Electron-phonon interaction effects on the dielectric response of Si

L. F. Lastras-Martínez

Instituto de Investigación en Comunicación Óptica, Universidad Autónoma de San Luis Potosí Álvaro Obregón 64, 78000, San Luis Potosí SLP, México

M. Cardona

Max-Planck-Institut für Festkörperforschung Heisenbergstrasse 1, 70569 Stuttgart, Germany (Recibido: 26 de julio de 2005; Aceptado: 8 de diciembre de 2005)

The availability of isotopically pure semiconductors in the last fifteen years has triggered their scientific and technological interest. The effects of the electron-phonon interaction on the band structure can be experimentally investigated by measuring the temperature or the isotopic composition dependence of energy gaps. In this article, we discuss the effects of isotopic composition on the dielectric function of silicon by using spectroscopic ellipsometry in the energy range from 3.1 to 3.7 eV. The silicon crystals investigated are the isotopically pure 28 Si and 30 Si, and the natural Si (nat Si, M_{nat} = 28.14 amu). At low temperatures, the energies of the interband transitions become mass-dependent through the dependence of the electron-phonon interaction and the lattice parameter on the average isotopic mass. We determine the mass dependence of critical point energies and other optical parameters as accurately as possible by analyzing the ellipsometric data in reciprocal (Fourier-inverse) rather than direct (frequency) space.

Keywords: Optical constants; Semiconductors; Theory; Models; Numerical simulation

1. Introduction

The effects of isotopic composition on lattice constants and electronic properties of semiconductors have an increasing scientific and technological interest due to the availability of isotopically pure materials. Experimental measurements of the isotopic mass dependence of lattice constant for diamond [1,2] Ge [3, 4], Si [5], GaAs [4] and ZnSe [4], as well as theoretical calculations for diamond [6], Si [6, 7], Ge [4, 6], GaAs [4, 8] and ZnSe [4, 8] have been reported in the literature. Likewise, there are several reports concerning the effects of the isotopic composition on the interband energies of the lowest direct or indirect gaps of GaAs [4], Ge [9, 10], ZnSe [11], CdS [12], and CuCl [13]. For higher energies, spectroscopic ellipsometry (SE) has been used to measure the isotopic dependences of the gaps and broadening parameters corresponding to the E_1 and $E_1+\Delta_1$ transitions of Ge [9, 10] and Si [14, 15].

Physical properties of semiconductors such as thermal conductivity, thermal expansion and gap energies of the interband transitions, are determined by the propagation of vibrations (phonons) through the lattice and their interaction with electronic states. These properties can be modified, for instance, by changing the mass of the atoms that constitute the lattice, i.e., by modifying the isotopic composition of the semiconductor.

In the case of the gap energies, the dependence on the isotopic composition has been attributed to three effects. The simplest one is related to the thermal expansion [16], which is explained by assuming anharmonic potentials for the corresponding oscillators. In this case, the time-average position of the oscillators is no longer equal to a_0 , but leads to a lattice parameter which is shifted to $a_0+\Delta a_0$. The formalism used to describe thermal expansion is the so-

called quasi-harmonic approximation, which assumes that the interatomic potential is still harmonic, with the energy levels of the vibrations given by $E_n = (n+1/2)\hbar\omega$, but with the lattice constant dependent on temperature. The effect on the critical point energy is given by [17]:

$$\Delta\omega_{g}(T) = aB_{0}(\frac{\Delta V(T)}{V_{0}})_{P}$$

$$= (\frac{\partial\omega_{g}}{\partial P})_{T}(V_{0}\frac{\partial P}{\partial V})_{T}(\frac{\Delta V(T)}{V_{0}})_{P'}$$
(1)

Where a is the volume deformation potential, B_0 the bulk modulus, P the pressure and $\Delta V(T)$ the volume expansion. In Eq. (1), the term with the strongest temperature dependence is the thermal expansion $(\Delta V(T)/V_0)_p$.

The other two effects that modify the gap energies, representing explicitly the electron-phonon interaction, are (a) the Fan terms, corresponding to a complex self-energy, and (b) the Debye-Waller terms. These terms are of second order in the phonon amplitude u and the temperature effect appears in the thermodynamical expectation value $\langle u^2 \rangle$, which is given by [17]:

$$\langle u^2 \rangle = \frac{\hbar}{4M_i \Theta_i} [1 + 2n_B] \tag{2}$$

where $n_B=1/(e^{\Theta_j/T}-1)$ is an average Bose-Einstein statistical factor, and M_j the mass of the j^{th} isotope with the corresponding phonon frequency Θ_i .

 $\begin{array}{l} \textbf{Table 1.} \ \ \text{renormalization parameters } \ a_B \ \text{and unrenormalized gap } E_B \\ \text{for the } E_1 \ \text{transitions of Si obtained in the present work and from the literature (based on the measured and calculated temperature dependence of the } E_1 \ \text{transition of } ^{nat} \text{Si}). \end{array} \ \ \text{The value of the renormalization parameter includes both components: the electron-phonon interaction and the volume effect.}$

a _B (meV)	E _B (eV)
110±5	3.562±0.005
39±13	3.562 ± 0.014
70.6	3.5266
	(meV) 110±5 39±13

^aRef. [18] ^bRef. [27]

The shifts with temperature and the isotopic mass dependence of a given interband energy gap $E_g(T,M_j)$ are proportional to $< u^2 >$ and can be described as [17, 18]

$$E_g(T, M_j) = E_B - a_B \left(\frac{M_{nat}}{M_j}\right)^{1/2} [1 + 2n_B],$$
 (3)

where E_B and a_B are the unrenormalized (bare) gap and the renormalization parameter, respectively, at T=0. Here, M_{nat} is the average mass of the natural isotopic composition of the semiconductor under consideration, $\Theta_j = \Theta(M_{nat}/M_j)^{1/2}$ and Θ is the Debye temperature of the semiconductor under consideration, in our case natural Si. The gap renormalization parameter a_B includes two contributions: that of the electron-phonon interaction and the effect of thermal expansion. Numerical values of E_B , Θ , and a_B can be obtained by fitting Eq. (3) to the temperature dependence of the gaps.

In the low temperature limit (T« Θ_i), Eq. (3) reduces to

$$E_g(T, M_j) = E_B - a_B \left(\frac{M_{nat}}{M_j}\right)^{1/2}.$$
 (4)

Note that a_B is the energy difference between the unrenormalized gap $(M_j \rightarrow \infty)$ and the renormalized value [1, 17]. A similar expression can be written for the Lorentzian broadening parameter Γ [9, 10].

In the high-temperature regime $(T * \Theta_j)$, Eq. (3) can be written as

$$E_g(T, M_j) = E_B - 2T \frac{a_B}{\Theta}.$$
 (5)

In this limit, $E_g(T,M_j)$ is independent of M_j [1, 17]. The extrapolation of Eq. (5) to T=0 can be used to determine the unrenormalized gap energy E_B , i.e., the value that corresponds to atoms in fixed lattice positions without

vibrations, from the measured temperature dependence of $E_g(T)$ in the high temperature (i.e., linear in T) region.

Using Eq. (4) the difference in energy ΔE_g between a given energy gap in isotopically pure and natural material can be written as

$$\Delta E_g = E_g(M_j) - E_g(M_{nat}) = a_B \left[1 - \left(\frac{M_{nat}}{M_j} \right)^{1/2} \right],$$
 (6)

with a similar expression for Γ . In principle, a reliable value of a_B for a given gap can be obtained by fitting Eq. (3) from low temperatures, where Eq. (4) is valid, to temperatures high enough for the slope of Eq. (5) to be well defined.

In the present work we review the isotopic effects on the dielectric function ε of Si around the E₁ direct interband transitions. We discuss low-temperature (25 K) ellipsometric measurements for natural silicon ^{nat}Si, M_{nat}= 28.14 amu) and for isotopically pure ²⁸Si and ³⁰Si samples from 3.1 to 3.7 eV. To determine the critical point parameters as accurately as possible we analyze the spectra in reciprocal (Fourier-inverse) [19-21] rather than direct (frequency) space. The reciprocal space approach has several advantages over the most commonly used direct space, because the information concerning the baseline, the physically relevant information, and the noise contributions appear separately in the low-, medium-, and high-index Fourier coefficients, respectively.

2. Reciprocal Space Analysis

We consider an optical spectrum consisting of N data points f_1 , f_2 ,..., f_N equally spaced in energy from E_i to E_f , and with a periodicity such that $f_j = f_{j+N}$. For simplicity, we take N to be an even number. Then the f_j 's can be represented by the Fourier series [19-21]

$$f(\theta) = A_0 + \operatorname{Re}(\sum_{n=1}^{N/2} C_n e^{-i\xi_n} e^{in\theta}), \tag{7}$$

where θ and E are related by

$$\theta = \frac{E - E_{i}}{E_{s}},\tag{8}$$

and the factor E_s is given by

$$E_{\rm s} = \frac{N(E_{\rm f} - E_{\rm i})}{2\pi(N - 1)}. (9)$$

We assume that the lineshape of the optical structure in ϵ due to the transitions associated with a critical point at energy E_g is described in "direct" space by the Lorentzian [19]

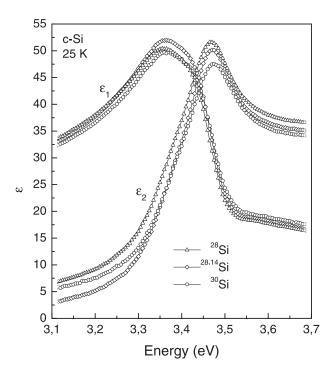


Figure 1. Spectra of the real (ϵ_1) and imaginary (ϵ_2) parts of the dielectric function for $\langle 111 \rangle$ - oriented ²⁸Si (triangles), ^{nat}Si (diamonds) and ³⁰Si (circles). The measurements were performed at 25 K. The ϵ spectra were obtained by correcting ellipsometrically measured $\langle \epsilon \rangle$ data for 35 Å of SiO₂.

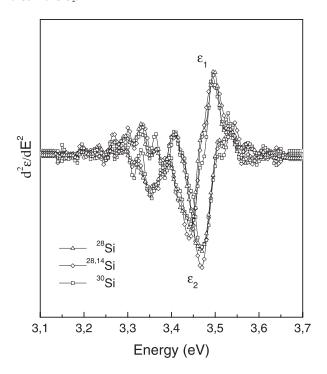


Figure 2. Second energy derivates calculated from spectra of Fig. 1. The spectra were smoothed by using the Savitzky-Golay filter method [26].

$$L(E) = B(E) + Ce^{i\beta}(E - E_g + i\Gamma)^{-\mu},$$
 (10)

where B(E) is a slowly varying background function, C and β are amplitude and phase factors, and μ is the order of the singularity. B(E) will only affect the lowest Fourier coefficients and can therefore be neglected in our analysis.

For the lineshape (Eq. (10)) and its κ^{th} derivative the complex Fourier coefficients are given by [19]

$$C_{n} = \frac{Cn^{\mu+\kappa-1}e^{-n\delta}}{E_{s}^{\mu+\kappa}\Gamma(\mu)},$$

$$\xi_{n} = \frac{(\mu-\kappa)\pi}{2} - \beta + n\theta_{g},$$
(11)

where $\delta = \Gamma/E_s$, $\theta_g = (E_g - E_i)/E_s$, and $\Gamma(\mu)$ is the gamma function

According to the shift theorem [22], if the Fourier inversion origin is moved from $\theta=0$ to $\theta=\theta_0$ the new Fourier coefficients are related to the unshifted ones by [19]

$$C'_{n} = C_{n},$$

$$\xi'_{n} = \xi_{n} - n\theta_{0}.$$
(12)

From Eqs. (11) and (12) it is clear that if $\theta_0 = \theta_g$, ξ_n will exhibit no dependence on n, in which case E_g is given by $E_g = \theta_0 E_s + E_i$.

The relative changes in E_g and Γ that result from a perturbation such as a change in isotopic mass can be expressed in a very simple way. From Eq. (11) we have

$$\Delta In(C_n) = -\frac{\Delta \Gamma}{E_s} = -\frac{\Gamma - \Gamma_{ref}}{E_s} n$$

$$\Delta \xi_n = \frac{\Delta E_g}{E_s} = \frac{E_g - E_{g,ref}}{E_s} n.$$
(13)

At this point, it is illustrative to compare the Lorentzian and Gaussian lineshapes in reciprocal space. In direct space a Gaussian is given by:

$$G(E) = \frac{-iCe^{i\beta}}{\Gamma}e^{-[(E-E_g)/\Gamma]^2},$$
(14)

and in reciprocal space by [21]

$$C_{n} = \frac{1}{\sqrt{\pi}} \frac{C n^{\mu+\kappa-1} e^{n^{2} \delta^{2}/2}}{E_{s}^{\mu+\kappa} \Gamma(\mu)},$$

$$\xi_{n} = \frac{(\mu-\kappa)\pi}{2} - \beta + n\theta_{g}.$$
(15)

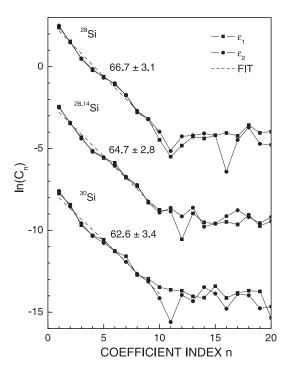


Figure 3. Natural logarithm of the amplitudes C_n of the Fourier transform coefficients of the data of Fig. 1, together with best fit straight lines to the average dependences over coefficients 1 through 10. The number given for each sample are the values of Γ (in meV) and their statistical uncertainties as calculated from the slopes of these lines. The white-noise region appears for $n \ge 11$.

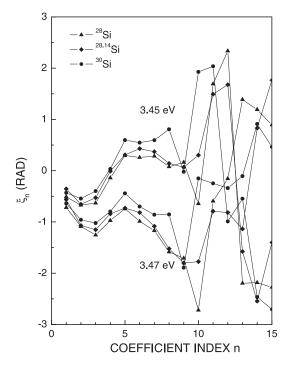


Figure 4. Same as Fig. 3, but for the averages of the phases ξ_n of ϵ_1 and ϵ_2 for inversion origins at 3.45 and 3.47 eV, as indicated.

From Eqs. (11) and (15) we note that $ln(C_n)$ is proportional to n and n^2 for Lorentzian and Gaussian lineshapes, respectively. Thus, by applying the inverse space formalism, the lineshape of a measured spectra can be determined directly from $ln(C_n)$.

Reciprocal space has several advantages over the real space for the analysis of optical spectra. Among them we mention:

- a) Base line effects, information and noise are separated in the low-, medium- and high-index Fourier coefficients, respectively.
- b) The order κ of a differentiation, appears in the reciprocal space as the factor $(n/E_s)^{\kappa}$ and the phase shift $-\kappa\pi/2$ in C_n and ξ_n , respectively.
- c) The energy of the gap can be estimated by shifting in energy the measured spectrum to obtain a phase ξ_n independent of n (Eq. (12)).
- d) Differences in gaps and broadening parameters between samples of different isotopic composition can be determined accurately using reciprocal space. This procedure can be applied also to the determination of relative gap shifts in systems such as quantum wells, mismatched heterostructures, composition in ternary systems, etc.
- e) The calculation of relative changes of gaps and broadening parameters is reduced to obtaining the slope of $\Delta \xi_n$ and $\Delta \ln(C_n)$, respectively. Equation (13) will be used to determine the relative dependence of E_g and Γ on isotopic composition.
- f) Line shapes such as Lorenzian and Gaussian can be directly recognized.
- g) The approach works fine only for one or two Lorenzians. The present formalism was specified to only one Lorenzian, however, if the spectrum under analysis consists of two Lorenzian transitions the formalism can be also applied [21].

3. Experimental results

In Ref. [14] three $\langle 111 \rangle$ oriented samples were investigated: natural Si (M_{nat} = 28.14) [23] and two nearly isotopically pure crystals, ²⁸Si and ³⁰Si. The details of the growth process can be found in Ref. [14]. The experiments were performed with a rotating analyzer ellipsometer [24], attached to a cryostat [14], with the samples at T=25±5K. At this temperature Eq. (4) is valid since $\Theta \approx 650$ K for Si. The surfaces were not prepared in any special manner prior to the measurement. Accordingly, the $\langle \epsilon \rangle$ spectra were corrected for the presence of SiO₂ layers [14] by applying the three-phase (bulk-oxide-ambient) model [25].

The ε spectra of the ²⁸Si, ^{nat}Si and ³⁰Si samples at 25 K from 3.1 to 3.7 eV are shown in Fig. 1 as triangles, diamonds, and circles, respectively. This energy range contains two critical points, E₀ and E₁[18], as evidenced by the dominant peak and the shoulder in the ε_1 spectrum of Fig. 1. Actually, the E_1 transition consists of an E_1 and an $E_1 + \Delta_1$ component, but since their separation, $\Delta_1 = 0.03$ eV, is far too small to be resolved, both transitions will be

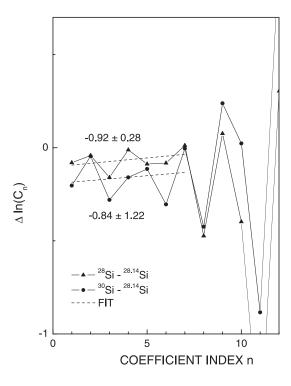


Figure 5. Differences between the ²⁸Si and ^{nat}Si (filled triangles) as well as the ³⁰Si and ^{nat}Si (filled circles) data of Fig. 2, along with best-fit straight (dashed) lines from coefficients 1 through 7. Values of $\Delta\Gamma$ (in meV) obtained from each fit are indicated.

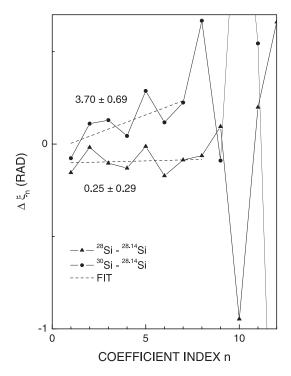


Figure 6. As Fig. 5, except for the phases. Values of ΔE_I (in meV) obtained from each fit are indicated.

treated as a single critical point. Considering also that the E_I transitions are stronger than the E_0 , the transitions can be treated as a single critical point and the reciprocal space formalism developed for only one critical point can be applied. The information obtained by this procedure should be mainly related to the E_I transitions.

The most common method to obtain precise values for parameters such E_g and Γ from optical spectra, is based on the fit of Eq. (10) to the second derivatives of the measured spectra with respect to energy. However, the noise/signal ratio increases in the derivative spectra and, in general, smoothing procedures are necessary to filter the noise with the consequence that the resolution and the information contained in the spectra deteriorate. This can be seen in Fig. 2, where we took the second energy derivatives of spectra of Fig. 1. The spectra were smoothed by using the Savitzky-Golay filter method [26]. The spectra look rather similar for the three samples and it is not possible, for instance, to see any difference among these gaps by fitting the second derivative of Eq. (10). Thus, in order to obtain reliable information we will analyze the spectra in reciprocal space.

Figure 3 shows the dependence of $ln(C_n)$ on n for each of the three samples investigated, as determined from both ε_1 (squares) and ε_2 (circles), along with best-fit straight lines (dashed) of the coefficients 1 through 10. It is seen that the data up to n=10 are rather well represented by straight lines, showing that, to the accuracy of the data, the critical points are indeed Lorentzian [21]. The Γ values determined from the slopes of the best-fit lines are 66.7 ± 3.1 , 64.7 ± 2.8 , and 62.6 ± 3.4 meV for the 28 Si, nat Si and 30 Si samples, respectively.

Results for ξ_n , taking inversion origins of 3.45 and 3.47 eV, are shown in Fig. 4. The presence of two critical points manifests itself in the beating oscillations, which have a period Δn of about 8, indicating a critical point separation of about 85 meV. Again, the white-noise region begins at about n=11 and extends to higher values. The average slopes of the phases are positive and negative for the inversion origins of 3.45 and 3.47 eV, respectively. This places, according to the shift theorem, E_g for the E_1 transitions somewhere between these energies. The gap energy of 3.454 \pm 0.002 eV at 30 K reported by Lautenschlager *et al* [18] is thus in good agreement with the value determined by using the inverse space method.

The uncertainty in E_g obtained from ξ_n is substantially larger than the expected isotope shift. However, this uncertainty results mainly from the interference oscillations between the contributions of the E_0 and E_1 constituents of the E_1 structure. Assuming that these have the same isotopic mass dependence, we can eliminate the associated inaccuracies by subtracting the $ln(C_n)$ data (see Fig. 3) of the natural silicon sample from those of the other two specimens.

We show in Figs. 5 and 6 the dependence of $\Delta In(C_n)$ and $\Delta \xi_n$ on n. According to Eq. (13) the slope of $\Delta In(C_n)$ and $\Delta \xi_n$ is proportional to the difference in Γ and E_g , respectively. From the slope of Fig. 5 determined by the

points 1 through 8 (dashed straight lines) we find that the broadening parameter of the 28 Si sample is smaller than that of the natural sample by 0.9 ± 0.3 meV. The broadening of the 30 Si and nat Si samples is the same to within experimental uncertainty.

Repeating the ξ_n analysis for points 1 through 8 (Fig. 6), we found that the critical point energy of the 30 Si sample is higher than that of the nat Si sample by 3.7 ± 0.7 meV. The 28 Si and nat Si samples exhibit essentially the same value of Eg; the higher value of 0.25 ± 0.3 meV obtained for the 28 Si sample is smaller than the statistical uncertainty. A theoretical value for $\Delta E_1 = 1.2$ meV has been obtained from pseudopotential calculations of the temperature dependence of the gap [27]. This value is close to our experimental result. Experiment and theory support the fact that the differences between the results for the 28 Si and natural samples should be negligible. We note that it would be impossible to obtain this accuracy by direct-space analysis.

From the values of ΔE_1 and Eq. (6), we can calculate values of a_B and E_B . These results and those calculated from literature data for $E_1(T)$, are given in table 1. The values of ΔE_I include both components of the gap shift, the electron-phonon interaction and the thermal expansion effect. The thermal expansion component can be estimated by writing Eq. (1) in the form $\Delta E_I^{vol} = (dE_I/dlnV)(dlnV/dM) \Delta M$ and the values of $dE_I/dlnV = 5.1$ eV [28], $dlnV/dM = 6.6 \times 10^5/amu$ [7], and $\Delta M = 1.8$ amu. This gives a value of $\Delta E_I^{vol} = 0.61$ meV.

4. Conclusions

We have discussed the effects of isotopic mass on the E_1 critical point parameters of Si by analyzing ellipsometric data for the dielectric function taken from 3.1 to 3.7 eV at 25 K. To obtain the maximum possible accuracy we determined the critical point parameters by reciprocal space analysis rather than using the more common direct (frequency) space formalism. We show that the gap shift with isotopic mass between the ^{30}Si and ^{nat}Si crystals is 3.7 ± 0.7 meV.

Acknowledgments

We would like to thank D.E. Aspnes for having introduced us to the reciprocal space method of analysis. Thanks are also due to D. Rönnow, T. Ruf, M. Konuma and S.D. Yoo for their collaboration in earlier versions of this work.

References

- M. Cardona, in Festkörperprobleme/ Advances in Solid State Physics, edited by R. Helbig, (Vieweg, Braunschweig, Wiesbaden 1994) p. 35.
- [2] H. Holloway, K.C. Hass, M.A. Tamor, T. R. Anthony and W. F. Banholzer, Phys. Rev. B 44, 7123 (1991).
- [3] R. C. Buschert, A. E. Merlini, S. Pace, S. Rodriguez, and M. H. Grimsditch, Phys. Rev. B 38, 5219 (1988).
- [4] N. Garro, A. Cantarero, M. Cardona, A. Göbel, T. Ruf and K. Eberl, Phys. Rev. B 54, 4732 (1996).
- [5] Kazimirov, J. Zegenhagen, and M. Cardona, Science, 282, 930 (1998).
- [6] P.Pavone, and S. Baroni, Solid State Commun. 90 295 (1994).
- [7] C. P. Herrero, Solid State Commun. 110, 243 (1999).
- [8] Debernardi and M. Cardona, Phys. Rev. B 54, 11305 (1996).
- [9] S. Zollner, M. Cardona and S. Gopalan, Phys. Rev. B 45,
- [10] D. Rönnow, L.F. Lastras-Martínez and M. Cardona, Eur. Phys. J. B 5, 29 (1998).
- [11] Göbel, T. Ruf, J. M. Zhang, R. Lauck, and M. Cardona, Phys. Rev. B 59, 2749 (1999).
- [12] J. M. Zhang, T. Ruf, R. Lauck, and M. Cardona, Phys. Rev. B 57, 9716 (1998).
- [13] Göbel, T. Ruf, C.T. Lin, M. Cardona, J.C. Merle, and M. Joucla, Phys. Rev. B 56, 210 (1997).
- [14] L. F. Lastras-Martínez, T. Ruf, M. Konuma and M. Cardona, Phys. Rev. B 61, 12946 (2000).
- [15] S. D. Yoo, D. E. Aspnes, L. F. Lastras-Martínez, T. Ruf, M. Konuma and M. Cardona, phys. stat. sol. (b) 220, 117 (2000).
- [16] J. Bardeen, W. Schockley, Phys. Rev. 80, 72 (1950).
- [17] M. Cardona, Solid State Commun. 133, 3 (2005).
- [18] P. Lautenschlager, M. Garriga, L. Viña and M. Cardona, Phys. Rev. B 36, 4821 (1987).
- [19] D. E. Aspnes, Surface Sci. 135, 284 (1983).
- [20] D. E. Aspnes, and H. Arwin, J. Opt. Soc. Am. 73, 1759 (1983).
- [21] D. E. Aspnes, Solar Energy Materials and Solar Cells **32**, 413 (1994)
- [22] R. N. Bracewell, *The Fourier transform and its applications* (McGraw-Hill, Singapore, 1986).
- [23] J. Emsley, The Elements (Clarendon Press, Oxford, 991).
- [24] D.E. Aspnes and A.A. Studna, Applied Optics **14**, 220 (1975).
- [25] R. M. A. Azzam and N.M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1997).
- [26] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, 1997), p. 650.
- [27] P. Lautenschlager, P. B. Allen and M. Cardona, Phys. Rev. B 31, 2163 (1985).
- [28] W. Kress in Landolt-Börnstein Tables, New Series, Vol. 17a, Physics of Group IV Elements and III-V Compounds, edited by O. Madelung, 64 (Springer-Verlag, Berlin-Heidelberg, 1982).