Many-Body Effects on the Zero-Point Renormalization of the Band Structure

G. Antonius,1,* S. Poncé,2 P. Boulanger,3 M. Côté,1 and X. Gonze2

1Département de physique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Canada H3C 3J7
2IMCN-NAPS, Université catholique de Louvain, Place Croix du Sud 1, B-1348 Louvain-la-Neuve, Belgium
3Institut Néel, 25 avenue des Martyrs, BP 166, 38042 Grenoble cedex 9, France

(Received 31 May 2013; revised manuscript received 16 May 2014; published 29 May 2014)

We compute the zero-point renormalization (ZPR) of the optical band gap of diamond from many-body perturbation theory using the perturbative $G_0W_0$ approximation as well as quasiparticle self-consistent GW. The electron-phonon coupling energies are found to be more than 40% higher than standard density functional theory when many-body effects are included with the frozen-phonon calculations. A similar increase is observed for the zero-point renormalization in GaAs when $G_0W_0$ corrections are applied. We show that these many-body corrections are necessary to accurately predict the temperature dependence of the band gap. The frozen-phonon method also allows us to validate the rigid-ion approximation which is always present in density functional perturbation theory.

DOI: 10.1103/PhysRevLett.112.215501 PACS numbers: 63.20.kd, 63.20.dk, 71.15.Mb, 78.20.-e

The coupling of electrons to a bosonic field generally causes a renormalization of the energy levels. Whereas in vacuum, the electromagnetic fluctuations lead to the Lamb shift observed in the hydrogen atom levels, in condensed matter, the phonon field renormalizes the band structure, even at zero temperature. Being as large as several hundreds of meV [1], this renormalization is critical to the predictive power of ab initio calculations when it comes to absorption spectra [2], photovoltaic materials [3], or topological insulators [4].

Following the early work of Fan and others in the 1950 s [5–7], the problem was addressed by Allen, Heine, and Cardona (AHC) [8,9], whose theory provides perturbative expressions in terms of the electron-phonon coupling; they find, at the lowest order, that two diagrams contribute to the renormalization, the Fan diagram coming from two first-order electron-phonon coupling vertices, and the Debye-Waller diagram coming from one second-order vertex. Using semiempirical methods, and later on, density functional theory (DFT), the temperature dependence of the band gap could be obtained for several semiconductors [10–18]. Among those, diamond has been a case study [19], both for the strong band gap renormalization it exhibits [20–25] and its phonon-driven superconductivity enabled by boron doping [26–29].

The reliability of DFT for the electron-phonon coupling has however been questioned in recent years. Since the scattering of an electron by a phonon probes the excited states of a system, a theory describing this process should rely on an accurate unrenormalized band structure, unlike the one of DFT. This has motivated the use of nonlocal DFT functionals [30–33]. These studies have shown that, as well as correcting the band gap, exact exchange functionals are also necessary to accurately describe the electron-phonon coupling. A truly ab initio scheme, however, would rely on many-body perturbation theory. As such, it was reported that $G_0W_0$ corrections led to a significant increase of the electron-phonon coupling in $C_{60}$ fullerene and in graphene [31,32]. In this work, we show that the $G_0W_0$ [34] and GW [35] treatments of the electron-electron interaction enhance the zero-point renormalization (ZPR) in diamond by more than 40% with respect to DFT in the local-density approximation (LDA), and that these corrections allow us to obtain the correct temperature dependence of the band gap.

We combine the frozen-phonon method and density functional perturbation theory (DFPT) [36,37] to compute the direct band gap renormalization. This allows us to revise an important approximation of the AHC theory, namely, the rigid-ion approximation. In all perturbative calculations, the Debye-Waller interaction term is simplified to allow its computation from linear response. The approximation breaks down in the case of diatomic molecules [38], but its reliability in solids has not been verified to our knowledge. Here we assert the validity of the rigid-ion approximation in diamond, and in general for three-dimensional systems.

Method.—The temperature dependence of the electronic eigenvalues originates from the phonon population and the thermal expansion of the lattice [1,13], the latter effect being relatively small. Neglecting dynamical effects (i.e., the phonon energy is assumed to be small with respect to electronic excitations), the phonon contribution gives

$$\varepsilon_\alpha(T) = \varepsilon_\alpha^0 + \frac{1}{N} \sum_j \frac{\partial \varepsilon_\alpha}{\partial n_j} \left[ n_j(T) + \frac{1}{2} \right],$$

where $\varepsilon_\alpha^0$ are the eigenvalues at equilibrium, and the sum over the $N$ phonon modes involves the electron-phonon coupling energies (EPCEs) $\partial \varepsilon_\alpha / \partial n_j$, and the Bose-Einstein occupation numbers $n_j$. With the frozen-phonon method,
and in the harmonic approximation, the EPCEs are obtained from the second-order derivatives of the eigenvalues at equilibrium:

$$\frac{\partial \epsilon_{\alpha}}{\partial n_j} = \frac{\hbar}{2M\omega_j} \frac{\partial^2}{\partial \varepsilon_j^2} \epsilon_{\alpha} [\mathbf{u}_{\tau}^j]' |_{\varepsilon_j = 0},$$

(2)

where $M$ is the reduced atomic mass, $\omega_j$ is the phonon frequency, and $\epsilon_{\alpha} [\mathbf{u}_{\tau}^j]'$ is an eigenvalue computed with the atoms displaced along the normalized polarization vector $\mathbf{u}_{\tau}^j$, with $\tau$ labeling the atoms of the unit cell. Each EPCE requires a supercell calculation to account for the wavelength of the phonon. Although this makes the technique more computationally demanding than DFPT, it offers several theoretical advantages. It makes no approximation for the Debye-Waller term, and it gives complete liberty on the method to compute the electronic structure.

We compute the ZPR of the top of the valence band of diamond ($\Gamma_{25v}$) and the first optically accessible state ($\Gamma_{15c}$) using a $4 \times 4 \times 4$ $\Gamma$-centered $q$-point grid. The lattice parameter (3.55 Å) was obtained by relaxation of the structure with a Trouiller-Martins LDA pseudopotential [39]. All calculations were done with the ABINIT code [40].

**Rigid-ion approximation.**—In the AHC theory, the EPCEs are expressed as

$$\frac{\partial \epsilon_{\alpha}}{\partial n_j} = \frac{\hbar}{2M\omega_j} \sum_{\tau, \tau'} \Phi_{\tau, \tau'} \left[ \mathbf{u}_{\tau}^j \mathbf{u}_{\tau'}^j - \frac{1}{2} (\mathbf{u}_{\tau}^j \mathbf{u}_{\tau'}^j + \mathbf{u}_{\tau}^j \mathbf{u}_{\tau'}^j) \right],$$

(3)

where $\Phi$ is the Hessian matrix of an eigenvalue derived with respect to all atomic positions, given by

$$\Phi_{\tau, \tau'} = (\langle \psi_{\alpha} | \nabla_{\tau} \nabla_{\tau'} H | \psi_{\alpha} \rangle + c.c.) + \langle \psi_{\alpha} | \nabla_{\tau} H | \psi_{\alpha} \rangle.$$  

(4)

The first term in Eq. (3) splits into the Fan and the Debye-Waller terms, when combined with the first and second terms of $\Phi$, respectively. The former describes the scattering of electronic states by a phonon, and the latter corresponds to a second-order electron-phonon interaction, which is difficult to compute in DFPT. By virtue of the acoustic sum rule [8], the two “rigid-ion” terms in Eq. (3) give a null contribution to the EPCE, but they allow for the desired approximation. Assuming the Debye-Waller term to be diagonal in atom sites ($\tau = \tau'$), the second-order derivative can be completely dropped, since Eq. (3) is explicitly off diagonal. The rigid-ion terms now assume the Debye-Waller contribution using only the first-order derivatives of the Hamiltonian and the wave functions, which are obtained self-consistently in DFPT. This simplification of the perturbation theory expressions is a major achievement of the AHC theory, albeit depending critically on the validity of the rigid-ion approximation.

The DFPT scheme allows for a fine sampling of the Brillouin zone (BZ), as shown on Fig. 1. Diamond being an indirect gap insulator, a divergence occurs in the $\Gamma_{15c}$ EPCEs when a phonon wave vector connects this state to another one with close energy. This is handled by adding a small imaginary part to the energy differences (0.1 eV) [41]. Although most of the divergences cancel out after integration, the broad peak at $\Gamma$ gives an important contribution to the renormalization of electronic energies. The central region covering one eighth of the BZ, accounts for 45% of the $\Gamma_{25v}$ ZPR and 20% of the $\Gamma_{15c}$ ZPR.

The rigid-ion approximation is not present in the frozen-phonon approach. Remarkably, using minute displacements (~0.01% of the bond length), the coupling energies computed from Eq. (2) fall closely on the DFPT curve. The total ZPR obtained with the frozen-phonon method differs by less than 3% from the DFPT value (see Table I), a discrepancy we attribute to the rigid-ion approximation. Since the neglected nondiagonal Debye-Waller term is a short-ranged interaction between neighboring atomic sites, it does not become any stronger near $\Gamma$ than it is at the BZ boundary. Therefore, this term should be of negligible importance in any three-dimensional crystal, where the BZ center area accounts for a significant portion of the ZPR.
The lowest conduction band reaches its minimum. The elements are positive in the $\Gamma$ negative renormalization of contracting bonds to the stretching bonds, in favor of lower of the bond length causes the electrons to move from the conduction band at the top of the valence band (lower panel) and the bottom of the conduction band at $\Gamma$ (top panel), in DFT (squares), in $G_0W_0$ (circles) and in $GW$ (stars).

**Many-body corrections.**—Displacing the atoms by 0.1% of the bond length causes the electrons to move from the contracting bonds to the stretching bonds, in favor of lower kinetic energy. This charge transfer pushes the $\Gamma_{25v}$ energy upward, causing a renormalization of 119 meV in the LDA. With the $G_0W_0$ corrections [41], the EPCE is uniformly increased all over the Brillouin zone by about 50 meV per $q$ point, as shown on Fig. 2. Although the $\Gamma_{15c}$ state shows a negative renormalization of $-318$ meV in LDA, the EPCE elements are positive in the $\Gamma-X$ direction, where the lowest conduction band reaches its minimum. The $G_0W_0$ corrections increase the amplitude of the negative EPCE elements while reducing the positive elements, bringing the $\Gamma_{15c}$ ZPR to $-477$ meV, and increasing the band gap renormalization by more than 40%, as reported in Table I. We observed the same trend in gallium arsenide, which is a direct band gap semiconductor. Whereas the LDA predicts a band gap renormalization of $-23$ meV in this material, the ZPR is increased by $-10$ meV when $G_0W_0$ corrections are applied.

The self-energy is thus more sensitive to the perturbation than the LDA exchange-correlation potential. This reflects the fact that the LDA is based on the jellium model. The electron-electron interaction is overly screened in the bonding region, where the charge density packs up. In contrast, retaining only the bare exchange corresponds to a Hartree-Fock calculation with the DFT wave functions. Such a calculation overestimates the renormalization of the $\Gamma_{25v}$ (207 meV), which is located in the bonds, while the $\Gamma_{15c}$ ZPR ($-473$ meV) does not suffer from the lack of correlation. This illustrates the importance of proper screening in the high density region. The same feature was reported in graphene to a higher extent, due to the shorter $s^p$ bonds [32].

Quasiparticle $GW$ calculations of the EPCE were performed on a subset of our $q$-point grid ($\Gamma$, $\Lambda$, $L$, $X$). The self-consistency increases slightly the EPCE of the $\Gamma_{25v}$ ($\sim8$ meV, on average) and the $\Gamma_{15c}$ ($\sim-22$ meV). Such a small change results from two cancelling effects. While the self-energy allows for a greater interstitial charge density, it also reduces the electron mobility by opening the band gap much more than $G_0W_0$ does [41]. Overall, the LDA + $G_0W_0$ calculation seems to agree well with the electron-phonon coupling obtained from self-consistent $GW$.

**Temperature dependence of the band gap.**—Our coarse sampling of the Brillouin zone might not capture the relative importance of the strongly coupling modes. Hence, we interpolate the many-body corrections on a dense $q$-point grid using a cubic polynomial fit for the $G_0W_0$ corrections to the EPCE [41]. We obtain the band gap renormalization and its temperature dependence from a DFT calculation on a $32 \times 32 \times 32$ $q$-point grid, as reported in Table II. This fine sampling reveals a bigger $G_0W_0$ correction to the $\Gamma_{15c}$ ZPR due to the strong positive contributions near $X$ being reduced and those near $\Gamma$ being enhanced. As a result, the band gap ZPR is increased from $-404$ meV to $-628$ meV when many-body corrections are applied.

**Table I.** Zero-point renormalization (meV) computed on a $4 \times 4 \times 4$ $q$-point grid in frozen-phonon (finite differences) and in DFPT (perturbative DFT).

<table>
<thead>
<tr>
<th>Approximation</th>
<th>$\Gamma_{25v}$</th>
<th>$\Gamma_{15c}$</th>
<th>Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rigid-ion approximation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perturbative DFT</td>
<td>113</td>
<td>$-314$</td>
<td>$-427$</td>
</tr>
<tr>
<td>Finite differences harmonic</td>
<td>117</td>
<td>$-320$</td>
<td>$-437$</td>
</tr>
<tr>
<td>Finite differences DFT</td>
<td>119</td>
<td>$-318$</td>
<td>$-437$</td>
</tr>
<tr>
<td>Finite differences $G_0W_0$</td>
<td>145</td>
<td>$-477$</td>
<td>$-622$</td>
</tr>
</tbody>
</table>

$^a$Phonon displacement: $\sim0.01\%$ of the bond length.

$^b$Phonon displacement: $\sim0.1\%$ of the bond length.

**Table II.** Zero-point renormalization (meV) computed on a $32 \times 32 \times 32$ $q$-point grid. The $G_0W_0$ and $GW$ corrections are interpolated with the models discussed in the text.

<table>
<thead>
<tr>
<th>Correction</th>
<th>$\Gamma_{25v}$</th>
<th>$\Gamma_{15c}$</th>
<th>Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_0W_0$</td>
<td>141</td>
<td>$-263$</td>
<td>$-404$</td>
</tr>
<tr>
<td>$\Delta G_0W_0$</td>
<td>$+26$</td>
<td>$-183$</td>
<td>$-209$</td>
</tr>
<tr>
<td>$\Delta GW-G_0W_0$</td>
<td>$+4$</td>
<td>$-11$</td>
<td>$-15$</td>
</tr>
<tr>
<td>Total</td>
<td>171</td>
<td>$-457$</td>
<td>$-628$</td>
</tr>
</tbody>
</table>
The quasiparticle peak in the electron-phonon self-energy causes a spreading of the band gap. Effects have a small impact, changing the band gap. It was checked, using the dynamical AHC theory, that those effects should include dynamical effects as well. We performed in the adiabatic approximation, whereas the full treatment was made. On the one hand, our calculations were consistent further increased the coupling, but the effect captures most of the many-body corrections. Those corrections increase the zero-point renormalization of the band gap as well as its high-temperature slope by more than 40%. Overall, we find that the DFPT + GW scheme reproduces accurately the temperature dependence of the direct band gap of diamond. While DFPT remains the most efficient method to treat the electron-phonon interaction, our results call for a more accurate scheme. Ideally, an improvement of the linear response would include nonlocal exchange and dynamical screening in the ways of the many-body perturbation theory. We hope this work motivates such development.

G. Antonius and M. Côté thank the NSERC and FRQNT for the financial support, and Calcul Québec for the computational resources. This work was also supported by the FRS-FNRS through a FRIA grant (S.P.). Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the Fonds de la Recherche Scientifique de Belgique (FRS-FNRS). The authors also acknowledge the technical support of Yann Pouillon and Jean-Michel Beuken, and many discussions about electron-phonon effects on band gap energies with A. Marini, F. Giustino, M. Verstraete, and C. Draxl.

**gabriel.antonius@gmail.com**

[19] We note important discrepancies between our DFPT results and those of Refs. [22,23]. These differences are well understood and are discussed in Ref. [25].
[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.112.215501 for computational details on DFPT calculations, $G_0W_0$ and GW calculations, finite differences calculations, interpolation scheme, and detailed results for GaAs.