Effective Numerical Calculation of Second Harmonic Generation Using the Linear Analytic Tetrahedron Method and Piezoelectric Coefficients Using Berry Phase Theory

Anders M. Westerkam, Jesper L. W. Sonne, and Karl G. Danielsen

Department of Materials and Production, Aalborg University Fibigerstraede 16, DK-9220 Aalborg East, Denmark Email: <u>aweste17@student.aau.dk</u>, jsonne17@student.aau.dk, <u>kdanie17@student.aau.dk</u>, Web page: <u>http://www.mechman.mp.aau.dk/</u>

Abstract

This paper examines how the computational load of numerical intergration over the first Brillouin zone (BZ) may be eased by employing the semianalytic integration scheme, namely the linear analytic tetrahedron method (LATM). Furthermore, it is examined how the integration may sometimes be reformulated to allow for faster convergence, here shown by calculating the piezoelectric coefficients utilising density functional theory (DFT). To show the power of the LATM calculations, the linear electronic susceptibility and the second harmonic generation (SHG) response of primarily zinc blende crystals was computed and compared with a simple point sampling scheme. The appropriate matrix elements and eigenvalues was calculated with the empirical pseudo potential (EMP) method. Here it was shown that the LATM can reduce the number of \mathbf{k} -points needed by up to two orders of magnitude for the linear susceptibility, while similar results can be seen for the lower frequencies of the SHG response.

Keywords: Linear Analytic Tetrahedron Method, Efficient Integration Schemes, Density Functional Theory, Second Order Harmonic Generation, Piezoelectrics

1. Introduction

With the advent of the laser in 1960 a resurgence in the study of nonlinear response in materials occurred, which has been ongoing to this day. Modern technology heavily utilises nonlinear properties of materials, from the DC-Pockels effect, used in the modulation of devices such as LCDs and pulsed lasers, to the different types of frequency mixing, such as SHG. These effects rely on the nonlinear electric susceptibility of materials, and as new materials join the roster of electronics in the continued development of the modern world it is crucial to map the properties these materials [1].

To determine the electric susceptibility of bulk crystals, an integral over the first Brillouin zone (BZ) has to be carried out. The integrand of the integral is a quotient consisting of appropriate matrix elements, energy differences, and the exciting frequency [2, 3, 4]. One approach to obtain these constituents is using EMP, where known band structure properties such as the band gap and effective mass is utilised to generate a set of Fourier coefficients from which the band structure and wave functions can be calculated [5]. A more robust alternative is to employ ab initio methods such as DFT to obtain the necessary constituents. However, because the integrand contains numerous poles, which makes it ill defined when tackled as a Riemann sum, hundreds of thousands of k-points have to be used to converge the integral, which is not at all feasible when using DFT to get the eigenvalues and matrix elements. It is thus of great interest to develop numerical integration methods which improves the convergence of these BZ integrals. One such method is the LATM, the effectiveness of which will be examined in this paper for both linear and second order response functions. In the case of the piezoelectric coefficients, BZ integration can be circumvented entirely. This is done by expressing the problem in terms of a Berry connection, which may be calculated along one dimension at a time.

2. Ab Initio Calculations

The corner stone of DFT is the Hohenberg Kohn theorem, which states that there is a one to one correspondence between a system, as described by the potential from the nuclei, and the spatially varying electron density of the ground state, implying that any observable of said system is a functional of the electronic density. In principle, this means that the cumbersome $N \times 4$ dimensional wave function, with N being the amount of electrons, which was previously required to evaluate observables of the system may be dispensed with in favour of the simpler three dimensional electron density [6]. Kohn and Sham continued the work, proposing that a system of non-interacting electrons could produce the same electronic density as the true interacting system, resulting in the KS-equation

$$\left[-\frac{1}{2}\nabla^2 + V_N + V_H + V_{xc}\right]\varphi_i = \epsilon_i\varphi_i,\qquad(1)$$

where φ_i are single electron wave functions. The exact form of the exchange correlation potential, V_{xc} , is unknown, but it must depend on the electronic density $\rho(\mathbf{r})$. An initial guess on the electronic density is therefore necessary to evaluate the potentials in eq. (1), after which φ_i and $\rho(\mathbf{r})$ can be updated. This procedure is then repeated until convergence [7]. This method, although precise, is computationally heavy for even relatively small k-points.

3. Linear Analytic Tetrahedron Method

The LATM was first proposed by Jepsen and Andersen as well as by Lehmann and Taut, in the beginning of the 1970s [8, 9], where the method proved effective at calculating physical properties such as the density of states and fermi-surface. Further work was done by D. J. Moss, J. E. Sipe, and H. M. van Driel in the late 80s, specifically with respect to the application of SHG response functions [10]. The method proposed by Sipe et al. is difficult to implement with a multitude of limiting cases, hence it is of interest to investigate to which degree the LATM developed for linear response functions may be employed for higher orders. Starting with a review of the method [10, 11].

By use of both time reversal symmetry, along with point group symmetry it is possible to reduce the BZ integrals to only be carried out over the IBZ, and the integrals in question will be on the form

$$I(\omega) = \int_{\Omega_{IBZ}} A(\mathbf{k}) \delta(\tilde{E} - \alpha \hbar \omega) d^3 \mathbf{k}, \qquad (2)$$

where $A(\mathbf{k})$ is some function of \mathbf{k} , and $\tilde{E}(\mathbf{k})$ is an appropriate energy difference. For readability, the energies dependence on \mathbf{k} is suppressed. The idea of the LATM is to subdivide the volume of integration into tetrahedral micro cells where \mathbf{k} and \tilde{E} are determined exactly at the corners. It is then assumed that \mathbf{k} and \tilde{E} varies linearly between the corners, after which the integral over each micro cell may be carried out analytically. To do this the integral is rewritten as a surface integral over surfaces of constant energy in the BZ, making the integral independent of the shape of the cell. Considering a surface of constant energy for a given difference $S(\tilde{E})$ and defining a vector perpendicular to this plane as

$$\hat{\mathbf{u}}_{\perp} = \frac{\nabla_{\mathbf{k}} \tilde{E}_n}{|\nabla_{\mathbf{k}} \tilde{E}_n|},\tag{3}$$

makes it possible to recast the volume element by integrating over each constant \tilde{E}_n plane inside the IBZ instead,

$$d^{3}\mathbf{k} = \hat{\mathbf{u}}_{\perp} \cdot d\mathbf{k}_{\perp} \, dS = \frac{1}{|\nabla_{\mathbf{k}} \tilde{E}_{n}|} d\tilde{E}_{n}(\mathbf{k}) dS. \quad (4)$$

Now with this result eq. (2) may for each micro cell be written as

$$I(\omega) = \frac{1}{\Omega_{BZ}} \sum_{T} \int_{\varepsilon} \int_{S \in T} A(\mathbf{k}) \delta(\tilde{E} - \alpha \hbar \omega) \times \frac{1}{|\nabla_{\mathbf{k}} \tilde{E}_{n}|} d\tilde{E}(\mathbf{k}) dS. \quad (5)$$

The response from each tetrahedron is then calculated and added together, to get the result for the entire IBZ zone. To this end the result from a single micro cell has to be calculated. Each tetrahedron is defined by its vertices, \mathbf{k}_0 , \mathbf{k}_1 , \mathbf{k}_2 and \mathbf{k}_3 , and at each vertex the energy and function $A(\mathbf{k})$ is calculated. The tetrahedron is usually constructed by some predefined algorithm, such as "generateMesh" in Matlab or "Delaunay" from the package scipy in python. The numbering of the **k**points are sorted such that

$$\tilde{E}_0 \le \tilde{E}_1 \le \tilde{E}_2 \le \tilde{E}_3,\tag{6}$$

Similarly the value of the function $A(\mathbf{k}_i)$ is also relabelled in accordance with eq. (6). The next step is then to linearly interpolate both $A(\mathbf{k})$ and $\tilde{E}(\mathbf{k})$ inside the tetrahedron,

$$\bar{A}(\mathbf{k}) = A_0 + \mathbf{b}_A \cdot \mathbf{k}',\tag{7}$$

where $\mathbf{k}' = \mathbf{k} - \mathbf{k_0}$, and $\mathbf{b}_A = [\nabla_{\mathbf{k}} A(\mathbf{k})]_{A_0}$ [10]. The linear interpolations may now be inserted into eq. (2), from which it can be seen that the delta function will be independent of *S* as the energy does not change for a given plane, hence the integral over *S* can be considered first

$$i = \frac{1}{|\mathbf{b}_{\tilde{E}}|} \left[\int_{S(\tilde{E})\in T} A_0 dS + \mathbf{b}_A \int_{S(\tilde{E})\in T} \mathbf{k}' dS \right].$$
(8)

The first integral in eq. (8) is simply the area of the constant energy surface, f, times A_0 , while the second integral can be seen as the centre of gravity \mathbf{k}_{cg} for an isotropic 2D medium. To calculate these properties it is of interest to split the energy surfaces into three different categories, namely $\tilde{E}_0 \leq S(\tilde{E}) < \tilde{E}_1$ which will be abbreviated S_0 , while the area $\tilde{E}_1 \leq S(\tilde{E}) < \tilde{E}_2$ will be abbreviated S_2 , and finally the region $\tilde{E}_2 \leq S(\tilde{E}) < \tilde{E}_3$ is abbreviated S_3 . From fig. 1 it can be seen that S_2 can be calculated as the difference between the areas S_0 and S_1 . To calculate the area of of the different



Fig. 1 Sketches the energy planes in each tetrahedron utilised in deriving the LATM.

surfaces an expression for the vertices of each surface, $\mathbf{k}_{i,j}^{(E)}$, has been constructed in terms of the energy \tilde{E} as well as the location of the vertices, using this the area can be calculated as the cross product of two of the corners i.e., $f_0 = \frac{1}{2} \left| \left(\mathbf{k}_{20}^{(E)} - \mathbf{k}_{10}^{(E)} \right) \times \left(\mathbf{k}_{30}^{(E)} - \mathbf{k}_{10}^{(E)} \right) \right|$. By doing this it is possible to express the area in terms of the energy difference at the vertices in the case of S_0 [11]

$$f_0 = \frac{3\Omega_T E_0^2}{\tilde{E}_{10}\tilde{E}_{20}\tilde{E}_{30}} |\mathbf{b}_{\tilde{E}}|.$$
 (9)

The same can be done for f_1 and f_3

$$f_1 = \frac{3\Omega_T E_1^2}{\tilde{E}_{10}\tilde{E}_{21}\tilde{E}_{31}} |\mathbf{b}_{\tilde{E}}|, \qquad (10)$$

$$f_3 = \frac{3\Omega_T E_3^2}{\tilde{E}_{30}\tilde{E}_{31}\tilde{E}_{32}} |\mathbf{b}_{\tilde{E}}|.$$
 (11)

Finally f_2 may be expressed as $f_2 = f_0 - f_1$. All these expressions are purely defined by the vertices of the

tetrahedron, and hence only the value at these vertices needs to be known.

The centre of gravity of a triangle can be expressed simply from the associated vertices

$$\mathbf{k}_{cg,j} = \frac{1}{3} \sum_{i=0}^{3} \mathbf{k}_{i,j}^{(E)}.$$
 (12)

Thus the dot product in eq. (8) can be calculated as

$$\mathbf{b}_{A} \cdot \mathbf{k}_{cg} = A_{j}0 + \frac{E_{j}}{3} \sum_{i=0}^{3} {}^{\prime} \frac{A_{i,j}}{\tilde{E}_{i,j}}.$$
 (13)

With these results it is now possible to construct a complete analytic expression for the integral over a tetrahedron micro-cell [11]. Starting with area S_0 i.e., $\tilde{E}_0 \leq \alpha \hbar \omega < \tilde{E}_1$

$$i_0 = \frac{\Omega_T E_0^2}{\tilde{E}_{10} \tilde{E}_{20} \tilde{E}_{30}} \left(3A_0 + \sum_{i=1}^3 \frac{A_{i,0}}{\tilde{E}_{i,0}} \right).$$
(14)

For area S_2 i.e., $\tilde{E}_1 \leq \alpha \hbar \omega < \tilde{E}_2$ this may be calculated as

$$i_{2} = \frac{\Omega_{T} E_{0}^{2}}{\tilde{E}_{10} \tilde{E}_{20} \tilde{E}_{30}} \left(3 A_{0} + \sum_{i=1}^{3} \frac{A_{i,0}}{\tilde{E}_{i,0}} \right) - \frac{\Omega_{T} E_{1}^{2}}{\tilde{E}_{10} \tilde{E}_{21} \tilde{E}_{31}} \left(3 A_{1} + \sum_{i \neq 1}^{3} \frac{A_{i,1}}{\tilde{E}_{i,1}} \right). \quad (15)$$

And lastly for S_3 , $\tilde{E}_2 \leq \alpha \hbar \omega \leq \tilde{E}_3$ the integral yields

$$i_3 = \frac{\Omega_T E_3^2}{\tilde{E}_{30} \tilde{E}_{31} \tilde{E}_{32}} \left(3 A_3 + \sum_{i \neq 0}^3 \frac{A_{i,3}}{\tilde{E}_{i,3}} \right).$$
(16)

Lastly two cases has to be considered, namely $\alpha \hbar \omega < \tilde{E}_0$ and $\tilde{E}_3 < \alpha \hbar \omega$. In these cases the energy surface will be completely outside the tetrahedron, and hence the surface area inside is zero, thus there will be no contributions from these areas. To complete the integration, the last integral over the different energies has to be carried out

$$I(\omega) = \int (i_0 + i_2 + i_3) \delta(\tilde{E} - \alpha \hbar \omega) d\tilde{E}.$$
 (17)

However, due to the delta function this intergration simply amounts to inserting $\alpha \hbar \omega$ into the E_j expression as $E_j = \alpha \hbar \omega - \tilde{E}_j$. This has already been done when the different energy cases were outlined to ease the readability. The expressions in eq. (14), eq. (15), and eq. (16) is also completely independent of the tetrahedral structure and hence very easy to implement, and throughout this project it will be shown that the implementation of this procedure will significantly help with the convergence of IBZ integrals by reducing the number of sampling points needed which is of great interest in DFT calculations as large point grids are unfeasible.

4. Methodology

Before testing the effectiveness and reliability of the LATM a brief overview of the methodology used in this paper will be presented here. As stated in section 3 the integrals of interest is on the form of eq. (2). Integrals like these appear when the imaginary part is extracted. This also means that the property calculated in this paper will be the imaginary part of the electric susceptibility. The real part may then be calculated through the Kramers-Kronig relations, thus an integration for the real part of the response function may be omitted which eases the calculation. Additionally the LATM for the real part of the response function is not independent of the tetrahedral geometries, which would significantly complicate the implementation [11].

All LATM calculations will be compared with EMP point sampling calculations with the form factors for wurtzite GaN taken from *Dielectric Properties of Wurtzite and Zincblende Structure Gallium Nitride* and the zinc blende form factors taken from *Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures* [5, 12]. For the wurtzite calculations a basis set of 197 G-vectors were used, while for zinc blende a basis set of 65 were used. DFT and EMP is used to get the matrix elements and eigenvalues of the linear response, while the nonlinear response only applies EMP. The calculation parameters for the DFT calculation can be seen in table I.

| # k | Mode | XC | FD width | PW |
|----------------|------|----------|----------|-----|
| $n_k^3 = 1000$ | PAW | PBE [13] | 0.01 | 600 |

Tab. I The settings which are constant for all DFT calculations of the linear susceptibility.

The focus of the SHG calculations is to test the reliability of the LATM when the $A(\mathbf{k})$ function is poorly behaved. To avoid numerical errors caused by the placement of \mathbf{k} , the same \mathbf{k} -point grid will be used for both the LATM and the point sampling calculations.

To quantify how quickly the calculations converges eq. (18) is used, which gives a total change in percent between two Im{ $\chi(\omega)$ },

$$\Delta_{\%} = \sum_{\omega} \frac{\left| \operatorname{Im}\{\chi_{(\#\mathbf{k}_{i})}\} \right| - \left| \operatorname{Im}\{\chi_{(\#\mathbf{k}_{i+1})}\} \right|}{\left| \operatorname{Im}\{\chi_{(\#\mathbf{k}_{i+1})}\} \right|} \times 100.$$
(18)

5. Linear Susceptibility with LATM and EMP

By means of time-dependent perturbation theory, the imaginary part of the susceptibility can be given as [14]

$$\operatorname{Im}\left\{\chi^{(1)}(\omega)\right\} = \frac{e^2}{4\pi^2 m^2 \varepsilon_0 \omega^2} \times \sum_{c,v} \int_{\mathbf{k}} |P_{cv}^j|^2 \delta(E_{cv}(\mathbf{k}) - \hbar\omega) d\mathbf{k}, \quad (19)$$

with ω being the angular frequency of the perturbing light source, $E_{cv}(\mathbf{k}) = E_c(\mathbf{k}) - E_v(\mathbf{k})$, and $|P_{cv}^{(j)}|^2 =$ $|\langle \varphi_{c,\mathbf{k}} | \hat{p}_j | \varphi_{v,\mathbf{k}} \rangle|^2$. The summation index v runs over the occupied bands, while c runs over the unoccupied and is truncated at some desired number, that being 20 for this paper. Because wurtzite is birefringent and symmetry operations were used to reduce the integration domain it is necessary to average the $P_{c,v}^x$ and $P_{c,v}^y$ matrix elements over the IBZ to get the correct ordinary susceptibility [5]. To validate the result of the LATM, an EMP calculation using the LATM and point sampling is carried out for GaAs, GaN and InSb in the zinc blende lattice using $\#\mathbf{k} = 5924$, which can be seen in fig. 2. It is readily observed the LATM and point sampling achieves the same result.



Fig. 2 Linear response of GaAs, GaP, and InSb using $\#\mathbf{k} = 5924$ with point sampling and LATM utilising EMP.

It is now of interest to apply the LATM to calculations based on DFT to see if the results can be converged for few #k relative to what is required using point sampling and EMP. Figure 3 shows Im $\{\chi^{(1)}(\omega)\}$ for wurtzite GaN using $\#\mathbf{k} = 98758$, and the change in % for increasing $\#\mathbf{k}$. For $\#\mathbf{k} > 14000$ the difference between two graphs is below 2 % but it can be seen that even at $\#\mathbf{k} > 50000$ there is still a change between two points meaning the method converges slowly.



Fig. 3 Linear response of GaN using $\#\mathbf{k} = 98758$ using point sampling. The insert shows the convergence of the linear response using point sampling.

In fig. 4 a similar calculation has been done using DFT along with the LATM method. When using DFT a selfconsistent routine is performed, using both point group symmetry, and time reversal symmetry, with $\#\mathbf{k} = 10^3$. Afterwards the electron density is kept fixed and the wave functions and eigenvalues are found using a new k-grid. In fig. 3 the response has been calculated using $\#\mathbf{k} = 2275$ which produces a graph of slightly inferior smoothness compared to fig. 3. The inserts in fig. 4 shows the convergence and from this it can be seen that for N = 13 the relative error is below 5% which is well below the points needed to converge the point sampling method. It is thus readily demonstrated that for the linear response the LATM converges faster than the sampling method, reducing the necessary k-points by two orders of magnitude.

6. Nonlinear Susceptibility with LATM and EMP

In the previous section it was shown that the LATM is very effective at reducing the number of k-points needed to achieve convergence of the linear susceptibility. It is now of interest to apply the method to the nonlinear response, which in this paper is SHG. SHG can be understood using the dipole approximation, where it is assumed that the wavelength of light is much larger



Fig. 4 Linear response for GaN with $\#\mathbf{k} = 2275$ using LATM. The insert shows the convergence of the linear response using LATM.

than the atoms, and hence the spacial variations in the electric field can be ignored. The electric field may then be approximated as [2]

$$\boldsymbol{\mathcal{E}} = \sum_{b,\omega} \mathcal{E}_b(\omega) e^{-i\omega t} \boldsymbol{\hat{b}},\tag{20}$$

where $\mathcal{E}_a(\omega)$ denotes the field strength for a field with frequency ω , and \hat{b} denotes the direction. The response from the system will be an induced polarisation, which in linear electromagnetic theory is proportional to the electric field with the proportionality factor χ . In nonlinear electromagnetism this factor is instead a series of factors, $\chi^{(1)}, \chi^{(2)}, \chi^{(3)}$... with each describing the response from a field of that order. $\chi^{(1)}$ is a second rank tensor while $\chi^{(2)}$ is a third rank tensor and so forth [15]. In the case of SHG it is $\chi^{(2)}$ that is of interest. Given two perturbing fields, the second order contribution to the polarisation is expressed as [2]

$$P_a^{(2)} = \varepsilon_0 \sum_{\omega_1,\omega_2} \sum_{b,c} \chi_{a,b,c}^{(2)} \mathcal{E}_b(\omega_1) \mathcal{E}_c(\omega_2) e^{-i(\omega_1 + \omega_2)t}.$$
(21)

SHG is then the special case where the two applied frequencies are identical $\omega_1 = \omega_2 = \omega$. In the length gauge, the dipole operator is normally split into two parts, namely an inter- and intraband part [16, 3, 2, 17]. This leads to four terms, namely a purely interband, a purely intraband, and two mixed terms. For a cold intrinsic semiconductor the intraband and one of the mixed terms disappear. The purely interband response may be seen as the transition of electrons between different bands due to the absorption of either one or two photons. The surviving mixed term results in a modulation of the linear susceptibility due to intraband movement of the electrons [4]. To test the LATM only the purely inter-band response will be considered, which in the velocity gauge, when symmetrised over polarisation indices, may be calculated as

$$\chi_{\alpha,\beta,\gamma}^{(2i)}(-2\omega;\omega,\omega) = \frac{ie^3}{(2\pi)^3 \varepsilon_0 \hbar^2 m^3 \omega^3} \times \sum_{nml} \int \frac{\left\langle p_{nm\mathbf{k}}^{(\alpha)} p_{ml\mathbf{k}}^{(\beta)} p_{ln\mathbf{k}}^{(\gamma)} \right\rangle}{\omega_{mn\mathbf{k}} - 2\omega} \times \left(\frac{f_{nl}}{\omega_{ln\mathbf{k}} - \omega} + \frac{f_{ml}}{\omega_{ml\mathbf{k}} - \omega} \right) d^3\mathbf{k}. \quad (22)$$

It is possible to extract the imaginary part to make the integral like the one in eq. (2) as [18]

$$\operatorname{Im}\left\{\chi_{\alpha,\beta,\gamma}^{(2i)}\right\} = -\frac{\pi e^{3}}{(2\pi)^{3}\varepsilon_{0}m^{3}\hbar^{2}\omega^{3}} \times \int \sum_{v,c} \left\{\Pi_{\alpha,\beta,\gamma}^{(\omega)}\delta(\omega_{cv}-\omega) + \Pi_{\alpha,\beta,\gamma}^{(2\omega)}\delta(\omega_{cv}-2\omega)\right\} d^{3}\mathbf{k}, \quad (23)$$

with

$$\Pi_{\alpha,\beta,\gamma}^{(w)} = \sum_{l} \left(\operatorname{Im} \left\{ \left\langle p_{vc}^{(\alpha)} p_{cl}^{(\beta)} p_{lv}^{(\gamma)} \right\rangle \right\} \times \frac{\omega_{cl} - \omega_{lv}}{(\omega_{cv} + \omega_{cl})(\omega_{cv} + \omega_{lv})} \right), \quad (24)$$

$$\Pi_{\alpha,\beta,\gamma}^{(2w)} = \sum_{l} \operatorname{Im}\left\{\left\langle p_{vc}^{(\alpha)} p_{cl}^{(\beta)} p_{lv}^{(\gamma)} \right\rangle\right\} \frac{2}{\omega_{cl} - \omega_{lv}}.$$
 (25)

Here, c denotes conduction bands, v denotes valence bands, and l is any given band. By comparison with eq. (2), $\Pi_{\alpha,\beta,\gamma}^{(\alpha'\omega)}$ would then correspond to $A(\mathbf{k})$, hence these will be linearly interpolated between each of the vertices of the tetrahedron. This linear interpolation was shown to yield great results for the linear susceptibility, however, the function that were interpolated in that case was the matrix elements which is known to vary slowly and contain no poles. This is not the case for the second order susceptibility, as both $\Pi^{(\omega)}_{\alpha,\beta,\gamma}$ and $\Pi^{(2\omega)}_{\alpha,\beta,\gamma}$ have denominators that may yield zero, which means that the linear approximation is obviously bad. However the pole may still be averted by the delta function, hence only a problem is expected when both the denominator and delta function encounters a resonance. This is a so-called double resonance which is hard to handle numerically, but it can be alleviated by adding a small broadening in both denominators. The result is that $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ will become complex in the calculations, where the real part represents the actual $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ and the imaginary part of $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ have no physical meaning.

Because EMP is utilised both for LATM and point sampling, the results may be compared one to one, which is done in fig. 5 for GaN using $\#\mathbf{k} = 5910$. The form and values are generally similar, though the numerical methods used to handle the delta functions with point sampling inhibits its graph from spiking as high as that of LATM.



Fig. 5 Im $\{\chi_{x,y,z}^{(2i)}\}$ using $\#\mathbf{k} = 5910$ with LATM and point sampling. The insert shows the tetrahedron mesh of the IBZ.

Testing the convergence of the methods eq. (18) is used, where the results for LATM is seen in fig. 6 with the insert being Im $\left\{\chi_{x,y,z}^{(2i)}\right\}$ for GaN using $\#\mathbf{k} =$ 5745, 5775 and 5910. It is clear that no appreciative convergence is taking place, with GaN being the most egregious. Slightly inferior, yet similar results are obtained when point sampling is employed, not shown here. The insert does however reveal that the three graphs are close to identical up until ~ 5.2 eV.

It is hypothesised that the divergence taking place at high energies are caused by the double resonance. To test this, the real and imaginary part of $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ is plotted in fig. 7 as the imaginary part is only appreciably different from zero when double resonances are present. It is clear that this correspondence is found, as the real



Fig. 6 Convergence of the LATM for various zinc blende semiconductors. The insert shows $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ using LATM.

part of $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ for GaN becomes unstable at the same energies that the imaginary part of $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ becomes nonzero.



Fig. 7 Plots the real and imaginary part of Im $\left\{\chi_{x,y,z}^{(2i)}\right\}$ using LATM and $\#\mathbf{k} = 5745, 5775$ and 5910.

Similar results are found for all materials tested here, with the trend being that nonlinear response for materials with larger band gaps are stable for higher energies. Ignoring Im $\{\chi_{x,y,z}^{(2i)}\}$ for energies above which the response of each materials becomes unstable, the convergence using eq. (18) is plotted in fig. 8 using LATM, where the insert shows Im $\{\chi_{x,y,z}^{(2i)}\}$ for all five materials using LATM with $\#\mathbf{k} = 5924$. It is seen that the graphs are nicely converged, and that $\operatorname{Im}\left\{\chi_{x,y,z}^{(2i)}\right\}$ is smooth for all materials.



Fig. 8 Convergence of Im $\left\{\chi_{x,y,z}^{(2i)}\right\}$ using LATM with the unstable energy range removed from each material. The insert plots Im $\left\{\chi_{x,y,z}^{(2i)}\right\}$ using LATM and $\#\mathbf{k} = 5924$.

To demonstrate the strength of LATM Im $\left\{\chi_{x,y,z}^{(2i)}\right\}$ is plotted in fig. 9 for InSb using both LATM and point sampling with $\#\mathbf{k} = 168, 776$, and 1890. Here LATM is convincing, especially at low energies, while the point sampling method is far from converged.

It has been shown that the LATM provides a strong tool for carrying out IBZ integration. It is very robust for linear response functions and may easily be implemented to reduce the number of k-points needed to converge a result. For the second order response it was hypothesised that the double resonance did not allow for convincing convergence graphs, but the method was stable in the low frequency range. Hence if it is the low frequency range that is of interest LATM has proven to be reliable. In the high frequency range, the method does not converge, but it is more stable than point sampling and may thus prove a useful tool.

7. Piezoelectric Coefficients

The study of piezo electric effects are of great interest in the context of III-V nitride semiconductors, as these show a large piezoelectric response, up to an order of magnitude larger than other III-V semiconductors [19, 20, 21]. As the electric field resulting from strain can significantly alter the electronic and optical effects of the device [22], it is of interest to understand the piezoelectric response, which will be done here for



Fig. 9 Plots Im $\left\{\chi_{x,y,z}^{(2i)}\right\}$ of InSb using point sampling and LATM with $\#\mathbf{k} = 168,776$ and 1890.

GaN in the wurtzite configuration. The piezoelectric coefficients are generally given by

$$\gamma_{i,j} = \frac{\partial P_i}{\partial \epsilon_j},\tag{26}$$

where Voigt notation is used. This then relates the change in polarisation to the change in strain, which yield a filled third rank tensor. However, due to wurtzite being a part of space group p63mc the only 3 unique, non-vanishing terms are $\gamma_{3,1} = \gamma_{3,2}$, $\gamma_{3,3}$, and $\gamma_{1,5} = \gamma_{2,4}$ [19]. To calculate the piezoelectric coefficient $\gamma_{3,3}$ one can consider a small applied strain, ϵ_3 , which would induce a polarisation along the *c*-axis described by

$$\delta P_3 = \gamma_{3,3} \epsilon_3. \tag{27}$$

To find the coefficient the polarisation is expanded to the first order in the lattice and internal parameter [22, 23]

$$\delta P_3 = \frac{\partial P_3}{\partial c}(c - c_0) + \frac{\partial P_3}{\partial u}(u - u_0).$$
(28)

The change in the lattice parameter c may be related to the strain along the c-axis as

$$c = c_0(1 + \varepsilon_3), \tag{29}$$

from which it follows that $\frac{\partial \epsilon_3}{\partial c} = \frac{1}{c_0}$. Additionally, the internal parameter is expressed as a function of the strain and then Taylor expanded around $\epsilon_3 = 0$, giving $u(\epsilon_3) = u(0) + \frac{\partial u}{\partial \epsilon_3}(\epsilon_3 - 0)$. Inserting these expressions in eq. (28) and equating with eq. (27) the piezoelectric coefficient can be expressed as

$$\gamma_{3,3} = \frac{\partial P_3}{\partial \epsilon_3} + \frac{\partial P_3}{\partial u} \left. \frac{\partial u}{\partial \epsilon_3} \right|_{u_0}.$$
 (30)

Similar expressions can be made for the other coefficients using the same method, yielding the following equations

$$\gamma_{3,1} = \frac{\partial P_3}{\partial \epsilon_1} + \frac{\partial P_3}{\partial u} \frac{\partial u}{\partial \epsilon_1} \bigg|_{u_0}, \qquad (31)$$

$$\gamma_{1,5} = \frac{\partial P_1}{\partial \epsilon_5} + \frac{\partial P_1}{\partial u} \frac{\partial u}{\partial \epsilon_5} \bigg|_{u_0}.$$
 (32)

Here the first term in each equation is referred to as the clamped ion term, $\gamma_{i,j}^{(0)}$, which details the polarisation response when the lattice parameters are fixed at their equilibrium values, while the second term is related to relaxation of the internal parameter. Both the polarisation and the internal parameter has a linear response for small strains, and hence the derivatives in eq. (30) through eq. (32) can easily be tackled through a finite difference approximation, if the polarisation as a function of strain can be calculated, up to modulus $e\mathbf{T}/V_{cell}$, with **T** being a translational lattice vector [24, 25]. The polarisation can be understood through an ionic and electronic contribution. The ionic term represents the movement of the atomic cores in relation to their equilibrium positions, and is given by

$$\mathbf{P}_{Ion} = \frac{e}{V_{cell}} \sum_{\mu} Z_{\mu} \boldsymbol{\tau}_{\mu}.$$
 (33)

The electronic contribution can qualitatively be understood as the displacement of the electrons as a response to the displacement of the ions, which may be viewed as a current flowing in the crystal [26]. The total change to the polarisation can be found by integrating this current. Rather than working with adiabatic currents, the problem is restated in terms of some unitless perturbation λ which varies between 0 and 1, with the former corresponding to the initial state of the system and the latter being the final state,

$$\Delta \mathbf{P} = \int_{0}^{\Delta t} \mathbf{J} dt = \int_{0}^{1} \frac{\partial \mathbf{P}}{\partial \lambda} \frac{d\lambda}{dt} dt$$
$$= \mathbf{P}(\lambda_{1}) - \mathbf{P}(\lambda_{0}). \quad (34)$$

The goal then, is to derive an expression for $\frac{\partial P}{\partial \lambda}$. The polarisation is the dipole moment per volume, whose operator is $\frac{-e\mathbf{r}}{\Omega}$. The issue here is that the position operator in the overlap integrals are ill defined. Therefore, the problem is recast in the form of the Berry phase, as it is done in "*Berry Phases in Electronic Structure Theory*" by David Vanderbilt [26]. The electronic contribution, can be shown to be related

| XC | $\frac{\partial P_3}{\partial u}$ | $\gamma_{3,1}^{(0)}$ | $\gamma_{3,3}^{(0)}$ | $\gamma_{1,5}^{(0)}$ | $\frac{\partial u}{\partial \epsilon_1}$ | $\frac{\partial u}{\partial \epsilon_3}$ | $\frac{\partial u}{\partial \epsilon_5}$ | $\gamma_{3,1}$ | $\gamma_{3,3}$ | $\gamma_{1,5}$ |
|---------|-----------------------------------|----------------------|----------------------|----------------------|--|--|--|----------------|----------------|------------------|
| LDA | -8.95 | 0.43 | -0.87 | 0.48 | 0.09 | -0.18 | 0.00 | -0.40 | 0.73 | 0.48 |
| PBE | -9.03 | 0.43 | -0.88 | 0.48 | 0.10 | -0.19 | 0.00 | -0.48 | 0.88 | 0.48 |
| LDA[23] | - | 0.45 | -0.84 | - | - | -0.16 | - | -0.49 | 0.73 | - |
| LDA[22] | - | - | - | - | - | - | - | -0.44 | 0.86 | 0.30^{\dagger} |

Tab. II †: Experimental results

to the Berry phase, and the total polarisation can then be calculated as [26, 27]

$$\mathbf{P}_{tot} = \frac{e}{V_{cell}} \left[\sum_{\mu} Z_{\mu} \boldsymbol{\tau}_{\mu} - \sum_{j} \frac{\bar{\phi}_{j}}{2\pi} \mathbf{a}_{j} \right].$$
(35)

Here, $\bar{\phi}$ is the average Berry phase for a plane perpendicular to a set of sufficiently dense strings of kpoints in the IBZ and \mathbf{a}_j is the corresponding direction of the lattice vector. The Berry phase calculation follows the method outlined by Kingsmith and Vanderbilt in their seminal paper, taking the form [24, 26]

$$\phi = -\mathrm{Im} \left\{ \ln \left[\prod_{i=0}^{N-1} \det \left(\begin{bmatrix} \langle u_{1,\mathbf{k}_{i}} | u_{1,\mathbf{k}_{i+1}} \rangle & \dots & \langle u_{1,\mathbf{k}_{i}} | u_{M,\mathbf{k}_{i+1}} \rangle \\ \vdots & \ddots & \vdots \\ \langle u_{M,\mathbf{k}_{i}} | u_{1,\mathbf{k}_{i+1}} \rangle & \dots & \langle u_{M,\mathbf{k}_{i}} | u_{M,\mathbf{k}_{i+1}} \rangle \end{bmatrix} \right) \end{bmatrix} \right\}.$$
(36)

Here $\langle u_{m,\mathbf{k}}|u_{n,\mathbf{k}}\rangle$ is the overlap of the cell periodic part of the wave function between band m and n. The resulting matrix of eq. (36) will be of size $(M \times M)$, with M being the number of occupied bands. The Nk-points in question has to be on a parallel string through the first BZ, where special care has to be taken for the \mathbf{k}_N point to preserve the periodicity of the BZ. In practice, this amounts to ensuring that the phase difference between the first and the last kpoint is exactly the phase aquired when transversing the unit cell, $|u_{n,\mathbf{k}_N}\rangle = e^{-i\mathbf{b}_j\cdot\mathbf{r}} |u_{n,\mathbf{k}_0}\rangle$, with \mathbf{b}_j being a reciprocal lattice vector corresponding to the polarisation direction in question. From this the average Berry phase is calculated. From eq. (36) it is clear that the polarisation is quantised as the logarithm only yields a definite answer mod 2π . Due to this fact further care has to be taken to wrap the answer onto a choice of branch, here that branch is $-\pi \leq \bar{\phi}_i \leq \pi$.

Continuing with calculating the derivatives in eq. (30), eq. (31), and eq. (32). The clamped ion term can be calculated by a finite difference approximation, where the internal parameter is kept constant. To do this, the wave functions has to be known for strings of **k**-points in the first BZ. For this paper, this was done by expanding the wave function using the augmented plane wave method, with a cut-off of 800 eV and using a $(4 \times 4 \times 4)$ Monkhorst-Pack grid for the self-consistent calculation. For the exchange correlation term both the Perdew

Wang local density approximation (LDA) and Perdew-Burke-Ernzerhof (PBE) method was used as the LDA can more readily be compared with existing literature, while PBE is considered more accurate [13, 28]. For the final calculation the wave functions are further expanded on a $(10 \times 10 \times 10)$ Monkhorst-pack grid keeping the electronic density fixed. The change in polarisation due to the change in internal parameter has been calculated using the same parameters, with the exception of the dependence of the internal parameter on the applied strain, where it was necessary to use a plane wave energy cutoff of 1600 eV. This derivative has been calculated by following the method outlined in [29]. Thus the strain in question is varied between $\pm 1\%$, and for each instance of strain the internal parameter is varied by ± 0.01 and the total energy is then calculated. A fourth degree polynomial is then fitted to the total energies and from this the optimal value of the uparameter can easily be found. The results of the calculations are gathered in the two upper rows in table II, while the two lower rows are values found in the literature. It is clear that values are in excellent agreement with those found in the literature, both when the LDA and PBE exhange correlation approximation is utilised. It has thus been shown that DFT can be utilised to compute nonlinear properties of three dimensional materials without sacrificing accuracy due to its computational intensity.

8. Conclusion

In this paper it has been demonstrated to which degree LATM may be applied to both the linear and second order electronic susceptibility. The LATM was initially applied to the linear susceptibility where it was seen that it could reduce the number of k-points needed by two orders of magnitude, and can hence be used to alleviate the computational load of expensive DFT calculations. The method was then applied to various zinc blende crystals to calculate their SHG response, where it was shown that the method could replicate the result obtained with a simple point sampling scheme. However, numerical noise proved to hinder the same degree of convergence as was the case for the linear susceptibility especially pronounced for the high band

gap semiconductors. It was theorised that this noise originates from double resonances found in the response function at high frequencies, and it was shown that for low frequencies the results did indeed converge. Hence the LATM may readily be employed to calculate the second order optical response, however, care has to be taken, if fine features in the spectrum at high frequencies are desired. Lastly the piezoelectric properties of GaN was calculated using an approach that circumvents the computationally expensive integrals over the BZ, and the calculations could therefore be performed using DFT. The obtained results was in good agreement with those found in the literature.

Acknowledgement

The authors of this work gratefully acknowledge Grundfos for sponsoring the 8th MechMan symposium. All DFT calculations was carried out utilising GPAW, which is an open source python library, implementing the PAW method to solve the kohnsham equation, which is available at GPAWs website [30, 31]. GPAW utilises the ASE library for atomic simulations and molecular dynamics and hence this libary has also been used, it may be found on ASEs website [32].

References

- [1] G. New, Introduction to Nonlinear Optics. Cambridge University Press, 2011.
- [2] A. Taghizadeh, K. S. Thygesen, and T. G. Pedersen, 'Two-dimensional materials with giant optical nonlinearities near the theoretical upper limit," ACS Nano, vol. 15, no. 4, pp. 7155-7167, 2021.
- [3] T. G. Pedersen, "Electric, Optical and Magnetic Properties of Nanostructures," 2020.
- [4] J. E. Sipe and E. Ghahramani, "Nonlinear optical response of semiconductors in the independent-particle approximation," Phys. Rev. B, vol. 48, pp. 11705-11722, Oct 1993.
- [5] R. Wang, P. Ruden, J. Kolnik, I. Oguzman, and K. Brennan, "Dielectric properties of wurtzite and zincblende structure gallium nitride," Journal of Physics and Chemistry of Solids, vol. 58, no. 6, pp. 913 - 918, 1997.
- [6] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Physical
- [6] P. Honenberg and w. Kom, Infomogeneous electre, gas, <u>review</u>, vol. 136, no. 3B, p. B864, 1964.
 [7] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," <u>Phys. Rev.</u>, vol. 140, pp. A1133-A1138, Nov. 1965.
- [8] O. Jepson and O. Anderson, "The electronic structure of h.c.p. ytterbium," Solid State Communications, vol. 9, no. 20, pp. 1763-1767, 1971.
- [9] G. Lehmann and M. Taut, "On the numerical calculation of the density of states and related properties," physica status solidi (b), vol. 54, no. 2, pp. 469-477, 1972.
- [10] D. J. Moss, J. E. Sipe, and H. M. van Driel, "Application of the linear-analytic tetrahedra method of zone integration to nonlinear response functions," Phys. Rev. B, vol. 36, pp. 1153-1158, Jul 1987.
- [11] V. Eyert, The Augmented Spherical Wave Method: A Comprehensive Treatment, vol. 849 of Lecture Notes in Physics. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2012.
- [12] M. L. Cohen and T. Bergstresser, "Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures," Physical review, vol. 141, no. 2, pp. 789-796, 1966.

- [13] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett., vol. 77, pp. 3865-3868, Oct 1996.
- [14] M. Trolle, "Microscopic theory of linear and nonlinear optical response: zinc-blende semiconductors," 2011.
- [15] Robert W. Boyd, Nonlinear Optics. Elsevier Science & Technology, 3rd ed., 2008.
- [16] E. Blount, "Formalisms of band theory," in Solid State Physics (F. Seitz and D. Turnbull, eds.), vol. 13, (New York), pp. 305-373, Academic Press, 1962.
- [17] C. Aversa and J. E. Sipe, "Nonlinear optical susceptibilities of semiconductors: Results with a length-gauge analysis," Phys. Rev. B, vol. 52, pp. 14636-14645, Nov 1995.
- [18] T. G. Pedersen and K. Pedersen, "Systematic tight-binding study of optical second-harmonic generation in carbon nanotubes," Phys. Rev. B, vol. 79, p. 035422, Jan 2009.
 [19] H. Morkoç, "Handbook of nitride semiconductors and devices:
- Electronic and optical processes in nitrides," Handbook of Nitride Semiconductors and Devices: Electronic and Optical Processes in Nitrides, vol. 2, pp. 1-846, 08 2009.
- [20] A. Beya-Wakata, P.-Y. Prodhomme, and G. Bester, "First- and second-order piezoelectricity in iii-v semiconductors," Phys. Rev. B, vol. 84, p. 195207, Nov 2011.
- [21] F. Bernardini, V. Fiorentini, and D. Vanderbilt, "Spontaneous polarization and piezoelectric constants of iii-v nitrides," Phys. Rev. B, vol. 56, pp. R10024-R10027, Oct 1997.
- [22] H. Morkoç, Handbook of nitride semiconductors and devices. Vol. 1, Materials properties, physics and growth. Handbook of Nitride Semiconductors and Devices (VCH), Weinheim: Wiley-VCH, 2009.
- [23] F. Bernardini, V. Fiorentini, and D. Vanderbilt, "Spontaneous polarization and piezoelectric constants of iii-v nitrides," Physical review. B, Condensed matter, vol. 56, no. 16, pp. R10024-R10027, 1997.
- [24] R. D. King-Smith and D. Vanderbilt, "Theory of polarization of crystalline solids," Physical review. B, Condensed matter, vol. 47, no. 3, pp. 1651–1654, 1993. R. Resta, "Theory of the electric polarization in crystals,"
- [25] Ferroelectrics, vol. 136, no. 1, pp. 51-55, 1992.
- [26] D. Vanderbilt, Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators. Cambridge University Press, 2018.
- [27] K. M. Rabe, C. H. Ahn, J.-M. Triscone, and K. Rabe, Physics of Ferroelectrics: A Modern Perspective, vol. 105 of Topics in applied physics. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2007
- [28] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," Phys. Rev. B, vol. 45, pp. 13244-13249, Jun 1992.
- [29] A. Dal Corso, M. Posternak, R. Resta, and A. Baldereschi, "Ab initio study of piezoelectricity and spontaneous polarization in zno," Phys. <u>Rev. B</u>, vol. 50, pp. 10715–10721, Oct 1994.
- [30] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, "Real-space grid implementation of the projector augmented wave method," Phys. Rev. B, vol. 71, no. 3, p. 035109, 2005.
- [31] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puska, T. T. Rantala, J. Schiøtz, K. S. Thygesen, and K. W. Jacobsen, "Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method," J. Phys.: Condens. Matter, vol. 22, no. 25, p. 253202, 2010.
- [32] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, "The atomic simulation environment - a python library for working with atoms," Journal of Physics: Condensed Matter, vol. 29, no. 27, p. 273002, 2017.