{(2)} (E_{\parallel}^{FW})^2 \cdot \vec{e}_r = \chi_{\perp\parallel\parallel}^{(2)} (a^{FW} M_{11}^{FW})^2 \cdot \vec{e}_r.$$

The jump of the tangential component of the SH electric field at the interface is now obtained as:

$$\begin{aligned} \Delta E_{\parallel} = -\nabla_{\parallel} P_{\perp}^{NL} &= -\frac{3\chi_{\perp\parallel\parallel}^{(2)}}{16\pi} \left(a^{FW} \cdot j_1(k_1^{FW} R) \right)^2 \cdot \nabla_{\parallel} (e^{2i\varphi} \cdot (-1 + \cos^2 \theta)) = \\ &= \frac{3\chi_{\perp\parallel\parallel}^{(2)}}{16\pi R} \left(a^{FW} \cdot j_1(k_1^{FW} R) \right)^2 \cdot 2 \sin \theta \cdot e^{2i\varphi} \cdot (i \cdot \vec{e}_{\varphi} + \cos \theta \cdot \vec{e}_{\theta}). \end{aligned} \quad (C.2)$$

From Eq. (C.2) one can see that the radiated field at the SH frequency contains again only one vector spherical harmonic N_{22}^{SH} .

If corresponding amplitude coefficients inside and outside (scattered field) the sphere of radius R are again a_{inside}^{SH} and a_{scat}^{SH} respectively, the boundary conditions yield:

$$\begin{aligned} a_{scat}^{SH} \cdot k^{SH} \cdot h_2(k^{SH} R) &= a_{inside}^{SH} \cdot k_1^{SH} \cdot j_2(k_1^{SH} R) \\ \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{a_{scat}^{SH}}{k^{SH} R} \frac{\partial(r h_2(k^{SH} r))}{\partial r} \Big|_{r=R} - \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{a_{inside}^{SH}}{k_1^{SH} R} \frac{\partial(r j_2(k_1^{SH} r))}{\partial r} \Big|_{r=R} &= \frac{3\chi_{\perp\parallel\parallel}^{(2)}}{8\pi R} \left(a^{FW} \cdot j_1(k_1^{FW} R) \right)^2, \end{aligned}$$

thus the sought amplitude a_{scat}^{SH} is obtained:

$$a_{scat}^{SH} = \frac{3\chi_{\perp\parallel\parallel}^{(2)}}{2R\sqrt{5\pi}} \left(a^{FW} \cdot j_1(k_1^{FW} R) \right)^2 \cdot \frac{j_2(k_1^{SH} R)}{\frac{j_2(k_1^{SH} R)}{k^{SH} R} \frac{\partial(r h_2(k^{SH} r))}{\partial r} \Big|_{r=R} - \frac{k^{SH} \cdot h_2(k^{SH} R)}{(k_1^{SH})^2 R} \frac{\partial(r j_2(k_1^{SH} r))}{\partial r} \Big|_{r=R}}.$$

Now let assume the case $\chi_{\perp\parallel\parallel}^{(2)} = 0$, $\chi_{\perp\perp\perp}^{(2)} \neq 0$. Since VSH M_{11} possesses zero normal component of the electric field, illumination with this VSH does not result in any SHG. Therefore one can proceed in this case only with the illumination by the VSH N_{11} . The normal component of the incident VSH N_{11} is given as:

$$N_{11,r}^{FW} = \sqrt{\frac{3}{4\pi}} \frac{j_1(k_1 r)}{k_1^{FW} r} e^{i\varphi} \sin \theta.$$

Expression for the surface nonlinear polarization is analogous to Eq. (C.1) and the jump of the SH electric field at the interface:

$$\Delta E_{\parallel} = -\nabla_{\parallel} P_{\perp}^{NL} = \frac{3\chi_{\perp\perp\perp}^{(2)}}{4\pi R} \left(\frac{a^{FW} \cdot j_1(k_1^{FW} R)}{k_1^{FW} R} \right)^2 \cdot 2 \sin \theta \cdot e^{2i\varphi} \cdot (i \cdot \vec{e}_{\varphi} + \cos \theta \cdot \vec{e}_{\theta}).$$

The boundary conditions in this case give:

$$\begin{aligned} a_{scat}^{SH} \cdot k^{SH} \cdot h_2(k^{SH} R) &= a_{inside}^{SH} \cdot k_1^{SH} \cdot j_2(k_1^{SH} R) \\ \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{a_{scat}^{SH}}{k^{SH} R} \frac{\partial(r h_2(k^{SH} r))}{\partial r} \Big|_{r=R} - \frac{1}{4} \sqrt{\frac{5}{\pi}} \frac{a_{inside}^{SH}}{k_1^{SH} R} \frac{\partial(r j_2(k_1^{SH} r))}{\partial r} \Big|_{r=R} &= \frac{3\chi_{\perp\perp\perp}^{(2)}}{4\pi R} \left(\frac{a^{FW} \cdot j_1(k_1^{FW} R)}{k_1^{FW} R} \right)^2, \end{aligned}$$

and for the sought amplitude a_{scat}^{SH} one finds:

$$a_{scat}^{SH} = \frac{5\chi_{\perp\perp\perp}^{(2)}}{R\sqrt{5\pi}} \left(\frac{a^{FW} \cdot j_1(k_1^{FW} R)}{k_1^{FW} R} \right)^2 \cdot \frac{j_2(k_1^{SH} R)}{\frac{j_2(k_1^{SH} R)}{k^{SH} R} \frac{\partial(r h_2(k^{SH} r))}{\partial r} \Big|_{r=R} - \frac{k^{SH} \cdot h_2(k^{SH} R)}{(k_1^{SH})^2 R} \frac{\partial(r j_2(k_1^{SH} r))}{\partial r} \Big|_{r=R}}.$$

Abbreviations and conventions

Abbreviations

BEM	boundary element method
BIC	bound states in continuum
BP	backward-propagating
CW	continuous wave
ED	electric dipole
FDTD	finite-difference time-domain (method)
FEM	finite-element method
FP	forward-propagating
FW	fundamental wave
MD	magnetic dipole
MIR	mid-infrared (range)
PhC	photonic crystal
PMC	phase-matching condition
PML	perfectly matched layer
Q	quality (factor)
QPM	quasi-phase-matching
SFG	sum-frequency generation
SH	second-harmonic
SHG	second-harmonic generation
SHW	second-harmonic wave
SSHG	surface second-harmonic generation
TE	transverse electric
TFGB	tightly-focused Gaussian beam
TH	third-harmonic
THG	third-harmonic generation
TM	transverse magnetic
TMM	transfer matrix method
UPA	undepleted-pump approximation
UV	ultraviolet (range)
VIS	visible (range)
VSH	vector spherical harmonic

Acknowledgement

My PhD project was mainly supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675745 (MSCA-ITN-EID NOLOSS) as a research collaboration between JCMwave GmbH, Berlin as industrial partner and Friedrich-Schiller-Universität Jena as academic partner. Thanks to funding of NOLOSS project I had a unique chance to take advantage of scientific environment both in Berlin and in Jena and to gain a really great research experience.

Firstly, I want to express my thanks to all my supervisors within the whole duration of the project both from Berlin and Jena. I am very grateful to Prof. Sven Burger for his significant scientific, organisational and administrative support during all time that I spent in Berlin. I am also very thankful to Prof. Thomas Pertsch for his all-embracing support during my employment in FSU Jena and to Dr. Frank Setzpfandt for fruitful scientific supervision and helping me with ideas on proceeding my research.

I also have to thank several people from research groups both in Berlin and Jena, who have especially contributed to my work. First of all I have to mention Dr. Martin Hammerschmidt from JCMwave GmbH for his coding of the developed FEM for the modeling of the surface SHG into JCMSuite software tool and helping me with the debugging and testing. I have to also thank Dr. Lin Zschiedrich from JCMwave GmbH for his useful advices and support with programming issues. Finally, I am thankful to all members of both Computational Nano Optics research group at Zuse Institute Berlin, team of JCMwave GmbH and Nanooptics group in Abbe Center of Photonics of FSU Jena for the great atmosphere and environment they provided. I was really pleased to meet a lot of very nice people there.

Special thanks must be given to all other members of the NOLOSS project. We had multiple joint meetings and workshops throughout the whole project in different locations across Europe and it was really a pleasure for me to meet all of you every time. I have to also mention here our project coordinator Prof. Toralf Scharf, who did his best trying to coordinate 15 PhD students in multiple countries simultaneously and has greatly helped us to solve whatever administrative issues we have faced.

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

- Dr. Martin Hammerschmidt von JCMwave GmbH war verantwortlich für die Programmierung des FEM in JCMsuite Software.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die geltende Promotionsordnung der Physikalisch-Astronomischen Fakultät ist mir bekannt.

Ich versichere ehrenwörtlich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Ort, Datum

Unterschrift

Curriculum Vitae

Personal information

Name Anton Pakhomov
Date of Birth 01.10.1991
Place of Birth Samara, UdSSR

School education

01.09.1998 - 30.06.2008 Primary and secondary general school № 131
Samara, Russia

University education

01.09.2008 - 30.06.2014 B.Sc. & M.Sc. in Applied Mathematics and Physics
Samara State Aerospace University, Russia
01.04.2017 - 31.03.2021 PhD in Optics
Abbe Center of Photonics
Friedrich Schiller University Jena, Germany

Employment

01.10.2011 - 31.03.2017 Research Fellow
Department of Theoretical Physics
Lebedev Physical Institute, Samara, Russia
01.04.2017 - 31.03.2019 Research Fellow
JCMwave GmbH, Berlin, Germany
01.04.2019 - 31.03.2021 Research Fellow
Abbe Center of Photonics
Friedrich Schiller University Jena, Germany