

Université de Montréal

Étude de la dépendance en température de la
structure électronique à l'aide de la théorie de la
fonctionnelle de la densité : effets non adiabatiques,
dilatation du point zéro, couplage spin-orbite et
application aux transitions de phase topologiques

par

Véronique Brousseau-Couture

Département de physique
Faculté des arts et des sciences

Thèse présentée en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Physique

Juillet 2022

Université de Montréal

Faculté des arts et des sciences

Cette thèse intitulée

**Étude de la dépendance en température de la structure
électronique à l'aide de la théorie de la fonctionnelle de la densité :
effets non adiabatiques, dilatation du point zéro, couplage
spin-orbite et application aux transitions de phase topologiques**

présentée par

Véronique Brousseau-Couture

a été évaluée par un jury composé des personnes suivantes :

Richard Leonelli

(président-rapporteur)

Michel Côté

(directeur de recherche)

Sébastien Francoeur

(membre du jury)

Samuel Poncé

(examineur externe)

Christian Pellerin

(représentant du doyen de la FESP)

Résumé

Les signatures de l'existence des phonons sont omniprésentes dans les propriétés des matériaux. En première approximation, on peut scinder l'effet des phonons sur la structure électronique en deux contributions. D'une part, l'interaction électron-phonon capture la réponse électronique aux vibrations des noyaux du cristal, et d'autre, l'énergie libre de la population de phonons modifie le volume cristallin à l'équilibre. En plus d'être responsables de la dépendance en température de la structure électronique, ces deux mécanismes affectent les niveaux d'énergie à température nulle, à travers le mouvement du point zéro et l'énergie du point zéro.

Cette thèse analyse l'apport de ces deux mécanismes à la renormalisation du point zéro (ZPR) de l'énergie de la bande interdite des semi-conducteurs. Une généralisation du modèle de Fröhlich prenant en compte l'anisotropie et les dégénérescences présentes dans les matériaux réels révèle que l'interaction non adiabatique entre les électrons et les noyaux domine le ZPR dans les matériaux polaires. La prise en compte de ce mécanisme dans l'évaluation de l'interaction électron-phonon est déterminante pour reproduire adéquatement les données expérimentales.

L'approche développée par Grüneisen, qui néglige communément les effets du point zéro, reproduit la dilatation du point zéro du réseau (ZPLE) et sa contribution au ZPR obtenues avec la méthode standard basée sur la minimisation de l'énergie libre à moindre coût numérique, y compris pour les matériaux anisotropes. La contribution du ZPLE au ZPR total, qui a reçu peu d'attention dans la littérature, peut atteindre de 20% à plus de 80% de la contribution de l'interaction électron-phonon, y compris dans des matériaux constitués de noyaux légers. Elle domine même le ZPR du GaAs dans le contexte de la DFT semi-locale. Il est donc essentiel de traiter les deux contributions sur le même pied d'égalité pour modéliser le ZPR avec précision.

L'inclusion du couplage spin-orbite (SOC) diminue le ZPR d'un ensemble substantiel de matériaux cubiques de structure zinc-blende, diamant et rock-salt. L'essentiel de cette variation tire son origine de l'effet du SOC sur les énergies électroniques statiques, qui provient en grande partie de la variation des masses effectives des bandes de valence au point Γ . La réduction du ZPR peut être estimée à partir d'un modèle de Fröhlich généralisé auquel

on a introduit le SOC. Les subtilités numériques liées au traitement de la séparation de Dresselhaus dans les matériaux non centrosymétriques sont discutées.

On démontre enfin comment l'effet combiné de l'interaction électron-phonon et de la dilatation thermique affecte le diagramme de phase topologique du BiTeI. L'augmentation de la température repousse l'apparition de la phase d'isolant topologique \mathbb{Z}_2 vers des pressions plus élevées et élargit la plage de pressions correspondant à la phase intermédiaire de type semi-métal de Weyl. Le caractère orbital dominant des extrema de bande influence significativement leur sensibilité à la pression et au changement de topologie. Pour guider la recherche expérimentale de phases topologiquement non triviales dans les matériaux de façon adéquate, les études numériques doivent donc considérer l'effet de la température.

Mots clés : interaction électron-phonon, renormalisation du point zéro, dilatation thermique, dilatation du point zéro du réseau, théorie de la fonctionnelle de la densité perturbative, modèle de Fröhlich , effets non adiabatiques, couplage spin-orbite, isolants topologiques, transitions de phase topologiques.

Abstract

Phonon signatures are ubiquitous in material properties. At first order, the effect of phonons on the electronic structure can be split into two contributions. On the one hand, the electron-phonon interaction captures the electronic response to the vibrations of the nuclei. On the other hand, the free energy of the phonon population modifies the crystalline volume at equilibrium. In addition to driving the temperature dependence of the electronic structure, these two mechanisms affect the energy levels at zero temperature through zero-point motion and zero-point energy.

This thesis investigates the contribution of these two mechanisms to the zero point renormalization (ZPR) of the band gap energy of semiconductors. A generalized Fröhlich model taking into account the anisotropy and degeneracies occurring in real materials reveals that the non-adiabatic interaction between electrons and nuclei dominates the ZPR in polar materials. Taking this mechanism into account when evaluating the electron-phonon interaction is crucial to reproduce experimental data adequately.

The Grüneisen formalism, which commonly neglects zero-point effects, reproduces the zero-point lattice expansion (ZPLE) and its contribution to the ZPR obtained from the standard method based on free energy minimization at lower numerical cost, including for anisotropic materials. The ZPLE contribution to the total ZPR, which has received little attention in the literature, can reach from 20% to more than 80% of the contribution of the electron-phonon interaction, including in materials containing light atoms. It even dominates the ZPR of GaAs within semilocal DFT. Therefore, both contributions should be treated on an equal footing to model the ZPR accurately.

The inclusion of spin-orbit coupling (SOC) decreases the ZPR of a substantial set of cubic materials of zincblende, diamond and rocksalt structure. This variation originates mostly from the effect of SOC on the static electronic eigenvalues, which comes largely from the variation of the effective masses of the valence bands at the Γ point. The reduction of the ZPR can be estimated from a generalized Fröhlich model in which SOC has been introduced. Numerical subtleties related to the treatment of Dresselhaus separation in non-centrosymmetric materials are discussed.

We finally show how the combination of electron-phonon interaction and thermal expansion affects the topological phase diagram of BiTeI. An increase in temperature pushes the \mathbb{Z}_2 topological insulator phase towards higher pressures and widens the pressure range corresponding to the Weyl semi-metal intermediate phase. The leading orbital character of the band extrema significantly influences their sensitivity to variations in pressure and topology. To adequately guide the experimental search for topologically non-trivial phases in materials, numerical studies must therefore consider the effect of temperature.

Keywords : electron-phonon interaction, zero-point renormalization, thermal expansion, zero-point lattice expansion, density functional perturbation theory, Fröhlich model, non-adiabatic effects, spin-orbit coupling, topological insulators, topological phase transitions.

Table des matières

Résumé	v
Abstract	vii
Liste des tableaux	xv
Liste des figures	xvii
Liste des sigles et des abréviations	xxi
Remerciements	xxv
Avant-propos	xxvii
Chapitre 1. Introduction	1
1.1. Principe d'émergence et quasi-particules	2
1.2. La température du point de vue atomique : une population de déplacements collectifs	4
1.3. Une multitude d'approximations	5
Chapitre 2. Structure électronique dans l'approximation Born-Oppenheimer	9
2.1. Un problème à la fois simple et insoluble	10
2.2. Approximation des noyaux statiques	11
2.3. Théorie de la fonctionnelle de la densité	12
2.3.1. Fonctions et fonctionnelles	13
2.3.2. Théorèmes de Hohenberg-Kohn	13
2.3.3. Équations de Kohn-Sham	15
2.3.4. Le problème de la bande interdite	18
2.4. Propriétés des états électroniques sous les opérations de symétrie	19
2.4.1. États de Bloch	19

2.4.2.	Opérations de symétrie générales	20
2.4.3.	Symétrie d'inversion	22
2.4.4.	Symétrie de renversement du temps et théorème de Kramers	22
2.5.	Couplage spin-orbite	24
2.5.1.	Interprétation semi-classique	25
2.5.2.	Précession de Thomas et émergence du SOC dans l'équation de Dirac	26
2.5.3.	Conséquences sur les niveaux d'énergie électroniques	27
2.5.4.	Interactions Dresselhaus et Rashba dans les matériaux sans symétrie d'inversion	30
2.6.	Conclusion	33
	Bibliographie	33
Chapitre 3.	Théorie de l'interaction électron-phonon	35
3.1.	Éléments fondamentaux de la théorie à N -corps	37
3.1.1.	Seconde quantification et opérateurs de champ fermioniques	37
3.1.2.	Fonction de Green et self-énergie	39
3.2.	Les phonons dans l'approximation harmonique	42
3.2.1.	Hamiltonien des électrons et des noyaux	42
3.2.2.	Hamiltonien des noyaux dans l'approximation harmonique	44
3.2.3.	Effets anharmoniques	46
3.3.	Dépendance en température des énergies électroniques	48
3.3.1.	Self-énergie de l'interaction électron-phonon	48
3.3.2.	Approximation des ions rigides	53
3.4.	Théorie de la fonctionnelle de la densité perturbative	56
3.4.1.	Développement perturbatif	57
3.4.2.	Factorisation de la phase	60
3.4.3.	Composante non analytique	61
3.5.	Modèle de Fröhlich	63
	Bibliographie	66
Chapitre 4.	Prédominance des effets non adiabatiques dans la renormalisation du point zéro de l'énergie de la bande interdite	71

4.0.1. Mise en contexte.....	71
4.0.2. Contributions des auteurs.....	74
Premier article. Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap.....	75
4.1. Introduction.....	77
4.2. Results.....	80
4.3. Discussion.....	86
4.4. Methods.....	88
Supplementary Information.....	94
Bibliography.....	120
Chapitre 5. Dilatation du point zéro du réseau cristallin.....	131
5.0.1. Origine anharmonique de la dilatation thermique.....	131
5.0.2. Mise en contexte.....	137
5.0.3. Contributions des auteurs.....	138
Deuxième article. Zero-point lattice expansion and band gap renormalization : Grüneisen approach versus free energy minimization.....	139
5.1. Introduction.....	141
5.2. Methodology.....	143
5.2.1. Zero-point renormalization of the band gap energy.....	143
5.2.2. Zero-point lattice expansion in the quasiharmonic approximation.....	145
5.2.3. Estimation of ZPLE correction to the band gap energy.....	151
5.2.4. Electron-phonon interaction.....	153
5.3. Computational details.....	155
5.3.1. Ground state calculations and lattice optimization.....	155
5.3.2. Lattice dynamics and electron-phonon interaction.....	156
5.3.3. Zero-point lattice expansion and band gap correction.....	156
5.4. Results and discussion.....	158
5.4.1. Zero-point lattice expansion.....	158
5.4.2. Free energy minimization vs Grüneisen approach.....	160
5.4.3. ZPLE renormalization of the band gap energy.....	163

5.4.4. Further discussion	175
5.5. Conclusion.....	176
Supplemental Material.....	178
Bibliography.....	188
Chapitre 6. Effet du couplage spin-orbite sur la renormalisation du point zéro de l'énergie de la bande interdite des matériaux cubiques	199
6.0.1. Mise en contexte.....	199
6.0.2. Contributions des auteurs.....	200
Troisième article. Effect of spin-orbit coupling on the zero-point renormalization of the electronic band gap in cubic materials : first principles and generalized Fröhlich model	201
6.1. Introduction	203
6.2. Methodology.....	206
6.2.1. AHC formalism.....	206
6.2.2. Spin-orbit interaction.....	209
6.2.3. Generalized Fröhlich model	210
6.2.4. Generalized Fröhlich model in the presence of Dresselhaus splitting.....	214
6.3. Computational details	216
6.3.1. First-principles calculations	216
6.3.2. Generalized Fröhlich model	217
6.4. Results and discussion	217
6.4.1. First-principles.....	217
6.4.2. Generalized Fröhlich model with SOC.....	226
6.5. Conclusion.....	231
Appendix.....	232
Supplemental Material.....	241
Bibliography.....	252
Intermède	261
Chapitre 7. Topologie des structures de bande	263

7.1. Concept mathématique	263
7.2. Modèle SSH du polyacétylène	264
7.2.1. Hamiltonien du cœur de la chaîne	267
7.2.2. Déformation adiabatique de l'hamiltonien et équivalence topologique	270
7.2.3. États de bord et correspondance cœur-frontière	272
7.2.4. Phases topologiques et phases de Landau	275
7.3. Isolants topologiques	277
7.3.1. États de surface conducteurs	278
7.3.2. Inversion du caractère des bandes	281
7.3.3. Invariant topologique \mathbb{Z}_2	283
7.3.4. Applications	289
7.4. Semi-métaux de Weyl	290
7.5. Transition de phase topologique	292
7.5.1. Transition de phase intermédiaire dans les matériaux sans symétrie d'inversion	294
7.6. Conséquences de l'interaction électron-phonon sur les phases topologiques	295
7.6.1. Une compétition entre les couplages interbandes et intrabandes	297
Bibliographie	299
Chapitre 8. Dépendance en température de la transition de phase topologique dans le BiTeI	307
8.0.1. Mise en contexte	307
8.0.2. Contributions des auteurs	308
Quatrième article. Temperature dependence of the topological phase transition of BiTeI from first principles	309
8.1. Introduction	311
8.2. Methodology	313
8.2.1. Electron-phonon interaction	313
8.2.2. Thermal expansion	316
8.3. Computational details	318
8.3.1. First-principles calculations	318
8.3.2. Structural properties	318

8.3.3. Lattice dynamics and electron-phonon coupling	319
8.3.4. Thermal expansion.....	320
8.4. Results and discussion	320
8.4.1. Topological phase transition in the static lattice	320
8.4.2. Temperature dependence of the topological phase transition.....	324
8.4.3. Weyl semimetal phase evolution.....	328
8.4.4. Discussion	331
8.5. Conclusion.....	332
Appendix.....	333
Bibliography.....	342
Conclusion	353

Liste des tableaux

4.1	List of the thirty materials considered in the present study, with parameters used in the first-principles computations.....	95
4.2	Experimental and theoretical zero-point renormalization of the band gap, for eighteen materials.....	96
4.3	Electronic gaps and corrections due to the zero-point motion, for thirteen materials.....	97
4.4	Experimental fundamental gap and first-principles zero-point renormalization of the band gap, from the AHC and from the ASC.....	98
4.5	For thirty materials: first-principles renormalization of the band edges using the AHC first-principles and generalized Fröhlich model methodologies.....	99
4.6	First-principles GGA-PBE parameters for the computation of the ZPR of band edges using the generalized Fröhlich model.....	100
5.1	Zero-point lattice expansion for cubic materials.....	158
5.2	Zero-point lattice expansion for wurtzite materials.....	159
5.3	Statistical errors on the ZPLE for the Grüneisen approach.....	162
5.4	Zero-point lattice expansion renormalization of the band gap.....	165
5.5	Statistical errors on the zero-point lattice expansion renormalization of the band gap.....	166
5.6	Estimation of the zero-point lattice expansion renormalization of the band gap..	167
5.7	Statistical errors on the estimation of the ZPLE-induced band gap ZPR.....	168
5.8	Zero-point renormalization of the band gap.....	171
5.9	Effect of exchange-correlation functional on the zero-point lattice expansion and zero-point renormalization of the band gap.....	173
S5.1	Calculations parameters for the twenty-two materials in our set.....	182
S5.2	Zero-point lattice expansion for cubic materials: absolute relative error of the Grüneisen approach with respect to FE minimization.....	184

S5.3	Zero-point lattice expansion for wurtzite materials: absolute relative error of the Grüneisen approach with respect to FE minimization	185
S5.4	Numerical values of $\Delta a(T = 0)/a^0$ and $\Delta c(T = 0)/c^0$ used to obtain estimations of the ZPLE band gap ZPR in Table 5.6 of the main text	186
S5.5	Comparison of the predicted band gap ZPR to experimental values.....	187
6.1	VBM, CBM and band gap ZPR of CdTe for different combinations of SOC strength	224
S6.1	VBM, CBM and band gap ZPR for different combinations of SOC strength	242
S6.2	Calculations parameters for twenty materials.....	245
S6.3	First-principles ZPR.....	246
S6.4	Ratio of eigenenergy renormalization with and without SOC evaluated at $T = 300$ K and $T = 0$ K	247
S6.5	Comparison between the total band gap ZPR and experimental values	248
S6.6	Generalized Fröhlich model ZPR	249
S6.7	Generalized Fröhlich model parameters	250
S6.8	Luttinger-Kohn and Dresselhaus model parameters.....	251
8.1	Static lattice properties at 0 and 5 GPa.	334
8.2	Optimized static lattice structure.....	334
8.3	Static lattice 0 GPa Rashba parameters along A-L.	334
8.4	Zone-center optical phonon frequencies ω (cm^{-1}) in the trivial phase	336
8.5	Zone-center optical phonon frequencies ω (cm^{-1}) in the topological phase	336
8.6	Temperature-dependent WSM phase boundaries.....	342

Liste des figures

1.1	Concept de quasi-particule	3
2.1	Cycle d'autocohérence et potentiel effectif de la DFT	18
2.2	Prédictibilité de l'énergie de la bande interdite des méthodes DFT et <i>GW</i>	19
2.3	Interprétation semi-classique du couplage spin-orbite et précession de Thomas ...	24
2.4	Structure électronique du CdTe avec et sans SOC	29
2.5	Représentation schématique des bandes de conduction en présence de SOC, avec et sans symétrie d'inversion	30
2.6	Séparations Rashba et Dresselhaus dans les matériaux non centrosymétriques ...	32
3.1	Dépendance en température de l'énergie de la bande interdite du silicium	36
3.2	Processus d'émission et d'absorption capturés par la self-énergie de Fan	50
3.3	Modes phononiques et composante non analytique du potentiel	62
4.1	Comparaison des méthodes adiabatiques perturbatives et non perturbatives pour la renormalisation des énergies électroniques	72
4.2	Lowest-order diagrams for the contribution of the electron-phonon interaction to the zero-point renormalization of the gap ZPR_g	78
4.3	Absolute values of first-principles band gap renormalization ZPR_g compared with experimental ones	80
4.4	Ratio between first-principles band gaps and experimental ones	81
4.5	Zero-point renormalization of the band gap: comparison between first-principles AHC values and generalized Fröhlich model	85
4.6	Schema of the long-wave phonon-induced potential and corresponding electronic wavefunction	87
4.7	Correction schemes for the ZPR_c^{AHC} of MgO	94
5.1	Origine anharmonique de la dilatation thermique	132

5.2	Renormalisation de l'énergie de la bande interdite du AIAs	136
5.3	Fan and Debye-Waller contributions entering the lowest order EPI self-energy in the AHC framework.....	154
5.4	Zero-point lattice expansion from FE minimization and Grüneisen formalism	161
5.5	Band gap ZPR induced by zero-point lattice expansion from FE minimisation and Grüneisen formalism	164
5.6	Estimation of the ZPLE-induced band gap renormalization from the empirical expression	168
5.7	Comparison of the predicted first-principles band gap ZPR to experimental values	172
6.1	Fan and Debye-Waller diagrams contributing to the AHC self-energy	206
6.2	Effect of SOC on the ZPR from first principles	218
6.3	Band structure of zincblende CdTe.....	220
6.4	Comparison between first-principles AHC band gap ZPR and experimental values	221
6.5	Total and mode contribution to ZPR_v for different \mathbf{q} -point norms, for polar materials	223
6.6	Contribution to the total ZPR_v of CdTe with varying SOC strength for different \mathbf{q} -point norms	225
6.7	Comparison between the generalized Fröhlich model ZPR_g and the first principles AHC ZPR_g , with and without SOC	227
6.8	Comparison of the SOC-induced reduction of the ZPR_v obtained with the generalized Fröhlich model and with first-principles AHC methodology	228
6.9	Dresselhaus splitting in CdS	237
6.10	Electronic dispersion from the Luttinger-Kohn and Dresselhaus models for CdTe	238
S6.1	Band structure, phonon dispersion and ZPR decomposition of zincblende CdTe, without SOC and with 1% SOC	242
S6.2	Contribution to the total ZPR_v of CdTe for different \mathbf{q} -point norms, with artificially reduced SOC	243
7.1	Équivalence topologique	264
7.2	Représentation schématique du modèle SSH du polyacétylène	265
7.3	Dispersion électronique et représentation bidimensionnelle de l'hamiltonien SSH .	269

7.4	Diagramme de phase topologique du modèle SSH	270
7.5	Déformations de l'hamiltonien	272
7.6	Chaîne entièrement dimérisée et solitons topologiques.	273
7.7	Distinction entre les phases topologiques et les phases de Landau.....	276
7.8	Représentation schématique des états de surface d'un isolant topologique.....	278
7.9	États de surface de l'isolant topologique Bi_2Se_3	279
7.10	Absence de rétrodiffusion dans les états de surface topologique	280
7.11	Protection topologique des états de surface.....	281
7.12	Inversion du caractère des bandes dans le puits quantique CdTe/HgTe	282
7.13	Correspondance cœur-frontière.....	283
7.14	Phase géométrique	285
7.15	Invariant topologique \mathbb{Z}_2 et trajectoires des centres de charge des fonctions de Wannier hybrides	286
7.16	Caractéristiques des semi-métaux de Weyl	290
7.17	Représentation schématique d'une transition de phase topologique	293
7.18	Représentation schématique d'une transition de phase topologique pour un matériau sans symétrie d'inversion	294
7.19	Modèle heuristique de la renormalisation du point zéro dans un isolant topologique.	298
8.1	The Fan and Debye-Waller diagrams describing the lowest-order EPI in the AHC framework.	314
8.2	Crystal structure and first Brillouin zone of BiTeI.	321
8.3	Band inversion process and critical pressures.	321
8.4	Variation of the indirect band gap with pressure.....	322
8.5	EPI contribution to the temperature-dependent renormalization.....	325
8.6	TE contribution to the temperature-dependent renormalization.....	327
8.7	Temperature-dependent band-gap energy.....	329
8.8	Phase diagram of BiTeI.....	329
8.9	Temperature-dependent renormalization of the band-gap edges at 1.5 GPa.....	338
8.10	Pressure and temperature-dependent Rashba parameters.....	340

Liste des sigles et des abréviations

AHC	Méthodologie de Allen, Heine et Cardona
CBM	Minimum de la bande de conduction, de l'anglais <i>conduction band minimum</i>
DFPT	Théorie de la fonctionnelle de la densité perturbative, de l'anglais <i>density-functional perturbation theory</i>
DFT	Théorie de la fonctionnelle de la densité, de l'anglais <i>density-functional theory</i>
EPI	Interaction électron-phonon, de l'anglais <i>electron-phonon interaction</i>
HWCC	Centre(s) de charge des fonctions de Wannier hybrides, de l'anglais <i>hybrid Wannier charge center(s)</i>
LO	(Mode phononique) optique longitudinal, de l'anglais <i>longitudinal optical (phonon mode)</i>
QHA	Approximation quasi harmonique, de l'anglais <i>quasi-harmonic approximation</i>

SOC	Couplage spin-orbite, de l'anglais <i>spin-orbit coupling</i>
TE	Dilatation thermique, de l'anglais <i>thermal expansion</i>
TPT	Transition de phase topologique
TRIM	Vecteur(s) d'onde invariant(s) sous renversement du temps, de l'anglais <i>time-reversal invariant momentum (momenta)</i>
VBM	Maximum de la bande de valence, de l'anglais <i>valence band maximum</i>
WSM	Semi-métal de Weyl (semi-métaux de Weyl), de l'anglais <i>Weyl semimetal (Weyl semimetals)</i>
ZPLE	Dilatation du point zéro du réseau cristallin, de l'anglais <i>zero-point lattice expansion</i>
ZPR	Renormalisation du point zéro, de l'anglais <i>zero-point renormalization</i>

À Jeanne, Léa et Estelle.
Suivez votre route, peu importe si elle change en chemin.

Remerciements

Mon doctorat est une aventure que je n'aurais pas pu traverser seule. J'aimerais tout d'abord remercier mes collègues présents et passés pour l'environnement de travail constructif et dynamique dans lequel j'ai eu l'occasion d'évoluer. À Emmanuel Bourret, Josiane Gaudreau, Daniel Gendron, Oliver Gingras, Vincent Gosselin, Félix-Antoine Goudreault, Julien Groulx, Azin Kazemi, Olivier Malenfant-Thuot, Maxime Morinière, Nicholas Michaud, Joris Paret, Bruno Rousseau, Kevin Ryczko, Hamutal Bary-Soroker, Surya Timilsina, Émile Godbout, Lucie-Maude Fournier, Marianne Lapierre, Alexandre Jean Perrollaz, Maria Sadikov et Tony Vuu, ainsi qu'à tous les autres collègues que j'ai côtoyés, merci pour tous ces échanges enrichissants. Je remercie tout spécialement Paul Boulanger pour cette conversation sur le syndrome d'imposteur lors d'un souper de Noël, qui a influencé profondément la manière dont j'ai abordé le reste de mon parcours.

Je remercie également le Conseil de recherches en sciences naturelles et génie du Canada, le Fonds de recherche du Québec - Nature et Technologies, Hydro-Québec et les Études supérieures et postdoctorales de l'Université de Montréal, dont le soutien financier m'a permis de me consacrer pleinement à mes activités de recherche sans avoir à me soucier de subvenir aux besoins de ma famille.

Merci à Gabriel Antonius d'avoir pris le temps de répondre à mes nombreuses questions au début de mon parcours. Je remercie sincèrement Xavier Gonze pour son soutien précieux à mes différents projets de recherche. Merci d'avoir partagé avec moi ton savoir et ta connaissance impressionnante de la littérature scientifique. Je suis ressortie de chacune de nos discussions riche d'un plus grand recul par rapport à mes travaux.

Je remercie de loin Josée Prévost et ses balados de course à pied, qui m'ont permis de conserver mon équilibre au cours du marathon particulièrement intense que fut cette dernière année, en me rappelant de toujours regarder en avant et qu'une thèse, ça s'écrit une phrase à la fois.

Je manque de mots pour exprimer combien je veux remercier Michel Côté, mon directeur de recherche. Pour ton écoute, ta rigueur scientifique et ton implication sincère dans mon cheminement de chercheuse, je te suis infiniment reconnaissante. Merci d'avoir partagé avec moi ton émerveillement face à la science. Merci également pour ton approche humaine et

ta compréhension lorsque j'ai traversé des périodes plus difficiles. Merci de m'avoir permis d'être complètement moi-même dans nos échanges, et d'avoir cru en moi et en mon travail lorsque je n'y arrivais pas. Je garde un souvenir précieux de mon passage dans ton groupe.

Un merci tout particulier à ma famille proche, plus spécialement à ma mère, pour son soutien indéfectible, de près ou de loin, toujours en accord avec ce dont j'avais besoin. Merci pour toutes les fois où tu as gardé les enfants pour me permettre de travailler efficacement. Ton aide a joué un rôle crucial dans ma réussite. Merci également de m'avoir supportée dans toutes les décisions que j'ai pu prendre dans ma vie, sans jamais me juger. Sache que j'aspire à être capable d'offrir la même chose à mes filles.

Un merci colossal à Marie Lafrance et son œil de lynx, pour toutes ces relectures de demandes de bourses et pour celle de ma thèse. Merci pour ta générosité sans bornes et pour ton écoute attentive et dévouée. Tu es sans contredit une des personnes les plus extraordinaires que j'ai eu l'occasion de rencontrer au cours de mon parcours en physique, et j'espère que nous avons encore de nombreux voyages de vélo devant nous.

À Dominique Rathel-Fournier, merci de partager ma vie. Merci d'avoir été présent autant dans mes bons moments que dans mes périodes de grand stress. Merci pour ta patience au cours de ces derniers mois et pour tous ces petits bonheurs dont tu parsèmes mon quotidien.

Je dédie enfin mon plus gigantesque remerciement à mes trois enfants, Jeanne, Léa et Estelle. Vous avez été partie prenante de cette aventure, au cours de laquelle je vous ai vues grandir et devenir vous-mêmes. Jeanne, merci de m'avoir laissé te raconter des histoires d'électrons pour t'endormir. Léa, merci pour ta douceur et ta patience, pour toutes ces fois où je t'ai entendue de loin rappeler à ta petite sœur de parler moins fort parce que maman travaille. Estelle, merci pour tous les dessins d'encouragement et les diplômes honorifiques que tu m'as offerts spontanément. Je vous remercie d'avoir navigué avec autant de résilience à travers toutes ces périodes où j'étais présente de corps, mais très peu d'esprit, accaparée par la préparation des conférences ou des examens. Votre joie de vivre, vos sourires magnifiques et votre curiosité sans bornes m'ont donné la force de continuer dans ces moments où j'aurais ô combien préféré aller m'amuser avec vous. J'espère sincèrement avoir pu être un exemple de persévérance, qui saura vous inspirer lorsque, dans trop peu longtemps, vous commencerez à tracer votre propre route. Je vous aime plus grand que la Terre, plus haut que la Lune et les étoiles, plus loin que la galaxie.

Avant-propos

Cette thèse est l'aboutissement de six années de travail aux études supérieures. La manière dont j'en ai abordé la rédaction a été fortement influencée par la pédagogue en moi, qui a accompagné quatre cohortes de bacheliers dans leurs premiers apprentissages de la mécanique quantique. Tout le long de mon cheminement tant d'étudiante que d'enseignante, j'ai eu à cœur d'être à la fois en mesure de bien comprendre les développements théoriques ayant mené aux outils que j'utilise dans mes travaux, que de chercher à voir au-delà des équations, de saisir ce qu'elles nous révèlent sur les mécanismes physiques qu'elles décrivent. Face aux notions plus difficiles ou abstraites, j'ai façonné ma compréhension des concepts physiques et des outils mathématiques en me construisant des analogies bâties sur des exemples simples qui, sans toujours être rigoureusement exactes, m'ont permis de développer une forme d'intuition des grandes idées en jeu, que j'ai ensuite laissée mûrir avant de m'attaquer aux détails techniques. Ainsi, dans les chapitres introductifs de cette thèse, j'ai souhaité prendre le temps de bien situer et expliquer les formalismes théoriques que j'ai utilisés dans le cadre de mes travaux, dans l'optique qu'une future étudiante ou un futur étudiant qui poursuivrait le projet ou qui souhaiterait simplement en apprendre plus sur le sujet puisse bénéficier d'une discussion à la fois globale et approfondie, une sorte de préambule à la lecture des articles de référence, qui lui permette d'en apprécier toute la richesse.

Dans un autre angle d'idées, au cours de mon doctorat, j'ai eu l'occasion de contribuer à 6 articles et un compte rendu de conférence (*Proceedings*) :

- (1) V. Brousseau-Couture, G. Antonius and M. Côté. Temperature dependence of the topological phase transition of BiTeI from first principles. *Phys. Rev. Research* **2**(2) : 023185, 2020. doi: [10.1103/PhysRevResearch.2.023185](https://doi.org/10.1103/PhysRevResearch.2.023185),
- (2) A. Miglio, V. Brousseau-Couture, E. Godbout, G. Antonius, Y.-H. Chan, S. G. Louie, M. Côté, M. Giantomassi and X. Gonze. Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap. *NPJ Comput. Mater.* **6**(1) : 167, 2020. doi: [10.1038/s41524-020-00434-z](https://doi.org/10.1038/s41524-020-00434-z),
- (3) V. Brousseau-Couture and M. Côté. Influence of the Electron-Phonon Interaction on the Topological Phase Transition in BiTeI. In M.B. Paranjape, R. Mackenzie, Z. Thomova, P. Winternitz and W. Witczak-Krempa, editors : *Quantum Theory and*

Symmetries, CRM Series in Mathematical Physics, p. 305-313, Cham, 2021. Springer International Publishing. doi: [10.1007/978-3-030-55777-5_29](https://doi.org/10.1007/978-3-030-55777-5_29),

- (4) B. Guster, P. Melo, B. A. A. Martin, V. Brousseau-Couture, J. C. de Abreu, A. Miglio, M. Giantomassi, M. Côté, J. M. Frost, M. J. Verstraete and X. Gonze. Fröhlich polaron effective mass and localization length in cubic materials : degenerate and anisotropic electronic bands. *Phys. Rev. B* **104**(23) : 235123, 2021. doi: [10.1103/PhysRevB.104.235123](https://doi.org/10.1103/PhysRevB.104.235123),
- (5) M. M. Piva, J. C. Souza, V. Brousseau-Couture, K. R. Pakuszewski, J. K. John, C. Adriano, M. Côté, P. G. Pagliuso and M. Nicklas : Tuning the nontrivial topological properties of the Weyl semimetal CeAlSi. Soumis à *Phys. Rev. Lett.* en octobre 2021. doi: [10.48550/arXiv.2111.05742](https://doi.org/10.48550/arXiv.2111.05742),
- (6) V. Brousseau-Couture, E. Godbout, X. Gonze and M. Côté. Zero-point lattice expansion and band gap renormalization : Grüneisen approach versus free energy minimization. *Phys. Rev. B* **106**(8) : 085137, 2022. doi: [10.1103/PhysRevB.106.085137](https://doi.org/10.1103/PhysRevB.106.085137)
- (7) V. Brousseau-Couture, X. Gonze and M. Côté. Effect of spin-orbit coupling on the zero-point renormalization of the electronic band gap in cubic materials : first-principles and generalized Fröhlich model. Soumis à *Phys. Rev. B* en juillet 2022.

Cette thèse regroupe les articles (1), (2), (6) et (7) de cette liste. Par souci de fluidité, les articles ne sont pas présentés en ordre chronologique. Je présente tout d'abord l'article traitant des effets non adiabatiques dans la renormalisation du point zéro (2), puis celui analysant la contribution de la dilatation du point zéro du réseau cristallin (6). L'article suivant reprend certaines questions explorées dans l'article (2) en mettant l'accent sur l'effet du couplage spin-orbite (7). Le premier article que j'ai publié en tant que première auteure (1), qui se veut une application concrète des méthodes étudiées dans les trois premiers articles au problème des transitions de phase topologiques, est le dernier à être présenté. Pour cette raison, certaines notions abordées dans les chapitres 4 à 6 ne sont pas utilisées dans les analyses décrites au chapitre 8. Les différents articles sont entrecoupés de chapitres développant les principaux éléments théoriques utilisés dans le cadre de mes travaux. Comme souligné auparavant, ces derniers ont été rédigés dans une optique pédagogique et de complétude, en ne supposant que des connaissances de base en physique du solide. Plus particulièrement, le chapitre 7 introduit les notions reliées aux phases topologiques des matériaux de la manière dont, rétrospectivement, j'aurais aimé qu'on me les présente au début de mon parcours. J'ai enfin écrit l'introduction en ayant en tête la curiosité naturelle de Jeanne, mon aînée, en

espérant que toute personne intéressée à en savoir plus sur le sujet de mes travaux soit en mesure d'en comprendre les grandes lignes.

Mon parcours au doctorat ne se limite pas au contenu de cet ouvrage. J'aimerais donc souligner quelques expériences connexes liées au partage de mes connaissances qui ont contribué à bâtir la chercheuse que je suis aujourd'hui. Je pense notamment aux présentations orales données au March Meeting de l'American Physical Society (New Orleans 2017, Los Angeles 2018, Boston 2019, en ligne 2021), au ABINIT Developers Workshop (en ligne, 2021) et au Symposium international Théorie quantique et symétries (Montréal, 2019), ainsi qu'aux présentations par affiche à la Grande conférence du Regroupement québécois sur les matériaux de pointe (Montréal, 2019, prix de la meilleure affiche), au ABINIT Developers Workshop (Louvain-la-Neuve, Belgique, 2019) et à Women in Physics Canada (Montréal, 2019). En parallèle, mes sept années d'expérience comme auxiliaire d'enseignement, récompensées par 3 prix Petit Nobel votés par les étudiantes et étudiants du baccalauréat et par le Prix d'excellence aux auxiliaires d'enseignement 2021 de l'Université de Montréal, ont profondément influencé ma façon de percevoir les liens entre les différents concepts et permis de développer tout ce bagage d'images et d'analogies qui ont renforcé ma propre compréhension.

Sur une note plus personnelle, je constate que ma thèse est finalement à l'image de la personne que je suis. Les gens qui me connaissent bien savent que je ne raconte jamais une histoire en ligne droite. Ma narration déambule le long de sentiers parallèles, plus ou moins sinueux, que j'emprunte au gré de l'évolution du récit, bifurquant parfois (souvent !) de ma cible initiale. J'ai plongé dans cette aventure en voulant démystifier les phases topologiques de la matière. J'ai voyagé au gré des questions qui se sont pointées sur mon chemin. Me voici maintenant, six ans plus tard, arrivée à une destination que je n'aurais pas imaginée au départ, mais qui, j'en suis convaincue, est d'autant plus riche qu'elle est celle où ma curiosité m'a menée.

Chapitre 1

Introduction

L'effet de la température sur les propriétés des dispositifs technologiques et des matériaux est omniprésent dans notre quotidien ; l'efficacité des batteries lithium-ion qui alimentent les voitures électriques et les téléphones cellulaires diminue lors des grands froids d'hiver, les grandes infrastructures routières sont conçues avec des joints de dilatation qui permettent d'absorber l'impact du changement de volume de la chaussée d'une saison à l'autre, et la performance de nos ordinateurs personnels diminue lorsque ceux-ci surchauffent. Dans tous ces cas, l'origine physique des propriétés macroscopiques observées réside dans des mécanismes qui se produisent à l'échelle atomique : la température influence la vitesse des réactions chimiques [1] dans les électrolytes, modifie la distance moyenne entre les noyaux [2], diminue la mobilité des électrons [3] et augmente le courant de fuite dans les transistors [4].

Certaines propriétés macroscopiques affectées par la température, quant à elles, dépendent explicitement des différents niveaux d'énergie accessibles aux électrons présents dans le matériau, c'est-à-dire de la structure électronique. Une quantité fondamentale liée à la structure électronique des matériaux semi-conducteurs et isolants est l'énergie de la bande interdite. Cette dernière correspond à la quantité minimale d'énergie nécessaire pour promouvoir un électron vers un état inoccupé qui lui permettrait de se propager dans le matériau et de conduire l'électricité. L'énergie de la bande interdite influence donc l'énergie de seuil pour l'absorption d'un photon ainsi que l'énergie des photons émis par une désexcitation d'un électron vers son état fondamental¹. En pratique, on observe par exemple que l'efficacité des panneaux solaires pour convertir le rayonnement solaire en courant électrique diminue progressivement lorsque la température de l'air dépasse $\sim 25^\circ\text{C}$ [5]. De même, la couleur émise par une diode électroluminescente monochromatique ou une diode laser va décaler vers le bleu si on la refroidit rapidement avec de l'azote liquide².

1. Quelques précisions seront apportées sur ce point au début du chap. 3.

2. Même lorsqu'on en comprend les rouages, il est toujours fascinant d'observer cette réaction en temps réel! <https://www.youtube.com/watch?v=B2XHQ3tqnZM>, https://www.youtube.com/watch?v=5PquJdIK_z8

La dépendance en température de la structure électronique a donc des conséquences directes sur la probabilité des différentes transitions optiques, qui vont altérer les spectres d'absorption optique et de photoluminescence. Elle affecte de plus les propriétés de transport telles que la mobilité électronique, à travers la variation de la vitesse de groupe des électrons et de leur taux de collision avec les ions du cristal. Une modélisation adéquate de la dépendance en température de la structure électronique est donc cruciale pour optimiser l'efficacité des dispositifs technologiques destinés à des applications à grande échelle, qui opèrent inévitablement dans des conditions où la température fluctue. Pour y parvenir, on doit d'abord déterminer comment se manifeste la température du point de vue atomique.

1.1. Principe d'émergence et quasi-particules

Avant même de considérer l'effet de la température, on peut se demander ce qui détermine fondamentalement les propriétés intrinsèques d'un matériau. A priori, n'importe quel matériau se résume à un assemblage organisé d'atomes, c'est-à-dire une combinaison structurée de protons, de neutrons et d'électrons. De ce point de vue, tous les assemblages d'atomes de carbone devraient posséder les mêmes propriétés. Or, selon la structure dans laquelle les noyaux se cristallisent, les atomes de carbone peuvent former à la fois du graphite, un matériau friable qui peut conduire l'électricité, et du diamant, un isolant électrique extrêmement dur. Les différentes liaisons chimiques propres à ces deux matériaux correspondent à différentes répartitions de la charge électronique dans le cristal, chacune minimisant l'énergie totale du système pour sa configuration cristalline propre. Par ailleurs, puisque les électrons sont des particules chargées, leur distribution spatiale dans le cristal est également influencée par la forte répulsion coulombienne entre chacun des électrons. En d'autres mots, les propriétés observées dépendent simultanément de la réponse des électrons au potentiel coulombien créé par l'ensemble des noyaux et de l'interaction répulsive qui tend à éloigner les électrons les uns des autres. Nous faisons donc face à un problème à N -corps.

L'information clé pour déterminer les propriétés fondamentales d'un matériau réside par conséquent dans l'interaction entre les différentes particules qui le constituent. Ces interactions donnent naissance à des propriétés inexplicables du point de vue des constituants eux-mêmes ; on parle alors de phénomènes émergents. Outre la distinction métal-isolant, l'un des exemples les plus frappants est sans contredit la supraconductivité. Dans ce cas, l'interaction entre les électrons et les vibrations atomiques crée une attraction effective entre des paires d'électrons appelées paires de Cooper, ce qui va à l'encontre de la répulsion coulombienne que l'on attribuerait à une paire d'électrons dans le vide. On peut également penser aux matériaux topologiques, qui sont simultanément isolants et conducteurs à différents niveaux de dimensionnalité, ainsi qu'aux isolants de Mott, où de fortes corrélations électroniques figent

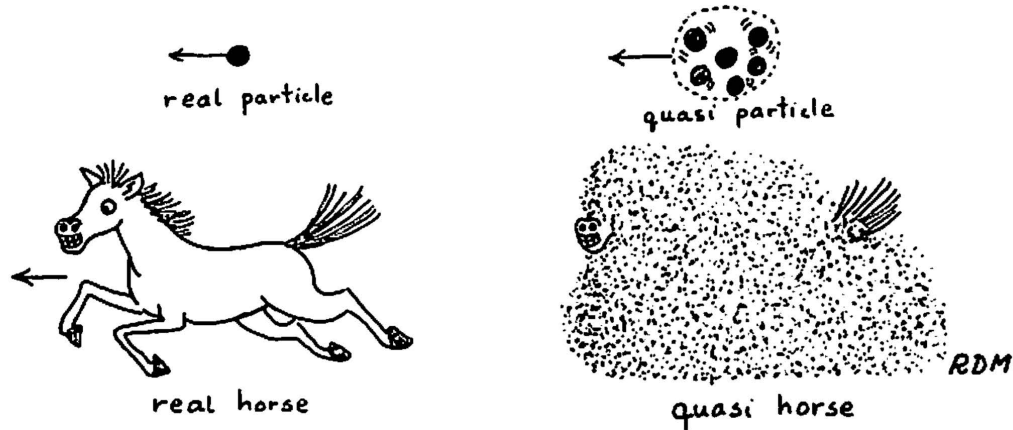


Figure 1.1 – Concept de quasi-particule. Le nuage de poussière, qui représente les interactions, masque en partie la silhouette du quasi-cheval. Du point de vue d’un observateur externe, ce dernier continue à se comporter comme un cheval, bien que certaines de ses propriétés semblent différentes de celles du cheval non-interagissant. Image tirée de la référence [6].

les électrons sur les différents sites du réseau, empêchant ainsi la conduction électrique dans un matériau qui devrait autrement être métallique.

Le traitement mathématique explicite d’un système à plusieurs corps est éminemment complexe. On peut néanmoins se demander si les propriétés émergentes d’un matériau peuvent être interprétées en des termes simples. En d’autres mots, on veut savoir si les excitations d’un système à N -corps en interaction s’apparentent encore à des particules individuelles qui n’interagiraient que faiblement (ou pas) entre elles. Dès lors, on va plutôt chercher à décrire le système en termes de particules fictives que l’on nomme *quasi-particules*. Si elles ne sont pas à proprement parler des particules réelles, les quasi-particules se comportent qualitativement comme des particules fondamentales, mais certaines de leurs propriétés sont modifiées par rapport à leur analogue sans interaction. L’analogie du quasi-cheval proposée par Richard Mattuck [6], qui est montrée à la figure 1.1, illustre bien ce concept. Tout comme le nuage de poussière masque une bonne partie du cheval tout en conservant une partie de sa silhouette, les interactions entre les différentes particules forment un nuage qui camoufle l’électron d’origine. Le quasi-électron semble alors posséder des propriétés physiques différentes du point de vue d’un observateur externe.

Si les travaux de Drude et Sommerfeld [7] au début des années 1900 suggéraient déjà que l’on pouvait négliger les interactions entre les électrons pour décrire les propriétés des métaux, c’est à la théorie du liquide de Fermi développée par Landau [8] que l’on doit la formalisation du concept de quasi-particule. Ce dernier a montré que les excitations décrivant un système d’électrons en interaction se comportent essentiellement comme des électrons indépendants ayant un temps de vie fini et une masse différente de celle des électrons dans le

vide, que l'on nomme masse effective. Ainsi, dans les matériaux semi-conducteurs, on parle de bandes de type quasi-électron ou de type quasi-trou, selon que l'excitation correspond à l'ajout d'une particule dans une bande initialement vide ou à son retrait dans une bande initialement occupée. Dans un même matériau, ces différentes bandes pourront être associées à des masses effectives différentes. De même, certains métaux dont les bandes partiellement occupées sont constituées d'orbitales localisées de type d ou f sont connus sous le nom de *fermions lourds*, en raison de la courbure très faible de leur dispersion électronique dans l'espace réciproque, qui est associée à une très grande « masse effective »³.

1.2. La température du point de vue atomique : une population de déplacements collectifs

La température est un concept à la fois très intuitif d'un point de vue empirique, mais étonnamment subtil à définir sur le plan formel. En effet, dans notre quotidien, la notion de température est intrinsèquement reliée aux concepts de chaud et de froid. Ainsi, si l'on juxtapose deux systèmes de températures différentes, on pourra distinguer le système « chaud » du système « froid » en identifiant la direction de propagation de la chaleur⁴. Bien que concrète, cette définition ne nous donne guère d'information sur ce que représente intrinsèquement la température. Puisque la chaleur est une forme d'énergie, une image instructive serait d'associer la température à une « densité d'énergie » caractéristique du système. On peut penser par exemple à un système simple formé d'une masse attachée à un ressort, dont l'énergie totale augmente avec l'amplitude du déplacement de la masse. De manière analogue, l'augmentation de la température se manifeste au niveau atomique par une augmentation de l'amplitude des oscillations des noyaux autour de leurs positions d'équilibre. Le potentiel cristallin ressenti par les électrons est alors constamment en mouvement et devient du même coup plus compliqué à décrire. Ainsi, si nous étions en mesure d'observer en temps réel les déplacements des différents noyaux à la température de la pièce, la résultante nous semblerait assurément chaotique et complètement désordonnée, ce qui nous indique que nous n'abordons pas le problème sous le bon angle. Dans l'esprit de la section précédente, on voudra décomposer ces vibrations en termes de quasi-particules nommées *phonons*, qui représentent une déformation collective structurée de l'ensemble des noyaux du réseau cristallin. Nous sommes donc a priori bien loin de l'idée qu'on se fait habituellement d'une « particule » ! Néanmoins, dans cette représentation abstraite, la réponse dynamique des électrons à un changement continu du potentiel cristallin est remplacée par l'image beaucoup plus intuitive de deux quasi-particules qui interagissent entre elles. L'augmentation des vibrations

3. Le concept de masse effective n'a toutefois de réel sens physique que dans un contexte où l'approximation parabolique des bandes électroniques est valide.

4. Par exemple, la chaleur voyage du rond de poêle allumé vers la marmite et irradie du feu de camp vers la guimauve que l'on souhaite griller.

des noyaux associée à une hausse de la température se traduit alors par une population de phonons plus élevée, qui multiplie les probabilités d'interaction avec les électrons.

En somme, le problème de la dépendance en température de la structure électronique se résume conceptuellement à introduire de nouvelles interactions dans notre système. Dans l'esprit du quasi-cheval de la figure 1.1, les quasi-particules qui décrivent les excitations électroniques en présence des phonons sont différentes de celles obtenues en considérant uniquement l'interaction entre les électrons. Tout comme la population de phonons, les propriétés du matériau vont donc maintenant dépendre de la température. Par ailleurs, la nature quantique des noyaux implique qu'ils ne sont jamais entièrement au repos, y compris à $T = 0$ K, lorsque la population moyenne des phonons est nulle. Ce *mouvement du point zéro* est un effet purement quantique. On peut l'interpréter comme un analogue des fluctuations électromagnétiques du vide, qui sont responsables du décalage de Lamb entre les niveaux $2s$ et $2p$ de l'atome d'hydrogène, ainsi que de l'effet Casimir, où on mesure une force électromagnétique non nulle entre deux plaques conductrices non chargées placées dans un milieu sous vide. Ainsi, les oscillations des noyaux interagissent avec les électrons et affectent leurs propriétés même à température nulle. On parle alors de renormalisation du point zéro (ZPR).

En plus d'interagir avec les électrons, la présence des phonons dans un matériau affecte son volume, qui augmente avec la température. On parle alors de dilatation thermique. Un analogue de la dilatation thermique existe également à température nulle, sous le nom de dilatation du point zéro du réseau, et contribue à modifier les niveaux d'énergie des électrons. Décrire correctement la structure électronique à une température donnée nécessite donc de considérer conjointement les contributions de ces deux mécanismes.

1.3. Une multitude d'approximations

Au prochain chapitre, nous verrons qu'il est impossible de modéliser la structure électronique des matériaux sans recourir à de nombreuses approximations qui sont essentielles pour rendre le calcul numériquement résoluble. Parmi celles-ci, l'approximation Born-Oppenheimer, qui consiste à considérer les noyaux du cristal comme des particules fixes, constitue sans contredit la pierre angulaire des méthodes numériques de structure électroniques couramment utilisées aujourd'hui. Cette approximation peut être évaluée à différents niveaux de théorie, chacun raffinant le traitement des corrélations électroniques. On pourra dès lors choisir de les modéliser à l'aide d'un potentiel effectif, comme il est fait dans le cadre de la théorie de la fonctionnelle de la densité utilisée dans cette thèse, avec des méthodes basées sur la théorie à N -corps ou même considérer la contribution des effets excitoniques. On utilisera ensuite cette solution statique comme point de départ pour obtenir la dépendance en température et la renormalisation du point zéro de la structure électronique au chapitre 3.

Naturellement, même à niveau de théorie donné, il existe toute une panoplie d'approximations moins fondamentales que l'on peut, en tant que théoriciennes et théoriciens, choisir d'appliquer ou non. Dans l'esprit du rasoir d'Occam, on souhaite bénéficier d'un modèle aussi simple que possible, mais qui contienne néanmoins toute la physique nécessaire pour produire des prédictions précises et nous permettre de développer une intuition sur les mécanismes d'interaction dominants dans les différents matériaux. On pourrait ainsi, par exemple, choisir de négliger l'interaction spin-orbite si on s'intéresse aux propriétés du diamant ou du silicium, puisque ces matériaux se trouvent dans le haut du tableau périodique, où il est connu que l'interaction spin-orbite est faible. Au contraire, un tel choix serait difficilement justifiable pour un matériau de type Rashba ou dans un isolant topologique, où cette interaction est la force motrice de leurs propriétés physiques particulières. Dans tous les cas, ces approximations auront des répercussions sur l'interaction électron-phonon, qui pourra être affectée soit dans sa formulation explicite ou de façon indirecte, à travers une modification des états électroniques statiques.

À ce propos, les approches théoriques et numériques permettent de contrôler quand et comment sont introduites les différentes approximations. On peut ainsi isoler leur effet sur la structure électronique en comparant deux résultats de calcul ne différant que par la présence de l'interaction qui nous intéresse. Il s'agit là d'un atout indéniable qui peut permettre de désenchevêtrer l'effet d'interactions ayant des signatures expérimentales similaires⁵. En étudiant un nombre suffisamment grand de matériaux, on peut également parvenir à extraire des tendances générales qui mettent en lumière la façon dont une approximation affecte les propriétés des quasi-particules. De plus, en comparant la prédictibilité des calculs avec les données expérimentales, nous sommes en mesure de développer une intuition sur les contextes physiques qui permettent ou non de justifier une approximation, ainsi que de mieux estimer l'impact de ces approximations sur les prédictions numériques.

Le cœur de cette thèse consiste donc à analyser comment différents choix d'approximation influencent les prédictions numériques de la renormalisation du point zéro (ZPR) de l'énergie de la bande interdite dans les semi-conducteurs. Parmi les questions qui seront explorées, on cherchera notamment à déterminer si l'utilisation de l'approximation adiabatique, qui sera introduite au chapitre 2, est adéquate pour décrire l'interaction électron-phonon dans les semi-conducteurs. On s'intéressera de plus à évaluer l'importance relative de la contribution de la dilatation du point zéro du réseau cristallin au ZPR par rapport à celle de l'interaction électron-phonon, qui est perçue par la communauté comme le mécanisme de renormalisation dominant à basse température. On examinera ensuite si l'inclusion du couplage spin-orbite est nécessaire pour modéliser correctement l'interaction électron-phonon. Enfin, on se demandera

5. Précisons néanmoins que ce ne sont pas toutes les approximations qui peuvent être contrôlées de cette manière d'un point de vue théorique, puisque certaines d'entre elles se situent aux limites des méthodes de pointe actuelles.

si la température, à travers l'effet combiné de l'interaction électron-phonon et de la dilatation thermique, est capable d'influencer la phase topologique des matériaux.

Pour la suite, les chapitres 2 et 3 introduisent respectivement les différentes notions théoriques et les méthodologies utilisées dans le cadre de cette thèse, tout d'abord dans l'approximation du cristal statique (chapitre 2), puis pour le traitement de l'interaction électron-phonon (chapitre 3). Le chapitre 4 se concentre sur l'importance de considérer le temps de réponse des électrons face au déplacement des noyaux, également appelés effets non adiabatiques, pour prédire correctement le ZPR des matériaux polaires. On y développe une généralisation du modèle de Fröhlich qui permet de traiter des matériaux dont les bandes électroniques sont anisotropes et dégénérées. Le chapitre 5 se veut ensuite une étude approfondie de la renormalisation du point zéro du réseau et de l'importance relative de sa contribution à la variation de l'énergie de la bande interdite. Le chapitre 6 reprend quant à lui certaines des idées développées au chapitre 4, mais focalise sur la manière dont l'introduction du couplage spin-orbite affecte le ZPR dans les matériaux cubiques. Le chapitre 8 présente finalement une application des formalismes étudiés dans les chapitres précédents au problème spécifique des transitions de phase topologiques. Par souci pédagogique, une introduction à la topologie des structures de bande (chapitre 7) précède cet article.

Bibliographie

- [1] S. Ma, M. Jiang, P. Tao, C. Song, J. Wu, J. Wang, T. Deng, and W. Shang. Temperature effect and thermal impact in lithium-ion batteries : A review. *Progress in Natural Science : Materials International*, 28(6) :653–666, 2018. doi: [10.1016/j.pnsc.2018.11.002](https://doi.org/10.1016/j.pnsc.2018.11.002).
- [2] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College, Philadelphia, 1976.
- [3] P. Y. Yu and M. Cardona. *Fundamentals of Semiconductors*. Springer-Verlag, Berlin Heidelberg, 2nd edition, 1999.
- [4] I. Kwon, H. I. Kwon, and I. H. Cho. Development of high temperature operation silicon based MOSFET for harsh environment application. *Results in Physics*, 11 :475–481, 2018. doi: [10.1016/j.rinp.2018.09.035](https://doi.org/10.1016/j.rinp.2018.09.035).
- [5] O. Dupré, R. Vaillon, and M. A. Green. Physics of the temperature coefficients of solar cells. *Solar Energy Materials and Solar Cells*, 140 :92–100, 2015. doi: [10.1016/j.solmat.2015.03.025](https://doi.org/10.1016/j.solmat.2015.03.025).
- [6] R. D. Mattuck. *A Guide to Feynman Diagrams in the Many-Body Problem*. Dover Books on Physics and Chemistry. Dover, New York, 2nd edition, 1992.
- [7] M. P. Marder. *Condensed Matter Physics*. John Wiley & Sons, Hoboken, 2nd edition, 2010.

- [8] L. Landau. The Theory of a Fermi Liquid. *Journal of Experimental and Theoretical Physics*, 3(6) :920, 1957, http://jetp.ras.ru/cgi-bin/dn/e_003_06_0920.pdf.

Chapitre 2

Structure électronique dans l'approximation Born-Oppenheimer

La modélisation *ab initio* des propriétés d'un cristal se résume conceptuellement à résoudre un système de noyaux et d'électrons en interaction en se basant sur les équations fondamentales de la mécanique quantique. Ce problème quantique vit néanmoins à une échelle spatiale qui est radicalement petite en comparaison avec les dimensions typiques de notre quotidien : on peut par exemple montrer que le nombre de molécules présentes dans un seul verre d'eau est environ 1000 fois supérieur au nombre de verres d'eau contenus dans l'ensemble des océans terrestres¹. Dans cette optique, considérant qu'il y a environ 8×10^{19} atomes dans un seul grain de sable, on peut aisément déduire que la dimensionnalité du problème explose bien avant que l'on atteigne des échelles de l'ordre du centimètre ou du millimètre. Le recours aux approximations devient dès lors inévitable pour parvenir à modéliser les propriétés macroscopiques des matériaux de manière efficace. Ce chapitre se concentre sur la description de la structure électronique que l'on obtient en appliquant une approximation fondamentale de la physique de la matière condensée, à savoir l'approximation Born-Oppenheimer. Nous introduisons de plus la théorie de la fonctionnelle de la densité (DFT), que nous avons utilisée dans le cadre de nos travaux. Nous démontrons ensuite brièvement certaines propriétés des états électroniques cristallins sous différentes opérations de symétrie, qui seront utiles pour comprendre les conséquences du couplage spin-orbite sur la structure électronique ainsi que pour interpréter les propriétés topologiques de certains matériaux. Nous présentons enfin l'origine relativiste du couplage spin-orbite et discutons de la manière dont cette interaction affecte les états électroniques, en mettant un accent particulier sur les matériaux non centrosymétriques.

1. Selon <https://hypertextbook.com/facts/2001/SyedQadri.shtml>, les océans terrestres contiennent environ $1.3 \times 10^9 \text{ km}^3 \simeq 1.3 \times 10^{21} \text{ L}$ d'eau, ce qui représente grossièrement 5×10^{21} verres d'eau de 250 ml. En comparaison, un verre d'eau de 250 ml contient $250\text{g}/(18\text{g/mol}) \sim 14 \text{ moles} \sim 10^{24}$ molécules d'eau.

2.1. Un problème à la fois simple et insoluble

Le problème quantique que l'on veut résoudre pour prédire les propriétés des solides se pose d'une manière remarquablement simple. En effet, l'hamiltonien d'un système formé de N_e électrons et N_{at} noyaux s'écrit [1]

$$\begin{aligned}\hat{H} &= -\sum_i^{N_e} \frac{1}{2} \nabla_{\mathbf{r}_i}^2 - \sum_i^{N_{at}} \frac{1}{2M_i} \nabla_{\mathbf{R}_i}^2 + \sum_{\substack{i \neq j \\ i < j}}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{\substack{i \neq j \\ i < j}}^{N_{at}} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} - \sum_i^{N_{at}} \sum_j^{N_e} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_j|} \\ &= \hat{T}^e(\{\mathbf{r}_i\}) + \hat{T}^n(\{\mathbf{R}_i\}) + \hat{V}^{e-e}(\{\mathbf{r}_i\}) + \hat{V}^{n-n}(\{\mathbf{R}_i\}) + \hat{V}^{e-n}(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}),\end{aligned}\quad (2.1.1)$$

où \mathbf{r}_i représente les coordonnées du $i^{\text{ième}}$ électron et \mathbf{R}_j , celles du $j^{\text{ième}}$ noyau, de masse M_j . Les symboles $\{\mathbf{r}_i\} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e})$ et $\{\mathbf{R}_j\} = (\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{N_{at}})$ dénotent l'ensemble des coordonnées des électrons et des noyaux. Les deux premiers termes de cet hamiltonien sont les opérateurs d'énergie cinétique des électrons et des ions, alors que les trois derniers décrivent respectivement le potentiel coulombien entre les électrons, entre les noyaux et entre les électrons et les noyaux. Sauf indication contraire, nous travaillons dans le système d'unités atomiques, tel que $\hbar = c = m_e = |e| = 1$, où \hbar est la constante de Planck réduite, c la vitesse de la lumière, et où m_e et $|e|$ sont respectivement la masse et la valeur absolue de la charge de l'électron. On néglige pour l'instant le spin de l'électron de même que les effets relativistes, tels que le couplage spin-orbite. Nous travaillons également dans la représentation de Schrödinger.

L'ensemble des propriétés physiques du système d'électrons et de noyaux est alors encodé dans la fonction d'onde $\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}, t)$, dont l'évolution est décrite par l'équation de Schrödinger :

$$\hat{H}\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}, t) = i \frac{\partial \Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}, t)}{\partial t}. \quad (2.1.2)$$

Les propriétés physiques du système s'obtiennent ensuite en évaluant la valeur attendue² des différents observables \mathcal{O} pour la fonction d'onde $|\Phi(t)\rangle$,

$$\langle \mathcal{O}(t) \rangle = \langle \Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}, t) | \mathcal{O} | \Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}, t) \rangle. \quad (2.1.3)$$

On pourra également exprimer $|\Phi(t)\rangle$ comme une combinaison linéaire des états propres de l'hamiltonien, $\{|\Psi_l(t)\rangle\}$, qui vérifient

$$\begin{aligned}\hat{H}\psi_l(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}) &= E_l \psi_l(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}), \\ \Psi_l(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}, t) &= e^{-iE_l t} \psi_l(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}),\end{aligned}\quad (2.1.4)$$

où E_l est l'énergie totale du système d'électrons et de noyaux dans l'état propre d'indice l .

2. Il s'agit ici d'une traduction libre d'*expectation value*.

L'essence de la physique d'un matériau est donc intégralement contenue dans les quatre équations précédentes³. Ce problème, d'une effarante simplicité en apparence, demeure néanmoins insoluble même pour des petites molécules comprenant à peine quelques atomes. La source de sa complexité réside dans la corrélation entre l'ensemble des particules du système. En effet, $3N_e \times 3N_{at}$ coordonnées spatiales sont nécessaires pour encoder entièrement la fonction d'onde. Pour un système formé de seulement 10 particules et en supposant que l'on se limite à une faible résolution spatiale de 10 points pour chaque coordonnée cartésienne, la quantité de mémoire nécessaire serait de l'ordre de $10^{30} \times 16$ octets $\sim 10^{22}$ Go, soit presque une mole de gigaoctets ! Ceci nous indique que la solution exacte de l'équation de Schrödinger est totalement inaccessible du point de vue des ressources numériques si l'on considère n'importe quel matériau macroscopique [4]. Il devient donc impératif de formuler des approximations qui permettront d'accéder aux propriétés physiques qui nous intéressent tout en maintenant un coût numérique raisonnable.

Pour comprendre de manière plus intuitive ce qu'implique une fonction d'onde entièrement corrélée, imaginons que le système d'électrons et de noyaux se trouve dans son état fondamental. La corrélation implique qu'à tout moment, la répartition de la charge électronique dans le cristal dépend à la fois de l'interaction coulombienne avec l'ensemble des noyaux et de celle entre les différents électrons. Ainsi, si les noyaux se déplacent, tous les électrons seront affectés par ce changement de « potentiel »⁴. Or, toute variation de la distribution électronique va non seulement modifier le potentiel coulombien ressenti instantanément par les noyaux, qui vont en réponse ajuster leurs déplacements respectifs, mais aussi altérer l'interaction coulombienne entre les électrons eux-mêmes. Cette réponse des électrons et des noyaux engendre dès lors une nouvelle modification de la distribution électronique, ce qui nous ramène à notre point de départ. Nous pourrions répéter cet argument du point de vue des noyaux, qui répondent simultanément aux déplacements des électrons et à celui des autres noyaux. C'est cette boucle de rétroaction dynamique entre des systèmes d'électrons interagissants et de noyaux interagissants qui interagissent entre eux qui rend le problème si difficile à modéliser.

2.2. Approximation des noyaux statiques

La première approximation à appliquer à l'équation (2.1.1), appelée approximation Born-Oppenheimer [1, 5], consiste à décorrélérer la dynamique des électrons de celle des noyaux.

3. On réfère ici aux interactions fondamentales entre les électrons et les noyaux. D'autres termes pourront être ajoutés à cet hamiltonien, par exemple les contributions décrivant la structure fine et la structure hyperfine de l'atome d'hydrogène [2]. De même, cet hamiltonien ne permet pas la création et la destruction de particules, pour lesquelles il faut recourir à la théorie des champs quantiques [3].

4. L'utilisation du mot « potentiel » constitue ici un abus de notation puisque nous n'avons pas encore appliqué l'approximation Born-Oppenheimer...

Considérant que la masse d'un noyau est environ 1000 fois plus grande que celle d'un électron, le temps de réponse des électrons dans la boucle de rétroaction est nettement plus rapide que celui des noyaux. Suivant le théorème adiabatique [6], on peut alors argumenter que la densité électronique du cristal va systématiquement s'ajuster à la configuration des noyaux de manière à minimiser son énergie totale. Ce faisant, on suppose que la réponse des électrons à un petit changement de la configuration atomique est *instantanée*; on pourra donc se limiter à évaluer la configuration atomique qui minimise les forces sur les différents noyaux, en supposant que ces derniers sont immobiles du point de vue des électrons. Ainsi, on traitera les noyaux comme des particules ponctuelles dont les positions d'équilibre font office de paramètres dans la fonction d'onde des électrons. Cette simplification tout aussi fondamentale que drastique nous permet de ramener le problème à celui d'un système d'électrons en interaction plongés dans un potentiel externe, $\hat{V}^{\text{ext}}(\{\mathbf{r}_i\})$,

$$\hat{H}\Psi_{l,\{\mathbf{R}_j\}}(\{\mathbf{r}_i\}) = \left(\hat{T}^e(\{\mathbf{r}_i\}) + \hat{V}^{e-e}(\{\mathbf{r}_i\}) + \hat{V}^{\text{ext}}(\{\mathbf{r}_i\})\right)\Psi_{l,\{\mathbf{R}_j\}}(\{\mathbf{r}_i\}) = E_l\Psi_{l,\{\mathbf{R}_j\}}(\{\mathbf{r}_i\}), \quad (2.2.1)$$

où les états propres $\Psi_{l,\{\mathbf{R}_j\}}(\{\mathbf{r}_i\})$ sont des fonctions antisymétriques sous échange de deux coordonnées électroniques et où E_l est l'énergie totale du système d'électrons dans l'approximation des noyaux statiques pour l'état propre d'indice l . Notons que cette énergie est différente de celle définie à l'équation (2.1.4) puisque nous avons éliminé les termes dépendant uniquement des coordonnées des noyaux. L'indice $\{\mathbf{R}_j\}$ des fonctions d'onde met en évidence leur dépendance implicite sur les coordonnées des noyaux, qui paramétrisent le potentiel externe.

La puissance de l'approximation Born-Oppenheimer consiste donc à diminuer la dimensionnalité du problème par un facteur $3N_{\text{at}}$, ce qui est tout sauf négligeable. Néanmoins, l'interaction coulombienne entre les électrons couple toujours l'ensemble des coordonnées électroniques, de sorte que le problème demeure numériquement insoluble. Cette constatation est à l'origine de la célèbre citation de Dirac, qui affirme en 1929 que « les lois physiques régissant la théorie mathématique d'une grande partie de la physique et de l'ensemble de la chimie sont ainsi parfaitement connues, et la difficulté est seulement que l'application exacte de ces lois conduit à des équations beaucoup trop compliquées pour être solubles » [7]. Le développement d'approximations supplémentaires sera donc nécessaire pour rendre accessible le calcul des propriétés physiques des cristaux.

2.3. Théorie de la fonctionnelle de la densité

Parmi les méthodes développées pour adresser ce problème, la théorie de la fonctionnelle de la densité (DFT), développée par Walter Kohn, Pierre Hohenberg et Lu J. Sham dans les années 1960 [8, 9], est rapidement devenue une méthode de prédilection pour le calcul *ab initio*

des propriétés électroniques des matériaux⁵. Rapide et peu coûteuse comparativement aux approches basées sur la théorie à N -corps, elle constitue un bon point de départ pour obtenir les propriétés de l'état fondamental d'un matériau, telles que l'énergie de formation, l'énergie d'ionisation et la longueur des liaisons atomiques [4]. La présente section survole de façon succincte les concepts clés de ce formalisme. Nous référons à l'article de Cappelle [4] pour une introduction pédagogique à cette méthodologie, ainsi qu'au livre de Richard Martin [10] pour un développement plus approfondi.

2.3.1. Fonctions et fonctionnelles

Commençons par définir le concept mathématique de fonctionnelle en construisant une analogie avec les objets plus familiers que sont les fonctions. Une fonction se définit comme une application qui relie un scalaire à un autre scalaire, $f(x) = y$: pour chaque valeur de x , la fonction $f(x)$ retourne une unique valeur y . Le concept est similaire dans le cas d'une fonctionnelle, à la différence près que l'argument de la fonctionnelle est une fonction plutôt qu'un scalaire. Ainsi, la fonctionnelle $F[g(x)] = y$ génère un scalaire y différent pour chaque fonction $g(x)$ qu'on lui donne en argument. Par exemple, si la distance entre deux villes A et B calculée à vol d'oiseau est une fonction, le temps requis pour voyager d'une ville à l'autre, qui dépend explicitement de la trajectoire empruntée, sera une fonctionnelle⁶. Dans le contexte de la DFT, c'est l'énergie totale que l'on exprimera comme une fonctionnelle de la densité électronique.

2.3.2. Théorèmes de Hohenberg-Kohn

L'approche standard de la mécanique quantique consiste à définir l'hamiltonien du système d'électrons, pour ensuite obtenir les états propres en résolvant l'équation de Schrödinger, puis évaluer les différents observables. En observant l'équation (2.2.1), on constate que la seule quantité pouvant changer d'un système à l'autre est le potentiel externe, \hat{V}^{ext} . Notre problème est donc déterminé de façon unique par la configuration cristalline spécifique du cristal. Ainsi, connaissant le potentiel externe, on peut diagonaliser \hat{H} pour obtenir l'ensemble des fonctions propres et valeurs propres, ce qui nous permet d'identifier l'état fondamental Ψ_0 , que l'on utilise enfin pour calculer la densité électronique qui y est associée,

5. Je reconnais que le recours aux fonctionnelles d'échange et corrélation qui est inhérent à l'approche DFT peut être perçu comme étant conceptuellement différent des méthodes basées uniquement sur les principes premiers de la mécanique quantique. La construction de certaines de ces fonctionnelles repose néanmoins sur les principes physiques fondamentaux du gaz d'électrons interagissants, dans l'esprit des principes premiers. Dans le cadre de cet ouvrage, on utilisera le terme *ab initio* pour désigner autant des approches issues des principes premiers que des méthodes basées sur la DFT.

6. Voyager de Montréal à Québec en passant par Sherbrooke sera de toute évidence beaucoup plus long que de rouler en longeant le fleuve...

$n_0(\mathbf{r})$. Cette stratégie se résume schématiquement par [4]

$$\hat{V}^{\text{ext}} \Rightarrow \hat{H} \Rightarrow \{\Psi_l\} \Rightarrow \Psi_0 \Rightarrow n_0(\mathbf{r}). \quad (2.3.1)$$

Les deux théorèmes de Hohenberg-Kohn [8] constituent les piliers centraux de la DFT. Le premier révèle une relation de bijection entre le potentiel externe et la densité électronique de l'état fondamental, $n_0(\mathbf{r})$. Ainsi, à un potentiel externe donné correspond une et une seule densité électronique qui minimise l'énergie totale du système. Ce faisant, la relation (2.3.1) devient

$$\begin{array}{ccc} n_0(\mathbf{r}) & \iff & \hat{V}^{\text{ext}} \\ \Downarrow & & \Downarrow \\ \Psi_0 & \iff & \hat{H} \end{array} . \quad (2.3.2)$$

La densité électronique devient donc une quantité centrale permettant de définir notre problème et prend le rôle auparavant accordé au potentiel externe. On réduit de plus substantiellement la dimensionnalité du problème : la valeur attendue d'une observable de l'état fondamental, qui était jusqu'à présent une fonctionnelle de la fonction d'onde $|\Psi(\{\mathbf{r}_i\})\rangle$ dépendant de $3N$ coordonnées, pourrait désormais, en principe, être exprimée comme une fonctionnelle de la densité électronique $n_0(\mathbf{r})$ qui, elle, dépend seulement de trois variables spatiales⁷.

Suivant l'approche de Levy [11–13], on pourra ainsi exprimer l'énergie totale du système comme une fonctionnelle unique de la densité électronique,

$$E[n(\mathbf{r})] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{V}^{e-e} + \hat{V}^{\text{ext}} | \Psi \rangle \quad (2.3.3a)$$

$$= \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{V}^{e-e} | \Psi \rangle + \int d^3r n(\mathbf{r}) v_{\text{ext}}(\mathbf{r}) \quad (2.3.3b)$$

$$= F[n(\mathbf{r})] + \int d^3r n(\mathbf{r}) v_{\text{ext}}(\mathbf{r}), \quad (2.3.3c)$$

où v_{ext} est le potentiel externe ressenti par les électrons,

$$v_{\text{ext}}(\mathbf{r}) = \frac{\delta}{\delta n(\mathbf{r})} \left(\min_{\Psi \rightarrow n} \langle \Psi | \hat{V}^{\text{ext}} | \Psi \rangle \right) \quad (2.3.4a)$$

$$= \frac{\delta E_{\text{ext}}[n(\mathbf{r}')] }{\delta n(\mathbf{r})}, \quad (2.3.4b)$$

que nous avons à la dernière ligne exprimé en termes de la fonctionnelle d'énergie externe $E_{\text{ext}}[n(\mathbf{r})]$, qui dépend de la densité électronique^{8,9}. La limite $\Psi \rightarrow n$ des équations (2.3.3)

7. Une telle formulation en termes de la densité électronique n'est toutefois pas connue.

8. À proprement parler, $E_{\text{ext}}[n(\mathbf{r})]$ correspond à l'énergie potentielle de la densité électronique $n(\mathbf{r})$ plongée dans un potentiel externe $v_{\text{ext}}(\mathbf{r})$.

9. La dérivée fonctionnelle $\delta/\delta n(\mathbf{r})$ de l'équation (2.3.4b) évalue le taux de variation de $E[n]$ lorsqu'on modifie infinitésimalement la densité électronique au voisinage du point \mathbf{r} , tel que détaillé dans les références [1, 10].

et (2.3.4) signifie que la procédure de minimisation se restreint à l'ensemble des fonctions d'onde Ψ reproduisant la densité électronique $n(\mathbf{r})$.

La fonctionnelle $F[n(\mathbf{r})]$ introduite à l'équation (2.3.3c), ou fonctionnelle d'énergie interne, regroupe quant à elle l'énergie cinétique et l'énergie potentielle coulombienne du système d'électrons en interaction. On peut de plus remarquer que cette quantité est une fonctionnelle *universelle* de la densité électronique, puisqu'elle est indépendante du potentiel externe créé par les noyaux. Ainsi, la formulation de la fonctionnelle d'énergie totale (éq. (2.3.3c)) d'une molécule et d'un cristal, de même que celle de deux cristaux différents, ne diffèrent que par l'expression du potentiel externe induit par la configuration des noyaux.

Une fois la fonctionnelle d'énergie totale posée, le second théorème de Hohenberg-Kohn stipule que nous pouvons obtenir l'énergie et la densité électronique de l'état fondamental en recourant au principe variationnel pour identifier la densité qui minimise l'équation (2.3.3c) :

$$\begin{aligned} E^{\text{BO}} &= \min_{n(\mathbf{r})} E[n(\mathbf{r})] \\ &= E[n_0(\mathbf{r})], \end{aligned} \tag{2.3.5}$$

où $E^{\text{BO}} = E_0$ porte le nom d'énergie de Born-Oppenheimer.

Le premier théorème de Hohenberg-Kohn nous garantit quant à lui que toute densité $n(\mathbf{r})$ différente de $n_0(\mathbf{r})$ sera associée à une énergie totale plus grande que celle de l'état fondamental.

2.3.3. Équations de Kohn-Sham

Bien que le formalisme de la DFT donné par les théorèmes de Hohenberg-Kohn soit exact, la minimisation directe de la fonctionnelle d'énergie définie à l'équation (2.3.3c) demeure ardue sur le plan computationnel, puisque nous ignorons la forme analytique de la fonctionnelle d'énergie interne pour un ensemble d'électrons en interaction. La contribution la plus problématique est celle de l'énergie cinétique, dont nous ne connaissons l'expression exacte que pour les systèmes non interagissants. En effet, les premières tentatives de paramétrisation de l'énergie cinétique d'un ensemble d'électrons en interaction uniquement en termes de la densité électronique, connues comme la théorie de Thomas-Fermi [1, 13], ne se sont pas révélées particulièrement prédictives pour des systèmes réels. Pour obtenir une meilleure approximation de la fonctionnelle d'énergie cinétique, Kohn et Sham ont proposé de remplacer le problème de N électrons en interaction par N problèmes d'électrons indépendants interagissant avec un potentiel effectif qui paramétrise l'effet des autres particules, sous la condition que les deux systèmes possèdent la même densité électronique $n_0(\mathbf{r})$ [9].

Pour ce faire, on décompose la fonctionnelle d'énergie interne en termes de fonctionnelles définies pour des états électroniques à une particule,

$$\begin{aligned}
F[n] &= T[n] + W[n] \\
&= T_s[n] + (T[n] - T_s[n]) + E_H[n] + (W[n] - E_H[n]) \\
&= T_s[n] + E_H[n] + E_{xc}[n],
\end{aligned} \tag{2.3.6}$$

où $T[n]$ et $W[n]$ correspondent respectivement aux fonctionnelles d'énergie cinétique et coulombienne du système en interaction. Dans cette expression, $T_s[n]$ est l'énergie cinétique d'un système de N électrons indépendants décrits par les fonctions d'onde à une particule $\{\phi_l(\mathbf{r})\}$:

$$T_s[n] = -\frac{1}{2} \sum_l^{\text{occ}} \int d^3r \phi_l^*(\mathbf{r}) \nabla^2 \phi_l(\mathbf{r}) = T_s[\{\phi_l\}], \tag{2.3.7}$$

où la somme est restreinte aux états occupés et où l'indice l réfère à l'ensemble des nombres quantiques nécessaires pour identifier uniquement un état à une particule. L'énergie de Hartree [1], $E_H[n]$, correspond quant à elle à l'énergie électrostatique coulombienne de la distribution de charge $n(\mathbf{r})$,

$$E_H[n] = \frac{1}{2} \iint d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.3.8}$$

Les termes entre parenthèses dans l'équation (2.3.6) regroupent enfin la portion de la fonctionnelle d'énergie interne qui ne peut pas être décrite par la notion de particules indépendantes et définissent l'énergie d'échange-corrélation, E_{xc} . Cette dernière corrige le terme de Hartree pour respecter le principe d'exclusion de Pauli et capture les contributions de la corrélation électronique à l'énergie totale du système en interaction. Les fonctions d'onde présentes dans l'équation (2.3.7) portent le nom d'*orbitales de Kohn-Sham* et forment la base d'états électroniques auxiliaires que nous allons utiliser pour minimiser la fonctionnelle d'énergie sous la contrainte qu'elles reproduisent la densité de l'état fondamental du système en interaction, sachant que

$$n(\mathbf{r}) = \sum_l^{\text{occ}} |\phi_l(\mathbf{r})|^2. \tag{2.3.9}$$

La fonction d'onde totale du système auxiliaire s'exprimera dès lors comme un déterminant de Slater des états à une particule,

$$\Phi(\{\mathbf{r}_i\}) = \frac{1}{\sqrt{N!}} \sum_{s(l_i)} (-1)^{\text{perm}(s(l_i))} \prod_{i=1}^{N_e} \phi_{s(l_i)}(\mathbf{r}_i), \tag{2.3.10}$$

où l_i indique l'état quantique dans lequel se trouve le $i^{\text{ième}}$ électron, alors que $s(l_i)$ réfère aux $N!$ différentes permutations des N_e indices l_i caractérisant l'occupation des états à une particule par les différents électrons. La notation $\text{perm}(s(l_i))$ compte quant à elle le nombre de permutations de paires de coordonnées électroniques dans la combinaison $s(l_i)$ et procure

le bon signe à chaque terme du déterminant de Slater pour que le principe d'exclusion de Pauli soit respecté.

Cette simplification peut sembler radicale au premier coup d'œil. Nous venons après tout de dissimuler toute la complexité du système interagissant à l'intérieur d'une petite boîte noire nommée énergie d'échange-corrélation ! Néanmoins, cette approche nous permet a priori d'obtenir la forme exacte de la densité électronique de l'état fondamental, dans la mesure où la fonctionnelle d'échange-corrélation universelle serait connue. C'est donc en choisissant une paramétrisation pour E_{xc} que nous introduisons des approximations dans le formalisme de la DFT. Parmi les plus utilisées, notons l'approximation de la densité locale (*local density approximation*, LDA) [14], où l'énergie d'échange-corrélation ne dépend que de la valeur de la densité au point \mathbf{r} , ainsi que l'approximation du gradient généralisé (*generalized gradient approximation*, GGA) [15], qui prend en plus en compte le gradient local de la densité à ce point.

Puisque la densité électronique s'exprime en termes des orbitales de Kohn-Sham, la minimisation de la fonctionnelle d'énergie (éq. (2.3.3c)) vérifiera [1]

$$\begin{aligned} 0 &= \frac{\delta}{\delta \phi_l^*(\mathbf{r})} \left(E[n] - \sum_{l'} \varepsilon_{l'} \langle \phi_{l'}^* | \phi_{l'} \rangle \right), \\ &= -\frac{1}{2} \nabla^2 \phi_l(\mathbf{r}) + \left(\int d^3 r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{\text{ext}}[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \right) \phi_l(\mathbf{r}) - \varepsilon_l \phi_l(\mathbf{r}), \end{aligned} \quad (2.3.11)$$

où $\{\varepsilon_{l'}\}$ sont des multiplicateurs de Lagrange que l'on a introduit pour assurer la normalisation des états auxiliaires à une particule. De la seconde ligne de l'équation précédente, on obtient les équations de Kohn-Sham,

$$\left(-\frac{1}{2} \nabla^2 + v_H(\mathbf{r}) + v_{\text{ext}}(\mathbf{r}) + v_{xc}(\mathbf{r}) \right) \phi_l(\mathbf{r}) = \varepsilon_l \phi_l(\mathbf{r}), \quad (2.3.12)$$

qui correspondent à l'équation de Schrödinger pour des électrons indépendants dans un potentiel effectif appelé *potentiel de Kohn-Sham*,

$$V^{\text{KS}}(\mathbf{r}) = v_H(\mathbf{r}) + v_{\text{ext}}(\mathbf{r}) + v_{xc}(\mathbf{r}). \quad (2.3.13)$$

On peut donc assimiler la DFT à une méthode de champ moyen, puisqu'elle remplace le traitement explicite des interactions entre les électrons par un potentiel à un corps ressenti indépendamment par les différentes particules auxiliaires, tel qu'illustré à la figure 2.1b.

Le cœur de la DFT consistera donc à résoudre les équations de Kohn-Sham à l'aide d'un processus itératif autocohérent, qui est illustré à la figure 2.1a. On commencera par estimer la densité électronique, $n_{i=0}(\mathbf{r})$, à partir du potentiel externe. À l'aide de cette densité initiale, on pourra ensuite construire le potentiel de Kohn-Sham et résoudre l'équation (2.3.12) pour obtenir un ensemble d'orbitales et une nouvelle densité, $\{\phi_l\}_{i=1}$ et $n_{i=1}(\mathbf{r})$, que l'on utilisera pour poser un nouveau problème de Kohn-Sham et obtenir une nouvelle solution.

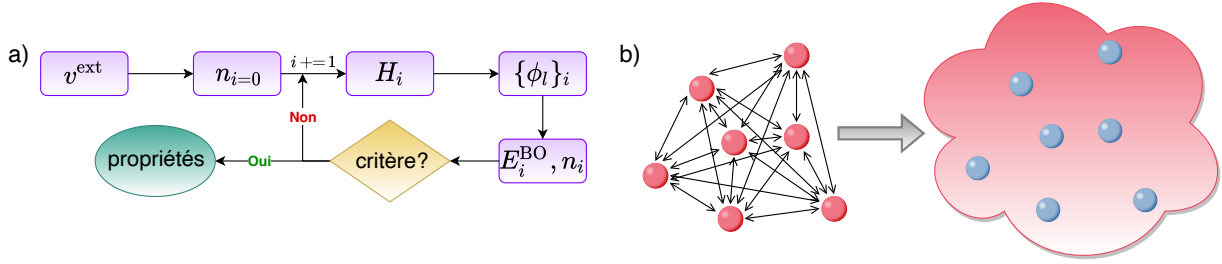


Figure 2.1 – Cycle d’autocohérence et potentiel effectif de la DFT. a) La DFT résoud les équations de Kohn-Sham de manière itérative, jusqu’à ce que les quantités physiques finales d’une boucle ne varient plus significativement par rapport aux quantités initiales. b) Le théorème de Kohn-Sham remplace le problème de N électrons en interaction (gauche) par N problèmes de fermions indépendants plongés dans un potentiel effectif (droite). Ces particules auxiliaires sont entièrement fictives, d’où leur couleur différente.

À la fin de chaque itération, les propriétés du système auxiliaire sont comparées à celles de l’itération précédente, jusqu’à ce que leur variation soit en deçà d’un critère d’arrêt préalablement déterminé. Par exemple, on pourra choisir d’arrêter l’algorithme lorsque la différence d’énergie totale entre deux itérations consécutives devient suffisamment faible, ou lorsque l’expression des orbitales de Kohn-Sham ou du potentiel de Kohn-Sham ne change plus de manière significative.

2.3.4. Le problème de la bande interdite

La forme des équations de Kohn-Sham suggère que l’on puisse assimiler les états $\{\phi_l\}$ aux fonctions d’onde des quasi-particules et les valeurs propres ε_l , à leurs énergies propres. Ces dernières sont d’ailleurs couramment utilisées pour visualiser la structure de bande des matériaux. Il ne faut néanmoins pas perdre de vue que le système auxiliaire a été introduit uniquement dans le but de reconstruire la densité de l’état fondamental. Les orbitales de Kohn-Sham et leurs valeurs propres décrivent donc des particules fictives et n’ont a priori aucune signification physique.

La littérature a cependant mis en évidence que les valeurs propres de Kohn-Sham s’avèrent une relativement bonne première approximation de la véritable structure électronique, dans la mesure où les corrélations électroniques sont suffisamment faibles [4]. Cet accord inattendu est d’ailleurs en grande partie à l’origine du succès de la DFT. Une exception de taille à cette observation est l’énergie de la bande interdite des semi-conducteurs et des isolants, qui est notoirement mal décrite par la DFT. En effet, même pour des matériaux faiblement corrélés comme le silicium, elle est très souvent sous-estimée par un facteur pouvant atteindre jusqu’à 50%, comme en témoignent les marqueurs bleus de la figure 2.2. Certains isolants dont les corrélations électroniques sont particulièrement fortes peuvent de

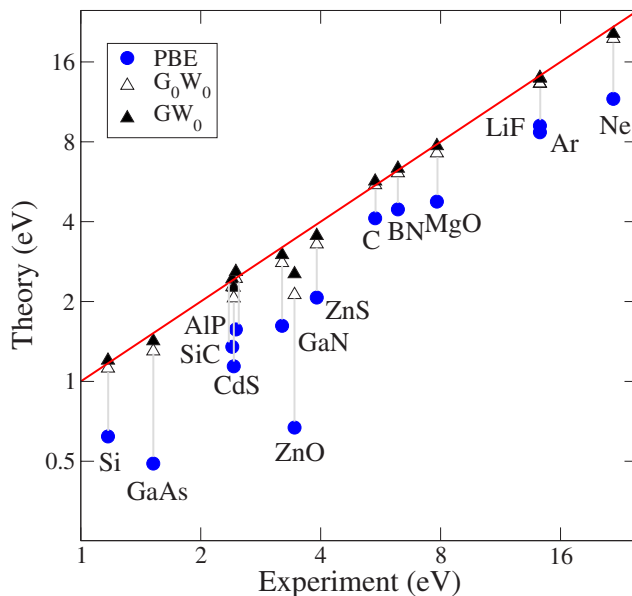


Figure 2.2 – Comparaison de la prédictibilité de l'énergie de la bande interdite des méthodes DFT (cercles bleus) et GW (triangles). Les marqueurs pleins et ouverts réfèrent à deux variations de la méthode GW . Pour les matériaux montrés, la DFT sous-estime systématiquement l'énergie de la bande interdite ; il s'agit là d'une manifestation du problème de la bande interdite discuté à la section 2.3.4. La GW traite explicitement les corrélations électroniques et reproduit beaucoup mieux les résultats expérimentaux. Figure tirée de la référence [16].

plus être décrits à tort comme des métaux. Une telle erreur systématique met en lumière que la DFT est en essence une théorie de l'état fondamental et n'est donc pas, en principe, en mesure de prédire correctement les propriétés des états excités. Pour obtenir une prédiction fiable de l'énergie de la bande interdite, une direction possible est l'utilisation des méthodes qui traitent explicitement les corrélations électroniques dans le cadre du problème à N -corps. Parmi les méthodes de pointe, la GW compte parmi les plus couramment utilisées et permet de prédire l'énergie de la bande interdite avec une erreur relative de l'ordre 5 à 10% par rapport aux valeurs expérimentales. Cette amélioration considérable par rapport à la DFT (fig. 2.2) ne s'obtient toutefois qu'au prix d'une augmentation significative du coût computationnel [16].

2.4. Propriétés des états électroniques sous les opérations de symétrie

2.4.1. États de Bloch

Dans un tout autre ordre d'idées, nous avons déjà souligné que les états électroniques dans un matériau dépendent de la manière dont sont assemblés les différents atomes qui le

constituent. Dans le cas des cristaux, cet arrangement est ordonné de façon périodique dans l'espace, de sorte que l'on peut décrire l'entièreté du cristal uniquement à partir d'une unité fondamentale appelée cellule unitaire, comprenant un très petit nombre d'atomes comparativement au nombre d'Avogadro. Cette cellule unitaire est ensuite répliquée dans l'espace en lui appliquant les opérations de translation appropriées. La périodicité spatiale des cristaux permet dès lors d'exprimer les états électroniques¹⁰ d'une manière remarquablement simple, grâce au théorème de Bloch. Ce dernier nous indique que toutes les solutions d'un hamiltonien à une particule dans un potentiel périodique peuvent s'exprimer sous la forme d'une onde plane modulée par une fonction possédant la même périodicité spatiale que le cristal. Tout comme pour l'approximation Born-Oppenheimer, la puissance du théorème de Bloch se manifeste dans la réduction drastique de la dimensionnalité du problème. En effet, la symétrie de translation qui relie les différentes cellules unitaires entre elles permet d'obtenir les propriétés du système de $\sim 10^{23}$ atomes uniquement à partir des N atomes formant la cellule unitaire [1].

Les états propres d'un système périodique, appelés également états de Bloch, prennent donc la forme [5]

$$\psi_{\mathbf{k}n}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}n}(\mathbf{r}), \quad u_{\mathbf{k}n}(\mathbf{r} + \mathbf{R}) = u_{\mathbf{k}n}(\mathbf{r}), \quad (2.4.1)$$

où $u_{\mathbf{k}n}(\mathbf{r})$ correspond à la portion périodique. La normalisation de $\psi_{\mathbf{k}n}(\mathbf{r})$ est implicite.

Par ailleurs, puisque les états électroniques reflètent la périodicité discrète du cristal, ils doivent être des états propres de l'opérateur de translation pour toute translation par un vecteur \mathbf{R} appartenant au réseau de Bravais. On peut montrer que cette condition, combinée à l'additivité des opérations de translation, impose la relation [5]

$$\psi_{\mathbf{k}n}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}} \psi_{\mathbf{k}n}(\mathbf{r}). \quad (2.4.2)$$

2.4.2. Opérations de symétrie générales

La formulation des états propres en termes d'états de Bloch est particulièrement utile pour caractériser la dispersion électronique en présence d'autres types de symétries. Nous allons maintenant généraliser la relation (2.4.2) à n'importe quel type de symétrie du cristal et relier ces symétries à la présence de dégénérescences dans le spectre de l'hamiltonien. Cette généralisation se montrera utile pour caractériser le comportement des états électroniques lorsque l'on considère l'interaction spin-orbite (section 2.5.3), ainsi que lorsqu'on s'intéressera aux propriétés des états de surface topologiques protégés par la symétrie de renversement du temps (section 7.3.1). Par souci de simplicité, nous traiterons l'exemple d'une rotation bidimensionnelle pour un système de symétrie cubique [17]. On négligera aussi le spin des électrons.

10. Dans le cadre de la DFT, les états électroniques correspondent aux orbitales de Kohn-Sham.

Considérons une symétrie $\hat{\mathbf{S}}$, qui appartient au groupe de symétrie ponctuelle du cristal¹¹. Pour un système cubique, on peut penser à une rotation par un multiple entier de $\pi/2$. Puisque $\hat{\mathbf{S}}$ est une symétrie du potentiel cristallin, ce dernier vérifie

$$V(\hat{\mathbf{S}}\mathbf{r}) = V(\mathbf{r}), \quad (2.4.3)$$

où $\hat{\mathbf{S}}\mathbf{r}$ correspond au vecteur \mathbf{r} transformé par la symétrie $\hat{\mathbf{S}}$. Puisque $\hat{\mathbf{S}}$ préserve la norme de \mathbf{r} , on peut déduire qu'il s'agit d'une transformation unitaire, vérifiant $\hat{\mathbf{S}}^\dagger = \hat{\mathbf{S}}^{-1}$ ¹².

Supposons maintenant que l'on connaît la solution à l'équation de Schrödinger pour l'état de Bloch $\psi_{\mathbf{k}n}(\mathbf{r})$, donnée par

$$\left(-\frac{\nabla_{\mathbf{r}}^2}{2} + V(\mathbf{r})\right) \psi_{\mathbf{k}n}(\mathbf{r}) = \varepsilon_{\mathbf{k}n} \psi_{\mathbf{k}n}(\mathbf{r}). \quad (2.4.4)$$

En appliquant la transformation $\hat{\mathbf{S}}$ à l'équation précédente, on obtient

$$\begin{aligned} \left(-\frac{\nabla_{\hat{\mathbf{S}}\mathbf{r}}^2}{2} + V(\hat{\mathbf{S}}\mathbf{r})\right) \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r}) &= \varepsilon_{\mathbf{k}n} \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r}), \\ \left(-\frac{\nabla_{\mathbf{r}}^2}{2} + V(\mathbf{r})\right) \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r}) &= \varepsilon_{\mathbf{k}n} \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r}), \end{aligned} \quad (2.4.5)$$

puisque le Laplacien est invariant sous toute transformation unitaire et que $\hat{\mathbf{S}}$ est une symétrie du potentiel vérifiant l'équation (2.4.3). Ce résultat nous indique que $\psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r})$ est un état de Bloch ayant la même énergie que $\psi_{\mathbf{k}n}(\mathbf{r})$. On doit également vérifier qu'il s'agit d'états linéairement indépendants. Pour ce faire, on utilise l'équation (2.4.2),

$$\begin{aligned} \psi_{\mathbf{k}n}(\hat{\mathbf{S}}(\mathbf{r} + \mathbf{R})) &= \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r} + \hat{\mathbf{S}}\mathbf{R}) \\ &= e^{i\mathbf{k} \cdot \hat{\mathbf{S}}\mathbf{R}} \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r}), \end{aligned} \quad (2.4.6)$$

puisque $\hat{\mathbf{S}}\mathbf{R}$, qui correspond à la transformation du vecteur du réseau \mathbf{R} par la symétrie du réseau $\hat{\mathbf{S}}$, appartient également au réseau de Bravais.

Sachant que les transformations unitaires préservent la norme et le produit scalaire, on peut distribuer une seconde transformation $\hat{\mathbf{S}}'$ sur les vecteurs formant le produit scalaire, $\hat{\mathbf{S}}'(\mathbf{k} \cdot \hat{\mathbf{S}}\mathbf{R}) = (\hat{\mathbf{S}}'\mathbf{k}) \cdot (\hat{\mathbf{S}}'\hat{\mathbf{S}}\mathbf{R})$. En choisissant judicieusement $\hat{\mathbf{S}}' = \hat{\mathbf{S}}^{-1}$, l'équation précédente devient

$$\psi_{\mathbf{k}n}(\hat{\mathbf{S}}(\mathbf{r} + \mathbf{R})) = e^{i(\hat{\mathbf{S}}^{-1}\mathbf{k}) \cdot \mathbf{R}} \psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r}), \quad (2.4.7)$$

ce qui nous indique que l'état de Bloch $\psi_{\mathbf{k}n}(\hat{\mathbf{S}}\mathbf{r})$ possède une impulsion cristalline $\mathbf{k}' = \hat{\mathbf{S}}^{-1}\mathbf{k}$.

11. Le groupe de symétrie ponctuelle regroupe les opérations de symétrie qui laissent le cristal invariant tout en maintenant fixe un des points du réseau.

12. Puisqu'elles agissent uniquement dans l'espace réel, les rotations correspondent plus exactement à des transformations orthogonales, qui forment une sous-classe des transformations unitaires.

En regroupant les deux observations précédentes, on obtient les relations suivantes entre les états de Bloch reliés par la symétrie \hat{S} ,

$$\psi_{\mathbf{k}n}(\hat{S}\mathbf{r}) = \psi_{(\hat{S}^{-1}\mathbf{k})n}(\mathbf{r}), \quad (2.4.8a)$$

$$\varepsilon_{(\hat{S}^{-1}\mathbf{k})n} = \varepsilon_{\mathbf{k}n}. \quad (2.4.8b)$$

2.4.3. Symétrie d'inversion

La symétrie d'inversion, \mathcal{I} , correspond à l'inversion des coordonnées spatiales, $\mathcal{I}\mathbf{r} = -\mathbf{r}$, $\mathcal{I}\mathbf{k} = -\mathbf{k}$ ¹³. Dans ces conditions, l'équation (2.4.8) devient

$$\psi_{\mathbf{k}n}(-\mathbf{r}) = \psi_{-\mathbf{k}n}(\mathbf{r}), \quad (2.4.9a)$$

$$\varepsilon_{-\mathbf{k}n} = \varepsilon_{\mathbf{k}n}. \quad (2.4.9b)$$

2.4.4. Symétrie de renversement du temps et théorème de Kramers

La symétrie de renversement du temps se distingue des transformations unitaires décrites précédemment. Cette opération correspond classiquement à inverser la direction de la propagation temporelle, $t \rightarrow -t$. Ainsi, les quantités physiques qui impliquent une direction de propagation, telles que l'impulsion et le champ magnétique, voient leur signe inversé, alors que les autres, comme la position et le champ électrique, demeurent invariantes. On s'attend donc à ce qu'une relation similaire existe pour les opérateurs.

Du point de vue de la mécanique quantique, la simple transformation $\mathbf{R} \rightarrow \mathbf{R}$, $\mathbf{P} \rightarrow -\mathbf{P}$ est néanmoins problématique, puisqu'elle ne conserve pas la relation de commutation canonique entre \mathbf{R} et \mathbf{P} . La solution, proposée originellement par Wigner, est de décrire le renversement du temps par un opérateur \mathcal{T} antiunitaire, ce qui implique d'ajouter la conjugaison complexe. Ainsi, on a bien $\mathcal{T}\mathbf{R}\mathcal{T}^{-1} = \mathbf{R}$, $\mathcal{T}\mathbf{P}\mathcal{T}^{-1} = (-i\nabla)^* = i\nabla = -\mathbf{P}$ et $\mathcal{T}[R_i, P_j]\mathcal{T}^{-1} = ([R_i, -P_j])^* = i\delta_{ij}$ [18, 19].

On cherche maintenant à identifier l'état de Bloch partenaire sous renversement du temps de l'état $\psi_{\mathbf{k}n}(\mathbf{r}, t)$ [17]. Sous transformation $t \rightarrow -t$, cet état devient

$$\begin{aligned} \psi_{\mathbf{k}n}(\mathbf{r}, -t) &= \psi_{\mathbf{k}n}(\mathbf{r})e^{-i\varepsilon_{\mathbf{k}n}(-t)} \\ &= \psi_{\mathbf{k}n}(\mathbf{r})e^{i\varepsilon_{\mathbf{k}n}t}. \end{aligned} \quad (2.4.10)$$

Le renversement du temps se manifeste donc par la conjugaison complexe de l'exponentielle, impliquant que l'état $\psi_{\mathbf{k}n}(\mathbf{r}, -t)$ vérifie la relation

$$\left(-\frac{\nabla_{\mathbf{r}}^2}{2} + V(\mathbf{r}) \right) \psi_{\mathbf{k}n}(\mathbf{r}, -t) = -i \frac{\partial \psi_{\mathbf{k}n}(\mathbf{r}, -t)}{\partial t}, \quad (2.4.11)$$

13. \mathbf{r} et \mathbf{k} sont ici des vecteurs et non des opérateurs, pour lesquels la transformation sous inversion est décrite par $\mathcal{I}\mathbf{R}\mathcal{I}^{-1} = -\mathbf{R}$

qui ne correspond pas à l'équation de Schrödinger dépendante du temps à cause du changement de signe $i \rightarrow -i$. On retrouve la forme correcte en appliquant la conjugaison complexe à l'équation précédente,

$$\left(-\frac{\nabla_{\mathbf{r}}^2}{2} + V(\mathbf{r})\right) \psi_{\mathbf{kn}}(\mathbf{r}, -t)^* = i \frac{\partial \psi_{\mathbf{kn}}(\mathbf{r}, -t)^*}{\partial t}, \quad (2.4.12)$$

ce qui nous indique que l'état partenaire que nous recherchons est

$$\mathcal{T} \psi_{\mathbf{kn}}(\mathbf{r}, t) = \psi_{\mathbf{kn}}(\mathbf{r}, -t)^* = \psi_{\mathbf{kn}}(\mathbf{r})^* e^{-i\varepsilon_{\mathbf{kn}} t}. \quad (2.4.13)$$

On obtient l'énergie de cet état en prenant la conjugaison complexe de l'équation de Schrödinger indépendante du temps pour $\psi_{\mathbf{kn}}(\mathbf{r})$,

$$\left(-\frac{\hbar^2 \nabla_{\mathbf{r}}^2}{2m_e} + V(\mathbf{r})\right) \psi_{\mathbf{kn}}(\mathbf{r})^* = \varepsilon_{\mathbf{kn}} \psi_{\mathbf{kn}}(\mathbf{r})^*, \quad (2.4.14)$$

puisque H est hermitien, impliquant que $\varepsilon_{\mathbf{kn}}^* = \varepsilon_{\mathbf{kn}}$.

Il nous reste à déterminer l'impulsion cristalline de l'état partenaire en utilisant l'équation (2.4.2) :

$$\begin{aligned} \psi_{\mathbf{kn}}(\mathbf{r} + \mathbf{R})^* &= \left(e^{i\mathbf{k} \cdot \mathbf{R}} \psi_{\mathbf{kn}}(\mathbf{r})\right)^* \\ &= e^{i(-\mathbf{k}) \cdot \mathbf{R}} \psi_{\mathbf{kn}}(\mathbf{r})^*, \end{aligned} \quad (2.4.15)$$

que l'on identifie comme étant $\mathbf{k}' = -\mathbf{k}$.

Par conséquent, la symétrie sous renversement du temps implique les relations [18, 19]

$$\psi_{\mathbf{kn}}(\mathbf{r})^* = \psi_{-\mathbf{kn}}(\mathbf{r}), \quad (2.4.16a)$$

$$\varepsilon_{-\mathbf{kn}} = \varepsilon_{\mathbf{kn}}. \quad (2.4.16b)$$

Cette conclusion ne dépend pas de l'existence de symétries spatiales. La double dégénérescence des états électroniques en présence de la symétrie de renversement du temps est connue sous le nom de théorème de Kramers [20]. Les paires de Kramers correspondent quant à elles aux états partenaires, $\psi_{\mathbf{kn}}(\mathbf{r})$ et $\psi_{-\mathbf{kn}}(\mathbf{r})$.

Notons finalement que l'antiunitarité de l'opérateur \mathcal{T} implique les relations suivantes [18] :

$$\mathcal{T}^2 = -\mathbb{1}, \quad (2.4.17a)$$

$$\mathcal{T} |\phi\rangle = |\phi\rangle^*, \quad (2.4.17b)$$

$$\mathcal{T} c |\psi\rangle = c^* \mathcal{T} |\psi\rangle, \quad (2.4.17c)$$

$$\langle \mathcal{T} \phi | \mathcal{T} \psi \rangle = \langle \phi | \psi \rangle^*, \quad (2.4.17d)$$

où c est un nombre complexe, ψ est un vecteur d'état appartenant à l'espace d'Hilbert et $\mathbb{1}$ est la matrice identité. La propriété 2.4.17d découle directement de la propriété 2.4.17b.

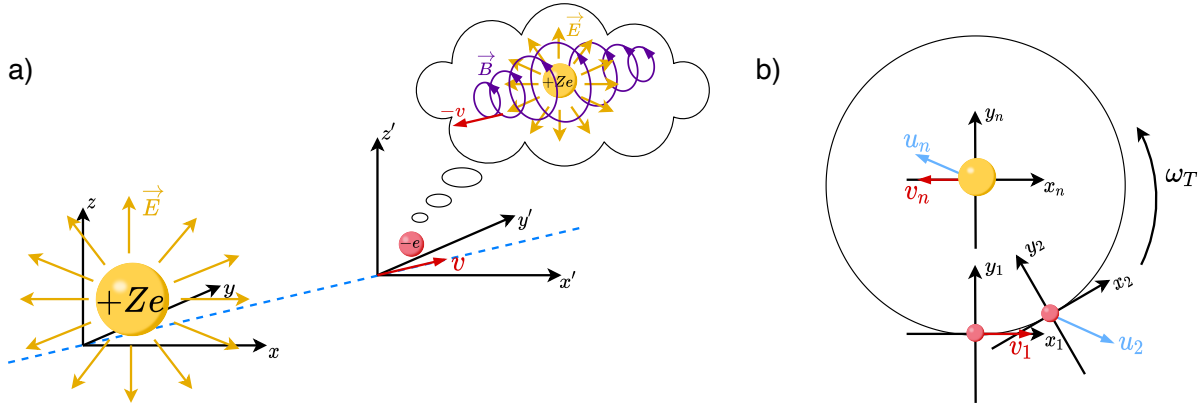


Figure 2.3 – Interprétation semi-classique du couplage spin-orbite et précession de Thomas.

a) Dans le référentiel inertiel du proton (jaune), l'électron (rouge) se déplace à vitesse \mathbf{v} (flèche rouge). Du point de vue de l'électron, c'est le proton qui se déplace à vitesse $-\mathbf{v}$. Par la transformation relativiste du champ électromagnétique, l'électron ressent un champ magnétique généré par le proton en mouvement¹⁴. b) En revenant au référentiel inertiel du proton, on doit tenir compte de la précession de Thomas, qui diminue d'un facteur 2 l'amplitude semi-classique du couplage spin-orbite. Au temps t_1 , le proton et l'électron se trouvent respectivement dans les référentiels (x_n, y_n) et (x_1, y_1) , et se déplacent à vitesse $v_1 = -v_n$ l'un par rapport à l'autre (flèches rouges). Après un temps Δt , l'électron se trouve dans le référentiel (x_2, y_2) . En appliquant la transformation relativiste des vitesses, on trouve que le référentiel du proton se déplace à vitesse u_n du point de vue de l'électron, alors que ce dernier se déplace à vitesse u_2 selon le proton (flèches bleues). Le postulat de la relativité restreinte impose que ces vitesses soient d'amplitude égale et de direction opposée. Puisque les grandeurs respectives de leurs composantes en \hat{x} et \hat{y} sont différentes, ceci n'est possible que si le référentiel de l'électron précède autour du noyau à la fréquence de Thomas, ω_T . Image adaptée de la référence [21].

2.5. Couplage spin-orbite

Jusqu'à présent, nous avons négligé le spin de l'électron et supposé que l'hamiltonien des électrons est diagonal dans l'espace des spins. Or, les états électroniques sont en réalité décrits par des spineurs. Ainsi, l'interaction du spin avec le moment cinétique orbital couple les fonctions d'onde des électrons de spins différents et modifie leurs énergies propres. Dans cette section, nous discutons de l'origine physique de cette interaction et analysons ses conséquences sur la structure électronique dans l'approximation des noyaux statiques. Pour les sections 2.5.1 et 2.5.2, nous passons temporairement aux unités SI.

14. Image adaptée de https://en.wikipedia.org/wiki/Lorentz_transformation#/media/File:Lorentz_boost_electric_charge.svg.

2.5.1. Interprétation semi-classique

Le couplage spin-orbite est une conséquence concrète de la relativité restreinte. On peut l'interpréter d'une manière profondément intuitive à l'aide d'un simple changement de référentiel, tout en conservant la description semi-classique d'un électron qui effectue une orbite circulaire autour d'un noyau [21]. En effet, puisque la masse du noyau est environ 1000 fois plus grande que celle de l'électron, on visualise en général l'orbite électronique dans le référentiel où le noyau est au repos. Or, du point de vue d'un électron qui orbite autour d'un noyau à vitesse \mathbf{v} , c'est le noyau positivement chargé qui se déplace autour de lui à vitesse $-\mathbf{v}$, ce qui, par la loi d'Ampère, génère un champ magnétique qui peut interagir avec son spin. Ces deux référentiels sont illustrés à la figure 2.3a. La transformation du champ électromagnétique sous un *boost* de Lorentz¹⁵ nous indique que ce champ s'exprime comme

$$\begin{aligned}\mathbf{B} &= \frac{1}{c^2}(-\mathbf{v}) \times \mathbf{E} \\ &= -\frac{1}{m_e c^2} \mathbf{P} \times \mathbf{E},\end{aligned}\tag{2.5.1}$$

où m_e et \mathbf{P} sont respectivement la masse et le moment cinétique de l'électron, c est la vitesse de la lumière et \mathbf{E} , le champ électrique produit par le noyau. Ce dernier peut être exprimé en termes du gradient du potentiel, $V(\mathbf{r})$,

$$\mathbf{E} = \frac{1}{e} \nabla V(\mathbf{r}).\tag{2.5.2}$$

Nous appliquons l'approximation faiblement relativiste, de sorte que le facteur de Lorentz $\gamma \approx 1$. On obtient donc

$$\mathbf{B} = \frac{1}{em_e c^2} \nabla V(\mathbf{r}) \times \mathbf{P},\tag{2.5.3}$$

où nous avons utilisé l'anticommutativité du produit vectoriel.

Ce champ magnétique interagit avec le moment magnétique de l'électron,

$$\boldsymbol{\mu} = -\frac{e}{m_e} \mathbf{S},\tag{2.5.4}$$

via l'hamiltonien

$$H^{\text{SOC}} = -\boldsymbol{\mu} \cdot \mathbf{B}.\tag{2.5.5}$$

En combinant les expressions précédentes, on obtient

$$H^{\text{SOC}} = \frac{\hbar}{2m_e^2 c^2} (\nabla V(\mathbf{r}) \times \mathbf{P}) \cdot \boldsymbol{\sigma},\tag{2.5.6}$$

par commutativité de l'opérateur \mathbf{S} avec les opérateurs de position et d'impulsion et où nous avons exprimé l'opérateur de spin de fermions de spin 1/2 en termes des matrices de Pauli, $\mathbf{S} = \hbar \boldsymbol{\sigma} / 2$, avec $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

¹⁵. On néglige pour l'instant la rotation du référentiel de l'électron et nous limitons à considérer ce dernier comme translaté à vitesse constante \mathbf{v} par rapport au noyau.

Puisque nous traitons le noyau comme une particule ponctuelle, on peut simplifier cette expression en y substituant le potentiel coulombien d'une particule de charge Ze , tel que

$$\nabla V(\mathbf{r}) = \nabla \left(-\frac{Ze^2}{r} \hat{\mathbf{r}} \right) = \frac{Ze^2}{r^2} \hat{\mathbf{r}} = \frac{Ze^2}{r^3} \mathbf{R}. \quad (2.5.7)$$

L'équation (2.5.6) prend alors la forme que l'on apprend dans les cours d'introduction à la mécanique quantique [2],

$$\begin{aligned} H^{\text{SOC}} &= \frac{\hbar Ze^2}{2m_e^2 c^2 r^3} (\mathbf{R} \times \mathbf{P}) \cdot \boldsymbol{\sigma} \\ &= \frac{Ze^2}{m_e^2 c^2 r^3} \mathbf{L} \cdot \mathbf{S}, \end{aligned} \quad (2.5.8)$$

puisque $\mathbf{r} \times \mathbf{P} = \mathbf{L}$ par définition du moment cinétique orbital.

2.5.2. Précession de Thomas et émergence du SOC dans l'équation de Dirac

L'expression obtenue à la section précédente, qui repose sur un argument heuristique semi-classique, capture correctement la dépendance en $\mathbf{L} \cdot \mathbf{S}$ du SOC, mais son préfacteur est incorrect par un facteur 1/2. En effet, nous avons supposé à tort que le référentiel de l'électron est inertiel. Or, l'électron effectuant une orbite circulaire, son référentiel est sans cesse accéléré. Un traitement adéquat demanderait de traiter le problème dans le référentiel inertiel du noyau ou, de manière équivalente, d'ajouter un terme supplémentaire au référentiel de l'électron, une « force fictive », pour prendre en compte l'effet des transformations de Lorentz successives impliquées dans son accélération, tel qu'illustré à la figure 2.3b. Cette correction est connue sous le nom de *précession de Thomas* [21]. On peut y voir une analogie avec la force de Coriolis, qui compense pour la rotation de la Terre, que l'on doit ajouter à la gravité pour décrire correctement la rotation du plan d'oscillation du pendule de Foucault.

L'argument de cette dérivation repose sur un postulat de la relativité restreinte, qui nous dit que si un référentiel inertiel se déplace à vitesse \mathbf{v} du point de vue d'un second référentiel inertiel, alors le second référentiel se déplace à vitesse $-\mathbf{v}$ du point de vue du premier. En appliquant ce postulat à notre électron en orbite autour du noyau, on démontre qu'il implique que le référentiel de l'électron précède autour du noyau avec une fréquence [21]

$$\omega_T = \frac{e}{2m_e c^2} \mathbf{v} \times \mathbf{E}, \quad (2.5.9)$$

appelée *fréquence de Thomas*. En comparant cette dernière avec la fréquence de Larmor [6], ω_L , qui est la fréquence de précession du moment magnétique de l'électron dans un champ magnétique externe, on trouve que la fréquence de précession nette du référentiel de l'électron

est

$$\begin{aligned}
\omega &= \omega_T - \omega_L \\
&= \frac{e}{2m_e c^2} \mathbf{v} \times \mathbf{E} - \frac{e}{m_e c^2} \mathbf{v} \times \mathbf{E} \\
&= -\frac{e}{2m_e c^2} \mathbf{v} \times \mathbf{E}.
\end{aligned} \tag{2.5.10}$$

La fréquence de Larmor apparaît avec un signe négatif puisqu'elle décrit la précession du spin \mathbf{S} de l'électron, qui pointe dans la direction opposée à son moment magnétique. On constate alors que l'amplitude de la fréquence de précession diminue d'un facteur 2 par rapport à la fréquence de Larmor. Ainsi, en retransformant l'équation (2.5.6) dans le référentiel du noyau, le préfacteur du calcul semi-classique diminue par la même proportion [21].

Il est d'ailleurs possible de montrer que le couplage spin-orbite émerge naturellement de l'équation de Dirac lorsqu'on lui applique la limite faiblement relativiste [21]. Au contraire de l'équation de Schrödinger qui est construite sur l'invariance galiléenne, l'équation de Dirac est invariante sous transformation de Lorentz, c'est-à-dire qu'elle traite la nature relativiste des électrons de façon explicite. Ses fonctions d'onde prennent la forme d'un bispineur, ce qui a mené à prédire l'existence du positron. Dans l'approximation faiblement relativiste, on retrouve un hamiltonien de dimension 2, s'appliquant à un seul spineur, qui s'exprime sous la forme

$$H = \frac{\mathbf{P}^2}{2m_e} + V(\mathbf{r}) + H^{\text{rel}}. \tag{2.5.11}$$

Les deux premiers termes correspondent à l'hamiltonien de Schrödinger. Le terme supplémentaire, H^{rel} , peut donc être interprété comme les corrections relativistes au premier ordre à l'équation de Schrödinger et contient trois contributions décrivant la structure fine de l'atome d'hydrogène [2]. Le terme de couplage spin-orbite y apparaît notamment avec le bon préfacteur,

$$H^{\text{SOC}} = \frac{\hbar}{4m_e^2 c^2} (\nabla V(\mathbf{r}) \times \mathbf{P}) \cdot \boldsymbol{\sigma}. \tag{2.5.12}$$

Pour un potentiel central coulombien, cette expression se simplifie à

$$H^{\text{SOC}} = \frac{Ze^2}{2m_e^2 c^2 r^3} \mathbf{L} \cdot \mathbf{S}. \tag{2.5.13}$$

2.5.3. Conséquences sur les niveaux d'énergie électroniques

De manière générale, l'ajout d'un terme supplémentaire à l'hamiltonien peut mener à des levées de dégénérescences du spectre d'énergie. En l'absence de SOC, l'hamiltonien des électrons de spin \uparrow et \downarrow est identique, ce qui implique que tous les états électroniques sont doublement dégénérés. Puisque le terme $\mathbf{L} \cdot \mathbf{S}$ brise la symétrie entre les deux états de spin, il est en mesure de lever cette dégénérescence.

Néanmoins, la présence du couplage spin-orbite préserve la symétrie de renversement du temps. Pour le démontrer, il suffit de considérer l'action de l'opérateur \mathcal{T} sur l'opérateur $\mathbf{L} \cdot \mathbf{S}$,

$$\begin{aligned}
\mathcal{T}(\mathbf{L} \cdot \mathbf{S})\mathcal{T}^{-1} &= \sum_{j=i}^3 (\mathcal{T}L_j S_j \mathcal{T}^{-1}) \\
&= \sum_{j=i}^3 (\mathcal{T}L_j \mathcal{T}^{-1} \mathcal{T}S_j \mathcal{T}^{-1}) \\
&= (\mathcal{T}\mathbf{L}\mathcal{T}^{-1}) \cdot (\mathcal{T}\mathbf{S}\mathcal{T}^{-1}) \\
&= (-\mathbf{L}) \cdot (-\mathbf{S}), \\
&= \mathbf{L} \cdot \mathbf{S}.
\end{aligned} \tag{2.5.14}$$

L'indice de sommation j réfère aux trois directions cartésiennes. Le changement de signe de l'opérateur \mathbf{L} sous renversement peut être déduit des relations présentées à la section 2.4.4 pour les opérateurs de position et d'impulsion, sachant que $\mathbf{L} = \mathbf{R} \times \mathbf{P}$.

Par conséquent, les états électroniques dont le vecteur d'onde est invariant sous renversement du temps conserveront leur double dégénérescence due aux états de spin. Les autres verront leur dégénérescence levée, à moins qu'une symétrie spatiale ne la préserve.

Par ailleurs, puisque le SOC préserve la symétrie de renversement du temps, le théorème de Kramers demeure valide, de sorte que les paires d'états électroniques équivalentes par renversement du temps possèdent la même énergie propre. L'identification des paires de Kramers doit toutefois prendre en compte que la symétrie \mathcal{T} inverse l'orientation du spin. Les paires de Kramers en présence du SOC vérifient alors [17]¹⁶

$$\psi_{\mathbf{k}n\uparrow}(\mathbf{r})^* = \psi_{-\mathbf{k}n\downarrow}(\mathbf{r}), \tag{2.5.15a}$$

$$\varepsilon_{-\mathbf{k}n\downarrow} = \varepsilon_{\mathbf{k}n\uparrow}. \tag{2.5.15b}$$

À titre d'exemple, considérons le maximum de la bande de valence (VBM) des systèmes de symétrie cubique, qui décrit de nombreux semi-conducteurs tels que le silicium, le diamant ou le GaAs. En l'absence du SOC, les dégénérescences de la structure électroniques aux différents points \mathbf{k} découlent des représentations irréductibles du groupe de symétrie ponctuelle, \mathcal{G} . Pour le centre de la zone de Brillouin, Γ , ces groupes sont $\mathcal{G} = O_h$ pour la structure diamant et $\mathcal{G} = T_d$ pour la structure zinc-blende. Rappelons que la structure zinc-blende diffère de la structure diamant simplement par l'inéquivalence des atomes de la base, qui brise la symétrie

16. La notation de l'équation (2.5.15) suppose que les états de spin \uparrow et \downarrow sont indexés par le même indice de bande n , c'est-à-dire que, comme en l'absence de SOC, chaque indice n est associé à deux états électroniques de spin différent. Dans nos calculs numériques, nous utiliserons plutôt une notation où chaque indice de bande n n'est associé qu'à un seul état électronique pour chaque \mathbf{k} , dont la composition en spin est une combinaison linéaire de \uparrow et \downarrow .

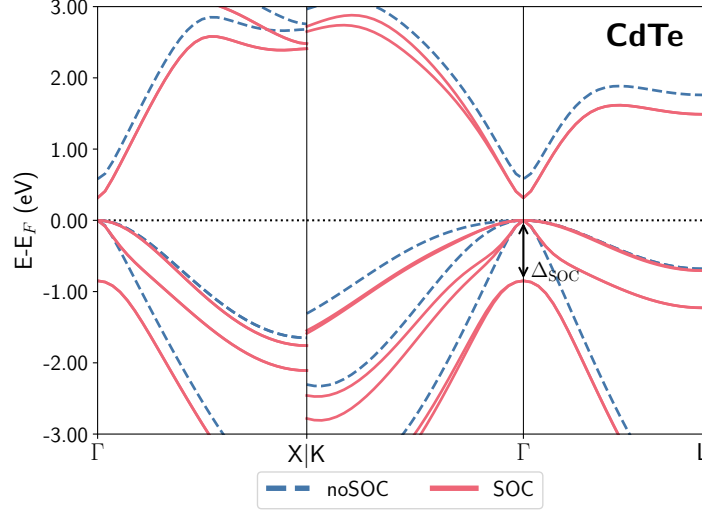


Figure 2.4 – Structure électronique du CdTe avec et sans SOC. En l’absence du SOC, le VBM situé à $\mathbf{k} = \Gamma$ est triplement dégénéré (tirets bleus). Chacun de ces états est implicitement dégénéré deux fois en spin, donc la dégénérescence totale est $d = 6$. Le SOC lève en partie cette dégénérescence, $6 \rightarrow 4 + 2$ (lignes pleines rouges). Le VBM est désormais formé des bandes de type trou lourd et trou léger, chacune doublement dégénérée. La bande *split-off*, également doublement dégénérée, est maintenant séparée en énergie du VBM par l’énergie du *split-off*, Δ_{SOC} .

d’inversion. Le VBM est triplement dégénéré, comme on peut voir à la figure 2.4 (tirets bleus) pour le CdTe de structure zinc-blende. En tenant compte de la dégénérescence implicite en spin, la dégénérescence totale du VBM est $d = 6$.

Lorsqu’on introduit le SOC, on doit prendre en compte que le signe de la portion de la fonction d’onde due au spin change de signe lorsqu’on effectue une rotation par 2π . Du point de vue de la théorie des groupes, on dira que l’état de spin se transforme comme une représentation irréductible du groupe $D_{\frac{1}{2}}$. Ainsi, les états électroniques seront décrits par le groupe double, qu’on obtient en prenant le produit direct $\mathcal{G} \otimes D_{\frac{1}{2}}$ [22]. Si la décomposition d’une représentation irréductible du groupe \mathcal{G} dans la base des représentations irréductibles du groupe double contient plus d’un terme, l’inclusion du SOC va provoquer une levée de dégénérescence. Pour le VBM des matériaux cubiques, on observe une levée de dégénérescence partielle (fig. 2.4, lignes pleines rouges) $d = 6 \rightarrow d = 4 + 2$. Le VBM est maintenant formé des bandes de type trou lourd et trou léger, qui sont chacune doublement dégénérées. Les deux autres états électroniques correspondent aux bandes *split-off*, qui sont séparées du VBM par une énergie appelée *énergie du split-off* (*split-off energy*), Δ_{SOC} . L’amplitude de Δ_{SOC} augmente avec la force du SOC.

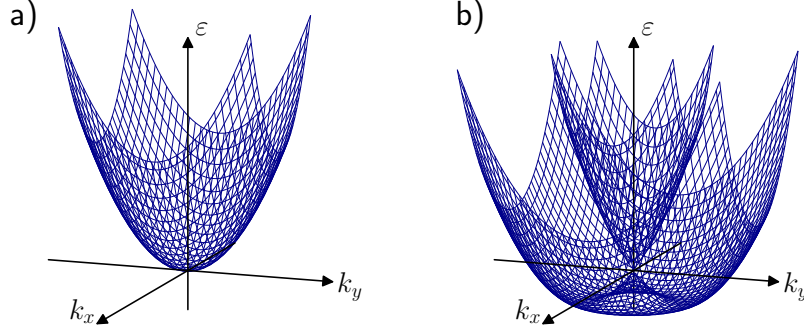


Figure 2.5 – Représentation schématique des bandes de conduction en présence de SOC. a) Les bandes d’un matériau centrosymétrique ont une courbure approximativement parabolique autour d’un extrema bien défini. b) En l’absence de symétrie d’inversion, les extrema sont séparés en impulsion et en énergie par rapport à la position de l’extrema en l’absence de SOC dans l’espace réciproque. Les bandes électroniques demeurent dégénérées au TRIM (*time-reversal invariant momenta*, voir le point de croisement à $k_x = k_y = 0$), comme imposé par le théorème de Kramers.

2.5.4. Interactions Dresselhaus et Rashba dans les matériaux sans symétrie d’inversion

Pour les matériaux sans symétrie d’inversion, le SOC agit donc comme un champ magnétique effectif qui, à l’instar de l’effet Zeeman [6], lève les dégénérescences des différents états de spin pour les valeurs de \mathbf{k} non nulles. D’un point de vue phénoménologique, on peut donc écrire [23, 24]

$$H^{\text{SOC}} = \mathbf{w}(\mathbf{P}) \cdot \boldsymbol{\sigma}, \quad (2.5.16)$$

où $\mathbf{w}(\mathbf{P})$ est un champ effectif dépendant du moment cinétique linéaire qui couple au spin de l’électron. Puisque le SOC préserve la symétrie de renversement du temps, le champ effectif doit être impair par rapport à \mathbf{P} ,

$$\mathbf{w}(-\mathbf{P}) = -\mathbf{w}(\mathbf{P}), \quad (2.5.17)$$

ce qui assure que l’effet du champ n’est observable que dans les systèmes sans symétrie d’inversion [23]. On obtiendra donc un hamiltonien qui dépend linéairement de l’impulsion cristalline \mathbf{k} , qui aura pour conséquence de déplacer la position des extrema des bandes électroniques par rapport à leur position dans l’espace réciproque en l’absence de SOC. Au lieu de bandes approximativement paraboliques autour d’un extrema bien défini, tel qu’on retrouve dans les systèmes centrosymétriques (fig. 2.5a), la dispersion électronique dans un plan 2D de la zone de Brillouin correspond schématiquement à la surface de révolution de paraboles décalées en \mathbf{k} et en énergie par rapport à la position de l’extrema sans SOC

(fig. 2.5b). La dispersion est donc extrême à un plus grand nombre de points qu'en l'absence de SOC. Notons que la symétrie de renversement du temps impose que les bandes électroniques demeurent dégénérées aux points \mathbf{k} invariants sous renversement du temps (TRIM, *time-reversal invariant momenta*). Puisque la séparation des bandes électroniques est intimement liée à leur polarisation en spin, le SOC crée une texture de spin bien définie dans l'espace réciproque, dont l'orientation est corrélée avec le moment cinétique linéaire (*spin-momentum locking*) [23].

La littérature distingue deux origines de l'absence de symétrie d'inversion dans les matériaux. Dans un premier cas, la séparation des bandes par le SOC est induite par l'absence de symétrie d'inversion intrinsèque du matériau, c'est-à-dire qu'elle résulte du groupe de symétrie volumique du cristal. On parle alors de *séparation Dresselhaus*. Dans ce cas, l'hamiltonien d'un système modèle à 2 bandes peut s'écrire comme [23, 24]

$$H_{Dress}^{\text{SOC}} = \gamma_D \left((k_y^2 - k_z^2) k_x \sigma_x + \text{perm.cycliques} \right), \quad (2.5.18)$$

où γ_D est un paramètre qui quantifie la force du SOC.

Dans l'autre cas, la brisure de la symétrie d'inversion tire son origine d'une asymétrie dans le potentiel cristallin le long d'un axe cristallographique particulier, qui joue le rôle d'un champ électrique effectif dans le matériau. Ce type d'interaction est décrit par l'hamiltonien de Rashba [24],

$$H_{Rashba}^{\text{SOC}} = \alpha_R (\mathbf{k} \times \hat{\mathbf{n}}) \cdot \boldsymbol{\sigma}, \quad (2.5.19)$$

avec $\hat{\mathbf{n}}$ la direction cristallographique associée à l'asymétrie du potentiel. La force du couplage est donnée par le paramètre de Rashba, α_R , que l'on peut exprimer en termes de la séparation en impulsion cristalline (k_R) et en énergie (E_R) par rapport au TRIM,

$$\alpha_R = \frac{2E_R}{k_R}. \quad (2.5.20)$$

De par sa nature intrinsèquement axiale, la séparation Rashba est souvent interprétée comme un effet de surface et peut être générée dans des hétérostructures de semi-conducteurs et dans des puits quantiques asymétriques [24]. Ce type d'asymétrie peut néanmoins être présent dans le groupe de symétrie du cristal. On peut penser, par exemple, aux structures wurtzite et trigonale et, de manière générale, aux matériaux quasi-2D dont les différentes couches atomiques sont faiblement liées par l'interaction Van der Waals. On parle alors d'effet Rashba volumique (*bulk Rashba effect*).

Bien que leurs origines physiques soient similaires, les séparations Rashba et Dresselhaus demeurent deux mécanismes d'interaction distincts qui peuvent coexister dans un même matériau. Ainsi, on peut déduire des équations (2.5.18) et (2.5.19) que la séparation Dresselhaus peut a priori être observée dans n'importe quel plan de la zone de Brillouin, alors que la séparation Rashba se limite au plan orthogonal à l'orientation de l'asymétrie du potentiel.

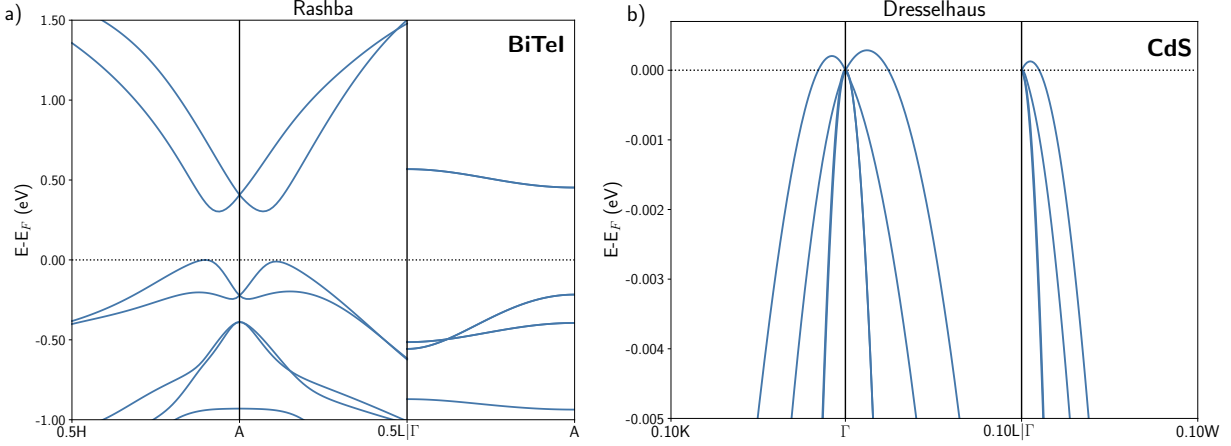


Figure 2.6 – Séparations Rashba et Dresselhaus dans les matériaux non centrosymétriques. a) La séparation Rashba observée dans le BiTeI se limite aux plans de l’espace réciproque orthogonaux à $\hat{\mathbf{k}}_z$ (chemin H-A-L). b) Structure de bande du CdS de structure zinc-blende au voisinage de Γ . La séparation Dresselhaus s’observe pour la grande majorité des directions cristallographiques, mais demeure très faible en amplitude. L’ordre de grandeur de la séparation en énergie et en impulsion par rapport au TRIM diffère significativement de celui caractérisant la séparation Rashba du BiTeI.

Cette différence est illustrée à la figure 2.6a, qui compare la dispersion électronique du BiTeI dans un plan orthogonal à l’axe cristallographique privilégié (chemin H-A-L), ainsi que dans un plan parallèle (chemin Γ -A). En comparaison, la séparation Dresselhaus est présente dans les trois directions cristallographiques montrées à la figure 2.6b pour le CdS cubique.

En comparant les échelles d’impulsion et d’énergie de la figure 2.6, on peut de plus constater que l’ordre de grandeur des séparations Rashba et Dresselhaus est considérablement différent, tant en énergie qu’en impulsion. Cette nuance qualitative provient beaucoup du préjugé favorable de la communauté scientifique envers les matériaux Rashba dont l’amplitude du couplage α_R est grande, puisqu’une séparation prononcée entre les bandes électroniques de spins différents renforce les nombreux effets physiques exotiques qui y sont associés [25]. Par ailleurs, contrairement à la séparation Dresselhaus, l’amplitude de la séparation Rashba peut être contrôlée par l’application d’un champ électrique externe dans la direction cristallographique appropriée [24], qui va renforcer artificiellement l’asymétrie du potentiel axial. Pour cette raison, les matériaux Rashba sont des candidats de prédilection pour les applications de spintronique, où on cherche à utiliser les degrés de liberté des spins pour contrôler les propriétés de transport des électrons. Ainsi, de manière générale, les matériaux Rashba connus présentent une séparation des bandes électroniques plus grande que ce qu’on observe pour la séparation Dresselhaus. Néanmoins, des travaux récents ont révélé une séparation Dresselhaus dans le IrBiSe cubique [26] dont l’ordre de grandeur est similaire à celui typiquement observé dans les matériaux Rashba.

2.6. Conclusion

Dans ce chapitre, nous avons présenté comment l'approximation Born-Oppenheimer rend accessible la modélisation de la structure électronique des cristaux. Nous avons également présenté certaines propriétés des états électroniques pouvant être déduites des différentes symétries qui caractérisent un cristal et discuté des conséquences de l'inclusion de couplage spin-orbite. Notre prochaine étape consistera à aller au-delà de l'approximation des noyaux statiques.

Bibliographie

- [1] M. P. Marder. *Condensed Matter Physics*. John Wiley & Sons, Hoboken, 2nd edition, 2010.
- [2] N. Zettili. *Quantum Mechanics : Concepts and Applications*. John Wiley & Sons, Chichester, U.K., 2nd edition, 2009.
- [3] P. Coleman. *Introduction to Many-Body Physics*. Cambridge University Press, Cambridge, 2015. doi: [10.1017/CBO9781139020916](https://doi.org/10.1017/CBO9781139020916).
- [4] K. Capelle. A bird's-eye view of density-functional theory. *arXiv : cond-mat/0211443*, 2002. doi: [10.48550/arXiv.cond-mat/0211443](https://doi.org/10.48550/arXiv.cond-mat/0211443).
- [5] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College, Philadelphia, 1976.
- [6] D. J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, Upper Saddle River, 2nd edition, 2005.
- [7] P. A. M. Dirac and R. H. Fowler. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 123(792) :714–733, 1929. doi: [10.1098/rspa.1929.0094](https://doi.org/10.1098/rspa.1929.0094).
- [8] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136(3B) :B864–B871, 1964. doi: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864).
- [9] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140(4A) :A1133–A1138, 1965. doi: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133).
- [10] R. M. Martin. *Electronic Structure : Basic Theory and Practical Methods*. Cambridge University Press, Cambridge, 2004. doi: [10.1017/CBO9780511805769](https://doi.org/10.1017/CBO9780511805769).
- [11] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v -representability problem. *Proceedings of the National Academy of Sciences*, 76(12) :6062–6065, 1979. doi: [10.1073/pnas.76.12.6062](https://doi.org/10.1073/pnas.76.12.6062).
- [12] M. Levy. Electron densities in search of Hamiltonians. *Phys. Rev. A*, 26(3) :1200–1208, 1982. doi: [10.1103/PhysRevA.26.1200](https://doi.org/10.1103/PhysRevA.26.1200).

- [13] R. M. Dreizler and E. K. U. Gross. *Density Functional Theory : An Approach to the Many-Body Problem*. Springer-Verlag, Berlin Heidelberg, 1990. doi: [10.1007/978-3-642-86105-5](https://doi.org/10.1007/978-3-642-86105-5).
- [14] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45(23) :13244–13249, 1992. doi: [10.1103/PhysRevB.45.13244](https://doi.org/10.1103/PhysRevB.45.13244).
- [15] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18) :3865–3868, 1996. doi: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [16] M. Shishkin and G. Kresse. Self-consistent GW calculations for semiconductors and insulators. *Physical Review B*, 75(23), 2007. doi: [10.1103/PhysRevB.75.235102](https://doi.org/10.1103/PhysRevB.75.235102).
- [17] F. Rana. Symmetry and Energy Bands : Lecture Notes from ECE407 (Physics of Semiconductors and Nanostructures), Cornell University. https://courses.cit.cornell.edu/ece407/Lectures/handout_symmetries.pdf, 2009.
- [18] B. W. Roberts. Three Myths about Time Reversal in Quantum Theory. *Philos. of Sci.*, 84(2) :315–334, 2017. doi: [10.1086/690721](https://doi.org/10.1086/690721).
- [19] B. W. Roberts. Time Reversal. <http://philsci-archive.pitt.edu/15033/>, 2019. Prepared for the Routledge Handbook of Philosophy of Physics, Eleanor Knox and Alistair Wilson (Eds).
- [20] P. Y. Yu and M. Cardona. *Fundamentals of Semiconductors*. Springer-Verlag, Berlin Heidelberg, 2nd edition, 1999.
- [21] P. Strange. *Relativistic Quantum Mechanics : With Applications in Condensed Matter and Atomic Physics*. Cambridge University Press, Cambridge, 1998. doi: [10.1017/CBO9780511622755](https://doi.org/10.1017/CBO9780511622755).
- [22] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio. *Group theory : application to the physics of condensed matter*. Springer-Verlag, Berlin, 2008.
- [23] A. Manchon, H. C. Koo, J. Nitta, S. M. Frolov, and R. A. Duine. New perspectives for Rashba spin–orbit coupling. *Nature Materials*, 14(9) :871–882, 2015. doi: [10.1038/nmat4360](https://doi.org/10.1038/nmat4360).
- [24] I. Žutić, J. Fabian, and S. Das Sarma. Spintronics : Fundamentals and applications. *Rev. Mod. Phys.*, 76(2) :323–410, 2004. doi: [10.1103/RevModPhys.76.323](https://doi.org/10.1103/RevModPhys.76.323).
- [25] M. S. Bahramy and N. Ogawa. Bulk Rashba Semiconductors and Related Quantum Phenomena. *Advanced Materials*, 29(25) :1605911, 2017. doi: [10.1002/adma.201605911](https://doi.org/10.1002/adma.201605911).
- [26] Z. Liu, S. Thirupathaiah, A. N. Yaresko, S. Kushwaha, Q. Gibson, W. Xia, Y. Guo, D. Shen, R. J. Cava, and S. V. Borisenko. A Giant Bulk-Type Dresselhaus Splitting with 3D Chiral Spin Texture in IrBiSe. *Phys. Status Solidi (RRL) – Rapid Research Letters*, 14(4) :1900684, 2020. doi: [10.1002/pssr.201900684](https://doi.org/10.1002/pssr.201900684).

Chapitre 3

Théorie de l'interaction électron-phonon

Les notions de structure électronique abordées jusqu'à présent reposent entièrement sur la supposition que les noyaux des atomes sont des particules ponctuelles localisées à leur position d'équilibre. Ainsi, bien que cette approximation permette de prédire adéquatement plusieurs propriétés de l'état fondamental des matériaux, elle ne prend pas en compte la nature quantique des noyaux, qui en réalité oscillent en permanence autour de leur position d'équilibre. Puisque la température constitue une mesure de l'amplitude de la vibration des noyaux, les méthodes statiques ne sont donc pas en mesure de décrire la dépendance en température des énergies électroniques. Ces fluctuations quantiques sont par ailleurs présentes même à température nulle. On parle alors du *mouvement du point zéro*. La description de la structure électronique statique donnée par la DFT est donc incomplète.

Par ailleurs, les vibrations atomiques sont responsables de la variation des propriétés optiques des matériaux en fonction de la température. Puisqu'elles modifient l'énergie respective des différents extrema locaux des bandes électroniques, l'énergie requise pour l'absorption ou l'émission d'un photon va changer avec la température. Ainsi, on peut observer un décalage des pics de photoluminescence et une modification du seuil d'absorption et d'émission optique des semi-conducteurs et des isolants. Dans ce dernier cas, l'énergie du seuil est en grande partie déterminée par celle de la bande interdite, qui varie avec la température. Pour obtenir une description complète des propriétés optiques, on devra toutefois adjoindre les effets de nature excitonique, c'est-à-dire l'interaction entre l'électron photoexcité et le trou laissé derrière, ainsi que tenir compte des règles de sélection qui déterminent les transitions optiquement permises. À haute température, on observe typiquement que l'énergie de la bande interdite des semi-conducteurs diminue linéairement avec la température, tel qu'illustré à la figure 3.1 pour le silicium. Ce comportement correspond à celui d'une population de phonons dans le régime classique, décrite par la distribution de Maxwell-Boltzmann. Les effets quantiques se manifestent à basse température : la discrétisation de l'énergie des phonons implique que certains modes ne peuvent plus être excités thermiquement en deçà d'une

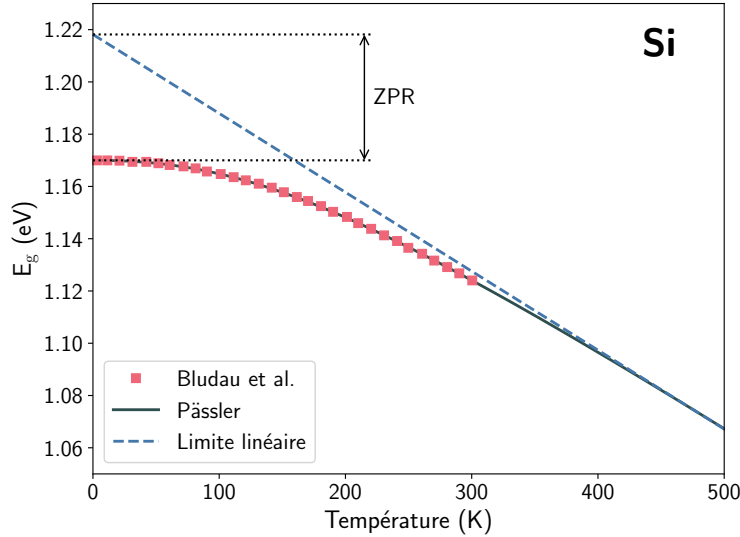


Figure 3.1 – Dépendance en température de l’énergie de la bande interdite du silicium. Les marqueurs carrés rouges correspondent aux données expérimentales de Bludau *et al.* [1]. La ligne pleine grise prolonge les données disponibles à plus haute température à l’aide du modèle de Pässler [2]. La renormalisation du point zéro (ZPR) équivaut à la différence d’énergie entre la valeur expérimentale et l’extrapolation de la limite linéaire (tirets bleus) à $T = 0$ K. Les incertitudes expérimentales sont au plus de l’ordre de la grosseur des marqueurs.

certaine température. L’énergie de la bande interdite dévie alors progressivement du comportement linéaire et atteint un plateau, qui permet de déterminer sa valeur à température nulle.

Du point de vue du calcul numérique, l’approximation des noyaux statiques revient à considérer ces derniers comme des particules classiques. Ainsi, l’énergie de la bande interdite prédite par un calcul statique devrait correspondre à l’extrapolation de la limite linéaire à $T = 0$ K (tirets bleus). La différence d’énergie entre ce résultat et la valeur réelle de la bande interdite à $T = 0$ K est appelée la *renormalisation du point zéro* (ZPR) et capture l’effet du mouvement du point zéro sur les énergies électroniques. Pour les matériaux comportant des noyaux légers, cette correction peut atteindre plusieurs centaines de meV et peut même s’avérer presque aussi importante que les corrections induites par l’interaction électron-trou, comme il sera démontré au chapitre 4. Dans cette optique¹, il est donc crucial de considérer la contribution de l’interaction électron-phonon pour prédire correctement les propriétés optiques des matériaux.

Ce chapitre présente le formalisme permettant de déterminer la dépendance en température des énergies électroniques induite par l’interaction entre les électrons et les vibrations du réseau cristallin. On introduira tout d’abord le formalisme à N -corps, pour ensuite définir

1. Sans vouloir faire de mauvais jeu de mots...

le cristal harmonique et déterminer les contributions à l'ordre le plus bas qui permettent d'évaluer le changement des énergies propres en fonction de la température. On discutera ensuite brièvement de la théorie de la fonctionnelle de la densité perturbative, que nous avons utilisée dans nos travaux, et on présentera enfin le modèle de Fröhlich, qui capture l'interaction non adiabatique entre un électron et un mode de phonon longitudinal optique dans les matériaux polaires.

3.1. Éléments fondamentaux de la théorie à N -corps

Le traitement de l'interaction électron-phonon dépasse le cadre des théories du champ moyen et nécessite de recourir au formalisme des fonctions de Green. La fonction de Green d'une particule va décrire l'amplitude de propagation de cette dernière entre deux coordonnées de l'espace-temps, d'où son appellation courante de *propagateur*. Dans un contexte à plusieurs corps, elle capture l'amplitude de probabilité qu'une particule dans un état initial donné se retrouve dans un certain état final après un certain temps. Ce faisant, elle encode l'ensemble des interactions entre les différentes particules et nous permettra d'évaluer leur effet sur les différents observables du système. Dans le cas qui nous intéresse, la fonction de Green nous informera sur les corrections à appliquer aux énergies électroniques obtenues dans l'approximation des noyaux statique pour prendre en compte les interactions avec les vibrations des noyaux. Dans cette section, on introduira les concepts fondamentaux de la méthode perturbative par fonctions de Green à température nulle. Notre objectif est de favoriser une compréhension conceptuelle de la méthodologie en se basant sur l'approche diagrammatique développée par Richard Feynman, tout en maintenant au minimum les développements mathématiques. Ce formalisme est discuté en détail dans les références [3-5].

3.1.1. Seconde quantification et opérateurs de champ fermioniques

Rappelons-nous tout d'abord que la première quantification de la mécanique consiste à remplacer les observables classiques, tels que l'énergie, la position et l'impulsion, par des opérateurs dont les valeurs propres déterminent les résultats possibles d'une mesure. L'ensemble de l'information sur un système quantique est alors entièrement contenue dans la fonction d'onde, que l'on obtient en résolvant l'équation de Schrödinger. Pour des particules isolées, on peut donc diagonaliser l'hamiltonien et obtenir une base complète de fonctions d'onde qui nous permet d'exprimer n'importe quel état quantique.

Pour un système composé de plusieurs particules identiques et non interagissantes, la fonction d'onde totale peut s'exprimer en termes d'un produit d'états à une particule, qui sera symétrique ou antisymétrique sous permutation de deux particules selon que l'on traite des fermions ou des bosons. Plus N devient grand, plus la notation de ce type de fonction d'onde devient lourde, puisque l'on doit expliciter l'ensemble des permutations possibles des

N particules. Or, la seule information dont nous avons réellement besoin pour caractériser l'état quantique est la répartition des particules dans les différents états quantiques. L'approche par fonction d'onde ne permet également pas de traiter des systèmes pour lesquels le nombre de particules n'est pas conservé.

La seconde quantification consiste donc à introduire des opérateurs qui vont respectivement créer et détruire des particules dans l'espace-temps. En imposant à ces derniers des relations de commutation ou d'anticommutation selon qu'ils agissent sur des bosons ou des fermions, on s'assure que l'état à N -corps respecte implicitement les bonnes propriétés de symétrie ou d'antisymétrie sous permutation de deux particules identiques.

On définit les opérateurs de champ fermioniques, $\hat{\psi}^\dagger(\mathbf{r}\sigma)$ et $\hat{\psi}(\mathbf{r}\sigma)$, dont l'action respectivement crée et détruit un fermion de spin σ au point \mathbf{r} . Pour respecter le principe d'exclusion de Pauli, ces opérateurs doivent respecter les relations d'anticommutation [3]

$$\begin{aligned}\{\hat{\psi}(r), \hat{\psi}^\dagger(r')\} &= \delta(r - r'), \\ \{\hat{\psi}^\dagger(r), \hat{\psi}^\dagger(r')\} &= \{\hat{\psi}(r), \hat{\psi}(r')\} = 0,\end{aligned}\tag{3.1.1}$$

où nous avons introduit la notation compacte $r = (\mathbf{r}\sigma)$. La notation $\delta(x)$ réfère à un delta de Dirac pour les variables continues contenues dans x et un delta de Kroenecker pour les variables discrètes. Ainsi, dans l'équation précédente, $\delta(r - r') = \delta(\mathbf{r} - \mathbf{r}') \delta_{\sigma\sigma'}$.

En choisissant une base complète d'états à une particule, $\{\phi_s(r)\}$, on peut également exprimer ces opérateurs comme²

$$\begin{aligned}\hat{\psi}(r) &= \sum_s \phi_s(r) c_s, \\ \hat{\psi}^\dagger(r) &= \sum_s \phi_s^*(r) c_s^\dagger,\end{aligned}\tag{3.1.2}$$

où l'indice s regroupe l'ensemble des nombres quantiques nécessaires pour caractériser l'état $\phi_s(r)$, typiquement $s = (\mathbf{k}n)$, respectivement le vecteur d'onde et l'indice de bande de l'électron. Les opérateurs c_s^\dagger et c_s créent et annihilent un fermion dans l'état à une particule $\phi_s(r)$. Des relations énoncées à l'équation (3.1.1), on déduit que ces derniers obéissent à

$$\begin{aligned}\{c_s, c_{s'}^\dagger\} &= c_s c_{s'}^\dagger + c_{s'}^\dagger c_s = \delta_{ss'}, \\ \{c_s^\dagger, c_{s'}^\dagger\} &= \{c_s, c_{s'}\} = 0.\end{aligned}\tag{3.1.3}$$

2. Cette décomposition s'obtient en analogie avec la décomposition de la fonction d'onde $\psi(r)$ dans la base complète $\{\phi_s(r)\}$, $\psi(r) = \langle r|\psi\rangle = \sum_s \langle r|s\rangle \langle s|\psi\rangle = \sum_s \phi_s(r) c_s$, où $\phi_s(r)$ est la fonction d'onde de l'état s et c_s est le coefficient de cet état dans la combinaison linéaire. En passant en seconde quantification, $\psi(r)$ et c_s deviennent des opérateurs [4].

La dépendance en temps des opérateurs de création et d'annihilation s'obtient finalement dans la représentation de Heisenberg [4],

$$\begin{aligned} c_s(t) &= e^{i\hat{H}t} c_s e^{-i\hat{H}t}, \\ c_s^\dagger(t) &= e^{i\hat{H}t} c_s^\dagger e^{-i\hat{H}t}, \end{aligned} \quad (3.1.4)$$

où H est l'hamiltonien du problème à N -corps.

3.1.2. Fonction de Green et self-énergie

La quantité fondamentale de la théorie à N -corps est la fonction de Green, qu'on définit comme [3, 4],

$$G(rt, r't') = -i \langle \hat{T} \hat{\psi}(rt) \hat{\psi}^\dagger(r't') \rangle. \quad (3.1.5)$$

La valeur attendue de $\langle \dots \rangle = \langle \Omega_0 | \dots | \Omega_0 \rangle$ est prise sur l'état fondamental du système en interaction. L'opérateur d'ordonnement dans le temps, \hat{T} , a pour effet de permuter les différents opérateurs fermioniques de façon à ce que le temps augmente de la droite vers la gauche :

$$\hat{T} \hat{\psi}(rt) \hat{\psi}^\dagger(r't') = \hat{\psi}(rt) \hat{\psi}^\dagger(r't') \theta(t - t') - \hat{\psi}^\dagger(r't') \hat{\psi}(rt) \theta(t' - t), \quad (3.1.6)$$

où $\theta(t)$ est la fonction de Heavyside. Ainsi, dans l'expression précédente, on peut distinguer l'amplitude de probabilité qu'une particule créée à r' au temps t' soit détruite à r au temps t (terme $\theta(t - t')$) du processus où la destruction de la particule se produit avant sa réapparition (terme $\theta(t' - t)$). Dans le premier cas, on interprétera l'amplitude de probabilité comme la propagation d'un électron et dans le second, celle d'une absence d'électron, c'est-à-dire d'un trou. On note également que pour un système à l'équilibre, la fonction de Green ne dépendra que de la différence entre les temps, $t - t'$. Dans l'approche diagrammatique de Feynman, la fonction de Green interagissante est représentée par une ligne double qui relie les points (r, t) et (r', t') ,

$$G(rt, r't') = \begin{array}{c} \mathbf{r}', t' \\ \xrightarrow{\hspace{1.5cm}} \\ \mathbf{r}, t \end{array}. \quad (3.1.7)$$

Bien que cette définition de la fonction de Green dans l'espace-temps réel soit conceptuellement intuitive, il est généralement beaucoup plus pratique de travailler avec sa représentation dans l'espace de Fourier, puisqu'elle nous permet d'interpréter les amplitudes de probabilité en termes de l'impulsion et de l'énergie de la quasi-particule. Pour un système d'électrons non interagissants, on trouve

$$G^0(rr', \omega) = \sum_{ss'} \phi_s^0(r) \phi_{s'}^{0*}(r') G_{ss'}^0(\omega), \quad (3.1.8a)$$

$$G_{ss'}^0(\omega) = -i \int_{-\infty}^{\infty} d(t - t') e^{i\omega(t-t')} \langle \hat{T} \hat{c}_s(t) \hat{c}_{s'}^\dagger(t') \rangle = \frac{\delta_{ss'}}{\omega - \varepsilon_s^0 + i\eta_s}, \quad (3.1.8b)$$

où ϕ_s^0 est l'état propre de l'hamiltonien sans interactions H^0 ayant une énergie ε_s^0 , alors que $\eta_s = \text{sgn}(\varepsilon_s^0 - E_F)$ est une quantité réelle infinitésimale qui positionne correctement les pôles de la fonction de Green par rapport à l'axe réel pour préserver la causalité³.

Pour distinguer le propagateur non interagissant de la fonction de Green interagissante, on le représentera à l'aide d'une ligne simple,

$$G_{ss'}^0(\omega) = \overset{s'}{\longrightarrow} \overset{\omega}{\longrightarrow} \overset{s}{\longrightarrow} . \quad (3.1.9)$$

Considérons maintenant ce qui change lorsqu'on inclut explicitement les interactions dans la fonction de Green. On voudra trouver une manière de décrire toute la complexité de notre système en interaction à l'aide de la notion d'un quasi-électron dont les propriétés sont affectées par les interactions. Pour ce faire, imaginons conceptuellement un système de N particules auquel on ajoute un électron supplémentaire. La présence de ce dernier va générer une réponse de l'ensemble des particules du système, qui vont modifier leur configuration pour s'ajuster à l'arrivée du nouvel électron. On va considérer cette réponse du système à N -corps comme un « nuage » qui « habille » la quasi-particule. Par exemple, si le nuage est formé d'électrons, la densité électronique au voisinage de l'électron ajouté va diminuer en réponse à la répulsion coulombienne. De la même façon, les noyaux chargés positivement vont vouloir se rapprocher de la charge négative. Le changement de configuration des particules du nuage va ensuite modifier le potentiel ressenti par l'électron, ce qui affecte en retour son énergie et son équation du mouvement. En quelque sorte, on peut considérer que l'électron interagit avec la réponse du système à sa propre présence, à travers le nuage à N -corps.

L'objet mathématique qui décrit l'ensemble de ces interactions porte le nom de *self-énergie*⁴. Pour déterminer son expression analytique, on travaillera dans la représentation d'interaction [3], qui consiste à exprimer l'hamiltonien du système comme

$$H = H^0 + V_I, \quad (3.1.10)$$

où H^0 est un hamiltonien non interagissant dont nous connaissons déjà la solution et V_I est un potentiel d'interaction que l'on considérera comme une petite perturbation par rapport à H^0 . Ce faisant, on pourra factoriser la dépendance temporelle des opérateurs due à H^0 pour se concentrer uniquement sur celle découlant des interactions entre les différentes particules. C'est donc V_I qui fera office de quantité d'intérêt dans notre problème.

3. Ainsi, lorsqu'on évalue la densité électronique en intégrant la fonction de Green en fermant le contour dans le demi-plan supérieur du plan complexe, seuls les pôles correspondant aux états occupés, dont l'énergie est inférieure à l'énergie de Fermi, E_F , contribuent à la densité.

4. On préférera cette forme anglicisée à la traduction plus littérale d'« auto-énergie »

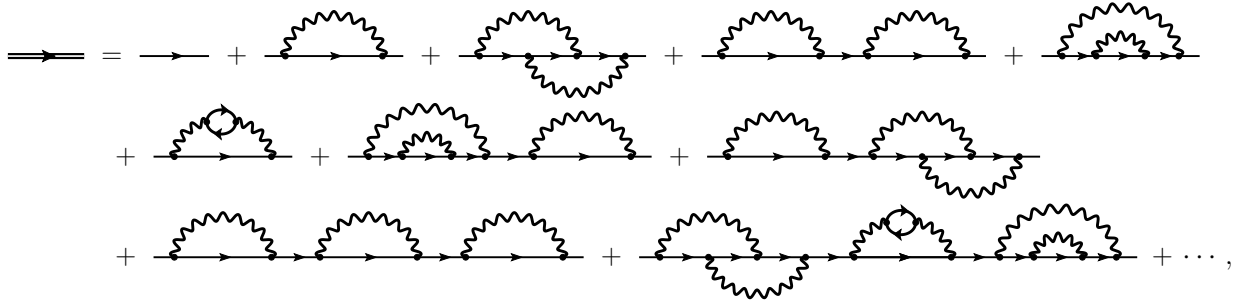
À l'aide du formalisme décrit dans les références [3–5], on pourra exprimer la fonction de Green interagissante comme une série de puissances du potentiel d'interaction,

$$-i \langle \hat{T} \hat{\psi}(rt) \hat{\psi}^\dagger(r't') \rangle_H = \frac{\langle \Omega_0 | \hat{T} S(-\infty, \infty) \hat{\psi}(rt) \hat{\psi}^\dagger(r't') | \Omega_0 \rangle_I}{\langle \Omega_0 | S(-\infty, \infty) | \Omega_0 \rangle_I}, \quad (3.1.11a)$$

$$S(-\infty, \infty) = \hat{T} \left(1 - i \int dt_a V_I(t_a) + \frac{(-i)^2}{2!} \iint dt_a dt_b V_I(t_a) V_I(t_b) + \dots \right), \quad (3.1.11b)$$

où $|\Omega_0\rangle$ est l'état fondamental du système en interaction. Les indices H et I réfèrent respectivement aux représentations de Heisenberg et d'interaction.

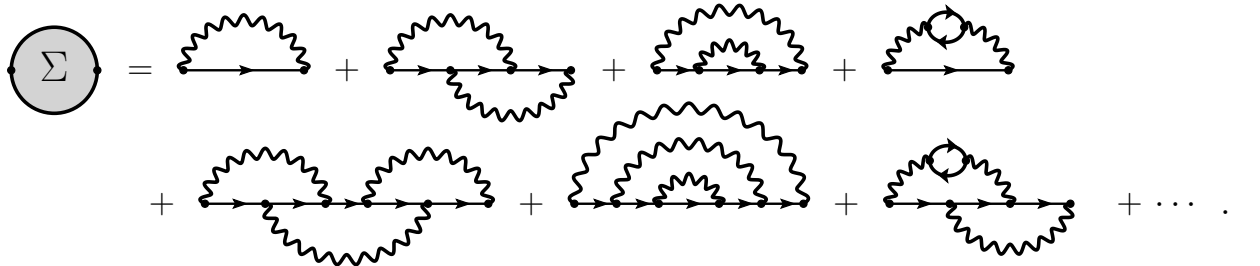
Grâce au théorème de Wick et au théorème des diagrammes liés (*linked cluster theorem*), on peut montrer que la fonction de Green peut être simplifiée par la somme des diagrammes d'interaction connectés⁵,



$$\begin{aligned} \Rightarrow &= \rightarrow + \text{diagram 1} + \text{diagram 2} + \text{diagram 3} + \text{diagram 4} + \text{diagram 5} \\ &+ \text{diagram 6} + \text{diagram 7} + \text{diagram 8} + \text{diagram 9} + \text{diagram 10} \\ &+ \text{diagram 11} + \text{diagram 12} + \text{diagram 13} + \dots, \end{aligned} \quad (3.1.12a)$$

$$V_I = \text{wavy line} . \quad (3.1.12b)$$

En observant l'équation (3.1.12), on peut constater que plusieurs de ces diagrammes sont en fait des séquences formées de processus d'interaction reliés entre eux par des propagateurs libres. On définira donc la self-énergie comme la somme des diagrammes irréductibles⁶,



$$\begin{aligned} \Sigma &= \text{diagram 1} + \text{diagram 2} + \text{diagram 3} + \text{diagram 4} \\ &+ \text{diagram 5} + \text{diagram 6} + \text{diagram 7} + \text{diagram 8} + \dots \end{aligned} \quad (3.1.13)$$

5. Un diagramme connecté est un diagramme où tous les propagateurs sont reliés entre eux.

6. Un diagramme irréductible ne peut pas être séparé en deux en coupant un seul propagateur.

Avec cette définition, on peut réécrire la fonction de Green sous la forme d'une équation de Dyson,

$$\begin{aligned}
\text{---} \text{---} &= \text{---} + \text{---} \circ \Sigma \text{---} + \text{---} \circ \Sigma \circ \Sigma \text{---} \\
&+ \text{---} \circ \Sigma \circ \Sigma \circ \Sigma \text{---} + \dots, \\
\text{---} \text{---} &= \text{---} + \text{---} \circ \Sigma \text{---} \text{---} \text{---},
\end{aligned}
\tag{3.1.14a}$$

c'est-à-dire

$$G_{ss'}(\omega) = G_{ss}^0(\omega)\delta_{ss'} + \sum_{s''} G_{ss}^0(\omega)\Sigma_{ss''}(\omega)G_{s''s'}(\omega). \tag{3.1.14b}$$

Contrairement aux valeurs propres de l'équation de Schrödinger, la self-énergie est une quantité complexe. La partie réelle gouverne la variation des énergies électroniques, tandis que la partie imaginaire va conférer un temps de vie fini aux quasi-particules, $\tau_{\mathbf{k}n} \propto (\Im[\Sigma^{\text{EPI}}])^{-1}$. En pratique, un temps de vie fini implique que les niveaux d'énergie électroniques obtenus avec la théorie des bandes vont acquérir une largeur spectrale, qui est mesurable expérimentalement. L'analogie de la structure de bande sera donné par la fonction spectrale,

$$A(\mathbf{k},\omega) = -\frac{1}{\pi} \sum_n \Im[G_{nn}(\mathbf{k},\omega)]. \tag{3.1.15}$$

Cette dernière, que l'on peut interpréter comme une mesure de la densité d'états disponible à une énergie ω et une impulsion \mathbf{k} données, est proportionnelle à l'intensité du spectre ARPES [6, 7].

À la lumière de notre discussion, le problème de l'interaction électron-phonon consiste donc à obtenir la fonction de Green interagissante des électrons, dans laquelle les interactions entre les électrons et les quantas de vibration du cristal seront capturés par la self-énergie. Ayant cette dernière en main, on pourra ensuite en extraire la renormalisation des énergies électroniques à n'importe quelle température. Pour ce faire, nous devons tout d'abord caractériser l'hamiltonien des noyaux et déterminer l'expression de notre potentiel d'interaction.

3.2. Les phonons dans l'approximation harmonique

3.2.1. Hamiltonien des électrons et des noyaux

Au-delà de l'approximation des noyaux statiques, on doit explicitement inclure l'énergie cinétique et potentielle des noyaux dans l'hamiltonien total du système,

$$H = T^e + T^n + V^{e-e} + V^{n-n} + V^{e-n}, \tag{3.2.1}$$

où les exposants e et n réfèrent respectivement aux électrons et aux noyaux. Les opérateurs T et V correspondent quant à eux à l'énergie cinétique et au potentiel coulombien.

Dans cette équation, seul le terme V^{e-n} couple les coordonnées des électrons et des noyaux. Puisque les noyaux oscillent autour de leur position d'équilibre, on peut écrire leur déplacement au temps t comme

$$\mathbf{R}_j(t) = \mathbf{R}_j^0 + \boldsymbol{\tau}_j(t), \quad (3.2.2)$$

où $\boldsymbol{\tau}_j$ représente le déplacement du noyau j par rapport à sa position d'équilibre \mathbf{R}_j^0 . On peut ensuite approximer V^{e-n} à l'aide d'un développement perturbatif sur les coordonnées atomiques. Pour la coordonnée électronique \mathbf{r}_i , ce développement s'écrit

$$\begin{aligned} V^{e-n}(\mathbf{r}_i) &= \sum_j V^{e-n}(|\mathbf{r}_i - \mathbf{R}_j|) \\ &= \sum_j V^{e-n}(|\mathbf{r}_i - \mathbf{R}_j^0|) - \sum_j \nabla_j V^{e-n}(|\mathbf{r}_i - \mathbf{R}_j^0|) \cdot \boldsymbol{\tau}_j \\ &\quad + \frac{1}{2} \sum_{jj'} \nabla_j \nabla_{j'} V^{e-n}(|\mathbf{r}_i - \mathbf{R}_j^0|, |\mathbf{r}_i - \mathbf{R}_{j'}^0|) \boldsymbol{\tau}_j \boldsymbol{\tau}_{j'} + \mathcal{O}(\boldsymbol{\tau}^3) \\ &= V_{\text{fixe}}^{e-n} + \Delta V^{e-n} + \mathcal{O}(\boldsymbol{\tau}^3). \end{aligned} \quad (3.2.3)$$

Le premier terme de la dernière ligne correspond simplement au potentiel externe des ions dans l'approximation des noyaux fixes, alors que le second, qui regroupe les deux derniers termes de la ligne précédente, capture l'interaction entre les électrons et les déplacements atomiques au deuxième ordre en théorie des perturbations. Avec cette approximation, l'hamiltonien du système peut être séparé en trois contributions,

$$H \approx (T^e + V^{e-e} + V_{\text{fixe}}^{e-n}) + (T^n + V^{n-n}) + \Delta V^{e-n} = H^0 + H^n + H^{\text{EPI}}, \quad (3.2.4)$$

où H^0 est l'hamiltonien électronique dans l'approximation des noyaux statiques, que l'on peut obtenir à partir de la DFT. Les deux autres termes capturent l'effet des vibrations atomiques. Puisque les électrons sont beaucoup plus légers que les ions, leur évolution temporelle est considérablement plus rapide que celle des noyaux. Ceci nous permet de découpler le terme purement électronique de celui dépendant uniquement des coordonnées ioniques, H^n , ainsi que de négliger temporairement le terme décrivant explicitement l'interaction électron-phonon, H^{EPI} , que nous traiterons à la section 3.3.1.

En appliquant l'approximation Born-Oppenheimer sur le système d'électrons et de noyaux découplés, $H^0 + H^n$, on peut écrire la fonction d'onde totale sous forme séparable :

$$\Psi(\{\mathbf{R}_j\}, \{\mathbf{r}_i\}) = \psi^e(\{\mathbf{r}_i\}) \Phi^{\text{ions}}(\{\mathbf{R}_j\}), \quad (3.2.5)$$

dont la solution électronique statique est supposée connue,

$$H^0 \psi^e(\{\mathbf{r}_i\}) = E^{\text{stat}}(\{\mathbf{R}_j\}) \psi^e(\{\mathbf{r}_i\}). \quad (3.2.6)$$

On peut ensuite intégrer sur l'ensemble des coordonnées électroniques pour obtenir l'expression de l'hamiltonien des ions :

$$\begin{aligned} \left(\prod_{i=1}^{N_e} \int dr_i (\psi^e\{\mathbf{r}_i\})^* \right) (H^0 + H^n) \psi^e(\{\mathbf{r}_i\}) \Phi^{\text{ions}}(\{\mathbf{R}_j\}) &= (E^{\text{stat}}(\{\mathbf{R}_j\}) + H^n) \Phi^{\text{ions}}(\{\mathbf{R}_j\}) \\ &= H^{\text{ions}} \Phi^{\text{ions}}(\{\mathbf{R}_j\}). \end{aligned} \quad (3.2.7)$$

La présence de l'énergie électronique statique dans l'expression précédente nous indique que le potentiel coulombien répulsif des noyaux est écranté par la densité électronique, qui constitue, en quelque sorte, une « colle » assurant la cohésion du système cristallin.

3.2.2. Hamiltonien des noyaux dans l'approximation harmonique

Commençons par réécrire les coordonnées des noyaux en termes de leur localisation dans les différentes cellules unitaires,

$$\begin{aligned} \mathbf{R}_j(t) &\rightarrow \mathbf{R}_{l\kappa}(t), \\ R_{l\kappa\alpha}(t) &= R_{l\kappa\alpha}^0 + \tau_{l\kappa\alpha}(t), \end{aligned} \quad (3.2.8)$$

où $R_{l\kappa\alpha}$ représente la position de l'atome κ situé dans la cellule unitaire l dans la direction cartésienne α . Dans l'approximation harmonique, nous supposons que l'énergie totale du cristal varie quadratiquement avec les déplacements atomiques [8],

$$E = E^{\text{stat}} + \frac{1}{2} \sum_{ll'} \sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \tau_{l\kappa\alpha} C_{\kappa\kappa'}^{ll'} \tau_{l'\kappa'\alpha'}. \quad (3.2.9)$$

Le terme de premier ordre du développement en série est nul puisque l'énergie est évaluée à la position d'équilibre des noyaux. E^{stat} est l'énergie totale du cristal dans l'approximation des noyaux fixes. La matrice C correspond quant à elle aux forces interatomiques dans l'espace réel,

$$C_{\kappa\kappa'}^{ll'} = \frac{\partial^2 E}{\partial R_{l\kappa\alpha} \partial R_{l'\kappa'\alpha'}}. \quad (3.2.10)$$

Notons que les expressions (3.2.9) et (3.2.10) couplent a priori le déplacement d'un noyau donné à celui de tous les autres noyaux du cristal, peu importe la cellule unitaire où ceux-ci sont situés. Cette situation diffère d'un modèle purement classique où on interpréterait le cristal comme un ensemble de masses-ressorts qui ne couple explicitement que les déplacements de deux voisins immédiats. Dans ce contexte, deux noyaux éloignés n'interagiraient alors qu'indirectement, à travers le déplacement des noyaux qui se trouvent entre eux.

Classiquement, la force de rappel sur le noyau situé à $\mathbf{R}_{l\kappa}$ est donnée par la loi de Hooke, $-\partial E / \partial \tau_{l\kappa}$, de sorte que les déplacements des noyaux vérifient l'équation du mouvement

$$M_\kappa \ddot{\tau}_{l\kappa\alpha} = - \sum_{l'\kappa'\alpha'} C_{\kappa\kappa'}^{ll'} \tau_{l'\kappa'\alpha'}. \quad (3.2.11)$$

En analogie avec le système classique, on suppose maintenant que le déplacement des différents noyaux du cristal est décrit par une onde plane. On peut donc réécrire les déplacements atomiques en termes des modes propres des phonons,

$$\tau_{l\kappa\alpha}(t) = \frac{1}{\sqrt{N_{\mathbf{q}}}} \sum_{\mathbf{q}\nu} U_{\kappa\alpha}^{\nu}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{R}_l} \Re \left[e^{-i\omega_{\mathbf{q}\nu}t} \right]. \quad (3.2.12)$$

Dans cette expression, \mathbf{q} et ν réfèrent respectivement au vecteur d'onde et à l'indice de branche du mode de phonon. Ce dernier permet d'identifier les $3N_{\text{at}}$ différents modes de vibration associés à un même vecteur d'onde, avec N_{at} le nombre de noyaux dans la cellule unitaire. La somme sur les $N_{\mathbf{q}}$ vecteurs d'onde \mathbf{q} couvre l'ensemble de la zone de Brillouin. Le vecteur $\mathbf{U}_{\kappa}^{\nu}$, appelé vecteur de polarisation, décrit le déplacement de l'atome κ dans le mode phononique $\mathbf{q}\nu$, alors que $\omega_{\mathbf{q}\nu}$ est la fréquence d'oscillation de ce mode.

Avec cette définition, l'équation (3.2.11) devient

$$M_{\kappa}\omega_{\mathbf{q}\nu}^2 U_{\kappa\alpha}^{\nu}(\mathbf{q}) = \sum_{\kappa'\alpha'} \tilde{C}_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q}) U_{\kappa'\alpha'}^{\nu}(\mathbf{q}), \quad (3.2.13)$$

où $\tilde{C}_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q})$ est la transformée de Fourier de la matrice des forces interatomiques. On peut ramener cette équation à un problème de valeurs propres conventionnel en absorbant les termes de masse dans la matrice des forces et les vecteurs de polarisation :

$$D_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q}) = \frac{1}{\sqrt{M_{\kappa}}} \tilde{C}_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q}) \frac{1}{\sqrt{M_{\kappa'}}}, \quad (3.2.14a)$$

$$\xi_{\kappa\alpha}^{\nu}(\mathbf{q}) = \sqrt{M_{\kappa}} U_{\kappa\alpha}^{\nu}(\mathbf{q}), \quad (3.2.14b)$$

où la matrice \mathbf{D} porte le nom de matrice dynamique. Les vecteurs $\xi_{\kappa}^{\nu}(\mathbf{q})$ sont les vecteurs propres des modes de phonons, que l'on obtient en résolvant l'équation dynamique,

$$\omega_{\mathbf{q}\nu}^2 \xi_{\kappa\alpha}^{\nu}(\mathbf{q}) = \sum_{\kappa'\alpha'} D_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q}) \xi_{\kappa'\alpha'}^{\nu}(\mathbf{q}). \quad (3.2.15)$$

L'orthonormalité des vecteurs propres, $\{\xi_{\kappa}^{\nu}(\mathbf{q})\}$, impose enfin la relation

$$\sum_{\kappa\alpha} M_{\kappa} (U_{\kappa\alpha}^{\nu}(\mathbf{q}))^* U_{\kappa\alpha}^{\nu'}(\mathbf{q}) = \delta_{\nu\nu'}. \quad (3.2.16)$$

On veut maintenant quantifier l'hamiltonien des noyaux,

$$H^{\text{ions}} = \sum_{l\kappa} \frac{\hat{\mathbf{P}}_{l\kappa} \cdot \hat{\mathbf{P}}_{l\kappa}}{2M_{\kappa}} + \sum_{\substack{l'l' \\ \kappa\kappa'}} \hat{\mathbf{t}}_{l\kappa} \cdot \mathbf{C}_{\kappa\kappa'}^{ll'} \cdot \hat{\mathbf{t}}_{l'\kappa'}. \quad (3.2.17)$$

L'indice α a été absorbé dans la notation vectorielle, à l'instar de la première ligne de l'équation (3.2.8).

On définit les opérateurs de déplacement et de position de manière analogue à l'oscillateur harmonique quantique unidimensionnel [3, 6, 9],

$$\hat{\tau}_{l\kappa}(t) = \sum_{\mathbf{q}\nu} \left(\frac{1}{2N_{\mathbf{q}}\omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} \mathbf{U}_{\kappa}^{\nu}(\mathbf{q}) \left(a_{\mathbf{q}\nu} e^{-i\omega_{\mathbf{q}\nu}t} + a_{-\mathbf{q}\nu}^{\dagger} e^{i\omega_{\mathbf{q}\nu}t} \right) e^{i\mathbf{q}\cdot\mathbf{R}_l}, \quad (3.2.18a)$$

$$\hat{\mathbf{P}}_{l\kappa}(t) = \sum_{\mathbf{q}\nu} -iM_{\kappa} \left(\frac{\omega_{\mathbf{q}\nu}}{2N_{\mathbf{q}}} \right)^{\frac{1}{2}} \mathbf{U}_{\kappa}^{\nu}(\mathbf{q}) \left(a_{\mathbf{q}\nu} e^{-i\omega_{\mathbf{q}\nu}t} - a_{-\mathbf{q}\nu}^{\dagger} e^{i\omega_{\mathbf{q}\nu}t} \right) e^{i\mathbf{q}\cdot\mathbf{R}_l}, \quad (3.2.18b)$$

où nous avons introduit les opérateurs de création et d'annihilation des phonons, $a_{\mathbf{q}\nu}^{\dagger}$ et $a_{\mathbf{q}\nu}$, qui respectivement créent et détruisent un quanta de vibration (un phonon) ayant un vecteur d'onde \mathbf{q} et une polarisation décrite par le vecteur $\mathbf{U}_{\kappa}^{\nu}(\mathbf{q})$. Ces opérateurs obéissent aux relations de commutation

$$\begin{aligned} [a_{\mathbf{q}\nu}, a_{\mathbf{q}'\nu'}^{\dagger}] &= \delta(\mathbf{q} - \mathbf{q}')\delta_{\nu\nu'}, \\ [a_{\mathbf{q}\nu}, a_{\mathbf{q}'\nu'}] &= [a_{\mathbf{q}\nu}^{\dagger}, a_{\mathbf{q}'\nu'}^{\dagger}] = 0, \end{aligned} \quad (3.2.19)$$

de sorte que la relation de commutation canonique des opérateurs de position et d'impulsion est vérifiée, $[\hat{\tau}_{l\kappa}, \hat{\mathbf{P}}_{l'\kappa'}] = i\delta_{ll'}\delta_{\kappa\kappa'}$.

On utilise également la convention de Born et Huang [10],

$$(\mathbf{U}_{\kappa}^{\nu}(\mathbf{q}))^* = \mathbf{U}_{\kappa}^{\nu}(-\mathbf{q}), \quad (3.2.20)$$

qui nous assure que H^{ions} soit hermitien.

En substituant les équations (3.2.18a) et (3.2.18b) dans l'équation (3.2.17) et en utilisant la relation d'orthonormalité des vecteurs de polarisation (éq. (3.2.16)), on obtient un hamiltonien décrivant un ensemble de $3N_{\text{at}} \times N_{\mathbf{q}}$ oscillateurs harmoniques indépendants,

$$H^{\text{ions}} = \sum_{\mathbf{q}\nu} \omega_{\mathbf{q}\nu} \left(a_{\mathbf{q}\nu}^{\dagger} a_{\mathbf{q}\nu} + \frac{1}{2} \right), \quad (3.2.21)$$

où

$$a_{\mathbf{q}\nu}^{\dagger} a_{\mathbf{q}\nu} = n_{\mathbf{q}\nu} \quad (3.2.22)$$

est l'opérateur de nombre phononique, dont l'action sur un état quantique compte le nombre de phonons dans le mode $\mathbf{q}\nu$. L'état fondamental des noyaux prend donc la forme d'un produit de fonctions gaussiennes [6]. Les différents états excités peuvent ensuite être créés par l'application successive des opérateurs $a_{\mathbf{q}\nu}$ et $a_{\mathbf{q}\nu}^{\dagger}$, de façon complètement analogue à celle dont sont construits les états propres de l'oscillateur harmonique quantique unidimensionnel abordé dans les cours d'introduction à la mécanique quantique [9].

3.2.3. Effets anharmoniques

Dans la section précédente, nous avons tronqué le développement de l'énergie totale autour de sa configuration à l'équilibre au deuxième ordre, ce qui nous a permis de construire la

matrice dynamique, que nous avons diagonalisée pour obtenir les fréquences et les vecteurs de polarisation des phonons, pour enfin réécrire l’hamiltonien des ions comme une somme d’oscillateurs harmoniques indépendants. Par conséquent, l’approximation harmonique suppose implicitement que les phonons n’interagissent pas entre eux et se propagent sans restriction dans le cristal, ce qui implique qu’ils ont un temps de vie infini.

Ainsi, dans l’approximation harmonique, il est tout aussi coûteux en énergie d’augmenter que de réduire la distance relative entre deux noyaux. Si cette supposition se justifie pour de petits déplacements autour de la configuration à l’équilibre, elle ne rend pas compte de la forte répulsion coulombienne qui se manifestera si l’on essaie de rapprocher deux noyaux très près l’un de l’autre. À l’opposé, la répulsion coulombienne sera beaucoup plus faible si l’on augmente la distance relative entre les noyaux par la même proportion. Par conséquent, la courbure de l’énergie libre du cristal sera plus prononcée lorsque la distance moyenne entre les atomes diminue par rapport à la configuration à l’équilibre que lorsqu’elle augmente.

La modélisation de cette asymétrie nécessite d’introduire des termes d’ordre supérieur dans le développement de l’énergie en termes des déplacements atomiques (éq. (3.2.9)), qui dépendront cette fois-ci d’au moins trois déplacements atomiques [11]⁷. Dans ce contexte, les fréquences des phonons obtenues dans l’approximation harmonique seront perturbées par une self-énergie d’interaction phonon-phonon [12] qui, comme discuté à la section 3.1.2, va à la fois modifier les fréquences des phonons et conférer à ces derniers un temps de vie fini, qui se manifeste par une largeur de bande qui est mesurable expérimentalement par spectroscopie de diffraction de neutrons [11, 13]. Ce faisant, les phonons non harmoniques peuvent désormais échanger de l’énergie entre eux, tant par des processus de fusion que de désintégration. Puisque ces processus de diffusion dépendent de l’occupation des différents modes phononiques, la correction anharmonique provenant de la self-énergie phonon-phonon variera avec la température. Une autre conséquence de l’anharmonicité de l’énergie du cristal est la dilatation thermique, qui sera traitée plus en détail au chapitre 5.

Certains travaux ont proposé des généralisations du formalisme présenté à la section 3.2.2 qui incluent explicitement des contributions anharmoniques, entre autres pour obtenir la dépendance en température du spectre des phonons anharmoniques [14] et ainsi que la conductivité thermique intrinsèque [13, 15]. Ces approches requièrent cependant de recourir à la théorie à N -corps et deviennent donc aussi coûteuses numériquement que l’évaluation de l’interaction électron-phonon qui sera abordée à la prochaine section. Dans le cadre de cette thèse, nous nous intéresserons spécifiquement à l’impact de la dilatation thermique sur l’énergie de la bande interdite. Bien que l’anharmonicité soit nécessaire pour expliquer l’origine physique de ce mécanisme, nous verrons au chapitre 5 qu’il existe une autre méthode qui permette d’évaluer l’effet au premier ordre de la dilatation thermique sans devoir utiliser l’approche à N -corps.

7. Les approches anharmoniques se limitent habituellement aux termes cubiques et quartiques.

3.3. Dépendance en température des énergies électroniques

Ayant désormais en main la solution à l'hamiltonien des électrons non interagissants, H^0 , ainsi que celle de l'hamiltonien des noyaux, H^{ions} , nous sommes en mesure d'évaluer la dépendance en température des énergies électroniques pour notre système d'électrons et de noyaux en interaction. Comme discuté à la section 3.1.2, cette dernière nous sera donnée par la partie réelle de la self-énergie entrant dans l'expression de la fonction de Green du système en interaction. Rappelons qu'à $T = 0$ K, la fonction de Green interagissante permet d'obtenir une approximation de l'état fondamental du système couplé, qui est bien défini. À température finie⁸, la situation n'est a priori pas aussi limpide : le système fluctue en permanence entre les différentes configurations accessibles thermiquement, incluant l'état fondamental et les états excités. Nous pouvons néanmoins recourir au fait que la température est une mesure de ces fluctuations thermiques, et donc de l'énergie moyenne du système, qui est la quantité physique qui nous intéresse. Nous pourrions dès lors évaluer simplement la moyenne thermodynamique sur l'ensemble des configurations possibles [4]. Dans la représentation d'Heisenberg, notre fonction de Green prendra alors la forme [3]

$$G_{ss'}(t - t') = \frac{\text{Tr} \left[e^{-\beta H} c_s(t) c_{s'}^\dagger(t') \right]}{\text{Tr} \left[e^{-\beta H} \right]}, \quad (3.3.1)$$

où $\beta = (k_B T)^{-1}$, k_B est la constante de Boltzmann, H est l'hamiltonien du système d'électrons et de phonons en interaction donné à l'équation (3.2.4) et où le symbole Tr dénote la trace de l'opérateur entre crochets. Nous avons de plus réintroduit la notation compacte $s = (\mathbf{k}n)$ pour identifier les nombres quantiques des électrons. La dépendance en temps des opérateurs fermioniques $c_s(t)$ et $c_s^\dagger(t)$ est donnée par l'équation (3.1.4).

3.3.1. Self-énergie de l'interaction électron-phonon

Nous allons donc construire notre hamiltonien en combinant les termes d'origine purement électronique et ionique au terme de couplage explicite présents dans l'équation (3.2.4)⁹,

$$\begin{aligned} H &= H^0 + H^{\text{ions}} + H^{\text{EPI}} \\ &= H^0 + H^{\text{ions}} - \sum_{i=1}^{N_e} \sum_{l\kappa} \nabla_{l\kappa} V_{e-n}^0(\mathbf{r}_i) \cdot \boldsymbol{\tau}_{l\kappa} + \frac{1}{2} \sum_{i=1}^{N_e} \sum_{\substack{l\kappa \\ l'\kappa'}} \nabla_{l\kappa} \nabla_{l'\kappa'} V_{e-n}^0(\mathbf{r}_i) \boldsymbol{\tau}_{l\kappa} \boldsymbol{\tau}_{l'\kappa'}, \end{aligned} \quad (3.3.2)$$

8. On entend évidemment ici « température non nulle ».

9. Nous reconnaissons que l'hamiltonien écrit sous cette forme est légèrement différent de celui énoncé à l'équation (3.2.4), puisque la définition de l'hamiltonien des ions obtenu à l'équation (3.2.7) ajoute l'énergie totale du système d'électrons dans l'approximation du cristal statique à l'hamiltonien impliquant seulement les noyaux. Cette subtilité, qui découle de l'approximation Born-Oppenheimer, est analysée en profondeur dans la thèse de doctorat de Paul Boulanger [16].

que nous avons réécrit avec la notation introduite à la section 3.2.2. On traitera les termes d'interaction comme une perturbation par rapport à $H^0 + H^{\text{ions}}$ dans la fonction de Green.

Réécrivons tout d'abord l'expression précédente en termes des opérateurs de création et d'annihilation des électrons et des phonons,

$$H = \sum_s \varepsilon_s^0 c_s^\dagger c_s + \sum_\lambda \omega_\lambda \left(a_\lambda^\dagger a_\lambda + \frac{1}{2} \right) + \sum_{ss'\lambda} g_{ss'\lambda}^{(1)} c_s^\dagger c_{s'} \hat{\phi}_\lambda + \frac{1}{2} \sum_{ss'\lambda\lambda'} g_{ss'\lambda\lambda'}^{(2)} c_s^\dagger c_{s'} \hat{\phi}_\lambda \hat{\phi}_{\lambda'}, \quad (3.3.3)$$

où nous avons introduit la notation $\lambda = (\mathbf{q}\nu)$ pour les nombres quantiques des phonons ainsi qu'un opérateur de champ phononique¹⁰,

$$\begin{aligned} \hat{\phi}_\lambda &= a_\lambda + a_{-\lambda}^\dagger \\ &= a_{\mathbf{q}\nu} + a_{-\mathbf{q}\nu}^\dagger. \end{aligned} \quad (3.3.4)$$

Les termes $g^{(1)}$ et $g^{(2)}$ décrivent quant à eux la force du couplage électron-phonon pour les deux ordres d'interaction,

$$g_{ss'\lambda}^{(1)} = - \langle s | \nabla_\lambda V^0 | s' \rangle, \quad (3.3.5a)$$

$$g_{ss'\lambda\lambda'}^{(2)} = \langle s | \nabla_\lambda \nabla_{\lambda'} V^0 | s' \rangle, \quad (3.3.5b)$$

où l'opérateur ∇_λ exprimé dans la base des positions est

$$\nabla_\lambda = \left(\frac{1}{2N_{\mathbf{q}}\omega_\lambda} \right)^{\frac{1}{2}} \sum_{\kappa\alpha} U_{\kappa\alpha}(\lambda) \sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial}{\partial R_{l\kappa\alpha}}. \quad (3.3.6)$$

Rappelons que α est une direction cartésienne et que la dépendance sur le vecteur d'onde \mathbf{q} des phonons est absorbée dans λ . Notons finalement que, dans les équations (3.3.3) et (3.3.4), la dépendance temporelle est implicitement contenue dans les opérateurs de création et d'annihilation.

À un ingrédient près, nous avons maintenant en main tous les éléments nécessaires pour évaluer la fonction de Green à température finie. Une subtilité importante par rapport à la section 3.1.2 est que nous devons recourir aux fréquences de Matsubara [3–5] et définir la fonction de Green à température finie en termes d'un temps imaginaire, $it \rightarrow \tau$. La plage sur laquelle τ varie dépendra de la température. On obtient alors

$$G_{ss'}(\tau - \tau') = - \langle \hat{T}_\tau c_s(\tau) c_{s'}^\dagger(\tau') \rangle, \quad (3.3.7)$$

où nous avons remplacé l'opérateur d'ordonnement temporel \hat{T} par son équivalent \hat{T}_τ , qui agit sur la fonction de Green de Matsubara. Les détails techniques de ce formalisme sont explicités dans les références [3–5].

10. Plus exactement, l'opérateur de champ qui agit dans l'espace réel prend la forme $\hat{\phi}(\mathbf{r}) = \sum_\lambda (a_\lambda + a_{-\lambda}^\dagger) e^{i\mathbf{q}\cdot\mathbf{r}}$. L'opérateur que nous avons défini à l'équation (3.3.4) correspond alors à la composante de Fourier λ de $\hat{\phi}(\mathbf{r})$.

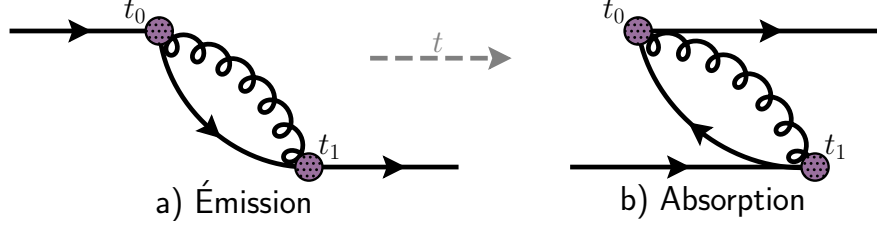


Figure 3.2 – Processus d’émission et d’absorption capturés par la self-énergie de Fan. Dans l’espace temporel, la flèche en tirets gris représente la direction de propagation du temps. a) Émission spontanée d’un phonon par un électron suivie de sa réabsorption. b) L’électron incident absorbe un phonon préalablement émis. L’état électronique intermédiaire est alors de type trou, d’où sa propagation « à rebours » dans le temps.

Il ne nous reste plus qu’à définir le propagateur des phonons indépendants, $D_{\lambda\lambda}$, qui entrera dans la représentation diagrammatique de la self-énergie. Ce dernier s’obtient de manière analogue au propagateur des électrons non interagissants [3],

$$D_{\lambda\lambda}(\tau - \tau') = - \langle \hat{T}_\tau \hat{\phi}_\lambda(\tau) \hat{\phi}_{-\lambda}(\tau') \rangle = \text{~~~~~} . \quad (3.3.8)$$

Une fois ramenée à l’espace de Fourier réel par continuation analytique, la fonction de Green électronique du système en interaction sera décrite par une équation de Dyson,

$$G_{ss'}(\omega, T) = G_{ss}^0 \delta_{ss'} + \sum_{s''} G_{ss}^0 \Sigma_{ss''}^{\text{EPI}}(\omega, T) G_{s''s'}(\omega, T), \quad (3.3.9)$$

où G_{ss}^0 est le propagateur des électrons non interagissants, que l’on obtient à partir de H^0 (éq. (3.1.8b)).

Le développement en série de la self-énergie de l’interaction électron-phonon tronquée au plus bas ordre non trivial en théorie des perturbations contient deux termes,

$$\Sigma_{ss''}^{\text{EPI}}(\omega, T) = \Sigma_{ss''}^{\text{Fan}}(\omega, T) + \Sigma_{ss''}^{\text{DW}}(T),$$

$$\Sigma^{\text{EPI}} = \text{~~~~~} + \text{~~~~~} . \quad (3.3.10)$$

Avant d’expliciter leur formulation mathématique, arrêtons-nous un instant pour apprécier toute l’information physique contenue dans ces diagrammes. Le premier terme, appelé terme de Fan, comporte deux vertex d’interaction d’ordre 1 (cercles violets à pois), reliés entre eux par un propagateur électronique et un propagateur phononique. On peut d’ores et déjà déduire que ce terme décrit un processus d’émission spontanée d’un phonon, qui sera ensuite réabsorbé par l’électron. Le processus inverse, soit l’absorption d’un phonon préalablement émis suivi de sa réémission, serait représenté visuellement en inversant à quels vertex sont connectés les propagateurs électroniques initiaux et finaux, comme on peut voir

à la figure 3.2. Nous verrons plus tard que la formulation mathématique de Σ^{Fan} capture naturellement ces deux processus équivalents. Par ailleurs, la conservation de l'impulsion impose que les états électroniques initiaux et finaux de l'équation (3.3.9) vérifient $s = (\mathbf{k}n)$ et $s' = (\mathbf{k}n')$. La périodicité de la variation du potentiel induite par le phonon implique de plus que $s'' = (\mathbf{k}''n'') = (\mathbf{k} + \mathbf{q}n'')$. La dépendance en fréquence de l'état électronique intermédiaire est enfin une conséquence de la nature dynamique de l'interaction, c'est-à-dire du temps nécessaire à l'électron pour passer d'un vertex à l'autre dans l'espace réel.

Le second terme, ou terme de Debye-Waller, comporte quant à lui un seul vertex d'interaction d'ordre 2 (cercle jaune clair rayé), la boucle du propagateur phononique se refermant sur elle-même. Cette forme de diagramme est généralement utilisée pour décrire une densité. Comme il n'y a aucun état électronique intermédiaire, on peut en déduire que ce terme capture la réponse de l'électron au changement du potentiel externe causé par les fluctuations de la densité de charge des noyaux autour de leur position d'équilibre à une température donnée [6]. En d'autres mots, la périodicité du potentiel atomique devient plus floue à haute température, ce qui affecte en retour l'énergie des électrons. C'est le type de mécanisme qui permet d'expliquer la perte d'intensité et l'élargissement spectral des pics de Bragg observés dans les spectres de diffraction des rayons X et de diffraction de neutrons [17].

Ces deux contributions sont connues et étudiées depuis les années 1950, grâce aux travaux respectifs de Fan [18, 19] et Antončík [20]. Néanmoins, elles ont longtemps été considérées comme deux approches équivalentes pour le calcul de l'interaction électron-phonon. Il faudra attendre les travaux de Allen, Heine et Cardona au début des années 1980 [21–23] pour réaliser que ces deux contributions sont en fait complémentaires et doivent être traitées sur un même pied d'égalité. Ainsi, la self-énergie décrite par l'équation (3.3.10) est aujourd'hui connue sous le nom de théorie d'Allen-Heine-Cardona (AHC) et constitue l'une des méthodes de prédilection pour évaluer la dépendance en température de la structure électronique.

Retournons maintenant à l'équation (3.3.10). L'expression du terme de Fan est

$$\Sigma_{ss'}^{\text{Fan}}(\omega, T) = \sum_{s''\lambda} g_{ss''\lambda}^{(1)} g_{s''s'\lambda}^{(1)} \left[\underbrace{\frac{n_\lambda(T) + 1 - f_{s''}(T)}{\omega - \varepsilon_{s''}^0 - \omega_\lambda + i\eta_\omega}}_{\text{émission}} + \underbrace{\frac{n_\lambda(T) + f_{s''}(T)}{\omega - \varepsilon_{s''}^0 + \omega_\lambda + i\eta_\omega}}_{\text{absorption}} \right], \quad (3.3.11)$$

où $n_\lambda(T)$ et $f_s(T)$ sont respectivement les fonctions d'occupation Bose-Einstein et Fermi-Dirac des phonons et des électrons, ε_s^0 correspond aux énergies électroniques non interagissantes, et $\eta_\omega = |\eta| \text{sgn}(\omega - E_F)$ est une quantité réelle infinitésimale qui assure que la localisation des pôles de la fonction de Green dans le plan complexe respecte la causalité, tout comme nous avons fait pour le propagateur libre des électrons. Les deux termes entre crochets décrivent respectivement les processus d'émission et d'absorption des phonons.

Par ailleurs, nous avons déjà soulevé que les états électroniques initiaux et finaux possèdent la même impulsion, sans restriction sur les indices de bande. A priori, la self-énergie de Fan peut donc promouvoir des électrons dans un état excité, ce qui aurait pour conséquence de modifier la densité électronique par rapport au système sans interaction [24]. Dans la mesure où l'énergie de la bande interdite est beaucoup plus grande que celle des phonons, ce qui est en général le cas pour les semi-conducteurs et les isolants à pression ambiante, on pourra justifier d'évaluer seulement les éléments diagonaux de $\Sigma_{ss'}^{\text{Fan}}$,

$$\Sigma_{ss'}^{\text{Fan}}(\omega) \approx \Sigma_{ss'}^{\text{Fan}}(\omega)\delta_{ss'}. \quad (3.3.12)$$

Dans le cas contraire, l'évaluation de la dépendance en température des états électroniques nécessitera de diagonaliser $\Sigma_{ss'}^{\text{Fan}}(\omega)$ pour le sous-espace des bandes pour lesquelles de telles excitations électroniques sont statistiquement possibles.

Dans l'approximation diagonale, le terme de Fan devient finalement

$$\begin{aligned} \Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega, T) &= \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu}^{\text{BZ}} \sum_{n'} |\langle \mathbf{k} + \mathbf{q}n' | \nabla_{\mathbf{q}\nu} V^{\text{KS}} | \mathbf{k}n \rangle|^2 \\ &\times \left[\frac{n_{\mathbf{q}\nu}(T) + 1 - f_{\mathbf{k}+\mathbf{q}n'}(T)}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 - \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} + \frac{n_{\mathbf{q}\nu}(T) + f_{\mathbf{k}+\mathbf{q}n'}(T)}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 + \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} \right], \end{aligned} \quad (3.3.13)$$

où nous avons explicité les différents nombres quantiques contenus dans s et λ . L'expression de l'opérateur $\nabla_{\mathbf{q}\nu}$ découle de l'équation (3.3.6) :

$$\nabla_{\mathbf{q}\nu} = \frac{1}{\sqrt{2\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha} U_{\kappa\alpha}^{\nu}(\mathbf{q}) \sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial}{\partial R_{l\kappa\alpha}} = \frac{1}{\sqrt{2\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha} U_{\kappa\alpha}^{\nu}(\mathbf{q}) \partial_{\kappa\alpha}(\mathbf{q}). \quad (3.3.14)$$

Rappelons que le préfacteur $1/N_{\mathbf{q}}$ vient pondérer la somme sur les états phononiques par le nombre de vecteurs d'onde \mathbf{q} considérés. La limite $N_{\mathbf{q}} \rightarrow \infty$ est implicite.

Le terme de Debye-Waller s'écrit quant à lui

$$\Sigma_{\mathbf{k}n}^{\text{DW}}(T) = \frac{1}{2} \sum_{\mathbf{q}\nu} g_{\mathbf{k}n}^{(2)} \delta_{\mathbf{q}', -\mathbf{q}} (2n_{\mathbf{q}\nu}(T) + 1), \quad (3.3.15)$$

où nous avons appliqué l'approximation diagonale comme pour le terme de Fan. Notons que la dépendance en température est entièrement contenue dans la distribution de Bose-Einstein des phonons.

Pour obtenir la renormalisation des énergies électroniques, il nous reste à déterminer comment traiter la dépendance en fréquence de la self-énergie de Fan. Nous devrions, en principe, évaluer sa partie réelle à la fréquence ω correspondant à l'énergie renormalisée de la quasi-particule, c'est-à-dire solutionner l'équation implicite

$$\begin{aligned} \varepsilon_{\mathbf{k}n}(T) &= \varepsilon_{\mathbf{k}n}^0 + \Re \left[\Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega = \varepsilon_{\mathbf{k}n}(T), T) + \Sigma_{\mathbf{k}n}^{\text{DW}}(T) \right] \\ &= \varepsilon_{\mathbf{k}n}^0 + \Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}(T), T) \right]. \end{aligned} \quad (3.3.16)$$

En pratique, il est raisonnable de supposer que l'énergie des quasi-particules est très proche de leurs énergies statiques. On peut alors appliquer l'approximation *on-the-mass-shell*, qui consiste à évaluer ω aux pôles de la fonction de Green, qui correspondent aux énergies électroniques non perturbées [25],

$$\varepsilon_{\mathbf{k}n}(T) \simeq \varepsilon_{\mathbf{k}n}^0 + \Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}^0, T) \right]. \quad (3.3.17)$$

Une autre approximation couramment utilisée consiste à faire le développement en série au premier ordre de la partie réelle de la self-énergie autour de l'énergie non perturbée,

$$\Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}(T), T) \right] \simeq \Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}^0, T) \right] + \Re \left[\left. \frac{\partial \Sigma_{\mathbf{k}n}^{\text{EPI}}}{\partial \omega} \right|_{\omega = \varepsilon_{\mathbf{k}n}^0} \right] (\varepsilon_{\mathbf{k}n}(T) - \varepsilon_{\mathbf{k}n}^0). \quad (3.3.18)$$

Ce faisant, on introduit à l'équation (3.3.17) un facteur de renormalisation [3],

$$\varepsilon_{\mathbf{k}n}(T) \simeq \varepsilon_{\mathbf{k}n}^0 + Z_{\mathbf{k}n} \Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}^0, T) \right], \quad (3.3.19)$$

$$Z_{\mathbf{k}n} = \left(1 - \Re \left[\left. \frac{\partial \Sigma_{\mathbf{k}n}^{\text{EPI}}}{\partial \omega} \right|_{\omega = \varepsilon_{\mathbf{k}n}^0} \right] \right)^{-1}.$$

Le facteur $Z_{\mathbf{k}n}$ est interprété comme le poids de la quasi-particule et permet, dans une certaine limite, d'estimer jusqu'à quel point la quasi-particule « ressemble » encore à un électron. L'approximation *on-the-mass-shell* correspond alors au cas limite $Z_{\mathbf{k}n} \rightarrow 1$, c'est-à-dire à l'expansion à l'ordre 0 de l'équation (3.3.16). On pourra justifier son utilisation lorsque l'interaction électron-phonon n'est pas suffisamment forte pour localiser la quasi-particule, ce qui est le cas dans plusieurs semi-conducteurs de référence tels que le Ge ou le GaAs. Dans ce cas, la fonction spectrale à un vecteur d'onde \mathbf{k} donné (éq. (3.1.15)) montrerait un pic bien défini autour d'une fréquence correspondant à l'énergie de la quasi-particule. Les deux approximations peuvent toutefois dévier l'une de l'autre en présence d'un très fort couplage entre les électrons et les phonons. Une proportion importante du poids spectral de la quasi-particule est alors transférée vers la portion incohérente du spectre et peut même former des structures satellites, ou *bandes répliques*, qui sont la signature d'états fortement couplés combinant la quasi-particule et un mode de phonon [26]. De telles signatures sont généralement interprétées en termes du concept de polaron, qui sera discuté à la section 3.5.

3.3.2. Approximation des ions rigides

Nous avons maintenant en main l'expression de la self-énergie de l'interaction électron-phonon, contenue dans les équations (3.3.13) et (3.3.15) qui nous permet d'évaluer la dépendance en température de la structure électronique à l'ordre le plus bas en théorie des perturbations. Des expressions explicitant la force du couplage, éq. (3.3.5), on trouve que la mise en application de cette méthodologie requiert l'évaluation des premières et deuxièmes

dérivées du potentiel par rapport aux différents déplacements des noyaux. Ces quantités peuvent, en principe, être obtenues dans le cadre de la théorie de la fonctionnelle de la densité perturbative (DFPT). Néanmoins, le calcul des deuxièmes dérivées présentes dans le terme de Debye-Waller nécessite de résoudre des équations de Sternheimer du second ordre, qui sont particulièrement coûteuses du point de vue numérique. Pour simplifier la mise en application du problème, nous allons donc utiliser l'invariance sous translation du cristal pour réexprimer ces deuxièmes dérivées en termes des premières dérivées du potentiel que l'on retrouve dans l'expression de la self-énergie de Fan. Cette stratégie a été développée dans les premiers travaux de Allen et Heine [21]. Dans la présente section, nous résumons la dérivation présentée dans la référence [27], en utilisant une notation similaire.

Notre point de départ est l'observation que les propriétés physiques d'un cristal sont invariantes sous une translation uniforme de l'ensemble du cristal. Par conséquent, la variation des énergies électroniques à une température donnée doit vérifier

$$\Delta\varepsilon_{\mathbf{k}n}(T) [\{\mathbf{R}_{l\kappa}\}] = \Delta\varepsilon_{\mathbf{k}n}(T) [\{\mathbf{R}_{l\kappa} + \boldsymbol{\delta}\}], \quad (3.3.20)$$

où $\{\mathbf{R}_{l\kappa}\}$ représente l'ensemble des positions des noyaux du cristal et le vecteur $\boldsymbol{\delta}$ est une translation rigide de l'ensemble des noyaux. Par conséquent, la contribution des modes de phonons dont les vecteurs propres décrivent une telle translation rigide à la variation des énergies électroniques doit être nulle. Ces modes correspondent aux phonons acoustiques à $\mathbf{q} = \Gamma$.

Avant de poursuivre, nous allons momentanément appliquer l'approximation adiabatique à la self-énergie de Fan (éq. (3.3.13)), ce qui revient à négliger les fréquences phononiques présentes dans les dénominateurs. Ce faisant, on peut factoriser une dépendance en température de la même forme que celle du terme de Debye-Waller, de sorte que

$$\Delta\varepsilon_{\mathbf{k}n}^{\text{adiab}}(T) = \sum_{\lambda} (\mathcal{D}_{\mathbf{k}n}(\lambda) + \mathcal{F}_{\mathbf{k}n}(\lambda)) (2n_{\lambda}(T) + 1), \quad (3.3.21)$$

où

$$\mathcal{D}_{\mathbf{k}n}(\lambda) = \frac{g_{\mathbf{k}n\lambda-\lambda}^{(2)}}{2} = \sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \mathcal{D}_{\kappa\kappa'}(\lambda) U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda), \quad (3.3.22a)$$

$$\mathcal{F}_{s=(\mathbf{k}n)}(\lambda) = \sum_{s'=(\mathbf{k}+\mathbf{q}n')} \frac{g_{ss'\lambda}^{(1)} g_{s's-\lambda}^{(1)}}{\varepsilon_s^0 - \varepsilon_{s'}^0} = \sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \mathcal{F}_{\kappa\kappa'}(\lambda) U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda). \quad (3.3.22b)$$

$$(3.3.22c)$$

Les expressions explicites pour $\mathcal{D}_{\kappa\kappa'}(\lambda)$ et $\mathcal{F}_{\kappa\kappa'}(\lambda)$ correspondent aux différents éléments des sommes détaillées aux équations (3.3.5) et (3.3.6), desquelles nous avons factorisé les vecteurs de polarisation phononiques. À la dernière ligne, la somme sur l'état électronique

s' , qui exclut implicitement le terme pour lequel le dénominateur est nul, a été absorbée dans la définition de $\mathcal{F}_{\kappa\kappa'}(\lambda)$.

Avec cette notation, l'invariance sous translation nous permet d'établir la règle de somme acoustique,

$$\sum_{\kappa} \left(\mathcal{D}_{\alpha\alpha'}^{\kappa\kappa'}(\Gamma) + \mathcal{F}_{\alpha\alpha'}^{\kappa\kappa'}(\Gamma) \right) = 0 \quad \forall \kappa', \alpha, \alpha', \quad (3.3.23)$$

avec $\Gamma = (\mathbf{\Gamma}\nu)$.

On profite maintenant du fait que l'expression précédente peut être multipliée par n'importe quel terme indépendant de κ tout en demeurant nulle. On peut donc réécrire la contribution de Debye-Waller en y ajoutant zéro sous une forme stratégique,

$$\begin{aligned} \mathcal{D}_{\mathbf{kn}}(\lambda) &= \sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \mathcal{D}_{\alpha\alpha'}^{\kappa\kappa'}(\lambda) U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda) \\ &\quad - \frac{1}{2} \left(\mathcal{D}_{\alpha\alpha'}^{\kappa\kappa'}(\Gamma) + \mathcal{F}_{\alpha\alpha'}^{\kappa\kappa'}(\Gamma) \right) (U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda) + U_{\kappa'\alpha}(\lambda) U_{\kappa\alpha}^*(\lambda)). \end{aligned} \quad (3.3.24)$$

Attardons-nous maintenant aux éléments de matrice contenus dans le terme de Debye-Waller, qui sont de la forme

$$\mathcal{D}_{\alpha\alpha'}^{\kappa\kappa'}(\lambda) \propto \langle \mathbf{kn} | \partial_{\kappa\alpha}(\lambda) \partial_{\kappa'\alpha'}(-\lambda) V^0 | \mathbf{kn} \rangle, \quad (3.3.25)$$

dont les dérivées sont définies dans le terme de droite de l'équation (3.3.14). On peut appliquer l'approximation des ions rigides (*rigid-ion approximation*, RIA), qui consiste à supposer que le potentiel V^0 est une somme de potentiels indépendants provenant de chacun des noyaux,

$$V^0(\hat{\mathbf{r}}) \approx \sum_{l\kappa} V_{l\kappa}(\hat{\mathbf{r}} - \mathbf{R}_{l\kappa}). \quad (3.3.26)$$

Dans ce cas, les deuxièmes dérivées du potentiel associées aux déplacements de deux atomes distincts d'une même cellule unitaire ou d'atomes situés dans des cellules unitaires différentes sont nulles, de sorte que

$$\mathcal{D}_{\alpha\alpha'}^{\text{RIA}\kappa\kappa'}(\lambda) = \mathcal{D}_{\alpha\alpha'}^{\kappa\kappa'}(\lambda = \Gamma) \delta_{\kappa\kappa'}. \quad (3.3.27)$$

En substituant cette expression dans l'équation (3.3.24), on constate qu'on peut séparer ce terme en deux contributions distinctes,

$$\begin{aligned} \mathcal{D}_{\mathbf{k}n}(\lambda) = & \left[\sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \mathcal{D}_{\kappa\kappa'}(\lambda) U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda) - \frac{1}{2} \mathcal{D}_{\kappa\kappa'}(\Gamma) (U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda) + U_{\kappa'\alpha}(\lambda) U_{\kappa\alpha'}^*(\lambda)) \right] \\ & - \left[\frac{1}{2} \sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \mathcal{F}_{\kappa\kappa'}(\Gamma) (U_{\kappa\alpha}(\lambda) U_{\kappa'\alpha'}^*(\lambda) + U_{\kappa'\alpha}(\lambda) U_{\kappa\alpha'}^*(\lambda)) \right], \\ & \triangleq \mathcal{D}_{\mathbf{k}n}(\lambda)^{\text{non-RIA}} + \mathcal{D}_{\mathbf{k}n}(\lambda)^{\text{RIA}}. \end{aligned} \quad (3.3.28)$$

Le premier terme, qui devient nul si on lui applique l'approximation des ions rigides, est celui que l'on négligera dans nos calculs. En imposant la condition énoncée à l'équation (3.3.27), on obtient finalement

$$\Sigma_{\mathbf{k}n}^{\text{DW,RIA}}(T) = -\frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu}^{\text{BZ}} \sum_{n' \neq n} \frac{1}{2\omega_{\mathbf{q}\nu}} \frac{|g_{\mathbf{k}n n'}^{\text{DW}}(\mathbf{q}\nu)|^2}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}n'}^0 + i\eta} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right), \quad (3.3.29)$$

où

$$\begin{aligned} |g_{\mathbf{k}n n'}^{\text{DW}}(\mathbf{q}\nu)|^2 = & \sum_{\substack{\kappa\kappa' \\ \alpha\alpha'}} \langle \mathbf{k}n | (\partial_{\kappa\alpha}(\mathbf{q} = \mathbf{0}) V^{\text{KS}})^* | \mathbf{k} + \mathbf{q}n' \rangle \langle \mathbf{k} + \mathbf{q}n' | \partial_{\kappa'\alpha'}(\mathbf{q} = \mathbf{0}) V^{\text{KS}} | \mathbf{k}n \rangle \\ & \times (U_{\kappa\alpha}^{\nu}(\mathbf{q}) U_{\kappa\alpha'}^{\nu*}(\mathbf{q}) + U_{\kappa'\alpha}^{\nu}(\mathbf{q}) U_{\kappa'\alpha'}^{\nu*}(\mathbf{q})). \end{aligned} \quad (3.3.30)$$

Nous avons désormais en mains des expressions nous permettant d'évaluer les éléments de matrice contenus dans le terme de Debye-Waller uniquement en termes des premières dérivées du potentiel qui entrent dans le terme de Fan.

3.4. Théorie de la fonctionnelle de la densité perturbative

Les équations (3.2.14), (3.3.13), (3.3.29) et (3.3.30) définissent donc entièrement le problème de la renormalisation des énergies électroniques en fonction de la température. Il nous reste maintenant à déterminer comment évaluer les dérivées présentes dans ces expressions dans un contexte de calcul *ab initio*.

Un point important à souligner ici est que le potentiel V^0 qui entre dans les différents éléments de matrice (voir éq. (3.3.5)) dépend des degrés de liberté électroniques¹¹. La variation de V^0 par rapport aux déplacements des noyaux devra, par conséquent, capturer la

11. Rappelons-nous que nous avons initialement défini V^0 à partir du potentiel coulombien entre les électrons et les noyaux.

réponse de la densité électronique au changement du potentiel externe causé par ces déplacements. Dans le contexte de la DFT, nous avons vu que le potentiel effectif à un corps ressenti par chaque électron indépendant décrit par les équations de Kohn-Sham tient implicitement compte des interactions entre l'ensemble des électrons, à travers les termes de Hartree et d'échange-corrélation (éq. (2.3.13)). C'est donc le potentiel de Kohn-Sham qui tiendra le rôle de V^0 .

La manière la plus simple d'aborder numériquement ce problème est d'évaluer les dérivées à l'aide de différences finies. Rappelons-nous que nous avons déjà utilisé l'approximation Born-Oppenheimer pour décorrélérer les degrés de liberté des électrons et des phonons lorsque nous avons exprimé l'hamiltonien des ions, supposant ainsi que la réponse de la densité électronique à un changement de la position des noyaux était instantanée du point de vue des électrons. Dans cette optique, on peut déformer le cristal par une perturbation périodique le long d'un mode de phonon donné et obtenir l'énergie de l'état fondamental du système perturbé à l'aide de la DFT. On sera alors en mesure d'évaluer les dérivées numériquement en échantillonnant différentes amplitudes de la perturbation. Cette approche porte judicieusement le nom de *méthode des phonons gelés* (*frozen phonon*) puisqu'on fige la perturbation et on laisse les électrons relaxer vers leur état d'énergie minimale. Ce faisant, on néglige toutefois la nature dynamique des vibrations atomiques et de la réponse des électrons à celles-ci.

L'apparente simplicité de la méthode des phonons gelés cache néanmoins une limitation sur la taille du système perturbé. En effet, pour capturer complètement la périodicité d'une perturbation dont le vecteur d'onde \mathbf{q} est non-nul, il est nécessaire de recourir à des supercellules, puisque la déformation est modulée par un facteur $e^{i\mathbf{q}\cdot\mathbf{r}}$ dans l'espace réel. La taille de la supercellule croît donc rapidement avec l'inverse du vecteur d'onde et devient du même coup inaccessible du point de vue numérique si l'on veut échantillonner des vecteurs d'onde \mathbf{q} très petits ou peu symétriques. Pour pallier ces limitations, nous allons plutôt utiliser la méthode perturbative développée au tournant des années 1990 par Gonze [28] et Savrasov [29], qui repose sur l'application de la théorie des perturbations au formalisme de la DFT. Nous verrons que cette méthode a l'avantage de permettre le calcul d'une perturbation arbitraire à l'intérieur de la cellule unitaire.

3.4.1. Développement perturbatif

La théorie de la fonctionnelle de la densité perturbative (DFPT) repose sur le développement d'une observable X en une série perturbative par rapport à un paramètre λ ,

$$X(\lambda) = X^{(0)} + \lambda X^{(1)} + \frac{\lambda^2}{2!} X^{(2)} + \frac{\lambda^3}{3!} X^{(3)} + \dots \quad (3.4.1)$$

Pour notre problème, nous avons besoin d'évaluer la matrice hessienne de l'énergie totale par rapport aux déplacements atomiques, $E^{(2)}$, pour construire la matrice dynamique, de même que le développement au premier ordre du potentiel de Kohn-Sham, $V_{\text{KS}}^{(1)}$, pour calculer la self-énergie. Ce dernier s'exprime comme [30]

$$V_{\text{KS}}^{(1)} = v_{\text{ext}}^{(1)} + v_{\text{Hxc}}^{(1)} = v_{\text{ext}}^{(1)} + \frac{\delta v_{\text{Hxc}}}{\delta n(\mathbf{r})} n^{(1)}(\mathbf{r}) + \frac{\partial v_{\text{Hxc}}}{\partial \lambda} \triangleq H^{(1)}, \quad (3.4.2)$$

où v_{ext} est le potentiel externe dû aux noyaux et v_{Hxc} , le potentiel découlant de la fonctionnelle de l'énergie regroupant les termes de Hartree et d'échange-corrélation,

$$v_{\text{Hxc}} = v_{\text{H}} + v_{\text{xc}} = \frac{\delta E_{\text{Hxc}}[n(\mathbf{r})]}{\delta n(\mathbf{r})}. \quad (3.4.3)$$

La réponse de la densité électronique $n(\mathbf{r})$ aux déplacements atomiques est quant à elle capturée par la densité d'ordre 1, $n^{(1)}(\mathbf{r})$. Cette dernière nécessite de connaître les fonctions d'onde au premier ordre en théorie des perturbations, $\psi_s^{*(1)}$,

$$n^{(1)}(\mathbf{r}) = \sum_{s=1}^{\text{occ.}} \psi_s^{*(1)}(\mathbf{r}) \psi_s^{(0)}(\mathbf{r}) + \psi_s^{*(0)}(\mathbf{r}) \psi_s^{(1)}(\mathbf{r}). \quad (3.4.4)$$

La somme sur l'indice s couvre uniquement les états occupés, avec $s = (\mathbf{k}n)$. Les fonctions d'onde $\psi^{(m)}$ correspondent aux orbitales de Kohn-Sham d'ordre m .

En somme, nous avons donc besoin de déterminer $E^{(2)}$ et $\{\psi_s^{(1)}(\mathbf{r})\}$ pour pouvoir solutionner notre problème. Puisque l'énergie totale E vérifie un principe variationnel sous contrainte, le théorème $(2n + 1)$ nous indique que le développement en série de l'énergie totale à l'ordre $(2n + 1)$ en théorie des perturbations ne requiert que les fonctions d'onde à l'ordre n [31]. On peut alors définir la fonctionnelle [30]

$$\begin{aligned} E^{(2)}[V_{\text{KS}}^{(1)}, \psi^0, \psi^1] &= \sum_{s=1}^{\text{occ.}} \left[\langle \psi_s^{(1)} | H^{(0)} - \varepsilon_s^{(0)} | \psi_s^{(1)} \rangle + \langle \psi_s^{(1)} | v_{\text{ext}}^{(1)} | \psi_s^{(0)} \rangle + \langle \psi_s^{(0)} | v_{\text{ext}}^{(1)} | \psi_s^{(1)} \rangle \right. \\ &\quad \left. + \langle \psi_s^{(0)} | v_{\text{ext}}^{(2)} | \psi_s^{(0)} \rangle \right] + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\delta^2 E_{\text{Hxc}}}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') \Bigg|_{n^{(0)}} \\ &\quad + \int d\mathbf{r} \frac{d}{d\lambda} \frac{\delta E_{\text{Hxc}}}{\delta n(\mathbf{r})} n^{(1)}(\mathbf{r}) \Bigg|_{n^{(0)}} + \frac{1}{2} \frac{d^2 E_{\text{Hxc}}}{d\lambda^2} \Bigg|_{n^{(0)}}, \end{aligned} \quad (3.4.5)$$

que l'on voudra minimiser par rapport à $\psi^{(0)}$ et $\psi^{(1)}$ en respectant la contrainte

$$\langle \psi_s^{(0)} | \psi_{s'}^{(1)} \rangle = 0, \quad (3.4.6)$$

pour tous les états s et s' occupés. Cette condition découle de notre choix travailler dans la jauge du transport parallèle [32]. Notons également que le développement en série du potentiel externe dû aux noyaux, $v_{\text{ext}}(\lambda)$ est supposée connue à tous les ordres de perturbation. Il en va de même pour les différentes dérivées par rapport à λ .

Connaissant la solution à l'équation de Schrödinger du système non perturbé, c'est-à-dire $H^{(0)}$, $\{\varepsilon_s^{(0)}\}$ et $\{\psi_s^{(0)}\}$, il nous reste donc à déterminer $\{\psi_s^{(1)}\}$, que l'on obtiendra à l'aide de l'équation de Sternheimer,

$$\left(H^{(0)} - \varepsilon_s^{(0)}\right) |\psi_s^{(1)}\rangle = - \left(H^{(1)} - \varepsilon_s^{(1)}\right) |\psi_s^{(0)}\rangle. \quad (3.4.7)$$

Cette expression découle du développement au 1^{er} ordre de l'équation de Schrödinger en théorie des perturbations [32]. En projetant cette équation sur $\langle\psi_s^0|$, on retrouve le théorème de Hellman-Feynman,

$$\varepsilon_s^{(1)} = \langle\psi_s^0| H^{(1)} |\psi_s^0\rangle. \quad (3.4.8)$$

En substituant ce résultat dans l'équation (3.4.7), on pourra enfin résoudre l'équation de Sternheimer de manière autocohérente pour obtenir $\psi_s^{(1)}$. Ce faisant, nous pourrons construire la matrice dynamique, dont la diagonalisation nous fournira les fréquences et les vecteurs de polarisation des phonons.

Notons que nous aurions tout aussi bien pu choisir d'évaluer $\psi_s^{(1)}$ en utilisant le résultat bien connu de la théorie des perturbations,

$$|\psi_s^{(1)}\rangle = \sum_{s' \neq s} |\psi_{s'}^{(0)}\rangle \frac{\langle\psi_{s'}^{(0)}| H^{(1)} |\psi_s^{(0)}\rangle}{\varepsilon_s^{(0)} - \varepsilon_{s'}^{(0)}}, \quad (3.4.9)$$

où la somme sur s' contient en principe une infinité de termes, dont la grande majorité correspond à des états inoccupés de très haute énergie par rapport au niveau de Fermi. L'équivalence entre les définitions de $\psi_s^{(1)}$ issues des équations (3.4.7) et (3.4.9) met en évidence le grand avantage de l'approche développée par Sternheimer : elle permet de remplacer l'évaluation d'une somme sur un très grand nombre d'états inoccupés par la solution d'une équation différentielle linéaire, qui est moins coûteuse numériquement à obtenir.

Puisque nos expressions de la self-énergie de l'interaction électron-phonon (éq. (3.3.13) et (3.3.29)) comprennent des sommes de ce type, il serait tout à fait judicieux d'utiliser cette technique pour limiter le nombre d'états inoccupés à calculer explicitement [33]. Néanmoins, la méthode de Sternheimer nécessite d'appliquer l'approximation adiabatique à la self-énergie de Fan, ce qui nous ferait perdre la contribution de l'aspect dynamique de l'interaction électron-phonon. Si cette approximation n'est pas souhaitable lorsque les niveaux électroniques intermédiaires en jeu se trouvent près de l'énergie non perturbée de l'état électronique à corriger, elle est tout à fait justifiée pour les bandes de très haute énergie. Nous allons donc choisir un indice de bande seuil, M , qui est suffisamment éloigné en énergie de l'état $\varepsilon_{\mathbf{k}n}^0$ et définir un projecteur sur l'espace orthogonal aux M premiers états électroniques non perturbés, $P_{\perp}(M)$, qui vérifie

$$P_{\perp}(M) |\psi_s^{(1)}\rangle = \sum_{s' > M} |\psi_{s'}^{(0)}\rangle \frac{\langle\psi_{s'}^{(0)}| H^{(1)} |\psi_s^{(0)}\rangle}{\varepsilon_s^{(0)} - \varepsilon_{s'}^{(0)}}. \quad (3.4.10)$$

L'équation de Sternheimer à résoudre pour $P_{\perp}(M) |\psi_s^{(1)}\rangle$ s'écrira alors [32]

$$P_{\perp}(M) \left(H^{(0)} - \varepsilon_s^{(0)} \right) P_{\perp}(M) |\psi_s^{(1)}\rangle = -P_{\perp}(M) H^{(1)} |\psi_s^{(0)}\rangle. \quad (3.4.11)$$

Une fois la solution en main, il suffira de projeter le résultat sur l'état $\langle H^{(1)} \psi_s^{(0)} |$ pour obtenir la contribution adiabatique des bandes de haute énergie. On l'ajoutera ensuite à celle des bandes ayant $n' < M$, qui capture toujours les effets non adiabatiques. Ainsi,

$$\Sigma_{\mathbf{kn}}^{\text{EPI}}(T) = \sum_{n' \leq M} \Sigma_{\mathbf{kn}}^{\text{non adiab.}}(T) + \sum_{n' > M} \Sigma_{\mathbf{kn}}^{\text{adiab., Stern.}}(T), \quad (3.4.12)$$

où le premier terme du membre de droite est calculé avec l'équation (3.3.13) et le second est obtenu avec l'équation de Sternheimer que nous venons d'énoncer.

3.4.2. Factorisation de la phase

Nous avons souligné précédemment que l'une des limitations numériques fondamentales du calcul de l'interaction électron-phonon est que l'on doit modéliser explicitement une période complète de la perturbation atomique caractérisant un mode de phonon donné. Le cas très simple d'un mode de phonon à la frontière de la zone de Brillouin, $\mathbf{q} = (\pi/a, 0, 0)$, nécessite donc de recourir à une cellule primitive qui soit le double de la cellule unitaire. Dans le cas général, la méthode des phonons gelés est de ce fait restreinte à évaluer des perturbations qui sont commensurables avec la supercellule. La taille de la supercellule devient dès lors rapidement prohibitive lorsque $|\mathbf{q}|$ diminue. Bien que différentes approches aient été développées au cours des dernières années pour limiter le coût numérique engendré par les supercellules [34, 35], la contrainte de commensurabilité de la perturbation demeure présente.

L'approche perturbative que nous utilisons pour nos travaux s'affranchit de cette contrainte en remarquant que la périodicité de la perturbation module les différentes variables du problème par une phase $e^{i\mathbf{q}\cdot\mathbf{R}}$ d'une cellule unitaire à l'autre,

$$\psi_{\mathbf{kqn}}^{(1)}(\mathbf{r} + \mathbf{R}) = e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}} \psi_{\mathbf{kqn}}^{(1)}(\mathbf{r}). \quad (3.4.13)$$

On pourra alors s'inspirer du théorème de Bloch et factoriser la phase liée à la perturbation, ce qui permet de définir une fonction $u_{\mathbf{kqn}}^{(1)}(\mathbf{r})$ qui possède la même périodicité que la cellule unitaire [30],

$$\tilde{u}_{\mathbf{kqn}}^{(1)}(\mathbf{r}) = (N\Omega^0) e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}} \psi_{\mathbf{kqn}}^{(1)}(\mathbf{r}) = \tilde{u}_{\mathbf{kqn}}^{(1)}(\mathbf{r} + \mathbf{R}), \quad (3.4.14)$$

où N est le nombre de cellules unitaires dans la cellule périodique de Born-von-Kármán et Ω^0 , le volume de la cellule unitaire non perturbée.

Il suffit dès lors de redéfinir les quantités physiques de notre problème en termes de ces nouvelles fonctions périodiques,

$$\tilde{n}_{\mathbf{q}}^{(1)}(\mathbf{r}) = e^{-i\mathbf{q}\cdot\mathbf{r}}n_{\mathbf{q}}^{(1)}(\mathbf{r}), \quad (3.4.15a)$$

$$\tilde{v}_{\text{ext}}^{(1)}(\mathbf{q},\mathbf{r}) = e^{-i\mathbf{q}\cdot\mathbf{r}}v_{\text{ext}}^{(1)}(\mathbf{q},\mathbf{r}), \quad (3.4.15b)$$

pour que la fonctionnelle définie à l'équation (3.4.5) ne dépende que de quantités dont la périodicité est contenue dans la cellule unitaire. Notons que ces expressions sont indépendantes de la commensurabilité de \mathbf{q} avec le réseau réciproque. Ainsi, la DFPT permet d'évaluer n'importe quelle perturbation \mathbf{q} arbitrairement petite uniquement à l'intérieur de la cellule unitaire, avec un coût numérique du même ordre que la perturbation $\mathbf{q} = \mathbf{\Gamma}$.

3.4.3. Composante non analytique

Jusqu'à présent, nous avons défini la matrice dynamique uniquement en termes des perturbations correspondant à des déplacements atomiques. Ceci résulte de notre définition de la force de rappel entre les noyaux comme la loi de Hooke. Dans la limite de l'approximation harmonique adiabatique, cette description est complète pour un cristal dont la cellule unitaire ne contient qu'un atome, qui ne possède que des modes de phonons acoustiques. Dans le cas contraire, le cristal possède aussi des modes optiques, qui résultent des déplacements relatifs des différents atomes de la cellule unitaire. La différence entre les modes acoustiques et optiques est illustrée à la figure 3.3a pour le cas d'une chaîne diatomique. Lorsque les atomes qui composent la base sont différents, la répartition de la charge électronique n'est pas symétrique dans la cellule unitaire. Les déplacements atomiques selon un mode optique longitudinal (LO) vont alors créer des dipôles électriques qui vont s'aligner entre eux et générer un champ électrique macroscopique, de sorte que leur effet sera ressenti à longue portée. La force de rappel associée à un mode LO contiendra alors une contribution supplémentaire due à l'interaction coulombienne entre les densités de charge distinctes autour de chaque noyau, ou *charges effectives*, qui aura pour conséquence d'augmenter l'énergie du mode LO par rapport aux modes transverses optiques (TO) dans la limite $\mathbf{q} \rightarrow 0$. Ce phénomène est connu sous le nom de *séparation LO-TO (LO-TO splitting)* [36]. Pour les structures diamant et zinc-blende, dont la base contient deux atomes, la propagation d'un phonon optique dans la direction cartésienne [111] correspond à des déplacements relatifs de plans monoatomiques, illustrés à la figure 3.3b. On peut alors interpréter la séparation LO-TO en termes de l'énergie emmagasinée dans un condensateur à plaques parallèles, qui dépend de la distance entre les plaques. Dans cette analogie, que l'on doit à Yu et Cardona [36], les déplacements atomiques selon un mode TO équivalent à un glissement relatif des plaques, qui n'affecte ni la géométrie du condensateur, ni son énergie. En revanche, un mode LO accroît la distance

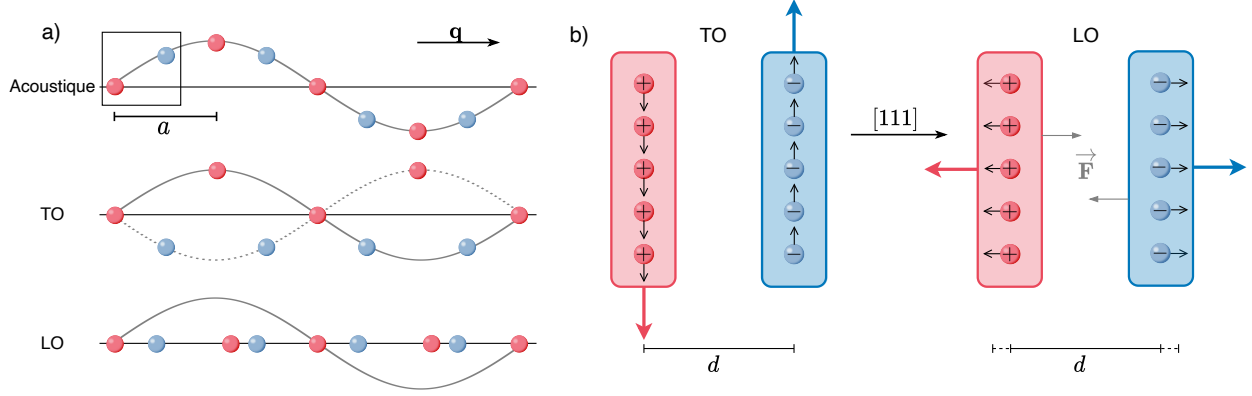


Figure 3.3 – Modes phononiques et composante non analytique du potentiel.

a) Déplacements relatifs des noyaux d’une chaîne diatomique selon différents types de modes de phonons ayant un même vecteur d’onde \mathbf{q} . Dans un mode acoustique, les atomes associés à une même cellule unitaire (encadré) se déplacent dans la même direction. Les modes optiques sont quant à eux associés à des déplacements relatifs de directions opposées. b) Les déplacements atomiques selon un mode optique se propageant dans la direction $[111]$ d’un matériau de structure zinc-blende peuvent être interprétés en analogie avec un condensateur à plaques parallèles. Le mode LO modifie la distance d entre les plans monoatomiques de charges effectives opposées, ce qui génère une force de rappel supplémentaire d’origine coulombienne, qui augmente l’énergie de ce mode comparativement aux modes TO, dont les déplacements atomiques conservent la distance entre les plans. Image adaptée de la référence [36].

entre les plaques de charge opposée, ce qui augmente la force de rappel coulombienne et, par conséquent, l’énergie du condensateur.

On peut inclure l’effet de ces dipôles à longue portée dans la matrice dynamique en y ajoutant un terme supplémentaire [37], $\tilde{C}^{\text{NA}}(\hat{\mathbf{q}})$,

$$\tilde{C}_{\alpha\alpha'}^{\kappa\kappa'}(\mathbf{q} \rightarrow \mathbf{0}) = \tilde{C}_{\alpha\alpha'}^{\kappa\kappa'}(\mathbf{q} = \mathbf{0}) + \tilde{C}_{\alpha\alpha'}^{\text{NA}}(\mathbf{q} \rightarrow \mathbf{0}), \quad (3.4.16)$$

où $\tilde{C}_{\alpha\alpha'}^{\kappa\kappa'}(\mathbf{q})$ est la transformée de Fourier de la matrice des forces interatomiques définie à l’équation (3.2.10). Le terme non analytique (NA) dépend de la direction \mathbf{q} par laquelle on approche le centre de la zone de Brillouin et capture la réponse du système à un champ électrique macroscopique en présence des déplacements atomiques selon un mode LO dans la limite des grandes longueurs d’onde.

Pour l’évaluer, nous aurons besoin de connaître les charges effectives de Born, $Z_{\kappa,\beta\alpha}^*$ qui décrivent le changement de la polarisation macroscopique par cellule unitaire, $\mathcal{P}^{\text{macro}}$, induit par un déplacement atomique. Une définition équivalente serait la variation de la force de rappel sur un atome donné dans la cellule unitaire, \mathbf{F}_{κ} , causée par un champ électrique macroscopique, \mathcal{E} . Ces deux définitions sont reliées par la propriété d’interchangeabilité des

secondes dérivées mixtes de l'énergie totale [37],

$$Z_{\kappa\beta,\alpha}^* = \Omega^0 \frac{\partial \mathcal{P}_\beta^{\text{macro}}}{\partial R_{\kappa\alpha}(\mathbf{q} = \mathbf{0})} = \frac{\partial F_{\kappa\alpha}}{\partial \mathcal{E}_\beta} = - \frac{\partial^2 E}{\partial R_{\kappa\alpha}(\mathbf{q} = \mathbf{0}) \partial \mathcal{E}_\beta}, \quad (3.4.17)$$

où α et β sont des directions cartésiennes. Pour un système isolant ou semi-conducteur, les charges effectives de Born vérifient la règle de somme

$$\sum_{\kappa} Z_{\kappa\beta,\alpha}^* = 0 \quad \forall \alpha, \beta, \quad (3.4.18)$$

qui assure la neutralité de la charge dans la cellule unitaire.

Connaissant les charges effectives de Born, on pourra enfin exprimer la contribution non analytique à la matrice dynamique comme [37]

$$\tilde{C}_{\kappa\kappa',\alpha\alpha'}^{\text{NA}}(\mathbf{q} \rightarrow \mathbf{0}) = \frac{4\pi}{\Omega^0} \frac{\sum_{\beta} q_{\beta} Z_{\kappa\beta,\alpha}^* \sum_{\beta'} q_{\beta'} Z_{\kappa'\beta',\alpha'}^*}{\sum_{\alpha\alpha'} q_{\alpha} \epsilon_{\alpha\alpha'}^{\infty} q_{\alpha'}}, \quad (3.4.19)$$

où q_{β} est la composante du vecteur d'onde \mathbf{q} dans la direction β et ϵ^{∞} est le tenseur diélectrique comprenant uniquement la contribution de la polarisation électronique [30, 37],

$$\epsilon_{\alpha\alpha'}^{\infty} = \delta_{\alpha\alpha'} - \frac{4\pi}{\Omega^0} \frac{\partial^2 E}{\partial \mathcal{E}_{\alpha}^* \partial \mathcal{E}_{\alpha'}}. \quad (3.4.20)$$

3.5. Modèle de Fröhlich

On pourrait a priori penser que la composante non analytique de la matrice dynamique n'a pour seul effet que de modifier la dispersion des phonons LO au voisinage du centre de la zone de Brillouin. Or, puisque la présence des dipôles alignés dans le matériau génère une polarisation à longue portée, les déplacements atomiques résultant des modes LO de grandes longueurs d'onde induisent un champ électrique macroscopique à l'intérieur du matériau. Ce dernier constitue par conséquent un mécanisme supplémentaire par lequel les électrons, qui sont des particules chargées, peuvent interagir avec les modes LO de grande longueur d'onde.

Ce type d'interaction est au cœur du modèle proposé par Fröhlich dans les années 1950 pour modéliser l'interaction d'un électron avec le champ de polarisation créé par les vibrations des noyaux dans un matériau polaire [38, 39]. Développé à l'origine pour fournir une explication théorique au phénomène de la rupture diélectrique dans les isolants [40], il décrit des quasi-particules appelées polarons, où le mouvement de l'électron est corrélé avec celui des noyaux. Le modèle de Fröhlich se limite au cas où la longueur de localisation de la fonction d'onde du polaron est grande par rapport à la distance interatomique, ce qu'on appelle également un « grand polaron ». Dans ce cas, on peut justifier d'appliquer l'hypothèse du continu et de négliger les détails microscopiques du réseau cristallin.

La formulation originale du modèle de Fröhlich considère un seul électron dans une bande parfaitement parabolique, isotrope et non dégénérée, caractérisée par une masse effective m^* .

Ce dernier interagit avec le champ électrique généré par un unique mode LO décrit par un phonon d'Einstein de fréquence ω_{LO} . L'hamiltonien de Fröhlich s'écrit alors

$$\begin{aligned} H^{\text{Fr}} &= \sum_{\mathbf{k}} \frac{k^2}{2m^*} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + \sum_{\mathbf{q}} \omega_{\text{LO}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + \sum_{\mathbf{k}\mathbf{q}} g^{\text{Fr}}(\mathbf{q}) c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} (a_{\mathbf{q}} + a_{-\mathbf{q}}^\dagger) \\ &= H^0 + H^{\text{ions}} + H^{\text{EPI}}. \end{aligned} \quad (3.5.1)$$

Cette expression a la même forme que l'hamiltonien général défini par les équations (3.2.21) et (3.3.3) en se limitant à une seule bande n et une seule branche ν , et où le terme d'interaction électron-phonon capture l'interaction diélectrique entre l'électron et le phonon LO. Nous avons supposé sans perte de généralité que l'énergie du minimum de la bande de conduction (CBM) est nulle.

Pour un matériau tridimensionnel, la constante de couplage, $g^{\text{Fr}}(\mathbf{q})$, s'écrit quant à elle

$$g^{\text{Fr}}(\mathbf{q}) = \frac{1}{q} \left[\frac{2\pi\omega_{\text{LO}}}{\epsilon^* V_{\text{BvK}}} \right]^{\frac{1}{2}}, \quad (3.5.2)$$

où V_{BvK} est le volume de normalisation de Born-von Kármán, q est la norme du vecteur d'onde \mathbf{q} , tandis que ϵ^* caractérise la permittivité du matériau et s'exprime en termes des constantes diélectriques optique et statique, ϵ^∞ et ϵ^0 ,

$$\frac{1}{\epsilon^*} = \frac{1}{\epsilon^\infty} - \frac{1}{\epsilon^0}. \quad (3.5.3)$$

Pour un matériau isotrope, ces dernières sont reliées à la fréquence des modes optiques par la relation de Lyddane-Sachs-Teller [17],

$$\frac{\omega_{\text{LO}}^2}{\omega_{\text{TO}}^2} = \frac{\epsilon^0}{\epsilon^\infty}. \quad (3.5.4)$$

Tout comme dans le cas général, on définit ensuite la self-énergie du modèle de Fröhlich à l'ordre le plus bas en théorie des perturbations dans l'approche de Rayleigh-Schrödinger. À $T = 0$ K, cette dernière prend la forme

$$\Sigma_{\mathbf{k}}^{\text{Fr}} \left(\omega = \frac{k^2}{2m^*} \right) = \frac{\Omega^0}{(2\pi)^3} \int d^3q \frac{|g^{\text{Fr}}(\mathbf{q})|^2}{\frac{k^2}{2m^*} - \frac{(\mathbf{k}+\mathbf{q})^2}{2m^*} - \omega_{\text{LO}}}, \quad (3.5.5)$$

où nous avons supposé par simplicité que le CBM est situé à $\mathbf{k} = \mathbf{0}$. Nous avons utilisé l'approximation *on-the-mass-shell* et appliqué l'hypothèse du continu,

$$\frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} \rightarrow \frac{\Omega^0}{(2\pi)^3} \int d^3q. \quad (3.5.6)$$

Le facteur $i\eta$ est absent de l'équation (3.5.5) puisque la self-énergie est réelle dans le proche voisinage du CBM¹².

¹². Le dénominateur sera réel si $k^2/2m^* < \omega_{\text{LO}}$, qui correspond à la fenêtre d'énergie où l'interaction Fröhlich est importante.

Après intégration de l'équation (3.5.5) [3], on obtient finalement l'énergie de formation du polaron, qui équivaut à la renormalisation des énergies électroniques (éq. (3.3.17)) induite par l'interaction Fröhlich [3],

$$\varepsilon_{\mathbf{k}} = \frac{k^2}{2m^*} - \alpha \frac{\omega_{\text{LO}}}{ka_{\text{LO}}} \arcsin(ka_{\text{LO}}), \quad (3.5.7)$$

où nous avons introduit la constante de Fröhlich ,

$$\alpha = \frac{1}{\epsilon^*} \left(\frac{m^*}{2\omega_{\text{LO}}} \right)^{\frac{1}{2}}, \quad (3.5.8)$$

qui quantifie la force du couplage, ainsi que la longueur caractéristique [41],

$$a_{\text{LO}} = (2m^*\omega_{\text{LO}})^{-\frac{1}{2}}. \quad (3.5.9)$$

On peut enfin développer l'arc sinus en série de Taylor, de sorte que

$$\varepsilon_{\mathbf{k}} = -\alpha\omega_{\text{LO}} + \frac{k^2}{2m^*} \left(1 - \frac{\alpha}{6} \right) + \mathcal{O}(k^4), \quad (3.5.10)$$

où le terme constant correspond au ZPR induit par l'interaction Fröhlich ,

$$\text{ZPR}^{\text{Fr}} = -\alpha\omega_{\text{LO}}. \quad (3.5.11)$$

En revenant à la définition de α (éq. (3.5.8)), on peut apprécier la remarquable simplicité de cette expression, qui capture la réponse diélectrique de l'électron aux vibrations atomiques uniquement à l'aide de quelques paramètres macroscopiques.

De l'équation (3.5.10), on peut également extraire la masse effective du polaron,

$$m_{\text{pol}}^* = m^* \left(1 - \frac{\alpha}{6} \right)^{-1}. \quad (3.5.12)$$

Cette expression nous indique d'une part que la masse effective du polaron est plus grande que celle de l'électron dans l'approximation des noyaux statiques. On peut interpréter tout naturellement ce résultat en considérant que, dans l'esprit du polaron, un électron qui se déplace dans le cristal doit « traîner » avec lui le nuage de polarisation des noyaux, ce qui ralentit sa propagation. Du point de vue d'un observateur externe, la quasi-particule semble donc plus lourde [3].

D'autre part, l'équation (3.5.12) nous renseigne sur la validité de notre développement perturbatif, qui suppose un couplage faible. On peut en effet remarquer que la masse effective du polaron diverge pour $\alpha = 6$, qui correspond à la valeur critique du couplage pour laquelle le polaron devient localisé dans une cellule unitaire. On entre alors dans le régime du couplage fort, lié au concept de « petit polaron », où la nature microscopique du cristal doit être prise en compte.

Avant de conclure cette section, soulignons que l'équation (3.5.5) ne contient que la contribution intrabande du terme de Fan. On peut justifier cette approximation, sachant que,

dans la limite $q \rightarrow 0$, la contribution du terme de Debye-Waller est entièrement annulée par celle des modes acoustiques du terme de Fan [21]. Il est néanmoins important de remarquer que le modèle de Fröhlich ne se veut pas une théorie complète de l'interaction électron-phonon au plus bas ordre en théorie des perturbations, à la différence des expressions développées à la section 3.3.1. Au contraire, il met précisément l'accent sur l'interaction entre l'électron et la polarisation du réseau cristallin pour un mode LO de grande longueur d'onde. La fréquence ω_{LO} y joue un rôle prépondérant, mettant en lumière la nature retardée, ou non adiabatique, de l'interaction. Ce faisant, le modèle de Fröhlich nous confère un accès direct aux contributions à la self-énergie qui vont au-delà de l'approximation adiabatique.

Connu depuis près de 70 ans, la force du modèle de Fröhlich réside dans la remarquable quantité d'information physique qu'il est possible d'en extraire à partir d'un nombre restreint de paramètres [42]. Sa simplicité et sa nature profondément intuitive font en sorte qu'il demeure toujours d'actualité, comme en témoigneront les deux prochains chapitres. Nous verrons entre autres que l'interaction de Fröhlich est cruciale pour décrire correctement la renormalisation de la bande interdite dans les matériaux dont le degré d'ionicté est important.

Bibliographie

- [1] W. Bludau, A. Onton, and W. Heinke. Temperature dependence of the band gap of silicon. *Journal of Applied Physics*, 45(4) :1846–1848, 1974. doi: [10.1063/1.1663501](https://doi.org/10.1063/1.1663501).
- [2] R. Pässler. Semi-empirical descriptions of temperature dependences of band gaps in semiconductors. *Phys. Status Solidi B*, 3 :710–728, 2003.
- [3] G. Mahan. *Many-Particle Physics*. Plenum, New York, 2nd edition, 1990.
- [4] P. Coleman. *Introduction to Many-Body Physics*. Cambridge University Press, Cambridge, 2015. doi: [10.1017/CBO9781139020916](https://doi.org/10.1017/CBO9781139020916).
- [5] R. D. Mattuck. *A Guide to Feynman Diagrams in the Many-Body Problem*. Dover Books on Physics and Chemistry. Dover, New York, 2nd edition, 1992.
- [6] F. Giustino. Electron-phonon interactions from first principles. *Reviews of Modern Physics*, 89(1) :015003, 2017. doi: [10.1103/RevModPhys.89.015003](https://doi.org/10.1103/RevModPhys.89.015003).
- [7] A. Damascelli, Z. Hussain, and Z. X. Shen. Angle-resolved photoemission studies of the cuprate superconductors. *Reviews of Modern Physics*, 75(2) :473–541, 2003. doi: [10.1103/RevModPhys.75.473](https://doi.org/10.1103/RevModPhys.75.473).
- [8] M. P. Marder. *Condensed Matter Physics*. John Wiley & Sons, Hoboken, 2nd edition, 2010.
- [9] D. J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, Upper Saddle River, 2nd edition, 2005.

- [10] M. Born and K. Huang. *Dynamical Theory of Crystal Lattices*. Oxford University Press, Oxford, 1954.
- [11] J. Breeze. Theory of Anharmonic Phonons. In *Temperature and Frequency Dependence of Complex Permittivity in Metal Oxide Dielectrics : Theory, Modelling and Measurement*, pages 111–133. Springer International Publishing, Cham, 2016. doi: [10.1007/978-3-319-44547-2_6](https://doi.org/10.1007/978-3-319-44547-2_6). Series Title : Springer Theses.
- [12] M. Lazzeri, M. Calandra, and F. Mauri. Anharmonic phonon frequency shift in MgB₂. *Phys. Rev. B*, 68(22) :220509, 2003. doi: [10.1103/PhysRevB.68.220509](https://doi.org/10.1103/PhysRevB.68.220509).
- [13] D. S. Kim, O. Hellman, N. Shulumba, C. N. Saunders, J. Y. Y. Lin, H. L. Smith, J. E. Herriman, J. L. Niedziela, D. L. Abernathy, C. W. Li, and B. Fultz. Temperature-dependent phonon lifetimes and thermal conductivity of silicon by inelastic neutron scattering and ab initio calculations. *Phys. Rev. B*, 102(17) :174311, 2020. doi: [10.1103/PhysRevB.102.174311](https://doi.org/10.1103/PhysRevB.102.174311).
- [14] L. Paulatto, I. Errea, M. Calandra, and F. Mauri. First-principles calculations of phonon frequencies, lifetimes, and spectral functions from weak to strong anharmonicity : The example of palladium hydrides. *Phys. Rev. B*, 91(5) :054304, 2015. doi: [10.1103/PhysRevB.91.054304](https://doi.org/10.1103/PhysRevB.91.054304).
- [15] L. Paulatto, F. Mauri, and M. Lazzeri. Anharmonic properties from a generalized third-order ab initio approach : Theory and applications to graphite and graphene. *Phys. Rev. B*, 87(21) :214303, 2013. doi: [10.1103/PhysRevB.87.214303](https://doi.org/10.1103/PhysRevB.87.214303).
- [16] P. Boulanger. *Études ab initio des effets de la température sur le spectre optique des semi-conducteurs*. PhD thesis, Université de Montréal et Université Catholique de Louvain, 2010.
- [17] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College, Philadelphia, 1976.
- [18] H. Y. Fan. Temperature Dependence of the Energy Gap in Monatomic Semiconductors. *Phys. Rev.*, 78(6) :808–809, 1950. doi: [10.1103/PhysRev.78.808.2](https://doi.org/10.1103/PhysRev.78.808.2).
- [19] H. Y. Fan. Temperature Dependence of the Energy Gap in Semiconductors. *Phys. Rev.*, 82(6) :900–905, 1951. doi: [10.1103/PhysRev.82.900](https://doi.org/10.1103/PhysRev.82.900).
- [20] E. Antončík. On the theory of temperature shift of the absorption curve in non-polar crystals. *Czechoslovakij fiziceskij zurnal*, 5(4) :449–461, 1955. doi: [10.1007/BF01687209](https://doi.org/10.1007/BF01687209).
- [21] P. B. Allen and V. Heine. Theory of the temperature dependence of electronic band structures. *J. Phys. C : Solid State Phys.*, 9(12) :2305, 1976. doi: [10.1088/0022-3719/9/12/013](https://doi.org/10.1088/0022-3719/9/12/013).
- [22] P. B. Allen and M. Cardona. Theory of the temperature dependence of the direct gap of germanium. *Physical Review B*, 23(4) :1495–1505, 1981. doi: [10.1103/PhysRevB.23.1495](https://doi.org/10.1103/PhysRevB.23.1495).

- [23] P. B. Allen and M. Cardona. Temperature dependence of the direct gap of Si and Ge. *Phys. Rev. B*, 27(8) :4760–4769, 1983. doi: [10.1103/PhysRevB.27.4760](https://doi.org/10.1103/PhysRevB.27.4760).
- [24] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12) :121102(R), 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).
- [25] A. Marini, S. Ponc e, and X. Gonze. Many-body perturbation theory approach to the electron-phonon interaction with density-functional theory as a starting point. *Phys. Rev. B*, 91(22) :224310, 2015. doi: [10.1103/PhysRevB.91.224310](https://doi.org/10.1103/PhysRevB.91.224310).
- [26] J. P. Nery, P. B. Allen, G. Antonius, L. Reining, A. Miglio, and X. Gonze. Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators : Cumulants applied to the full first-principles theory and the Fr ohlich polaron. *Phys. Rev. B*, 97(11) :115145, 2018. doi: [10.1103/PhysRevB.97.115145](https://doi.org/10.1103/PhysRevB.97.115145).
- [27] S. Ponc e, G. Antonius, Y. Gillet, P. Boulanger, J. Laflamme Janssen, A. Marini, M. C ot e, and X. Gonze. Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation. *Phys. Rev. B*, 90(21) :214304, 2014. doi: [10.1103/PhysRevB.90.214304](https://doi.org/10.1103/PhysRevB.90.214304).
- [28] X. Gonze, D. C. Allan, and M. P. Teter. Dielectric tensor, effective charges, and phonons in α -quartz by variational density-functional perturbation theory. *Phys. Rev. Lett.*, 68(24) :3603–3606, 1992. doi: [10.1103/PhysRevLett.68.3603](https://doi.org/10.1103/PhysRevLett.68.3603).
- [29] S. Y. Savrasov. Linear response calculations of lattice dynamics using muffin-tin basis sets. *Phys. Rev. Lett.*, 69(19) :2819–2822, 1992. doi: [10.1103/PhysRevLett.69.2819](https://doi.org/10.1103/PhysRevLett.69.2819).
- [30] X. Gonze. First-principles responses of solids to atomic displacements and homogeneous electric fields : Implementation of a conjugate-gradient algorithm. *Phys. Rev. B*, 55(16) :10337–10354, 1997. doi: [10.1103/PhysRevB.55.10337](https://doi.org/10.1103/PhysRevB.55.10337).
- [31] X. Gonze. Perturbation expansion of variational principles at arbitrary order. *Physical Review A*, 52 :1086, 1995. doi: <https://doi.org/10.1103/PhysRevA.52.1086>.
- [32] X. Gonze. Adiabatic density-functional perturbation theory. *Physical Review A*, 52 :1096, 1995. doi: <https://doi.org/10.1103/PhysRevA.52.1096>.
- [33] X. Gonze, P. Boulanger, and M. C ot e. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Annalen der Physik*, 523(1-2) :168–178, 2011. doi: [10.1002/andp.201000100](https://doi.org/10.1002/andp.201000100).
- [34] J. H. Lloyd-Williams and B. Monserrat. Lattice dynamics and electron-phonon coupling calculations using nondiagonal supercells. *Phys. Rev. B*, 92(18) :184301, 2015. doi: [10.1103/PhysRevB.92.184301](https://doi.org/10.1103/PhysRevB.92.184301).
- [35] M. Zacharias and F. Giustino. Theory of the special displacement method for electronic structure calculations at finite temperature. *Phys. Rev. Research*, 2(1) :013357, 2020. doi: [10.1103/PhysRevResearch.2.013357](https://doi.org/10.1103/PhysRevResearch.2.013357).
- [36] P. Y. Yu and M. Cardona. *Fundamentals of Semiconductors*. Springer-Verlag, Berlin Heidelberg, 2nd edition, 1999.

- [37] X. Gonze and C. Lee. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B*, 55(16) :10355–10368, 1997. doi: [10.1103/PhysRevB.55.10355](https://doi.org/10.1103/PhysRevB.55.10355).
- [38] H. Fröhlich. Interaction of electrons with lattice vibrations. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 215(1122) :291–298, 1952. doi: [10.1098/rspa.1952.0212](https://doi.org/10.1098/rspa.1952.0212).
- [39] H. Fröhlich. Electrons in Lattice Fields. *Advances in Physics*, 3 :325–361, 1954.
- [40] H. Fröhlich. Theory of Electrical Breakdown in Ionic Crystals. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 160(901) :230–241, 1937. doi: [10.1098/rspa.1937.0106](https://doi.org/10.1098/rspa.1937.0106).
- [41] J. P. Nery and P. B. Allen. Influence of Fröhlich polaron coupling on renormalized electron bands in polar semiconductors : Results for zinc-blende GaN. *Phys. Rev. B*, 94 (11) :115135, 2016. doi: [10.1103/PhysRevB.94.115135](https://doi.org/10.1103/PhysRevB.94.115135).
- [42] J. T. Devreese and A. S. Alexandrov. Fröhlich polaron and bipolaron : recent developments. *Rep. Prog. Phys.*, 72(6) :066501, 2009. doi: [10.1088/0034-4885/72/6/066501](https://doi.org/10.1088/0034-4885/72/6/066501).

Chapitre 4

Prédominance des effets non adiabatiques dans la renormalisation du point zéro de l'énergie de la bande interdite

4.0.1. Mise en contexte

Ce projet se veut une remise en question fondamentale de l'utilisation de l'approximation adiabatique pour le calcul de la renormalisation du point zéro. Rappelons que cette approximation consiste à supposer que la densité électronique s'ajuste instantanément à un changement du potentiel ionique, puisque les électrons sont beaucoup plus légers que les noyaux.

Dans une approche perturbative comme la méthode AHC, l'approximation adiabatique simplifie l'évaluation de la renormalisation des énergies électroniques et est en règle générale justifiée par le fait que l'énergie typique d'un mode phononique est nettement plus faible que la différence d'énergie entre les états électroniques en jeu (voir l'expression des dénominateurs de la self-énergie de Fan, éq. (3.3.13)). Appliquée avec succès pour prédire le ZPR du Si et du Ge dans les premiers travaux de Allen, Heine et Cardona [1, 2], l'utilisation systématique de cette approximation dans un contexte perturbatif a été fortement critiquée par Poncé *et al.* [3], qui ont démontré qu'elle cause une divergence du ZPR des matériaux actifs dans l'infrarouge dans la limite des grandes longueurs d'onde. Cette divergence n'apparaît pas dans le cadre d'un traitement non adiabatique de l'interaction électron-phonon, où la nature dynamique de la réponse des électrons aux oscillations atomiques est conservée, ce qui nous indique qu'elle n'est pas physique.

Dans un autre ordre d'idées, l'approximation adiabatique permet d'évaluer la renormalisation des énergies électroniques sans avoir recours à un développement perturbatif, avec la méthode des phonons gelés, que nous avons brièvement introduite à la section 3.4. Cette approche obtient le ZPR par l'entremise de la moyenne thermodynamique d'un ensemble

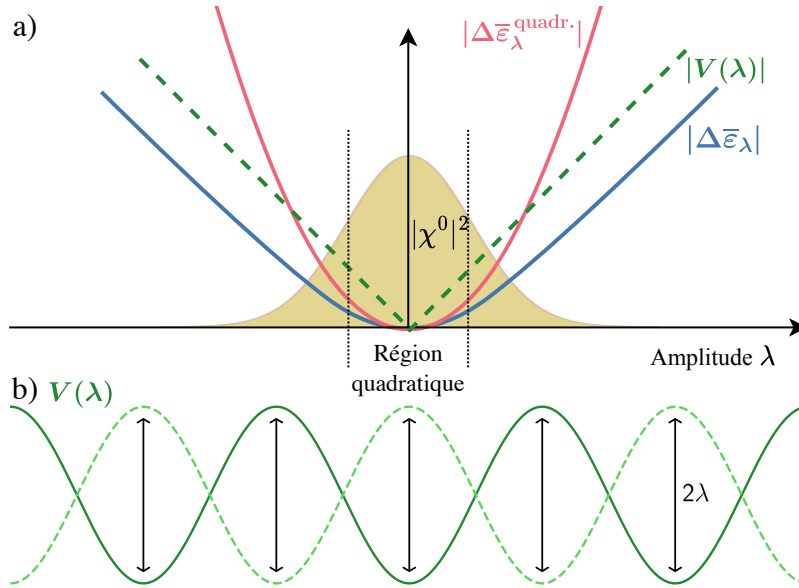


Figure 4.1 – Comparaison de méthodes adiabatiques perturbatives et non perturbatives pour la renormalisation des énergies électroniques. a) L’amplitude du potentiel induit par les déplacements atomiques (tirets verts) varie linéairement avec l’amplitude du déplacement. Les méthodes non perturbatives échantillonnent une vaste plage de déplacements atomiques et capturent le régime des grandes amplitudes, où la variation des énergies électroniques est linéaire ($\Delta\bar{\epsilon}$, bleu). En comparaison, les méthodes perturbatives se limitent à la région où la variation des énergies électroniques est quadratique ($\Delta\bar{\epsilon}^{\text{quadr.}}$, rouge). La gaussienne ($|\chi^0|^2$, sable) donne le poids de chacune des amplitudes λ dans la fonction d’onde des phonons. Image adaptée de la référence [4]. b) La profondeur du potentiel effectif dû aux phonons dépend de l’amplitude des oscillations atomiques, λ .

d’amplitudes de vibration atomiques¹. La réponse électronique à une configuration cristalline où les différents noyaux sont déplacés selon un des modes de phonon est évaluée à l’aide de la DFT, ce qui implique que la méthode est par construction adiabatique. Au contraire de la méthode AHC, l’échantillonnage plus vaste en amplitude des vibrations atomiques permet de capturer des effets allant au-delà de l’approximation harmonique, qui préviennent la divergence du ZPR des matériaux actifs dans l’infrarouge. La nature non perturbative de la méthode permet également de s’affranchir de l’approximation des ions rigides pour la contribution Debye-Waller.

La figure 4.1a illustre la différence fondamentale entre ces deux approches. On y voit que la profondeur du potentiel effectif généré par un mode de phonon et ressenti par l’électron (tirets verts) varie linéairement avec l’amplitude des déplacements atomiques, λ (fig. 4.1b). Le changement d’énergie moyenne de l’état électronique ($\Delta\bar{\epsilon}$, bleu), quant à lui, varie d’abord

1. Le formalisme des phonons gelés est discuté plus en détail à la section [Supplementary Discussion A](#).

quadratiquement avec λ pour de petites oscillations, puis atteint un régime linéaire. La distribution gaussienne des amplitudes d’oscillation pour l’état fondamental d’un oscillateur harmonique quantique ($|\chi^0|^2$, sable) montre le poids relatif de chaque amplitude dans la moyenne thermodynamique. Alors que la méthode des supercellules adiabatiques échantillonne l’ensemble de la gaussienne, l’approche perturbative adiabatique se limite aux amplitudes pour lesquelles la variation d’énergie propre est quadratique ($\bar{\varepsilon}^{\text{quadr.}}$, rouge), de sorte que l’on approxime la moyenne thermodynamique par la courbure de $\bar{\varepsilon}$ à $\lambda = 0$. Pour un mode phononique non actif dans l’infrarouge, cette courbure est proportionnelle à $1/q^2$, où q est la norme du vecteur d’onde du phonon, qui donne lieu à une divergence intégrable lorsque l’on intègre sur l’ensemble des vecteurs d’onde phononiques. Pour un mode LO d’un matériau polaire, la dépendance en q de la constante de couplage (éq. (3.5.2)) implique par contre une courbure proportionnelle à $1/q^4$; cette divergence non intégrable explique la divergence du ZPR observée dans la limite $q \rightarrow 0$.

Or, si l’approximation adiabatique donne lieu à une divergence non physique dans le cadre de l’approximation adiabatique, il advient de se demander s’il est réellement justifié d’appliquer cette approximation dans un contexte non perturbatif. L’objectif de ce projet était donc de comprendre le phénomène physique qui se cache derrière la divergence observée dans l’approximation adiabatique perturbative, ainsi que de caractériser l’importance des contributions non adiabatiques au ZPR pour les phonons LO de très grandes longueurs d’onde. Dans cette optique, nous avons développé une généralisation du modèle de Fröhlich, qui met précisément l’accent sur ce régime d’interaction tout en capturant toute la complexité de la structure électronique des matériaux réels. Ce faisant, nous sommes en mesure d’isoler les contributions non adiabatiques et de quantifier leur importance par rapport au ZPR total.

En plus de mettre en évidence la prédominance des effets non adiabatiques dans le ZPR des matériaux polaires, nos travaux ont révélé les conséquences majeures de cette contribution sur le pouvoir prédictif des calculs numériques. Ce faisant, nous avons remis en question l’utilisation des méthodes intrinsèquement adiabatiques pour les matériaux polaires et mis en lumière la nécessité de développer des stratégies correctives pour pallier à l’absence de ces contributions. Entre autres, l’approche proposée par Nery et Allen [5] permet d’estimer la contribution manquante à un calcul adiabatique due à l’absence de l’interaction Fröhlich. Des travaux récents ont également proposé une méthode d’interpolation permettant d’évaluer l’opérateur de couplage électron-phonon (éq. (3.3.6)) pour un vecteur d’onde \mathbf{q} arbitraire à partir des perturbations par rapport aux déplacements atomiques évaluées à l’aide de supercellules adiabatiques de taille raisonnable [6]. Cette interpolation est exacte si le vecteur d’onde \mathbf{q} est commensurable avec la supercellule. Ce faisant, il devient possible d’évaluer la self-énergie AHC de manière non adiabatique dans un contexte de supercellules adiabatique avec une prédictibilité similaire à celle de l’approche DFT+DFPT. Cependant, on perd du même coup l’accès aux interactions d’ordre supérieur et à la contribution de Debye-Waller

au-delà de l'approximation des ions rigides auxquels on accéderait si la renormalisation des énergies électroniques était elle aussi évaluée avec la méthode des phonons gelés.

Précisons finalement que le formalisme AHC non adiabatique utilisé dans cet article conserve néanmoins une certaine part d'adiabaticité. En effet, bien que la self-énergie du couplage électron-phonon prenne en compte le temps de réaction des électrons face aux oscillations des noyaux, le potentiel perturbé ressenti par les électrons demeure, pour sa part, évalué dans l'approximation adiabatique, puisqu'il est obtenu en figeant le déplacement des noyaux du cristal. La non adiabaticité entre donc en jeu uniquement au moment où on évalue la réponse d'un état électronique au changement de potentiel induit par un mode de phonon donné.

4.0.2. Contributions des auteurs

Anna Miglio a effectué les calculs pour la plupart des oxydes et le Si (10 matériaux), avec le soutien de Matteo Giantomassi. Gabriel Antonius et Yang-Hao Chan ont effectué les calculs du SrTiO_3 et du ZnO. Émile Godbout a calculé la dilatation du point zéro du réseau pour 6 matériaux. J'ai personnellement effectué les calculs EPI pour 15 matériaux, calculé le modèle de Fröhlich généralisé pour 17 matériaux et obtenu la dilatation du point zéro du réseau pour 11 matériaux. Xavier Gonze a dérivé le traitement perturbatif du modèle de Fröhlich généralisé et a rédigé la première version du manuscrit. Xavier Gonze et Michel Côté ont supervisé le projet. Tous les auteurs ont contribué à la révision du manuscrit.

Premier article.

Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap

par

Anna Miglio^{†1}, Véronique Brousseau-Couture^{†2}, Émile Godbout³, Gabriel Antonius⁴, Yang-Hao Chan⁵, Steven G. Louie⁶, Michel Côté⁷, Matteo Giantomassi⁸ et Xavier Gonze⁹

- (¹) Institute of Condensed Matter and Nanosciences, UCLouvain, B-1348 Louvain-la-Neuve, Belgium
- (²) Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Montreal, C.P. 6128, Succursale Centre-Ville, Montreal, Canada H3C 3J7
- (³) Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Montreal, C.P. 6128, Succursale Centre-Ville, Montreal, Canada H3C 3J7
- (⁴) Department of Physics, University of California at Berkeley, California 94720, USA and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA and Département de Chimie, Biochimie et Physique, Institut de recherche sur l'hydrogène, Université du Québec à Trois-Rivières, Trois-Rivières, Canada
- (⁵) Department of Physics, University of California at Berkeley, California 94720, USA and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
- (⁶) Department of Physics, University of California at Berkeley, California 94720, USA and Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
- (⁷) Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Montreal, C.P. 6128, Succursale Centre-Ville, Montreal, Canada H3C 3J7
- (⁸) Institute of Condensed Matter and Nanosciences, UCLouvain, B-1348 Louvain-la-Neuve, Belgium
- (⁹) Institute of Condensed Matter and Nanosciences, UCLouvain, B-1348 Louvain-la-Neuve, Belgium and Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel St. 3, Moscow, 143026, Russia

Cet article a été publié dans npj Computational Materials.

† Anna Miglio and Véronique Brousseau-Couture contributed equally to this work.

RÉSUMÉ. Les propriétés électroniques et optiques des matériaux sont affectées par le mouvement des atomes à travers l'interaction électron-phonon : non seulement l'énergie de la bande interdite change avec la température, mais même au zéro absolu, le mouvement du point zéro provoque une renormalisation de la bande interdite. Nous présentons une étude à grande échelle utilisant les principes premiers pour évaluer la renormalisation du point zéro des extrema de bande au-delà de l'approximation adiabatique. Pour les matériaux comprenant des éléments légers, la renormalisation de la bande interdite est souvent supérieure à 0.3 eV et peut atteindre 0.7 eV. Cet effet ne peut être ignoré si l'on cherche à évaluer l'énergie de la bande interdite avec précision. Pour les matériaux actifs dans l'infrarouge, un accord global avec les données expérimentales disponibles n'est obtenu que lorsque les effets non adiabatiques sont pris en compte. Ces derniers dominent même la renormalisation du point zéro pour de nombreux matériaux, comme le démontre un modèle Fröhlich généralisé qui tient compte de multiples branches de phonons et d'extrema de bande anisotropes et dégénérés, dont le domaine de validité est établi par comparaison avec les résultats obtenus par principes premiers.

Mots clés : Structure électronique, dépendance en température, théorie de la fonctionnelle de la densité, interaction électron-phonon

ABSTRACT. Electronic and optical properties of materials are affected by atomic motion through the electron-phonon interaction: not only band gaps change with temperature, but even at absolute zero temperature, zero-point motion causes band-gap renormalization. We present a large-scale first-principles evaluation of the zero-point renormalization of band edges beyond the adiabatic approximation. For materials with light elements, the band gap renormalization is often larger than 0.3 eV, and up to 0.7 eV. This effect cannot be ignored if accurate band gaps are sought. For infrared-active materials, global agreement with available experimental data is obtained only when non-adiabatic effects are taken into account. They even dominate zero-point renormalization for many materials, as shown by a generalized Fröhlich model that includes multiple phonon branches, anisotropic and degenerate electronic extrema, whose range of validity is established by comparison with first-principles results.

Keywords: Electronic structure, temperature dependence, Density-Functional Theory, electron-phonon interaction

4.1. Introduction

The electronic band gap is arguably the most important characteristic of semiconductors and insulators. It determines optical and luminescent thresholds, but is also a prerequisite for characterizing band offsets at interfaces and deep electronic levels created by defects [7]. However, accurate band gap computation is a challenging task. Indeed, the vast majority of first-principles calculations relies on Kohn-Sham Density-Functional Theory (KS-DFT), valid for ground state properties [8], that delivers a theoretically unjustified value of the band gap in the standard approach, even with exact KS potential [9–11].

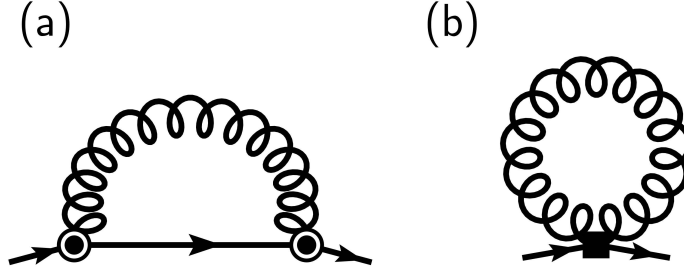


Figure 4.2 – Lowest-order diagrams for the contribution of the electron-phonon interaction to the zero-point renormalization of the gap ZPR_g . (a) Fan diagram, with two first-order screened electron-phonon vertices; (b) Debye-Waller diagram, with one second-order screened electron-phonon interaction vertex.

The breakthrough came from many-body perturbation theory, with the so-called GW approximation, first non-self-consistent (G_0W_0) by Hybertsen and Louie in 1986 [12], then twenty years later self-consistent (GW) [13] and further improved by accurate vertex corrections from electron-hole excitations (GWeh) [14]. The latter methodology, at the forefront for band-gap computations, delivers a 2%–10% accuracy, usually overestimating the experimental band gap. GWeh calculations are computationally very demanding, typically about two orders of magnitude more than G_0W_0 , itself two orders of magnitude more time-consuming than KS-DFT calculations, roughly speaking.

Despite being state-of-the-art, such studies ignored completely the electron-phonon interaction. The electron-phonon interaction drives most of the temperature dependence of the electronic structure of semiconductors and insulators, but also yields a zero-point motion gap modification at $T = 0$ K, often termed zero-point renormalization of the gap (ZPR_g) for historical reasons.

The ZPR_g had been examined 40 years ago, by Allen, Heine and Cardona (AHC) [1, 15] who clarified the early theories by Fan [16] and Antoncik [17]. Their approach is, like the GW approximation, rooted in many-body perturbation theory, where, at the lowest order, two diagrams contribute, see Fig. 4.2, the so-called “Fan” diagram, with two 1st-order electron-phonon vertices and the “Debye-Waller” diagram, with one 2nd-order electron-phonon vertex. In the context of semi-empirical calculations, the AHC method was applied to Si and Ge, introducing the adiabatic approximation, in which the phonon frequencies are neglected with respect to electronic eigenenergy differences and replaced by a small but non-vanishing imaginary broadening [1]. It was later extended without caution to GaAs and a few other III-V semiconductors [18, 19]. In this work, we present first-principles AHC ZPR_g calculations beyond the adiabatic approximation, for 30 materials. Comparing with experimental band gaps, we show that adding ZPR_g improves the GWeh first-principles band gap, and moreover, that the ZPR_g has the same order of magnitude as the G_0W_0 to GWeh correction for half of

the materials (typically materials with light atoms, e.g. O, N ...) on which GWeh has been tested. Hence, the GWeh level of sophistication misses its target for many materials with light atoms, if the ZPR_g is not taken into account. With it, the theoretical agreement with direct measurements of experimental ZPR_g is improved. This also demonstrates the crucial importance of phonon dynamics to reach this level of accuracy.

Indeed, first-principles calculations of the ZPR_g using the AHC theory are very challenging, and only started one decade ago, on a case-by-case basis (see Refs [3, 4, 20–28], and the [Supplementary Note 3](#)), usually relying on the adiabatic approximation, and without comparison with experimental data. An approach to the ZPR_g , alternative to the AHC one, relies on computations of the band gap at fixed, distorted geometries, for large supercells (see Refs [23, 29–36] and the [Supplementary Note 3](#)). As the adiabatic approximation is inherent in this approach, we denote it as ASC, for “adiabatic supercell”. A recent publication by Karsai and co-workers [37] presents ASC ZPR_g based on DFT values as well as based on G_0W_0 values for 18 semiconductors, with experimental comparison for 9 of them. Both AHC and ASC methodologies have been recently reviewed [38].

Although the adiabatic approximation had already been criticized by Poncé *et al.* in 2015 [3] ([Supplementary Note 5](#)) as causing an unphysical divergence of the adiabatic AHC expression for infrared-active materials with vanishing imaginary broadening, thus invalidating all adiabatic AHC calculations except for non-infrared-active materials, the full consequences of the adiabatic approximation have not yet been recognized for ASC. We will show that the non-adiabatic AHC approach outperforms the ASC approach, so that the predictions arising from the mechanism by which the ASC approach bypass the adiabatic AHC divergence problem for infrared-active materials are questionable. This is made clear by a generalized Fröhlich model with a few physical parameters, that can be determined either from first principles or from experimental data.

Although the full electronic and phonon band structures do not enter in this model, and the Debye-Waller diagram is ignored, for many materials it accounts for more than half the ZPR_g of the full first-principles non-adiabatic AHC ZPR_g . As this model depends crucially on non-adiabatic effects, it demonstrates the failure of the adiabatic hypothesis be it for the AHC or the ASC approach.

By the same token, we also show the domain of validity and accuracy of model Fröhlich large-polaron calculations based on the continuum hypothesis, that have been the subject of decades of research [39–45]. Such model Fröhlich Hamiltonian capture well the ZPR for about half of the materials in our list, characterized by their strong infrared activity, while it becomes less and less adequate for decreasing ionicity. In the present context, the Fröhlich large-polaron model provides an intuitive picture of the physics of the ZPR_g .

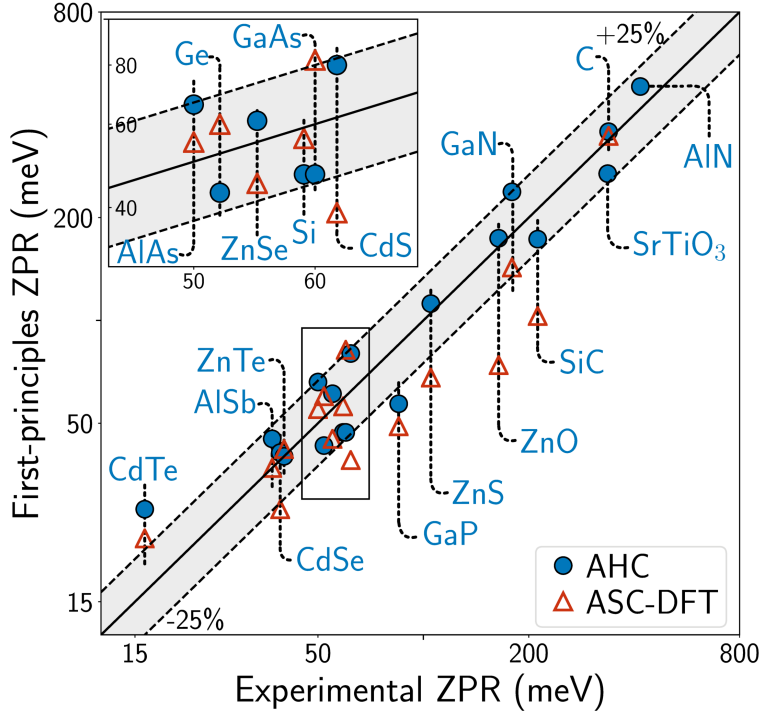


Figure 4.3 – Absolute values of first-principles band gap renormalization ZPR_g compared with experimental ones. Blue full circles: present calculations, using non-adiabatic AHC, based on KS-DFT ingredients; red empty triangles: adiabatic supercell KS-DFT results [37]. Dashed lines: limits at which the smallest of both ZPR is 25% smaller than the largest one (note that the scales are logarithmic). For consistency, the computed lattice ZPR_g^{lat} were added to both ASC and AHC. The majority of non-adiabatic AHC-based results fall within the 25% limits, while this is not the case for ASC-based ones. See numerical values in Supplementary Table 4.2.

4.2. Results

Zero-point renormalization: experiment versus first principles. Fig. 4.3 compares first-principles ZPR_g with experimental values. As described in the **METHODS** section, and in **Supplementary Note 2**, the correction due to zero-point motion effect on the lattice parameter, ZPR_g^{lat} , has been added to fixed volume results from both non-adiabatic AHC (present calculations) and ASC methodologies [37]. While for a few materials experimental ZPR_g values are well established, within 5–10%, globally, experimental uncertainty is larger, and can hardly be claimed to be better than 25% for the majority of materials, see **Supplementary Note 1**. This will be our tolerance.

Let us focus first on the ASC-based results. For the 16 materials present in both Ref. [37] and the experimental set described in the Supplementary Table 4.1, the ASC vs experimental discrepancy is more than 25% for more than half of the materials [46]. There is a global trend to underestimation by ASC, although CdTe is overestimated.

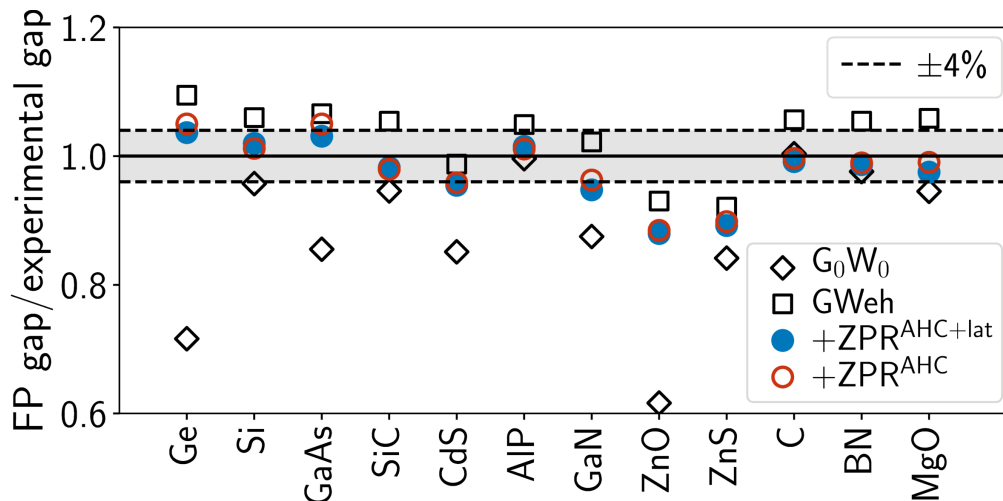


Figure 4.4 – Ratio between first-principles band gaps and experimental ones [47]. First-principles results without electron-phonon interaction[14] are based on G_0W_0 (diamonds) or on GWeh (squares). Blue full circles are the GWeh results to which AHC ZPR_g and lattice ZPR_g^{lat} have been added, while empty red circles are the GWeh results to which only AHC ZPR_g was added. GWeh usually overshoots, while adding the electron-phonon interaction gives better global agreement. See numerical values in Supplementary Table 4.3.

By contrast, the non-adiabatic AHC ZPR_g (blue full circles) and experimental ZPR_g agree with each other within 25% for 16 out of the 18 materials. The outliers are CdTe with a 43% overestimation by theory and GaP with a 33% underestimation. For none of these the discrepancy is a factor of two or larger. On the contrary, from the ASC approaches, several materials show underestimation of the ZPR_g by more than a factor of 2. The materials showing such large underestimation (CdS, ZnO, SiC) are all quite ionic, while more covalent materials (C, Si, Ge, AlSb, AlAs) are better described.

Therefore, Fig. 4.3 clearly shows that the non-adiabatic AHC approach performs significantly better than the ASC approach. AHC ZPR_g and ASC ZPR_g also differ by more than a factor of two for TiO_2 and MgO (see Supplementary Note 3), although no experimental ZPR_g is available for these materials to our knowledge. We now examine band gaps. Fig. 4.4 presents the ratio between first-principles band gaps and corresponding experimental values, for 12 materials. The best first-principles values at fixed equilibrium atomic position, from GWeh [14], are represented, as well as their non-adiabatic AHC ZPR corrected values.

For GWeh without ZPR_g , a 4% agreement is obtained only for two materials (CdS and GaN). There is indeed a clear, albeit small, tendency of GWeh to overestimate the band gap value, except for the 3 materials containing shallow core d-electrons (ZnO, ZnS and CdS) that are underestimated. By contrast, if the non-adiabatic AHC ZPR_g is added to the GWeh data (blue dots), a 4% agreement is obtained for 9 out of the 12 materials (8 if ZPR_g^{lat} is not included).

For ZnO and ZnS, with a final 10–12% underestimation, and CdS with a 5% underestimation, we question the GWeh ability to produce accurate fixed-geometry band gaps at the level obtained for the other materials, due to the presence of rather localized 3d electrons in Zn and 4d electrons in CdS.

As a final lesson from Fig. 4.4, we note that for 4 out of the 12 materials (SiC, AlP, C, and BN), the ZPR_g is of similar size than the G_0W_0 to GWeh correction, and it is a significant fraction of it also for Si, GaN and MgO. As mentioned earlier, GWeh calculations are much more time-consuming than G_0W_0 calculations, possibly even more time-consuming than ZPR calculations (although we have not attempted to make a fair comparison). Thus, for materials containing light elements, first row and likely second row (e.g. AlP) in the periodic table, GWeh calculations miss their target if not accompanied by ZPR_g calculations. A review of the variance and accuracy of G_0W_0 calculations for these materials is discussed in Ref. [48].

Supplementary Table 4.5 gathers our full set of ZPR results, beyond those present in the ASC or experimental sets. It also includes 10 oxides, while the experimental set only includes three materials containing oxygen (ZnO, MgO and SrTiO₃), and there are none in Karsai’s ASC set. The ZPR_g of the band gap for materials containing light elements (O or lighter) is between -157 meV (ZnO) and -699 meV (BeO), while, relatively to the experimental band gap, it ranges from -4.6% (ZnO) to -10.8% (TiO_{2-t}). This can hardly be ignored in accurate calculations of the gap.

Generalized Fröhlich model.

We now come back to the physics from which the ZPR_g originates, and explain our earlier observation that the ASC describes reasonably well the more covalent materials, but can fail badly for ionic materials. We argue that, for many polar materials, the ZPR_g is dominated by the diverging electron-phonon interaction between zone-center LO phonons and electrons close to the band edges, and the slow (non-adiabatic) response of the latter: the effects due to comparatively fast phonons are crucial. This was already the message from Fröhlich [39] and Feynman [40], decades ago, initiating large-polaron studies. However, large-scale assessment of the adequacy of Fröhlich model for real materials is lacking. Indeed, the available flavors of Fröhlich model, based on continuum (macroscopic) electrostatic interactions, do not cover altogether degenerate and anisotropic electronic band extrema or are restricted to only one phonon branch, unlike most real materials [49–52].

In this respect, we introduce now a generalized Fröhlich model, gFr, whose form is deduced from first-principles equations in the long-wavelength limit (continuum approximation), see [Supplementary Note 5](#). Such model covers all situations and still uses as input only

long-wavelength parameters, that can be determined either from first principles ([Supplementary Note 6](#)) or from experiment ([Supplementary Note 7](#)). We then find the corresponding ZPR_g from perturbation theory at the lowest order, evaluate it for our set of materials, and compares its results to the ZPR_g obtained from full first-principles computations. The corresponding Hamiltonian writes (see [Supplementary Note 5](#) for detailed explanations):

$$\hat{H}^{\text{gFr}} = \hat{H}_{\text{el}}^{\text{gFr}} + \hat{H}_{\text{ph}}^{\text{gFr}} + \hat{H}_{\text{EPI}}^{\text{gFr}}, \quad (4.2.1)$$

with (i) an electronic part

$$\hat{H}_{\text{el}}^{\text{gFr}} = \sum_{\mathbf{k}n} \frac{k^2}{2m_n^*(\hat{\mathbf{k}})} \hat{c}_{\mathbf{k}n}^\dagger \hat{c}_{\mathbf{k}n}, \quad (4.2.2)$$

that includes direction-dependent effective masses $m_n^*(\hat{\mathbf{k}})$, governed by so-called Luttinger parameters in case of degeneracy, electronic creation and annihilation operators, $\hat{c}_{\mathbf{k}n}^\dagger$ and $\hat{c}_{\mathbf{k}n}$, with \mathbf{k} the electron wavevector and n the band index, (ii) the multi-branch phonon part,

$$\hat{H}_{\text{ph}}^{\text{gFr}} = \sum_{\mathbf{q}j} \omega_{j0}(\hat{\mathbf{q}}) \hat{a}_{\mathbf{q}j}^\dagger \hat{a}_{\mathbf{q}j}, \quad (4.2.3)$$

with direction-dependent phonon frequencies $\omega_{j0}(\hat{\mathbf{q}})$, phonon creation and annihilation operators, $\hat{a}_{\mathbf{q}j}^\dagger$ and $\hat{a}_{\mathbf{q}j}$, with \mathbf{q} the phonon wavevector and j the branch index, and finally (iii) the electron-phonon interaction part

$$\hat{H}_{\text{EPI}}^{\text{gFr}} = \sum_{\mathbf{q}j, \mathbf{k}n'n} g^{\text{gFr}}(\mathbf{q}j, \mathbf{k}n'n) \hat{c}_{\mathbf{k}+\mathbf{q}n'}^\dagger \hat{c}_{\mathbf{k}n} (\hat{a}_{\mathbf{q}j} + \hat{a}_{-\mathbf{q}j}^\dagger). \quad (4.2.4)$$

The \mathbf{k} and \mathbf{q} sums run over the Brillouin zone. The sum over n and n' runs only over the bands that connect to the degenerate extremum, renumbered from 1 to n_{deg} . The generalized Fröhlich electron-phonon interaction [[53](#), [54](#)] is

$$g^{\text{gFr}}(\mathbf{q}j, \mathbf{k}n'n) = \frac{i}{q} \frac{4\pi}{\Omega_0} \left(\frac{1}{2\omega_{j0}(\hat{\mathbf{q}}) V_{\text{BvK}}} \right)^{1/2} \frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \times \sum_m s_{n'm}(\hat{\mathbf{k}}') (s_{nm}(\hat{\mathbf{k}}))^*. \quad (4.2.5)$$

This expression depends on the directions $\hat{\mathbf{k}}$, $\hat{\mathbf{q}}$, and $\hat{\mathbf{k}}'$ ($\mathbf{k}' = \mathbf{k} + \mathbf{q}$), but not explicitly on their norm (hence only long-wavelength parameters are used), except for the $\frac{1}{q}$ factor. The electron-phonon part also depends only on few quantities: the Born effective charges (entering the mode-polarity vectors \mathbf{p}_j), the macroscopic dielectric tensor ϵ^∞ , and the phonon frequencies ω_{j0} , the primitive cell volume Ω_0 , the Born-von Karman normalisation volume V_{BvK} corresponding to the \mathbf{k} and \mathbf{q} samplings. The s tensors are symmetry-dependent unitary matrices, similar to spherical harmonics. Eqs. (4.2.1–4.2.5) define our generalized Fröhlich Hamiltonian. Although we will focus on its $T = 0$ properties within perturbation theory, such Hamiltonian could be studied for many different purposes (non-zero T, mobility,

optical responses ...), like the original Fröhlich model, for representative materials using first-principles or experimental parameters.

The corresponding $\text{ZPR}_c^{\text{gFr}}$ can be obtained with a perturbation treatment (Supplementary Note 5), giving

$$\begin{aligned} \text{ZPR}_c^{\text{gFr}} &= -\sum_{jn} \frac{1}{\sqrt{2}\Omega_0 n_{\text{deg}}} \int_{4\pi} d\hat{\mathbf{q}} \left(m_n^*(\hat{\mathbf{q}})\right)^{1/2} \\ &\times \left(\omega_{j0}(\hat{\mathbf{q}})\right)^{-3/2} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})}\right)^2. \end{aligned} \quad (4.2.6)$$

A similar expression exists for the valence $\text{ZPR}_v^{\text{gFr}}$. The few material parameters needed in Eq. (4.2.6) can be obtained from experimental measurements, but are most easily computed from first principles, using density-functional perturbation theory with calculations only at $\mathbf{q} = \Gamma$ (e.g. no phonon band structure calculation). Eq. (4.2.6) can be evaluated for all band extrema in our set of materials, irrespective of whether the extrema are located at Γ or other points in the Brillouin Zone (e.g. X for the valence band of many oxides, with anisotropic effective mass), whether they are degenerate (e.g. the three-fold degeneracy of the top of the valence band of many III-V or II-VI compounds), and irrespective of the number of phonon branches (e.g. 3 different LO frequencies for TiO_2 , moreover varying with the direction along which $\mathbf{q} \rightarrow 0$).

Fig. 4.5 compares the band gap ZPR from the first-principles non-adiabatic AHC methodology and from the generalized Fröhlich model. The 30 materials can be grouped into five sets, based on their ionicity: 11 materials containing oxygen, rather ionic, for which the Born effective charges and the ZPR are quite large, 6 materials containing chalcogenides, also rather ionic, 4 materials containing nitrogen and 5 III-V materials, less ionic, and 4 materials from group-IV elements, non-ionic, except SiC.

For oxygen-based materials, the ZPR ranges from 150 meV to 700 meV, and the gFr model captures this very well, with less than 25% error, with only one exception, BeO. The chalcogenide materials are also reasonably well described by the gFr model, capturing at least two-third of the ZPR. Globally their ZPR is smaller (note the logarithmic scale).

For the nitride materials and for SiC, the gFr captures about 50% of the quite large ZPR (between 176 meV and 406 meV). The adequacy of the gFr model decreases still with the III-V materials and the three non-ionic IV materials. In the latter case, the vanishing Born effective charges result in a null ZPR within the gFr model. (These three materials are omitted from Fig. 4.5).

For the oxydes and chalcogenides, the ZPR is thus dominated by the zone-center parameters (including the phonon frequencies), and the physics corresponds to the one of the large-polaron picture [40], namely, the slow electron motion is correlated to a phonon cloud

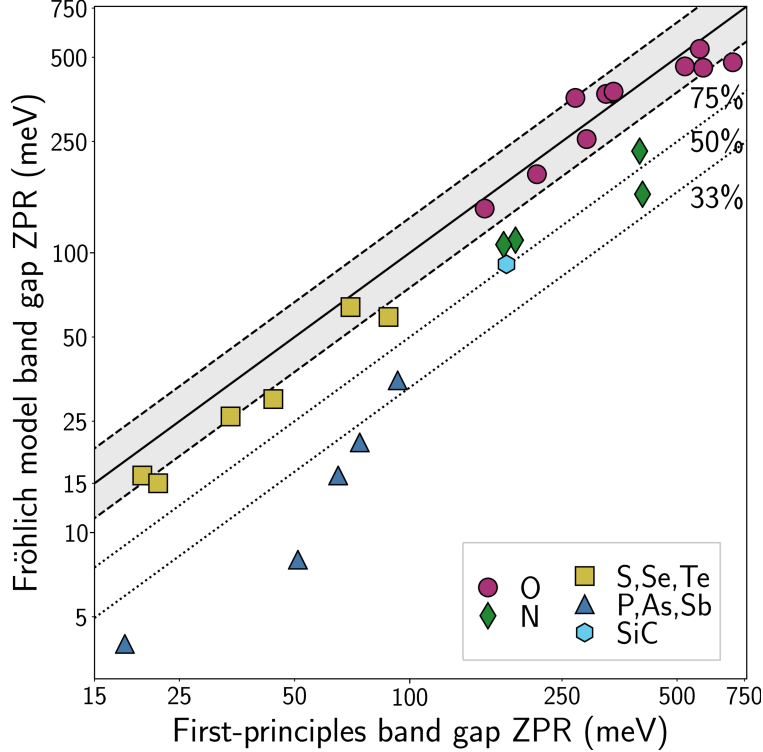


Figure 4.5 – Zero-point renormalization of the band gap: comparison between first-principles AHC values and generalized Fröhlich model. See Eq. (4.2.6). Data for 27 materials are represented. Non-ionic group IV materials are omitted. The markers identify materials of similar ionicity: oxides (purple circles), other II-VI materials containing S, Se or Te (yellow squares), nitrides (green diamonds), other III-V materials, containing P, As, or Sb (dark blue triangles), and group IV material SiC (light blue hexagon). Dashed lines: limits at which the smallest of both ZPR is 25% smaller than the largest one. Dotted lines: limits at which the Fröhlich model ZPR misses respectively 50% and 66% of the AHC values (note that the scales are logarithmic). See numerical values in Supplementary Table 4.5.

that dynamically adjusts to it. This physics is completely absent from the ASC approach. Even for nitrides, the gFr describes a significant fraction of the ZPR.

A perfect agreement between the non-adiabatic AHC first-principles ZPR and the generalized Fröhlich model ZPR is not expected. Indeed, differences can arise from different effects: lack of dominance of the Fröhlich electron-phonon interaction in some regions of the Brillouin Zone, departure from parabolicity of the electronic structure (obviously, the electronic structure must be periodic so that the parabolic behavior does not extend to infinity), interband contributions, phonon band dispersion, incomplete cancellation between the Debye-Waller and the acoustic phonon mode contribution.

It is actually surprising to see that for so many materials, the generalized Fröhlich model matches largely the first-principles AHC results. Anyhow, as a conclusion for this section, for a large number of materials, we have validated, a posteriori and from first principles, the

relevance of large-polaron research based on Fröhlich model despite the numerous approximations on which it relies.

4.3. Discussion

We focus on the mechanism by which the AHC divergence of the ZPR in the adiabatic case for infrared-active materials [3] is avoided, either using the ASC methodology or using the non-adiabatic AHC methodology. As Fröhlich and Feynman have cautioned us [39, 40], and already mentioned briefly in previous sections, the dynamics of the “slow” electron is crucial in this electron-phonon problem.

In the ASC approach, the bypass of this divergence can be understood as follows, see Fig. 4.6(a). Consider a long-wavelength fluctuation of the atomic positions, frozen in time. At large but finite wavelength, the potential is periodically lowered in some regions of space and increased in some other regions of space, in an oscillatory manner with periodicity $\Delta L \propto 1/q$, where \mathbf{q} is the small wavevector of the fluctuation, see Fig. 4.6(a) “LO phonon potential” part. Without such long-wavelength potential, the electron at the minimum of the conduction band has a Bloch type wavefunction, with an envelope phase factor characterized by the wavevector \mathbf{k}_c multiplying a lattice-periodic function. Its density is lattice periodic. With such long-wavelength potential, as a function of the amplitude of the atomic displacements, the corresponding electronic eigenenergy changes first quadratically (as the average of the lowering and increase of potential for this Bloch wavefunction forbids a linear behaviour except in case of degeneracy), but for larger amplitudes, it behaves linearly, as the electron localizes in the lowered potential region and the minimum of the potential is linear in the amplitude of the atomic displacements. This is referred to as “nonquadratic coupling” in Ref. [35]. A wavepacket is formed, by combining Bloch wavefunctions with similar lattice periodic functions but slightly different wavevectors (\mathbf{k}_c , $\mathbf{k}_c + \mathbf{q}$, $\mathbf{k}_c - \mathbf{q}$, etc), coming from a small interval of energy $\Delta\epsilon^0 \propto q^2 \propto 1/(\Delta L)^2$, see Fig. 4.6(a) “Wavefunction” part. This nonquadratic effect is actually illustrated in Fig. 4 of Ref. [4] (see the frozen-phonon eigenvalues), as well as in Fig. 2 of Ref. [35].

By contrast, in the time-dependent case, as illustrated in Fig. 4.6(b), the wavepacket will require a characteristic time $\Delta\tau$, to form or to displace. This will be given by the Heisenberg uncertainty relation, $\Delta\epsilon^0\Delta\tau \gtrsim \hbar$, hence $\Delta\tau \gtrsim \hbar/\Delta\epsilon^0 \propto \Delta L^2$. For long wavelengths, the characteristic time diverges. As soon as $\Delta\tau$ is larger than the phonon characteristic time $\tau_{\text{ph}} \sim 1/\omega_{\text{LO}}$, the “slow” electron will lag behind the phonon, and the static or adiabatic picture described above is no longer valid.

In all adiabatic approaches, either AHC or ASC, the electron is always supposed to have the time to adjust to the change of potential, in contradiction with the time-energy uncertainty principle. Furthermore, the adiabatic AHC approach only considers the quadratic

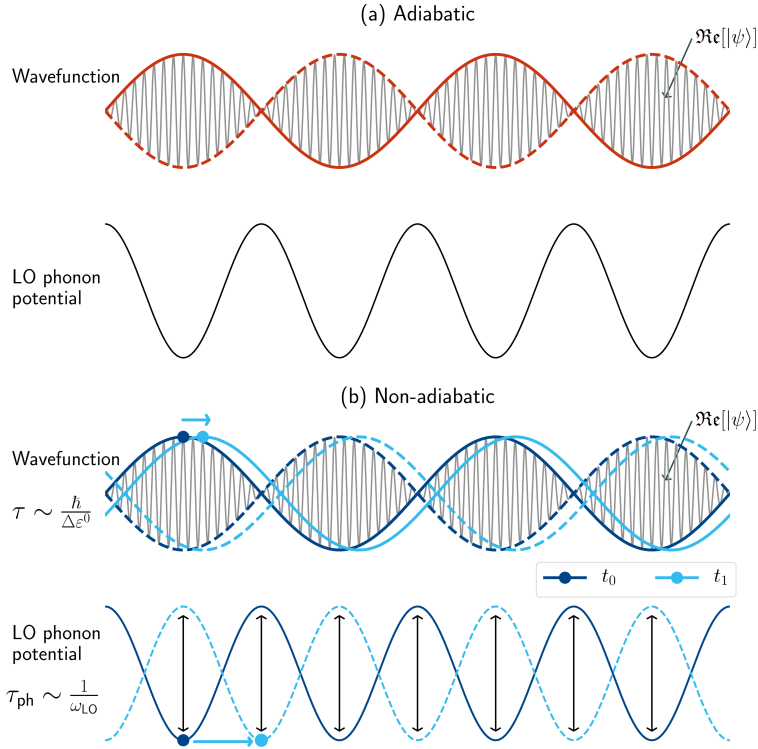


Figure 4.6 – Schema of the long-wave phonon-induced potential and corresponding electronic wavefunction. Real part (oscillating with lattice periodicity) and envelope are represented. (a) Adiabatic case, (b) non-adiabatic, time-dependent case. In the non-adiabatic case, the electron does not have the time to adjust to the change of potential, see text.

region for the above-mentioned dependence of eigenvalues with respect to amplitudes of displacements. This results in a diverging term [3]. At variance with the AHC case, the ASC approach samples a whole set of amplitudes, including the onset of the asymptotic linear regime, in which case the divergence does not build up. However, this ASC picture does not capture the real physical mechanism that prevent the divergence to occur, the impossibility for the electron to follow the phonon dynamics, that we have highlighted above. By contrast, such physical mechanism is present both in the non-adiabatic AHC approach and in the (generalized) Fröhlich model: the “slow” electron does not follow adiabatically (instantaneously) the atomic motion. The divergence of the adiabatic AHC is indeed avoided in the non-adiabatic picture by taking into account the non-zero phonon frequencies.

Thus the ASC avoids the adiabatic AHC divergence for the wrong reason, which explains its poor predictive capability for the more ionic materials emphasized by Fig. 4.3. To be clear, we do not pretend the nonquadratic effects are all absent, but the non-adiabatic effects have precedence, at least for materials with significant infrared activity, and the non-quadratic localization effects will be observed only if the electrons have the time to physically react. The shortcomings of the ASC approach are further developed in the [Supplementary](#)

Discussion, and as a consequence of such understanding, all the results obtained for strongly infrared-active materials using the adiabatic frozen-phonon supercell methodology should be questioned.

For non-infrared-active materials, the physical picture that we have outlined, namely the inability of slow electrons to follow the dynamics of fast phonons, is still present, but does not play such a crucial role: the electron-phonon interaction by itself does not diverge in the long-wavelength limit as compared to the infrared-active electron-phonon interaction, see Eq. (4.2.5), only the denominator of the Fan self-energy diverges, which nevertheless results in an integrable ZPR [3]. In such case, neglecting non-adiabatic effects, as in the ASC approach, is only one among many approximations done to obtain the ZPR.

Beyond the discovery of the predominance of non-adiabatic effects in the zero-point renormalization of the band gap for many materials, in the present large-scale first-principles study of this effect, we have established that electron-phonon interaction diminishes the band gap by 5% to 10% for materials containing light atoms like N or O (up to 0.7 eV for BeO), a decrease that cannot be ignored in accurate calculations of the gap. Our methodology, the non-adiabatic Allen-Heine-Cardona approach, has been validated by showing that, for nearly all materials for which experimental data exists, it achieves quantitative agreement (within 25%) for this property.

We have also shown that most of the discrepancies with respect to experimental data of the (arguably) best available methodology for the first-principles band-gap computation, denoted GWeh, originate from the first-principles zero-point renormalization: after including it, the average overestimation from GWeh nearly vanishes. There are some exceptions, materials in which transition metals are present, for which the addition of zero-point renormalization worsens the agreement of the band gap. For the latter materials, we believe that the GWeh approach is not accurate enough.

4.4. Methods

First-principles electronic and phonon band structures. Calculations have been performed using ABINIT [55] with norm-conserving pseudopotentials and a plane-wave basis set. Supplementary Table 4.1 provides calculation parameters: plane-wave kinetic cut-off energy, electronic wavevector sampling in the BZ, and largest phonon wavevector sampling in the BZ. For most of the materials, the GGA-PBE exchange-correlation functional [56] has been used and the pseudopotentials have been taken from the PseudoDojo project [57]. For diamond, BN-zb, and AlN-wz, results reported here come from Ref. [3], where the LDA has been used, with other types of pseudopotentials.

The calculations have been performed at the theoretical optimized lattice parameter, except for Ge for which the gap closes at such parameter, for GaP, as at such parameter the

conduction band presents unphysical quasi-degenerate valleys, and for TiO_2 , as the GGA-PBE predicted structure is unstable [58]. For these, we have used the experimental lattice parameter. The case of SrTiO_3 is specific and will be explained later.

Density-functional perturbation theory [59–61] has been used for the phonon frequencies, dielectric tensors, Born effective charges, effective masses, and electron-phonon matrix elements.

First-principles calculations of zero-point renormalization. We first detail the method used for the AHC calculations. In the many-body perturbation theory approach, an electronic self-energy Σ appears due to the electron-phonon interaction, with Fan and Debye-Waller contributions at the lowest order of perturbation [38]:

$$\Sigma_{\mathbf{k}n}(\omega) = \Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega) + \Sigma_{\mathbf{k}n}^{\text{DW}}. \quad (4.4.1)$$

The Hartree atomic unit system is used throughout ($\hbar = m_e = e = 1$). An electronic state is characterized by \mathbf{k} , its wavevector, and n , its band index, ω being the frequency. These two contributions correspond to the two diagrams presented in Fig. 4.2.

Approximating the electronic Green’s function by its non-interacting KS-DFT counterpart without electron-phonon interaction, gives the standard result for the $T = 0$ K retarded Fan self-energy [38]:

$$\Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega) = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}j}^{\text{BZ}} \sum_{n'} |\langle \mathbf{k} + \mathbf{q}n' | H_{\mathbf{q}j}^{(1)} | \mathbf{k}n \rangle|^2 \times \left[\frac{1 - f_{\mathbf{k}+\mathbf{q}n'}}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'} - \omega_{\mathbf{q}j} + i\eta} + \frac{f_{\mathbf{k}+\mathbf{q}n'}}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'} + \omega_{\mathbf{q}j} + i\eta} \right]. \quad (4.4.2)$$

In this expression, contributions from phonon modes with harmonic phonon energy $\omega_{\mathbf{q}j}$ are summed for all branches j , and wavevectors \mathbf{q} , in the entire Brillouin Zone (BZ). The limit for infinite number $N_{\mathbf{q}}$ of wavevectors (homogeneous sampling) is implied. Contributions from transitions to electronic states $\langle \mathbf{k} + \mathbf{q}n' |$ with KS-DFT electron energy $\varepsilon_{\mathbf{k}+\mathbf{q}n'}$ and occupation number $f_{\mathbf{k}+\mathbf{q}n'}$ (1 for valence, 0 for conduction, at $T = 0$ K) are summed for all bands n' (valence and conduction). The $H_{\mathbf{q}j}^{(1)}$ is the self-consistent change of potential due to the $\mathbf{q}j$ -phonon [38]. Limit of this expression for vanishing positive η is implied. For the Debye-Waller self-energy, $\Sigma_{\mathbf{k}n}^{\text{DW}}$, we refer to Refs. [26, 38].

In the non-adiabatic AHC approach, the ZPR is obtained directly from the real part of the self-energy, Eq. (4.4.1), evaluated at $\omega = \varepsilon_{\mathbf{k}n}$ [38]:

$$\text{ZPR}_{\mathbf{k}n}^{\text{AHC}} = \Re \Sigma_{\mathbf{k}n}(\omega = \varepsilon_{\mathbf{k}n}). \quad (4.4.3)$$

If the adiabatic approximation is made, the phonon frequencies $\omega_{\mathbf{q}j}$ are considered small with respect to eigenenergy differences in the denominator of Eq. (4.4.2) and are simply dropped, while a finite η , usually 0.1 eV, is kept. With a vanishing η , the adiabatic AHC ZPR at band edges diverges for infrared-active materials, see [Supplementary Note 5](#).

Summing the Fan and Debye-Waller self-energies, and working also with the rigid-ion approximation for the Debye-Waller contribution delivers the non-adiabatic AHC ZPR, given explicitly in Eqs. (16) and (17) of Ref. [3]. We do not work with the approach called dynamical AHC, also mentioned in Ref. [3]. It corresponds to Eq. (166) of Ref. [38]. Both non-adiabatic and dynamical AHC ZPR flavors were studied e.g. in Ref. [4], but the comparison with the diagrammatic quantum Monte Carlo results for the Fröhlich model, see e.g. Fig. 1 of Ref. [28], is clearly in favor of the non-adiabatic AHC approach. Actually, this also constitutes a counter-argument to the claim by Cannuccia and Marini, Ref. [22], that band theory might not apply to carbon-based nanostructures. As shown in Ref. [28], the cumulant expansion results for the spectral function demonstrates that the dynamical AHC spectral function is unphysical, with only one wrongly placed satellite. The physical content of the present non-adiabatic AHC theory, focusing on the crucial role of the LO phonons, is very different from the physical analysis of Ref. [22], based on the dynamical AHC theory.

The imaginary smearing of the denominator in the ZPR computation is 0.01eV, except for SiC, where 0.001 eV is used. Other technical details are similar to previous studies by some of ours, e.g. Refs. [3] and [28].

The dependence of the electronic structure on zero-point lattice parameter corrections is computed from

$$\text{ZPR}_{\mathbf{kn}}^{\text{lat}} = \varepsilon_{\mathbf{kn}}(\{\mathbf{R}_i^{T=0}\}) - \varepsilon_{\mathbf{kn}}(\{\mathbf{R}_i^{\text{fix}}\}), \quad (4.4.4)$$

where the lattice parameters $\{\mathbf{R}_i^{\text{fix}}\}$ minimize the Born-Oppenheimer energy without phonon contribution, while $\{\mathbf{R}_i^{T=0}\}$ minimizes the free energy that includes zero-point phonon contributions, see [Supplementary Note 2](#).

At variance with the AHC approach, in the ASC the temperature-dependent average band edges (here written for the bottom of the conduction band) are obtained from

$$\langle \varepsilon_c(T) \rangle = Z_I^{-1} \sum_m \exp(-\beta E_m) \langle \varepsilon_c \rangle_m, \quad (4.4.5)$$

where $\beta = k_B T$ (with k_B the Boltzmann constant), T the temperature, Z_I the canonical partition function among the quantum nuclear states m with energies E_m ($Z_I = \sum_m \exp(-\beta E_m)$), and $\langle \varepsilon_c \rangle_m$ the band edge average taken over the corresponding many-body nuclear wavefunction. At zero Kelvin, this gives an instantaneous average of the band edge value over zero-point atomic displacements, computed while the electron is NOT present in the conduction band (or hole in the valence band) thus suppressing all correlations between the phonons and the added (or removed) electron.

Convergence of the calculation. As previously noted [3, 28], the sampling of phonon wavevectors in the Brillouin zone is a delicate issue, and has been thoroughly analyzed in Sec. IV.B.2 of Ref. [3]. In particular, for infrared-active materials treated with the non-adiabatic effects, at the band structure extrema, a $N_{\mathbf{q}}^{-1/3}$ convergence of the value is obtained. We have taken advantage of the knowledge acquired in Ref. [3] to accelerate the convergence by three different methodologies. In the first one, we fit the $N_{\mathbf{q}}^{-1/3}$ behavior for grids of different sizes, and extrapolate to infinite $N_{\mathbf{q}}$. In the second one, we estimate the missing contribution to the integral around $\mathbf{q} = 0$, at the lowest order, see [Supplementary Methods B](#), using ingredients similar to those needed for the generalized Fröhlich model, except the effective masses. For ZnO and SrTiO₃, a third correcting scheme, further refining the region around the band edge with an extremely fine grid, is used.

The special case of SrTiO₃. While phonons in most materials in the present study are well addressed within the harmonic approximation, this is not the case for SrTiO₃. This material is found in the cubic perovskite structure at room temperature, and undergoes a transition to a tetragonal phase below 150 K, characterized by tilting of the TiO₆ octahedra [62]. Within our first-principles scheme in the adiabatic approximation, the cubic phase remains unstable with respect to tilting of the octahedra. Quantum fluctuations of the atomic positions actually play a critical role in stabilizing the cubic phase at high temperature [63], as well as suppressing the ferroelectric phase at low temperature [64, 65].

Since SrTiO₃ is however a material for which large polaron effects are clearly identified, see e.g. Ref. [51] and references therein, we decided to study it as well. We addressed the anharmonic stabilization problem using the state-of-the-art TDep methodology [66]. We used VASP molecular dynamics to generate 40 configurations in a 2x2x2 cubic cell of STO at 300 K, producing 20,000 steps with 2fs per step and sampling the 40 configurations out of the last 5000 steps. Then we computed the forces with ABINIT and performed TDep calculations with the ALAMODE code [67]. Our calculation stabilizes the acoustic phonon branches and yields a phonon band structure in good agreement with experimental data.

Sources of discrepancies between experiment and theory. The anharmonic corrections to phonon frequencies are not the only reasons for potential differences between the experimental $ZPR_{\mathbf{g}}$ and our non-adiabatic AHC $ZPR_{\mathbf{g}}$ values. The following phenomena may also play a role: (1) the rigid-ion approximation [23, 26]; (2) the nonquadratic behaviour of the eigenenergies with collective displacements of the nuclei, in reference to the ASC, especially emphasized in Ref. [35]; (3) the reliance on GGA-PBE eigenenergies and eigenfunctions, instead of more accurate (e.g. GW) ones, as in Ref. [32], also discussed in Ref. [37]; (4) self-trapping effects, overcoming the quantum fluctuations, yielding small

polarons, see e.g. Ref. [68]. There is still little knowledge about each of these effects when correctly combined to predict the ZPR_g beyond the AHC picture.

As an example, in Ref. [37], the difference between the ASC-PBE and the ASC-GW was argued to be only a few meV, but a more careful look at their values show that it is often bigger than 10% of the ASC-PBE. Unfortunately, the convergence of the ASC-GW results with respect to supercell size could not be convincingly achieved in Ref. [37]. It remains to be seen whether a non-adiabatic AHC treatment based on GW matrix elements would differ by such relative ratio, see the [Supplementary Discussion](#). Altogether, it would be hard to claim more than 25% accuracy with respect to experimental data, from our non-adiabatic AHC ZPR_g calculations. Together with the experimental uncertainties, this explains our choice for the 25% accuracy comparative limit used in Fig. 4.3.

Supplementary Information

In the [Supplementary Information](#), the tables with the numerical data used to create the figures in the main text have been collected, as well as information and comments about such data (including the corresponding bibliographical references for results that have not been computed in the present work). Experimental data for the ZPR are critically examined and discussed. The gap ZPR values are also discussed with respect to previously published values, highlighting the importance of the control of phonon wavevector grid convergence and imaginary broadening. The non-quadratic effects and GW approximation versus DFT approach are further discussed in the light of previously published data.

Data Availability

The numerical data used to create all the figures in the main text have been collected in the [Supplementary Tables](#).

Code Availability

ABINIT is available under GNU General Public Licence from the ABINIT web site (<http://www.abinit.org>).

Acknowledgements

We acknowledge fruitful discussions with Y. Gillet and S. Ponc e. This work has been supported by the Fonds de la Recherche Scientifique (FRS-FNRS Belgium) through the PdR Grant No. T.0238.13 - AIXPHO, the PdR Grant No. T.0103.19 - ALPS, the Fonds de Recherche du Qu ebec Nature et Technologie (FRQ-NT), the Natural Sciences and Engineering Research Council of Canada (NSERC) under grants RGPIN-2016-06666. Computational

resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL), the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the FRS-FNRS under Grant No. 2.5020.11, the Tier-1 supercomputer of the Fédération Wallonie-Bruxelles, infrastructure funded by the Walloon Region under the grant agreement No. 1117545, as well as the Canadian Foundation for Innovation, the Ministère de l'Éducation des Loisirs et du Sport (Québec), Calcul Québec, and Compute Canada. This work was supported by the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM) at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05CH11231, as part of the Computational Materials Sciences Program. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Author Contributions

A. Miglio has conducted calculations for most oxide materials, with help from M. Giantomassi. V. Brousseau-Couture has conducted calculations for most other materials. G. Antonius and Yang-Hao Chan have conducted calculations for SrTiO_3 and ZnO . E. Godbout has conducted calculations of the lattice ZPR of 6 materials. X. Gonze has worked out the generalized Fröhlich model and perturbative treatment. X. Gonze and M. Côté have supervised the work. All authors have contributed to the writing of the manuscript.

Competing Interests

The authors declare no competing financial or non-financial interests.

Supplementary Information

Supplementary Figures

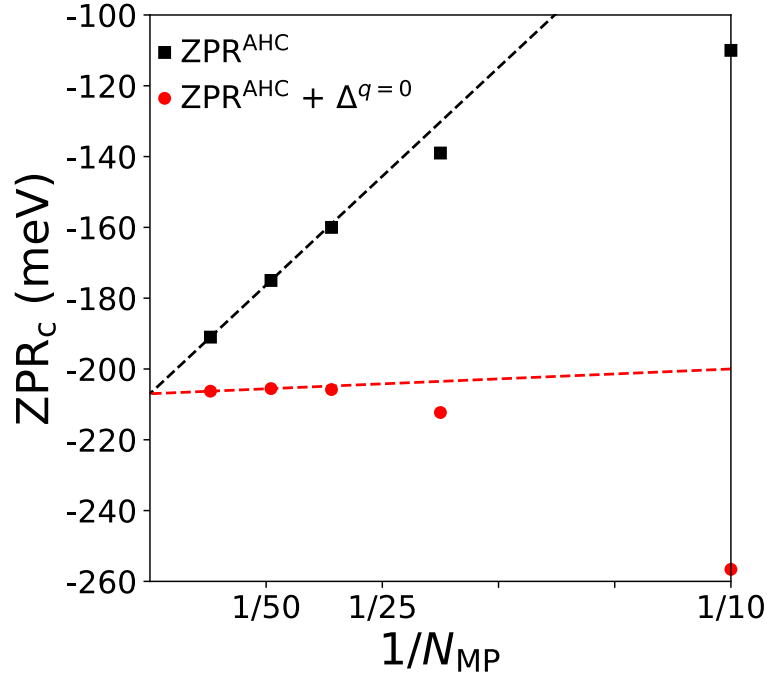


Figure 4.7 – The ZPR_c^{AHC} of MgO is computed without correction (black squares), and then with correction Supplementary Eq. (S4.3) (red circles), for different phonon wavevector Monkhorst-Pack \mathbf{q} -grid samplings, characterized by N_{MP} . The points correspond to $N_{\text{MP}} = 96, 48, 32, 20$ and 10. Dotted lines show linear extrapolation with respect to $1/N_{\text{MP}}$ from the $N_{\text{MP}} = 48$ and $N_{\text{MP}} = 96$ grids, to the infinitely large N_{MP} value.

Supplementary Tables

Table 4.1 – List of the thirty materials considered in the present study, with parameters used in the first-principles computations. From left to right: material; Materials Project ID [69]; energy cut-off for the planewave basis set; wavevector sampling for the electronic states; wavevector sampling for the phonon states. Abbreviations for crystal structures: sc=simple cubic, rs=rocksalt, zb=zinc-blende, w=wurtzite, dia=diamond, t=tetragonal, rh=rhombohedral. See [Supplementary Methods A](#).

Material	ID	E_{cut} (Ha)	k -point sampling	q -point sampling
Ge-dia	mp-32	40	6x6x6 (x4shifts)	48x48x48
Si-dia	mp-149	20	6x6x6 (x4shifts)	100x100x100
GaAs-zb	mp-2534	40	6x6x6 (x4shifts)	64x64x64
CdTe-zb	mp-406	50	6x6x6 (x4shifts)	48x48x48
AlSb-zb	mp-2624	40	6x6x6 (x4shifts)	48x48x48
CdSe-zb	mp-2691	50	6x6x6 (x4shifts)	48x48x48
AlAs-zb	mp-2172	40	6x6x6 (x4shifts)	48x48x48
ZnTe-zb	mp-2176	40	6x6x6 (x4shifts)	48x48x48
GaP-zb	mp-2690	40	6x6x6 (x4shifts)	48x48x48
SiC-zb	mp-8062	35	6x6x6 (x4shifts)	48x48x48
CdS-zb	mp-2469	45	6x6x6 (x4shifts)	48x48x48
AlP-zb	mp-1550	25	8x8x8 (x4shifts)	48x48x48
ZnSe-zb	mp-1190	40	6x6x6 (x4shifts)	48x48x48
TiO ₂ -t	mp-2657	40	6x6x8	20x20x32
SrTiO ₃ -sc	mp-5229	70	8x8x8	48x48x48
GaN-w	mp-804	40	8x8x8	64x64x64
GaN-zb	mp-830	40	6x6x6 (x4shifts)	48x48x48
ZnO-w	mp-2133	50	6x6x6	48x48x48
SnO ₂ -t	mp-856	40	6x6x8	20x20x32
ZnS-zb	mp-10695	40	6x6x6 (x4shifts)	48x48x48
BaO-rs	mp-1342	40	8x8x8	32x32x32
SrO-rs	mp-2472	40	8x8x8	32x32x32
C-dia	mp-66	30	6x6x6 (x4shifts)	125x125x125
AlN-w	mp-661	35	6x6x6	34x34x34
BN-zb	mp-1639	35	8x8x8 (x4shifts)	100x100x100
CaO-rs	mp-2605	40	8x8x8	32x32x32
Li ₂ O	mp-1960	50	8x8x8	32x32x32
MgO-rs	mp-1265	50	8x8x8 (x4shifts)	96x96x96
SiO ₂ -t	mp-6947	40	6x6x8	20x20x32
BeO-w	mp-2542	50	12x12x6	32x32x16

Table 4.2 – For eighteen materials: experimental fundamental gap and type (d=direct, i=indirect); experimental zero-point renormalization of the gap; lattice contribution to the zero-point renormalization of the gap; first-principles AHC zero-point renormalization of the gap; ratio R between sum of AHC and lattice values and experimental values $(\text{ZPR}_g^{\text{AHC}} + \text{ZPR}_g^{\text{lat}}) / \text{ZPR}_g^{\text{exp}}$; first-principles ASC zero-point renormalization of the gap. Experimental $\text{ZPR}_g^{\text{exp}}$ data are obtained from extrapolation of temperature dependence of the gap, or isotopic mass derivatives of the gap. The latter are indicated with an asterisk. For SiC, the experimental data refer to the 15R polymorph, while the theoretical data refer to the 3C (zb) polymorph. They have similar sp^3 topology, but differ by stacking sequence. For SrTiO_3 , the $\text{ZPR}_g^{\text{lat}}$ was hardly computable with the TDep methodology presented in the **METHODS** section of the main text. We have estimated it combining experimental data for the zero-point correction to the volume [70] and theoretical data for the derivative of the gap with volume [71]. The reference experimental value for AlN is the average of the reported values from Refs. [72] and [73].

Material	E_g^{exp} [74] (eV)	$\text{ZPR}_g^{\text{exp}}$ (meV)	$\text{ZPR}_g^{\text{lat}}$ (meV)	$\text{ZPR}_g^{\text{AHC}}$ present work (meV)	R	$\text{ZPR}_g^{\text{ASC}}$ PBE [37] (meV)
Ge-dia	0.74 (i)	-52 [75]*	-10	-33	0.83	-50
Si-dia	1.17 (i)	-52 [72] -59 [76]* -72 [72]	+9	-56	0.80	-65
GaAs-zb	1.52 (d)	-60 [72]	-29	-18	0.78	-53
CdTe-zb	1.61 (d)	-16 [72]	-8	-20	1/0.57	-15
AlSb-zb	1.69 (i)	-35 [72]	+6	-51	0.78	-43
CdSe-zb	1.85 (d)	-38 [72]	-7	-34	1/0.93	-21
AlAs-zb	2.23 (i)	-50 [72]	+8	-74	1/0.76	-63
GaP-zb	2.34 (i)	-85 [72]	+8	-65	0.67	-57
ZnTe-zb	2.39 (d)	-40 [72]	-18	-22	1.00	-24
SiC	2.40 (i)	-215 [72]	+6	-179	0.80	-109
CdS-zb	2.58 (d)	-62 [77]* -34 [72]	-10	-70	1/0.78	-29
ZnSe-zb	2.82 (d)	-55 [72]	-17	-44	1/0.90	-28
ZnS-zb	3.84 (d)	-105 [78]* -78 [72]	-24	-88	1/0.94	-44
C-dia	5.48 (i)	-338 [79]* -334 [72]	-27	-330	1/0.95	-320
SrTiO_3	3.25 (i) [80]	-336 [81]	$\approx +20$	-290	0.80	
ZnO-w	3.44 (d)	-164 [82]*	-17	-157	1/0.94	-57
GaN-w	3.47 (d)	-180 [72]	-49	-189	1/0.75	-94
AlN-w	6.20 (d)	-417 -350 [72] -483 [73]	-85	-399	1/0.88	

Table 4.3 – Electronic gaps and corrections due to the zero-point motion (ZPR_g^{AHC} and ZPR_g^{lat}), for thirteen materials: one-shot GW calculation of E_g , without ZPR ($E_g^{\text{G}_0\text{W}_0}$); self-consistent GW with electron-hole corrections (E_g^{GWeh}) which is the best theoretical E_g without ZPR ; first-principles ZPR_g^{AHC} and first-principles ZPR_g^{lat} (this work); ZPR-corrected E_g^{GWeh} ($E_g^{+\text{ZPR}}$), namely $E_g^{+\text{ZPR}} = E_g^{\text{GWeh}} + ZPR_g^{\text{AHC}} + ZPR_g^{\text{lat}}$; experimental gap (E_g^{exp}). The ratios between $E_g^{\text{G}_0\text{W}_0}$ and E_g^{exp} , between E_g^{GWeh} and E_g^{exp} , between $E_g^{\text{GWeh}} + ZPR_g^{\text{AHC}}$ and E_g^{exp} and between $E_g^{+\text{ZPR}}$ and E_g^{exp} , are shown in Fig. 3, but not reported here. Then, the ratio $R_{\text{GWeh}}^{\text{ZPR}}$ between $ZPR_g^{\text{AHC}+\text{lat}}$ and the correction brought by the scGWeh approximation E_g^{GWeh} with respect to $E_g^{\text{G}_0\text{W}_0}$, is mentioned. The values for $E_g^{\text{G}_0\text{W}_0}$ are taken from Ref. [83], except for Ge-dia, that comes from Ref. [84].

Material	$E_g^{\text{G}_0\text{W}_0}$ [83],[84] eV	E_g^{GWeh} [14] eV	ZPR_g^{AHC} eV	ZPR_g^{lat} eV	$E_g^{+\text{ZPR}}$ [74] eV	E_g^{exp} eV	$R_{\text{GWeh}}^{\text{ZPR}}$
Ge-dia	0.53	0.81	-0.033	-0.010	0.77	0.74	-0.15
Si-dia	1.12	1.24	-0.056	0.009	1.19	1.17	-0.39
GaAs-zb	1.30	1.62	-0.024	-0.029	1.57	1.52	-0.17
SiC-zb	2.27	2.53	-0.179	0.006	2.36	2.40	-0.67
CdS-zb	2.06	2.39	-0.070	-0.010	2.31	2.42	-0.24
AlP-zb	2.44	2.57	-0.093	0.010	2.49	2.45	-0.64
GaN-w	2.80	3.27	-0.189	-0.049	3.03	3.20	-0.51
ZnO-w	2.12	3.20	-0.157	-0.017	3.03	3.44	-0.16
ZnS-zb	3.29	3.60	-0.088	-0.024	3.49	3.91	-0.36
C-dia	5.50	5.79	-0.330	-0.027	5.43	5.48	-1.23
BN-zb	6.10	6.59	-0.406	-0.017	6.17	6.25	-0.86
MgO-rs	7.25	8.12	-0.524	-0.117	7.48	7.67	-0.74

Table 4.4 – Experimental fundamental gap E_g^{exp} and type (d=direct, i=indirect), first-principles zero-point renormalization of the gap ZPR_g^{AHC} : from the AHC (present work), and from the ASC.

Material	E_g^{exp} [74] (eV)	ZPR_g^{AHC} present work (meV)	ZPR_g^{ASC} KS-DFT (meV)		
Si-dia	1.17 (i)	-56	-60 [34]		
			-57 [85]		
			-65 [37]		
			-75 [86]		
			3.40 (d)	-42	\approx -28 [87]
GaAs-zb	1.52 (d)	-18	-44 [85]		
			-23 [32]		
			-32 [85]		
SiC	2.40 (i)	-179	-53 [37]		
			-109 [34]		
			-109 [37]		
TiO ₂ -t	3.03 (d)	-337	-145 [86]		
			\approx -150 [87]		
			-334 [34]		
C-dia	5.48 (i)	-330	-345 [85]		
			-320 [37]		
			-437 [86]		
			7.07 (d)	-416	-437 [32]
			\approx -410 [87]		
MgO-rs	7.67 (i)	-524	-450 [85]		
			\approx -220 [87]		
			-281 [86]		

Table 4.5 – For thirty materials: fundamental gap (E_g^{exp}); location and possible degeneracy of the band edges (valence/conduction) from Ref.[69] except for SrTiO₃; first-principles renormalization of the band edges (valence band maximum ZPR_v, conduction band minimum ZPR_c) and of the band gap (ZPR_g); and ratio between ZPR_g and E_g (R_g). Methodology: “exp” for experimental results [74], “gFr” for generalized Fröhlich model results, Supplementary Eq. (28), and “AHC” for AHC first-principles results (with correction Supplementary Eq. (3)).

Material	E_g^{exp} eV	Edge v/c deg.	ZPR _v gFr meV	ZPR _v AHC meV	ZPR _c gFr meV	ZPR _c AHC meV	ZPR _g AHC meV	R_g AHC %
Ge-dia	0.74	Γ/L	0	17	0	-16	-33	-4.5
Si-dia	1.17	$\Gamma(3)/$ $X^*(X-\Gamma)$	0	35	0	-21	-56	-4.8
GaAs-zb	1.52	$\Gamma(3)/\Gamma$	3	19	-1	1	-18	-1.2
CdTe-zb	1.61	$\Gamma(3)/\Gamma$	12	17	-4	-3	-20	-1.2
AlSb-zb	1.69	$\Gamma(3)/$ $X^*(X-\Gamma)$	4	24	-4	-27	-51	-3.0
CdSe-zb	1.85	$\Gamma(3)/\Gamma$	20	26	-6	-8	-34	-1.8
AlAs-zb	2.23	$\Gamma(3)/X$	12	39	-9	-35	-74	-3.3
GaP-zb	2.34	$\Gamma(3)/$ $X^*(X-\Gamma)$	9	28	-7	-37	-65	2.8
ZnTe-zb	2.39	$\Gamma(3)/\Gamma$	11	19	-4	-3	-22	-0.9
SiC-zb	2.40	$\Gamma(3)/X$	59	112	-32	-67	-179	-7.5
CdS-zb	2.42	$\Gamma(3)/\Gamma$	39	44	-15	-26	-70	-2.9
AlP-zb	2.45	$\Gamma(3)/X$	21	54	-14	-39	-93	-3.8
ZnSe-zb	2.82	$\Gamma(3)/\Gamma$	22	31	-8	-13	-44	-1.6
TiO ₂ -t	3.03	Γ/Γ	236	195	-135	-142	-337	-11.1
SrTiO ₃	3.25	$R(3) /$ $\Gamma(3)$ [80]	140	191	-115	-99	-290	-8.9
GaN-w	3.20	Γ/Γ	80	101	-31	-88	-189	-5.9
GaN-zb	3.29	Γ/Γ	77	85	-30	-91	-176	-5.3
ZnO-w	3.44	Γ/Γ	106	108	-38	-49	-157	-4.6
SnO ₂ -t	3.73	Γ/Γ	144	176	-47	-39	-215	-5.8
ZnS-zb	3.91	Γ/Γ	40	49	-19	-39	-88	-2.3
BaO-rs	4.80	X/X	225	189	-133	-82	-271	-5.6
SrO-rs	5.22	$\Gamma(3)/X$	229	214	-141	-112	-326	-6.2
C-dia	5.48	$\Gamma(3) /$ $X^*(X-\Gamma)$	0	134	0	-196	-330	-6.0
AlN-w	6.20	Γ/Γ	151	204	-80	-195	-399	-6.4
BN-zb	6.25	$\Gamma(3)/X$	94	201	-68	-205	-406	-6.5
CaO-rs	6.93	$\Gamma(3)/X$	223	211	-154	-130	-341	-4.9
Li ₂ O	7.55	$\Gamma(3)/X$	364	347	-172	-226	-573	-7.6
MgO-rs	7.83	$\Gamma(3)/\Gamma$	327	317	-137	-207	-524	-6.7
SiO ₂ -t	8.90	Γ/Γ	292	300	-167	-285	-585	-6.6
BeO-w	10.59	$\Gamma(2)/\Gamma$	289	369	-191	-330	-699	-6.6

Table 4.6 – First-principles GGA-PBE parameters for the computation of the ZPR of band edges using the (generalized) Fröhlich model. Only materials with one IR-active phonon branch, cubic space group, and non-degenerate band extremum (though possibly anisotropic effective masses) are included in this table. For AlSb and GaP, the edge is located close to the X point, along the X- Γ line, which is denoted X*.

Material	edge	ϵ^∞	ϵ^0	ω_{LO} (meV)	m_{xx}^* m_{yy}^*	m_{zz}^*	α	ZPR ^{gFr} (meV)
GaAs-zb	c Γ	15.31	17.55	33.5	0.009	0.009	0.016	-1
CdTe-zb	c Γ	8.89	12.37	19.1	0.052	0.052	0.192	-4
AlSb-zb	c X*	12.02	13.35	39.8	0.222	1.142	0.090	-4
CdSe-zb	c Γ	7.83	11.78	23.6	0.051	0.051	0.233	-6
AlAs-zb	c X	9.49	11.51	47.3	0.243	0.897	0.184	-9
GaP-zb	c X*	10.50	12.53	48.6	0.230	1.062	0.152	-7
ZnTe-zb	c Γ	9.05	11.99	24.1	0.076	0.076	0.178	-4
SiC-zb	c X	6.97	10.30	117.0	0.228	0.677	0.280	-32
CdS-zb	c Γ	6.21	10.24	34.4	0.118	0.118	0.432	-15
AlP-zb	c X	8.12	10.32	59.9	0.252	0.809	0.184	-14
ZnSe-zb	c Γ	7.35	10.73	29.3	0.089	0.089	0.276	-8
GaN-zb	c Γ	6.13	11.00	86.0	0.144	0.144	0.345	-30
ZnS-zb	c Γ	5.97	9.40	40.6	0.167	0.167	0.458	-19
BaO-rs	v X	4.21	92.43	47.3	4.035	0.431	4.757	225
BaO-rs	c X	4.21	92.43	47.3	0.380	1.197	2.812	-132
SrO-rs	c X	3.77	20.91	55.4	0.407	1.225	2.545	-141
BN-zb	c X	4.52	6.69	161.0	0.299	0.895	0.422	-68
CaO-rs	c X	3.77	16.76	66.8	0.443	1.424	2.305	-154
Li ₂ O	c X	2.90	7.80	86.3	0.437	0.850	1.993	-172
MgO-rs	c Γ	3.23	11.14	84.5	0.340	0.340	1.624	-137

Supplementary Methods

A. List of materials and calculation parameters

We have evaluated Eqs. (8) and (9) of the main text for thirty materials. The list of these materials, with Materials Project IDs [69], is presented in Supplementary Table 4.1. It includes three non-polar materials (C, Si, Ge), and one of their combinations (SiC), nine III-V compounds (GaAs, AlAs, AlP, GaN-w, GaN-zb, AlN, BN, AlSb, GaP) among which four nitrides, ten oxides (TiO₂, ZnO, SnO₂, BaO, SrO, CaO, Li₂O, MgO, SiO₂-stishovite, BeO) and SrTiO₃, as well as six non-oxide II-VI compounds (CdTe, CdSe, CdS, ZnS, ZnSe, ZnTe).

With this set of materials, we have (i) a large overlap with well established experimental ZPR_g values, as discussed in Supplementary Note 1, and complete coverage of the set of materials for which ASC ZPR_g calculations have been done by Karsai *et al.* [37]; (ii) nearly entire coverage of the set of materials for which GWeh computations have been performed, from Shishkin *et al.* [83]; and we include also (iii) a dozen of binary oxyde infrared-active materials for which the trends in ZPR can be analyzed.

Supplementary Table 4.1 also lists the parameters used in the first-principles computations. Within Density-Functional Perturbation Theory (see METHODS section), the computation of phonons at arbitrary q-points is done without any supercell calculation, thus allowing us to rely on the very fine q-point sampling grids mentioned in Supplementary Table 4.1 at affordable CPU time cost. Still the convergence with respect to such sampling is extremely slow, and is adressed below.

B. Accuracy of the Brillouin Zone sampling in the AHC case

As mentioned in the METHODS section of the main text, the convergence with respect to the \mathbf{q} -point sampling can be improved by three different techniques: either by a linear extrapolation based on the expected scaling $(N_{\mathbf{q}})^{-1/3}$, or by the computation of the coefficient of such scaling behaviour, as explained below, or by explicitly refining the region around Γ . For ZnO and SrTiO₃, we have used the latter, with a 192x192x192 fine grid.

The second methodology is the following. According to Ref. [3], the $(N_{\mathbf{q}})^{-1/3}$ behavior is associated with a $q = 0$ discarded (missing) part of the Brillouin zone in the Fan contribution, Eq. (8) of the main text. We denote $\Omega_{q=0}$ this region of the BZ, and estimate its contribution using the generalized Fröhlich electron-phonon interaction, however neglecting the electronic dispersion in this small region (unlike in the generalized Fröhlich model of the main text).

The estimation of the dominant correction to the ZPR_c^{AHC} thus reads

$$\Delta^{q=0}ZPR_c^{\text{AHC}} = -\frac{1}{\pi\Omega_0} \int_{\Omega_{q=0}} d\mathbf{q} \sum_j \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{q\omega_{j0}(\hat{\mathbf{q}})\epsilon^\infty(\hat{\mathbf{q}})} \right)^2. \quad (\text{S4.1})$$

Note that the sum over degenerate states has been cancelled by the n_{deg} denominator. The volume of the missing region is actually the BZ volume ($\Omega_{\text{BZ}} = (2\pi)^3/\Omega_0$) divided by $N_{\mathbf{q}}$, the number of points sampling the BZ. The equivalent sphere has a cut-off radius

$$q_c = 2\pi \left(\frac{3}{4\pi\Omega_0} \right)^{1/3} (N_{\mathbf{q}})^{-1/3}. \quad (\text{S4.2})$$

We then replace the integral over $\Omega_{q=0}$ by an integral in the sphere with radius q_c , evaluate the radial part of the integral, and obtain

$$\Delta^{q=0}ZPR_c^{\text{AHC}} \approx -\frac{8\pi}{\Omega_0} \left(\frac{3}{4\pi\Omega_0} \right)^{1/3} (N_{\mathbf{q}})^{-1/3} K_{\text{av}}, \quad (\text{S4.3})$$

where K_{av} is the angular average of the $\hat{\mathbf{q}}$ -dependent quantities in Supplementary Eq. (S4.1), namely,

$$K_{\text{av}} = \frac{1}{4\pi} \int_{4\pi} d\hat{\mathbf{q}} \sum_j \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\omega_{j0}(\hat{\mathbf{q}})\epsilon^\infty(\hat{\mathbf{q}})} \right)^2. \quad (\text{S4.4})$$

The correction is of opposite sign for the top of the valence band, due to the change of occupation number.

This technique and the linear extrapolation technique have been tested for most materials in our list. Supplementary Tables 4.2 and 4.5 report the values obtained with the correction from Supplementary Eq. (S4.3), except for the following materials: the non-infrared-active materials (C, Si and Ge), for which there is no such correction, and also AlN-w and BN-zb, for which the convergence study had been done in Ref. [3], based on the linear extrapolation [88].

Supplementary Figure 4.7 compares Supplementary Eq. (S4.3) and the linear extrapolation technique in the case of MgO. The \mathbf{q} -grids are $(N_{\text{MP}} \times N_{\text{MP}} \times N_{\text{MP}})$ Monkhorst-Pack samplings of the BZ, [89] thus cubic grids shifted four times, with a total number of points $N_{\mathbf{q}} = 4 \cdot N_{\text{MP}}^3$. A 10% accuracy (20 meV in this case) is obtained from corrected points already with a $N_{\text{MP}} = 20$ grid (without extrapolation), while, without correction, linear extrapolation from the $N_{\text{MP}} = 20$ and $N_{\text{MP}} = 32$ data gives it as well. Without correction neither extrapolation, such accuracy is only reached with the finest $N_{\text{MP}} = 96$ grid. Still, the correction Supplementary Eq. (S4.3) does not exactly removes the $1/N_{\text{MP}}$ behaviour, although it decreases it by more than a factor of ten. The coefficient to the $(N_{\mathbf{q}})^{-1/3}$ factor computed from this equation is only approximate.

Still another technique to speed up the convergence, based on Fröhlich model, but not restricted to lowest order, has been sketched in Ref. [5]. However, in this reference, it has

only fully been elaborated for materials with isotropic characteristics. It should be possible to develop it further using the generalized Fröhlich model introduced in the present work.

C. Comparing the AHC Brillouin Zone sampling and the ASC supercell size

One might wonder why even with the above-mentioned corrections, the AHC approach needs a grid with typically $N_{\text{MP}} = 20$ to obtain reasonably converged values, while for the ASC approach, results with $5 \times 5 \times 5$ or $6 \times 6 \times 6$ supercells, corresponding to grid $N_{\text{MP}} = 5$ or $N_{\text{MP}} = 6$ are apparently converged at the same level [37]. We argue that the needs of both methods are not identical, due to different physics, and the rate of convergence is quite different.

In the ASC methodology, for which the divergence of the harmonic approximation is removed by the inclusion of the nonquadratic coupling, the convergence is much faster. This is partly because, as argued in the **DISCUSSION** section of the main text, “for larger amplitudes, \langle the eigenenergy \rangle behaves linearly, as the electron localizes in the lowered potential region and the minimum of the potential is linear in the amplitude of the atomic displacements.”. This is at variance with the quadratic behaviour found from the AHC perturbative theory.

Let us examine a numerical proof of this difference. In Ref. [37], Karsai and coauthors quote their detailed convergence study for diamond (their Table 2) from $3 \times 3 \times 3$ to $6 \times 6 \times 6$ supercells, with apparent convergence within 6 meV for a ZPR around 330 meV (values: -0.278 , -0.365 , -0.315 , and -0.321 eV). With the AHC, we checked that fluctuations of ZPRs between corresponding \mathbf{q} -point grids are on the order of 100 meV or larger, which is in line with the results from Ref. [27] for random wavevector grids. Fluctuations becomes less than 100 meV only above $8 \times 8 \times 8$. Thus, converged values are much easier to reach with the ASC methodology than with the AHC methodology. Still, as emphasized in the main text, the ASC avoids the adiabatic AHC divergence for the wrong physical reason, and the better convergence rate of ASC is misleading, as the converged value is not the right one.

Supplementary Notes

1. Band gap zero-point renormalization from experiment and from the theoretical approaches (AHC and ASC)

Supplementary Table 4.2 compares $\text{ZPR}_{\mathbf{g}}$ data from experiment and from different computations. These data are used in Fig. 4.3 of the main text. Further computed $\text{ZPR}_{\mathbf{g}}$ data

are provided in Supplementary Table 4.3, although the latter focuses on band gaps, analyzed in the Supplementary Note 4. Gathering experimental ZPR_g data needed some care, as explained below. Also, theoretical data ought to be discussed in view of previous works. For the in-depth discussion of the theoretical data among themselves, we refer to the Supplementary Discussion. We also complete the list of references in which ZPR_g have been computed, that had been mentioned in the main text.

Experimental data As explained by Cardona and Thewalt, in relation to the Table 3 of their review Ref. [47], the experimental ZPR are obtained by two different techniques. The first one starts from measurements of the band gap energy for different samples of a material with varying isotopic content. The derivative(s) of the band gap with respect to the mass(es) are extracted, followed by extrapolation to infinite mass(es). The second one relies on measurements of the band gap energy in a large temperature range, including the low-temperature regime, but also the high-temperature regime, where a linear asymptote must be reached. In this case, the ZPR is the difference between the measured zero-temperature limit and the linear extrapolation of the high-temperature asymptote to zero temperature, as shown for Ge in Fig. 21 of Ref. [47]. When both techniques agree, the experimental result can be reasonably trusted.

For the latter technique (extrapolation of temperature-dependent data), we rely on the analysis of Pässler[72, 73], that superceeds his earlier work[90] on which Ref. [47] relied. The two publications by Pässler[72, 73] yield essentially equivalent results, except for AlN. For this material, we have taken the mean of the two reported values as reference. However, also in the latter case, our results are within the 25% limit. We have surveyed the literature after 2003, but were unable to find more reliable data.

For the isotopic extrapolation, to our surprise, a large fraction of the data provided in the related column of Table 3 of Ref. [47] do NOT report actual isotopic extrapolation analysis. We have thus extensively examined the primary literature, and have obtained that such isotopic extrapolation is available reliably for the following materials in our list: Ge, Si, C, CdS, ZnO and ZnS. They are clearly identified as such in Supplementary Table 4.2. Moreover, we have also realized that a non-negligible number of data available through other secondary references are misleading. We will not elaborate further on this problem, but generally speaking, we warn the reader about the unreliability of several publications in the field. The above-mentioned works of Pässler are apparently free of such problem.

Anyhow, the two techniques agree very well for Ge or C, or reasonably (within 25%) for Si and ZnS, but clearly differ for CdS (about a factor of 2). Pässler [72] does not provide data for ZnO. In Fig. 4.3 of the main text, we have taken the isotopic extrapolation result as reference, since the extrapolation of temperature-dependent data is often plagued with insufficient sampling of the high-temperature region, which is crucial for the correct extrapolation, as discussed at length in Ref. [90]. For the above-mentioned CdS discrepancy,

the global expected trend due to the change of mass, comparing with CdSe and CdTe, or with the Zn series, appears more reasonable from the isotopic technique than from extrapolation of temperature-dependent data. Also, the comparison of the different analyses by Pässler, in 1999, 2002, and 2003, Refs. [72, 73, 90], shows stability for a few materials, but, more generally, non-negligible variations, at the 10-25% level. On this basis, one can infer that, globally speaking, the reliability of the experimental data is on the order of 25%, which is the reason of our use of this value in the main text.

Note that Karsai et al. [37] quotes experimental data (most from secondary sources) for about half the materials they compute, but miss many available data, especially those that do not compare well with their results.

Theoretical AHC data The values quoted in Supplementary Tables 4.2 and 4.3 for the band gap ZPR of C, BN, and MgO, respectively -330 meV, -406 meV, and -524 meV, differ from those previously published in Ref. [4], authored by some of us, giving respectively -366 meV, -370 meV, and -341 meV. We emphasize that the same first-principles method as in our current manuscript (non-adiabatic AHC) was used to obtain these data. The differences with the present data are entirely due to the sampling of the Brillouin Zone and associated imaginary broadening factor.

In Ref. [4], the sampling of the Brillouin Zone was already emphasized as an important technical issue, especially considering the numerical noise. For the purpose of computing the self-energy (which was the focus of Sec. II of Ref. [4]), or computing the importance of “anharmonic effects” in the adiabatic approximation (Sec. III of Ref. [4]), such numerical noise was addressed by keeping a large finite imaginary broadening (between 0.1-0.4 eV) and a fixed wavevector sampling $32 \times 32 \times 32$. The choice of the value of the imaginary broadening was detailed in the Appendix of this paper, which was explicitly mentioned in the caption of this Table 1: “See Appendix for the values of the η parameters used”. Such imaginary broadening was determined to ensure removal of the noise in the self-energy, in order to produce Fig. 1 and 2, but however did not deliver an accurate absolute value for the ZPR in Table I of Sec. II.

Indeed, in another publication of some of ours (even slightly before Ref. [4]), namely Ref. [3], the proper limit of vanishing imaginary broadening with much improved wavevector sampling was performed, for C and BN (up to $125 \times 125 \times 125$ grid). The same values for C and BN as mentioned in our current manuscript, namely -330 meV and -406 meV, were found (with one more digit, see Table VII, column “Non-adiabatic”). MgO was not examined in Ref. [3], but the MgO band gap ZPR was then mentioned in Ref. [28], still another publication from some of ours, page 11, end of second paragraph of first column, 526 meV. Such value is very similar to the values in our manuscript. The extrapolation scheme, differing from the one chosen in the present paper, explains the difference of 2 meV for MgO. The analysis of LiF is not done in the current work because an estimate of the Fröhlich parameter for

the valence bands yields α bigger than 8, see Ref. [28], which is beyond the validity of the present perturbative treatment. By the way, the top of the valence band does not occur at Γ (see e.g. Ref. [69], which was not mentioned in Ref. [28]).

2. Modification of the ZPR due to lattice parameter changes

The $\text{ZPR}_g^{\text{AHC}}$ is computed at the fixed theoretical lattice parameters and internal coordinates minimizing the DFT energy (or fixed experimental lattice parameters in the case of Ge, GaAs and TiO_2). The phonon population effects on the lattice parameters, and induced change of band gap, are not taken into account in such procedure. This corresponds strictly to the harmonic approximation.

In the so-called quasi-harmonic approximation, the equilibrium lattice parameters and possibly internal coordinates are relaxed in order to minimize the free energy of the system. This makes such parameters depend on the temperature (thermal expansion), but also induces their zero-point change at $T = 0$ K. In turn, such changes modifies the gap, an effect to be accounted for in order to compare with the experimental ZPR_g . This effect has been investigated by Garro and coworkers, Ref. [91]. It is also included in several first-principles calculations of the ZPR, see e.g. Ref. [92].

In the present work, we have indeed added this effect beyond the harmonic approximation: we have computed the equilibrium lattice parameters with and without zero-point motion, then evaluated the change of gap. In the first case, we have minimized the Born-Oppenheimer energy without phonon contribution, giving $\{\mathbf{R}_i^{\text{fix}}\}$. In the second case, the free energy, including zero-point phonon contributions computed at different lattice parameters and interpolated between them, was minimized, giving $\{\mathbf{R}_i^{T=0}\}$. This procedure is similar to the one described in Ref. [93], albeit extended to two dimensions (uniaxial and basal lattice parameters) in the case of wurtzite materials. The differences in band gap values for such lattice parameters form the lattice ZPR modification of the electronic structure:

$$\text{ZPR}_{\mathbf{kn}}^{\text{lat}} = \varepsilon_{\mathbf{kn}}(\{\mathbf{R}_i^{T=0}\}) - \varepsilon_{\mathbf{kn}}(\{\mathbf{R}_i^{\text{fix}}\}). \quad (\text{S4.5})$$

In such approach, one supposes that there is no cross-influence of the modification of lattice parameters on the AHC ZPR. Also, one relies on the approximation that the internal parameters, if not determined by symmetry, are those that minimize the Born-Oppenheimer energy at the corresponding \mathbf{R}_i lattice parameters. In other words, the deviatoric thermal forces and their effect on the band gap are not included, but the deviatoric thermal stresses are taken into account, following Ref. [94]. In the list of materials of Supplementary Tables 4.2 and 4.3, deviatoric thermal forces only appear for wurtzite materials, that have a quite symmetric tetrahedral environment, and thus, very small expected effect.

3. A survey of zero-point renormalization computations

In complement to the references given in the introduction of the main text, we list additional ones in view of further comparisons, without pretending to be exhaustive. Not all mentioned references present computations of equal quality. In particular, the difficult sampling of the Brillouin Zone for AHC has rarely been tackled correctly before 2015. Similarly the supercell size convergence for ASC is also a concern.

Refs. [3–5, 20–28, 92, 95–105] report calculations of the ZPR_g based on the AHC approach with KS-DFT wavefunctions and eigenenergies. Refs. [23, 29–37, 85, 87, 96, 106–110] report calculations of the ZPR_g based on the ASC approach, usually based on KS-DFT, but also sometimes based on GW, see Refs. [32, 37, 87]. As analyzed by the present generalized Fröhlich model, the ASC might be reasonably predictive for purely covalent materials (e.g. C, Si, Ge), or weakly infrared-active ones (e.g. GaAs), however, when the ZPR^{gFr} is a significant fraction of the ZPR, the predictive capability of the ASC approaches should be questioned. Even in the covalent case, non-adiabatic corrections can be sizeable (5...15%), see Table VII of Ref. [3]. The ASC GW calculations performed until now [32, 37, 87] rely at most on 4x4x4 supercells, which is at the limit of convergence. Going from 4x4x4 to 5x5x5 supercells for KS-DFT indeed still brings some modification, as shown by Ref. [37]. For this reason, the effect of relying on GW corrections can hardly be clarified in the present context, which focuses on the non-adiabatic effects.

In a very recent publication [111], Engel and coworkers have proposed a new method to compute the ZPR, combining supercell calculations with an interpolation method, still in the adiabatic approximation. They applied it for 9 materials, a subset of those examined by Karsai and coworkers [37] (focusing only on covalent or weakly ionic materials). The results that they obtain do not change the overall assessment of the adiabatic approach that we present here: for some materials, the agreement with experiment is slightly improved or deteriorated.

Studies of strongly infrared-active materials using the ASC include e.g. hydrogen-containing materials HF, H₂O, NH₃ and CH₄ in Ref. [35], where a giant electron-phonon interaction in molecular crystal and related important nonquadratic coupling had been investigated, the study of ice in Ref. [36], and the study of LiF, MgO and TiO₂ in Ref. [87], and some of the results from Ref. [37] for IR-active materials.

In addition to these references, the electron-phonon interaction has been directly incorporated in GW calculations in Refs. [112–114]. These are further commented upon in the next section.

Supplementary Table 4.4 presents KS-DFT ASC data that were not reported in Supplementary Table 4.2, as the latter only mentioned such data from Ref. [37]. This Supplementary Table 4.4 still focuses on our set of materials, for which AHC data are available. For the

indirect gap of Si and of SiC, as well as both the direct and indirect gaps of C, the different ASC calculations agree very well. The agreement is less good for GaAs and the direct gap of Si. Anyhow, these data confirm that AHC and ASC calculations of the fundamental band gap reasonably agree for covalent materials, but strongly disagree for strongly IR-active materials (SiC-zb, TiO₂-t, MgO-rs).

4. Gap values from experiment and different computations

Supplementary Table 4.3 compares gap values from experiment and from different computations. These data are used in Fig. 4.4. The set of 12 materials in this table includes all those for which self-consistent GW calculations with electron-hole corrections have been performed in Ref. [14], except Ne, LiF and Ar. We did not include them in our study for the following reasons. Ne and Ar are strongly affected by weak van der Waals interactions. Much more than the other materials, they will be modified by the zero-point lattice corrections, discussed earlier. For LiF, as mentioned previously, the estimated value of α show that the present perturbative treatment is likely invalid. For all three materials, excitonic effects are strong, and should likely be included for meaningful comparison with experimental data.

Supplementary Table 4.3 quotes $E_g^{G_0W_0}$ values from Ref. [83], except for Ge-dia, that comes from Ref. [84]. The latter reference also presents $E_g^{G_0W_0}$ values for many materials, including more than half of those reported in Supplementary Table 4.3. Many other $E_g^{G_0W_0}$ values for those materials have been published. The spread in such values for one material can be as large as 0.5 eV. A comparative study of published $E_g^{G_0W_0}$ values is out-of-scope of the present publication, but is presented in Ref. [48]. The values $E_g^{G_0W_0}$ in Supplementary Table 4.3 have the goal to show the typical size of the $E_g^{G_0W_0}$ to $E_g^{GW_{eh}}$ correction, and point out that the ZPR correction is of the same order of magnitude for materials with light nuclei.

In addition to the AHC and ASC methodologies, there has also been attempts to incorporate directly the electron-phonon interaction inside the GW approach [112–114] to obtain modified gap values or ZPR_g . The first publication on the subject, by Botti and Marques [112] was later shown by Lambrecht, Bhandari and van Schilfgaarde [113] to ignore the range of the Fröhlich interaction, yielding an erroneous integration over the Brillouin Zone. Even in the latter work, there is a factor of 2 difference for the Fröhlich α compared to the usual definition, used in the present work. Their ZPR_g results (Table I of Ref. [113] and [114]) for MgO, GaN-zb and SrTiO₃, resp. (-219 meV, -67 meV, -404 meV) can be compared with ours (-517 meV, -171 meV, -290 meV), see Supplementary Table 4.5. They very clearly differ, with strong underestimation with respect to ours for MgO, GaN-zb, and strong overestimation for SrTiO₃. For SrTiO₃, our results agree better with the experimental data (-336 meV), see Supplementary Table 4.2. For MgO and GaN-zb, a direct measurement of ZPR_g is not available. Still, experimental value for the other polymorph of GaN,

namely GaN-w (-180 meV) is reported in Supplementary Table 4.2 and matches very well our ZPR_g^{AHC} value (-189 meV) for GaN-w.

5. From the adiabatic AHC approach to the generalized Fröhlich Hamiltonian

In order to help understand why the generalized Fröhlich Hamiltonian captures correctly the physics of the Fan diagram, at variance with the adiabatic approximation, we summarize first the discussion of the AHC case given in Ref. [3]. Then, we derive the generalized Fröhlich Hamiltonian from Eq. (4.4.2) of the main text.

The divergence of the adiabatic AHC for IR-active materials. Under the adiabatic approximation, the phonon frequencies $\omega_{\mathbf{q}j}$ are neglected when compared to differences between electronic eigenenergies. In other words, one supposes that the electron (or hole) responds instantaneously to the atomic vibrations, even for those electronic transitions with a characteristic time scale (inversely proportional to the eigenenergy difference) larger than the phonon oscillation period. This simplification is of course questionable for intraband transitions with small momentum transfer, as shown in Ref. [3]. Suppressing $\omega_{\mathbf{q}j}$ in the denominators of main text Eq. (4.4.2), and evaluating the AHC ZPR according to main text Eqs. (4.4.1) and (4.4.3) yields

$$ZPR_{\mathbf{k}n}^{\text{adiab}} = \Re e \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}j}^{\text{BZ}} \sum_{n'} \frac{|\langle \mathbf{k} + \mathbf{q}n | H_{\mathbf{q}j}^{(1)} | \mathbf{k}n' \rangle|^2}{\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k}+\mathbf{q}n'} + i\eta} + \Sigma_{\mathbf{k}n}^{\text{DW}}. \quad (\text{S4.6})$$

For band edges, in particular the top of the valence band or the bottom of the conduction band, and considering at present non-degenerate bands, the intraband ($n' = n$) contribution when $\eta = 0$ has two different types of divergences for $q \rightarrow 0$, reinforcing each others: (i) for IR-active materials, the electron-phonon interaction matrix elements of $H_{\mathbf{q}j}^{(1)}$ diverge like $\delta_{nn'}/q$, while (ii) the denominator behaves like $q^2/2m_{\hat{\mathbf{q}}}^*$. In these expressions, q is the norm of the \mathbf{q} wavevector, and $m_{\hat{\mathbf{q}}}^*$ is the effective mass along direction \mathbf{q} , represented by the unit vector $\hat{\mathbf{q}}$, with $\mathbf{q} = q \cdot \hat{\mathbf{q}}$. The effective mass is the inverse of the second derivative of the eigenenergy with respect to the wavevector along $\hat{\mathbf{q}}$ at the band extremum.

Thus, the integrand at small q behaves like $1/q^4$. Focusing on the ZPR_c^{adiab} of the bottom of the conduction band (the handling of the top of the valence band is similar) one finds, in spherical coordinates, and introducing a spherical cut-off q_c whose limit $q_c \rightarrow 0$ has to be taken:

$$ZPR_c^{\text{adiab}}(q_c < q) = \int_{4\pi} \int_{q_c} \frac{f^{\text{adiab}}(q\hat{\mathbf{q}})}{q^2 \left(\frac{q^2}{2m_{\hat{\mathbf{q}}}^*} + \mathcal{O}(q^3) \right)} q^2 dq d\hat{\mathbf{q}}, \quad (\text{S4.7})$$

where $f^{\text{adiab}}(q\hat{\mathbf{q}})$ is a smooth function of q (it does not diverge, but also does not tend to zero for $q \rightarrow 0$). As announced [3], $\text{ZPR}_c^{\text{adiab}}(q_c < q)$ diverges like $1/q_c$ when $q_c \rightarrow 0$.

By contrast, if phonon frequencies are not suppressed in main text Eq. (8), which is the non-adiabatic case, one falls back to an integral of the type (here considered for one isotropic phonon branch only, and one non-degenerate isotropic electronic band):

$$\text{ZPR}_c^{\text{non-adiab}}(q_c < q) = \int_{4\pi} \int_{q_c} \frac{f^{\text{non-adiab}}(q\hat{\mathbf{q}})}{q^2 \left(\frac{q^2}{2m_{\hat{\mathbf{q}}}^*} - \omega_0 + \mathcal{O}(q^2) \right)} q^2 dq d\hat{\mathbf{q}}, \quad (\text{S4.8})$$

where the $\mathcal{O}(q^2)$ contribution in the denominator comes from the curvature of $\omega_{\mathbf{q}}$ with respect to q , usually much smaller than the curvature of the electronic eigenenergies governed by $m_{\hat{\mathbf{q}}}^*$, explicitly mentioned in Eq. (S4.8). $\text{ZPR}_c^{\text{non-adiab}}(q_c < q)$ does not diverge for $q_c \rightarrow 0$, provided ω_0 is non-zero.

Acoustic modes should not be forgotten. Their eigenfrequency ω_0 tends to zero in the $q \rightarrow 0$ limit. However, the corresponding first-order electron-phonon matrix element does not diverge like $1/q$, and also, it has been shown by Allen and Heine [15] that, thanks to translational symmetry, the contribution to the Fan term from acoustic modes is exactly cancelled by the Debye-Waller term, in the $q \rightarrow 0$ limit. Thus no divergence arises from them.

To summarize, finite phonon frequencies at $q = 0$ remove the AHC divergence in the non-adiabatic case.

The generalized Fröhlich Hamiltonian. Instead of suppressing phonon frequencies appearing in the Fan self-energy Eq.(8) of the main text, like in the adiabatic approximation, a radically different strategy is followed in the generalized Fröhlich Hamiltonian: one uses only parameters at $q = 0$, including crucially the phonon frequencies, and extends them to the whole \mathbf{q} -space, using simple behavior for medium and large- q (i.e. parabolic electronic dispersion and no phonon dispersion). This corresponds to macroscopic electrostatic treatment and a continuum space hypotheses.

The model proposed by Fröhlich [39] in 1954 is the most simplified version of such hypotheses: one non-dispersive LO phonon branch, one electronic parabolic band governed by a single, isotropic effective mass, and electron-phonon interaction originating from the isotropic macroscopic dielectric interaction between the IR-active LO mode and the charged electronic carrier. No Debye-Waller contribution is included. Spin might be ignored, as only one electron is considered, without spin-phonon interaction. Explicitly, the Fröhlich Hamiltonian reads [39, 40, 115, 116],

$$\hat{H}^{\text{Fr}} = \hat{H}_{\text{el}}^{\text{Fr}} + \hat{H}_{\text{ph}}^{\text{Fr}} + \hat{H}_{\text{EPI}}^{\text{Fr}}, \quad (\text{S4.9})$$

with

$$\hat{H}_{\text{el}}^{\text{Fr}} = \sum_{\mathbf{k}} \frac{\mathbf{k}^2}{2m^*} \hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}}, \quad (\text{S4.10})$$

$$\hat{H}_{\text{ph}}^{\text{Fr}} = \sum_{\mathbf{q}} \omega_{\text{LO}} \hat{a}_{\mathbf{q}}^\dagger \hat{a}_{\mathbf{q}}, \quad (\text{S4.11})$$

$$\hat{H}_{\text{EPI}}^{\text{Fr}} = \sum_{\mathbf{q}, \mathbf{k}} g^{\text{Fr}}(\mathbf{q}) \hat{c}_{\mathbf{k}+\mathbf{q}}^\dagger \hat{c}_{\mathbf{k}} (\hat{a}_{\mathbf{q}} + \hat{a}_{-\mathbf{q}}^\dagger), \quad (\text{S4.12})$$

where $\hat{a}_{\mathbf{q}}^\dagger$ and $\hat{a}_{\mathbf{q}}$ are phonon creation and annihilation operators, $\hat{c}_{\mathbf{q}}^\dagger$ and $\hat{c}_{\mathbf{q}}$ are electron creation and annihilation operators, and the electron-phonon coupling parameter is given by

$$g^{\text{Fr}}(\mathbf{q}) = \frac{i}{q} \left[\frac{2\pi\omega_{\text{LO}}}{V_{\text{BvK}}} \left(\frac{1}{\epsilon^\infty} - \frac{1}{\epsilon^0} \right) \right]^{1/2}, \quad (\text{S4.13})$$

where V_{BvK} is the Born-von Karman normalisation volume corresponding to the \mathbf{k} and \mathbf{q} samplings, q is the norm of \mathbf{q} and $\hat{\mathbf{q}}$ is its direction, $\mathbf{q} = q \cdot \hat{\mathbf{q}}$, ϵ^∞ is the optical dielectric constant, and ϵ^0 is the low-frequency dielectric constant. The sums over \mathbf{k} and \mathbf{q} run over the whole reciprocal space.

At the lowest order of perturbation theory, this yields the well-known result [39, 40, 117] for the polaron binding energy (that we denote ZPR^{Fr}):

$$\text{ZPR}^{\text{Fr}} = -\alpha\omega_{\text{LO}}, \text{ with } \alpha = \left(\frac{1}{\epsilon^\infty} - \frac{1}{\epsilon^0} \right) \sqrt{\frac{m^*}{2\omega_{\text{LO}}}}. \quad (\text{S4.14})$$

Accurate diagrammatic Monte Carlo simulations [41] have demonstrated reasonable validity of this lowest-order perturbation result in a range that extends to about $\alpha \approx 8$. Beyond this, electronic self-trapping becomes too important, and the lowest-order result deviates strongly from the exact results.

Multiphonon generalisations of the Fröhlich model Hamiltonian have been considered in several occasions, but only in the cubic case, while the anisotropy of the electronic effective mass has been accounted for in an approximate way, and not altogether with the degeneracy of band edges, that has either been treated approximately, or for a spherical model, without warping, or only considering one phonon branch see e.g. Refs. [5, 49–52, 114].

Our generalized Fröhlich model includes all the features of real materials altogether: multiphonon, anisotropic, degenerate band extrema. It is derived by considering the full Eq. (4.4.2) of the main text and applying the same ansatz as the original Fröhlich model: we treat as exactly as possible the region in the Brillouin zone where the phonon frequency is large with respect to the eigenenergy differences, and, coherently, suppress the contributions from bands with a large energy difference with respect to phonon frequencies (interband contributions), keeping only intraband contributions. Like for the Fröhlich Hamiltonian, we use only parameters that describe the functions in the Brillouin zone around $q = 0$, and extend them to the whole reciprocal space, using the same simple behaviors for medium and

large- q : parabolic electronic dispersion and no phonon dispersion. We also retain only the macroscopic Fröhlich electron-phonon interaction, $\mathbf{G} = 0$ component, following Vogl [53]. This is also coherent with the first-principles Fröhlich vertex of Ref. [54], and the associated macroscopic electric field, derived in the appendix A of Poncé *et al.*, Ref. [3]. The possible degeneracy of the electronic states, causing the warping of the electronic structure, is addressed following Refs. [118] and [61]. It has been shown also by AHC that the contribution to the Fan term from acoustic modes, for which the frequency vanishes when $q \rightarrow 0$, is exactly cancelled by the Debye-Waller term, close to $q = 0$.

Having highlighted the basic ideas of the derivation of the generalized Fröhlich Hamiltonian, in what follows, we write down the mathematical expressions. We focus on the conduction band expression, ZPR_c , associated to one of the possibly degenerate states, labelled c , with energy ε_c and with wavevector \mathbf{k}_0 . Similar expressions for ZPR_v can be derived. For the sake of simplicity, we will suppose that \mathbf{k}_0 is the Brillouin zone center Γ , but the equations that follow can be straightforwardly adapted to another value of \mathbf{k}_0 . We also introduce $\mathbf{k}' = \mathbf{k} + \mathbf{q}$.

Under the above hypotheses, the matrix elements of $H_{\mathbf{q}j}^{(1)}$ for $q \rightarrow 0$ writes

$$\begin{aligned} \langle \mathbf{k}', n' | H_{\mathbf{q}j}^{(1)} | \mathbf{k}, n \rangle &= \frac{i}{q} \frac{4\pi}{\Omega_0} \left(\frac{1}{2\omega_{j0}(\hat{\mathbf{q}}) V_{\text{BvK}}} \right)^{1/2} \frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \\ &\times \langle \mathbf{k}', n' | \mathbf{k}, n \rangle_{\text{P}}, \end{aligned} \quad (\text{S4.15})$$

with Ω_0 the primitive cell volume, V_{BvK} the Born-von-Karman normalisation volume corresponding to the \mathbf{k} and \mathbf{q} samplings, and the scalar product $\langle \mathbf{k}', n' | \mathbf{k}, n \rangle_{\text{P}}$ is computed from the periodic part of the Bloch functions. Sums over \mathbf{k} and \mathbf{q} of the type $\sum_{\mathbf{q}} f(\mathbf{q})/V_{\text{BvK}}$ are to be replaced by $\Omega_0/(2\pi)^3 \int d^3\mathbf{q} f(\mathbf{q})$ in the macroscopic limit in what follows. We have used the notation $\omega_{j0}(\hat{\mathbf{q}})$ for the $q \rightarrow 0$ limit of the j -phonon branch ($\omega_{j0}(\hat{\mathbf{q}}) = \lim_{q \rightarrow 0} \omega_j(q\hat{\mathbf{q}})$), while the related mode-polarity vectors $\mathbf{p}_j(\hat{\mathbf{q}})$, see Eq. (41) of Ref. [119], are defined from the Born effective charges and the eigendisplacements of the phonon mode, summing over the atoms labelled by κ , and the Cartesian direction of displacement α ,

$$p_{\gamma,j}(\hat{\mathbf{q}}) = \lim_{q \rightarrow 0} \sum_{\kappa\alpha} Z_{\kappa\alpha,\gamma}^* U_{\kappa\alpha,j}(q\hat{\mathbf{q}}). \quad (\text{S4.16})$$

Similarly, the optic dielectric constant along $\hat{\mathbf{q}}$ is obtained from the optic dielectric tensor by summing over wavevector components, $\epsilon^\infty(\hat{\mathbf{q}}) = \sum_{\alpha\beta} \hat{q}_\alpha \epsilon_{\alpha\beta}^\infty \hat{q}_\beta$, as in Eq. (56) of Ref. [60]. The mode-polarity vector vanishes for all non-IR-active modes. Thus naturally only the LO modes contribute to the ZPR^{gFr} in this generalized Fröhlich approach (for acoustic modes and TO modes, the mode-polarity vector vanishes). Still, the directional dependency of the phonon frequencies, of the mode-polarity vectors, and of the dielectric constant, is correctly taken into account.

We then restrict ourselves to the bands that are degenerate with c , and choose a $\hat{\mathbf{k}}$ -independent complete basis for these states at $\mathbf{k}_0 = \Gamma$, denoted $|\Gamma, m\rangle$. We also define the $\hat{\mathbf{k}}$ -dependent overlap between states m at Γ and states n along $\hat{\mathbf{k}}$ in the $k \rightarrow 0$ limit:

$$s_{nm}(\hat{\mathbf{k}}) = \lim_{k \rightarrow 0} \langle k\hat{\mathbf{k}}, n | \Gamma, m \rangle_{\mathcal{P}}. \quad (\text{S4.17})$$

This definition supposes that the phase of the $|k\hat{\mathbf{k}}, n\rangle$ states is chosen continuously as a function of k , for vanishing k along the $\hat{\mathbf{k}}$ direction. The phase can always be chosen to be continuous, and this choice will have no bearing on the results obtained later. Still, it does not imply any continuity hypothesis for the phase between states n for different $\hat{\mathbf{k}}$ directions. The matrix $s_{nm}(\hat{\mathbf{k}})$ is unitary. Note that the $|\Gamma, m\rangle$ states are fixed before the $\hat{\mathbf{k}}$ direction is known. The coefficients $s_{nm}(\hat{\mathbf{k}})$ can be computed from first principles following the perturbative treatment of Refs. [118] and [61], which is the first-principles equivalent of the $\mathbf{k}\cdot\mathbf{p}$ approach, restricted to the subspace spanned by the bands that connect to the degenerate states. It delivers Luttinger-Kohn parameters $D_{jj'}^{\alpha\beta}$, [120] that allows one to find $s_{nm}(\hat{\mathbf{k}})$ for all $\hat{\mathbf{k}}$. The Luttinger-Kohn parameters can also be determined from experimental measurements. The s matrices can be thought as being the adaptation of spherical harmonics to the specific symmetry representation to which the degenerate wavefunctions belong.

With these definitions, we write the generalized Fröhlich electron-phonon interaction as

$$g^{\text{gFr}}(\mathbf{q}j, \mathbf{k}n'n) = \frac{i}{q} \frac{4\pi}{\Omega_0} \left(\frac{1}{2\omega_{j0}(\hat{\mathbf{q}})V_{\text{BvK}}} \right)^{1/2} \frac{\hat{\mathbf{q}}\cdot\mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \times \sum_m s_{n'm}(\hat{\mathbf{k}}')(s_{nm}(\hat{\mathbf{k}}))^*. \quad (\text{S4.18})$$

This expression depends on the directions $\hat{\mathbf{k}}$, $\hat{\mathbf{k}}'$, and $\hat{\mathbf{q}}$, but not explicitly on their norm, except for the $\frac{1}{q}$ factor.

We focus then on the denominators appearing in the self-energy expression Eq. (8) of the main text, evaluated at $\omega = \varepsilon_c$, following Eq. (9) of the main text. Again we consider only the intraband terms, for bands n' degenerate with c in the $q \rightarrow 0$ limit. These bands, similarly to the c state, are unoccupied, hence $f_{\mathbf{k}'n'} = 0$. Thus the numerator of the second term in Eq. (4.4.2) of the main text vanishes. We find:

$$\lim_{q \rightarrow 0} (\varepsilon_c - \varepsilon_{\mathbf{k}'n'} - \omega_{\mathbf{q}j}) = -\frac{q^2}{2m_{n'}^*(\hat{\mathbf{k}}')} - \omega_{j0}(\hat{\mathbf{q}}). \quad (\text{S4.19})$$

See Refs. [118] and [61] for the detailed treatment of $m_{n'}^*(\hat{\mathbf{k}}')$, the directional effective mass associated to one electronic band among a set of degenerate bands at an extremum.

Having clarified the electronic, phononic and electron-phonon interaction ingredients, we can write the corresponding Hamiltonian, Eqs. (4.2.1)–(4.2.5) of the main text that indeed generalizes properly Supplementary Eqs. (S4.9)–(S4.13).

Lowest-order perturbation treatment. Given the Hamiltonian, the lowest-order perturbation ZPR value can be inferred by integration over the radial q coordinate. One obtains for the conduction bands:

$$\begin{aligned} \text{ZPR}_c^{\text{gFr}} &= -\frac{1}{\pi\Omega_0} \int_{4\pi} d\hat{\mathbf{q}} \sum_{jn'} \frac{|s_{n'c}(\hat{\mathbf{q}})|^2}{2\omega_{j0}(\hat{\mathbf{q}})} \\ &\times \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2 \left(\frac{2m_{n'}^*(\hat{\mathbf{q}})}{\omega_{j0}(\hat{\mathbf{q}})} \right)^{1/2} \frac{\pi}{2}. \end{aligned} \quad (\text{S4.20})$$

Let us denote by n_{deg} the degeneracy of the band edge. In the non-degenerate case ($n_{\text{deg}} = 1$), only $n' = c$ has to be considered, then obviously $|s_{n'c}(\hat{\mathbf{q}})| = 1$ independently of $\hat{\mathbf{q}}$. Grouping terms adequately, we obtain

$$\begin{aligned} \text{ZPR}_c^{\text{gFr}} &= -\sum_j \frac{1}{\sqrt{2}\Omega_0} \int_{4\pi} d\hat{\mathbf{q}} (m_c^*(\hat{\mathbf{q}}))^{1/2} \\ &\times (\omega_{j0}(\hat{\mathbf{q}}))^{-3/2} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2. \end{aligned} \quad (\text{S4.21})$$

The link can then be established with Supplementary Eq. (S4.14) in case of isotropic effective mass, isotropic dielectric tensor, and single LO phonon branch.

Let us further examine the degenerate case ($n_{\text{deg}} > 1$). We suppose that the degeneracy is not accidental, but due to symmetry reasons. In this case, we can exploit the symmetry to show that the computation $|s_{n'c}(\hat{\mathbf{q}})|^2$ can be avoided, as this factor can be replaced by $1/n_{\text{deg}}$ in the integral. We will present the demonstration for the case of a three-fold degeneracy, that arises due to a cubic space group. This demonstration could be made general thanks to group theory.

At the bottom of the conduction band, located at the \mathbf{k}_0 wavevector, but that for convenience we again choose to be Γ , we set up a basis of three orthonormalized eigenfunctions, denoted $\{|X\rangle, |Y\rangle, |Z\rangle\}$, that form an irreducible representation of the symmetry group. The eigenfunction $|\Gamma, c\rangle$ for which we compute the $\text{ZPR}_c^{\text{gFr}}$, can be decomposed in this basis, with coefficients (u_{cX}, u_{cY}, u_{cZ}) :

$$|\Gamma, c\rangle = u_{cX}|X\rangle + u_{cY}|Y\rangle + u_{cZ}|Z\rangle. \quad (\text{S4.22})$$

The eigenfunctions $|q\hat{\mathbf{q}}, n'\rangle$ over which the n' sum and $\hat{\mathbf{q}}$ integral are done in Supplementary Eq. (S4.20) can be decomposed as well in this basis:

$$|q\hat{\mathbf{q}}, n'\rangle = s_{n'X}(\hat{\mathbf{q}})|X\rangle + s_{n'Y}(\hat{\mathbf{q}})|Y\rangle + s_{n'Z}(\hat{\mathbf{q}})|Z\rangle \quad (\text{S4.23})$$

Thus Supplementary Eq. (S4.20) becomes

$$\text{ZPR}_c^{\text{gFr}} = \sum_{m \in \{X, Y, Z\}} \sum_{m' \in \{X, Y, Z\}} u_{cm}^* T_{mm'} u_{cm'} \quad (\text{S4.24})$$

where

$$T_{mm'} = - \sum_{jn'} \frac{1}{\sqrt{2}\Omega_0} \int_{4\pi} d\hat{\mathbf{q}} s_{n'm}(\hat{\mathbf{q}}) s_{n'm'}^*(\hat{\mathbf{q}}) (m_{n'}^*(\hat{\mathbf{q}}))^{1/2} (\omega_{j0}(\hat{\mathbf{q}}))^{-3/2} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2. \quad (\text{S4.25})$$

The second-rank tensor $T_{mm'}$ does not depend on the state c . However the same $\text{ZPR}_c^{\text{gFr}}$ must be obtained from Supplementary Eq. (S4.24) for any symmetrically equivalent state c' . Hence, the tensor $T_{mm'}$ must be invariant under all the symmetry operations of the group, which means

$$T_{mm'} = t \delta_{mm'}, \quad (\text{S4.26})$$

where δ is the Kronecker symbol. t can be evaluated thanks to the trace of $T_{mm'}$,

$$n_{\text{deg}} \cdot t = - \sum_{jn'} \frac{1}{\sqrt{2}\Omega_0} \int_{4\pi} d\hat{\mathbf{q}} \left(\sum_{m \in \{X,Y,Z\}} s_{n'm}(\hat{\mathbf{q}}) s_{n'm}^*(\hat{\mathbf{q}}) \right) (m_{n'}^*(\hat{\mathbf{q}}))^{1/2} (\omega_{j0}(\hat{\mathbf{q}}))^{-3/2} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2, \quad (\text{S4.27})$$

yielding

$$\text{ZPR}_c^{\text{gFr}} = - \sum_{jn} \frac{1}{\sqrt{2}\Omega_0 n_{\text{deg}}} \int_{4\pi} d\hat{\mathbf{q}} (m_n^*(\hat{\mathbf{q}}))^{1/2} (\omega_{j0}(\hat{\mathbf{q}}))^{-3/2} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2, \quad (\text{S4.28})$$

since $s_{nm}(\hat{\mathbf{k}})$ is unitary. Similar equations can be derived for the top of the valence band, yielding $\text{ZPR}_v^{\text{gFr}}$.

6. Generalized Fröhlich model with parameters from first-principles calculations

Supplementary Table 4.5 presents the ZPR for the bottom of the conduction band and the top of the valence band from the present first-principles calculations: either based on the AHC approach, or from feeding the few relevant first-principles parameters in the generalized Fröhlich model. These data are used in Fig. 4.5.

Supplementary Table 4.6 presents the parameters computed from first principles, fed in the (generalized) Fröhlich model, only for materials with one phonon branch, with cubic symmetry and edges with non-degenerate electronic state, though including both isotropic or anisotropic effective masses. The parameters for other materials will be reported elsewhere. Most of such edges are found at the bottom of the conduction band, as most top of valence bands are degenerate. The information in Supplementary Table 4.6 allows one to compute

the Fröhlich α parameter for these edges, ranging from 0.016 to 4.757. The lowest-order perturbation results are rather accurate in this range. A similar parameter to measure the breakdown of the perturbation theory in the case of degenerate electronic states is still to be established to our knowledge.

7. Generalized Fröhlich model with parameters from experimental data

We discuss how a few-parameter Fröhlich Hamiltonian, of the type we have introduced, can be obtained from experimental data, thus constituting a few-parameter model of the ZPR of real materials.

In the Fröhlich model [39, 40, 115, 116] four parameters are needed: the LO phonon frequency ω_{LO} , the electronic dielectric constant ϵ^∞ , the low-frequency dielectric constant ϵ^0 , and the effective mass m^* . These four parameters combine to give the α parameter mentioned in Supplementary Eq. (S4.13). The Fröhlich Hamiltonian can be expressed in the natural units of electronic effective mass and phonon frequency, to make appear α as the single parameter defining the Fröhlich Hamiltonian [115, 117]. Despite its simplicity, this Fröhlich Hamiltonian has resisted analytical treatments.

One can consider the generalization of such Fröhlich Hamiltonian to arbitrary electronic dispersion, phonon dispersion or electron-phonon interaction, as mentioned in the introduction of Ref. [43]. However, one has a model for a real material only if the parameters describing such dispersions and such interaction are available. Ideally, there should be as few parameters as possible if the goal is to extract the relevant physics.

In this respect, the Hamiltonian in Eqs. (4.2.1)–(4.2.5) of the main text is the proper generalization of Supplementary Eqs. (S4.9)–(S4.14) to all band extrema (degenerate or not, isotropic or anisotropic) in all crystal types (cubic or not), because it can be parameterized with only a few numbers, and all such parameters can be found, in principle, from experiment, even if, in the present work, they were deduced from first principles.

Indeed, the electronic part, Eq. (4.2.2) of the main text, is fully determined if the so-called Luttinger-Kohn parameters [120] are available. Taking the example of the top of the valence band of a cubic material, with three-fold degeneracy, only three Luttinger-Kohn parameters, called A , B , and C in Ref. [120] are needed to obtain the direction-dependent effective masses, $m_n^*(\hat{\mathbf{k}})$, of the three electronic bands, for all directions \mathbf{k} . Similarly, the direction-dependent phonon frequencies $\omega_{j0}(\hat{\mathbf{q}})$ and their mode-polarity vectors $p_{\gamma,j}(\hat{\mathbf{q}})$ can be obtained from neutron and infra-red spectra measurements and parametrized with a few numbers. For a cubic material with two atoms per unit cell, there is only one LO frequency, with mode-polarity vector directly determined from the LO-TO splitting. The optic dielectric

constant along any direction \mathbf{q} comes from the knowledge of a 3x3 symmetric tensor, easily obtained by optical measurements.

Conversely, in Ref. [54], by Verdi and Giustino, a model *for the electron-phonon interaction (or vertex)* is proposed (at variance with our model of the full Hamiltonian, with a simplified vertex). It is used for a very different purpose. In this very nice study, Verdi and Giustino used the Fröhlich idea and the microscopic theory developed by Vogl [53] to model the non-analytic behaviour of the electron-phonon vertex in the $\mathbf{q} \rightarrow \mathbf{0}$ limit. This approach was instrumental to make the electron-phonon interaction amenable to Fourier interpolation in an attempt to diminish the computational resources needed to capture the long-range behaviour of the scattering potentials and perform mobility calculations using Wannier functions. However, their work does not propose a ZPR model or Fröhlich Hamiltonian per se. Indeed, the electronic dispersion in Verdi and Giustino is the full first-principles dispersion in the entire Brillouin Zone and similarly for the phonon dispersion, unlike the one of the original Fröhlich approach (one parameter for each of these dispersions), and the present one, which relies on few parameters, the number of which depends on the crystal type and symmetries, and on the Luttinger-Kohn parameters.

Supplementary Discussion

A. ASC Formalism and relation with the AHC

In the adiabatic supercell approach, the temperature-dependent average band edges (here written for the bottom of the conduction band) are obtained from

$$\langle \varepsilon_c(T) \rangle = Z_I^{-1} \sum_m \exp(-\beta E_m) \langle \varepsilon_c \rangle_m, \quad (\text{S4.29})$$

where $\beta = k_B T$ (with k_B the Boltzmann constant), T the temperature, Z_I the canonical partition function among the quantum nuclear states m with energies E_m ($Z_I = \sum_m \exp(-\beta E_m)$), and $\langle \varepsilon_c \rangle_m$ the band edge average taken over the corresponding many-body nuclear wavefunction. See e.g. Eq. (4) of Ref. [121].

Such approach allows one to perform nonperturbative calculations, and thus includes effects beyond the adiabatic AHC approach, as emphasized in Sec. XI.A.2 of Ref. [38]. It also allows one to use as starting point the GW eigenenergies, instead of the DFT ones, as proposed and used by some of us [32], and further exploited by Monserrat [87] and Karsai et al. [37].

As detailed in Ref. [4], and illustrated in Fig. 4 of this reference, one can dissociate the harmonic approximation for the nuclear positions from the (non)quadratic behaviour of the eigenenergies as a function of the nuclear positions. While phonon anharmonicities can be safely ignored for most of the materials in the present context (except SrTiO₃, see the

METHODS section of the main text, the (non)quadratic behaviour might very important in the ASC formalism. Within the harmonic approximation, one finds Eq. (19) of Ref. [4]:

$$\langle \varepsilon_c(T) \rangle = Z_I^{-1} \sum_{\mathbf{n}} \exp(-\beta E[\mathbf{n}]) \int d\mathbf{z} |\chi_{\mathbf{n}}(\mathbf{z})|^2 \varepsilon_c(\mathbf{z}). \quad (\text{S4.30})$$

In this equation, \mathbf{z} denotes the ensemble of all phonon coordinate amplitudes, while \mathbf{n} denotes the ensemble of all phonon occupation numbers. As such, this expression does not diverge for IR-active materials, thanks to the sampling over coordinate amplitudes and the $\varepsilon_c(\mathbf{z})$ asymptotic linear behaviour, see Fig. 4 of Ref. [4] or Fig. 2 of Ref. [35]. By contrast, making the quadratic approximation for the $\varepsilon_c(\mathbf{z})$ yields the AHC result for the Fan term, and hence, a divergence: the curvature becomes indeed infinite for IR-active longwavelength phonons. An extreme nonquadratic behavior is also observed in the degenerate case, as illustrated in Fig. 1 of Ref. [37]. The connection with the adiabatic AHC result has been analyzed in detail in Ref. [26].

However, obviously, Supplementary Eq. (S4.29) as well as Supplementary Eq. (S4.30) are based on the adiabatic approximation: the quantum nuclear state is not responding to the presence of an added or removed electron. Instead, Supplementary Eq. (S4.29) should have reflected such dependence to be correct. The correlation between electron and nuclei is not present in Supplementary Eq. (S4.29): the electron is affected by the phonons, but the phonons are incorrectly not affected by the addition of an electron. For IR-active materials, as explained in the main text, the lack of such correlation has important consequences for the predictive capability of the method.

As another consequence of the time-dependent atomic motion, the nonquadratic behaviour, present in the adiabatic treatment, loses its importance. The behaviour observed for degenerate electronic state, in the adiabatic approximation, exemplified by Fig. 1 of Ref. [37] might nevertheless still be relevant in the context of the Jahn-Teller effect. Anyhow, the transition between states, due to the first-order electron-phonon interaction, always involves a modification of the energy by the emission or absorption of a phonon. Hence, the corresponding initial and final energies are not identical.

B. The non-quadratic effect for IR-inactive materials

Coming back to the C-diamond case, the ‘‘anharmonic effect’’ in the adiabatic approximation is argued to be large in the above-mentioned Sec. III of Ref. [4], that some of us co-authored. At present, we call it more properly ‘‘non-quadratic effect’’ following Ref. [35]. However, this study was performed within the adiabatic approximation. In the case of infrared active materials, without such ‘‘non-quadratic effects’’, the ZPR is infinite in the adiabatic approximation. The non-quadratic effects succeed to bring the adiabatic ZPR to

a finite value, i.e. it has “infinite” impact. For C, without such “non-quadratic effects”, the ZPR is finite in the adiabatic approximation, but nevertheless comes from the integration of a diverging integrand. Then the non-quadratic effects modify the adiabatic harmonic value by 40% by suppressing this diverging integrand. Still, despite yielding a finite value, the adiabatic approximation is incorrect to start with: if atomic motion effects are taken into account (and they should be taken into account from the very start, since the physics is the one of electrons not being able to follow the atomic motions), there is no diverging integrand that the non-quadratic effects would reduce afterwards. Thus, the crucial role of atom dynamics cannot be ignored even in diamond. We actually think that the estimation of non-quadratic effects in our publication Ref. [4] or in Ref. [35] cannot be transposed to the non-adiabatic treatment, even for the case of C. The non-quadratic effects in Ref. [4] or Ref. [35] remove a non-existing divergence, either global (IR-active materials), or integrable (non-IR-active materials), spuriously introduced by the adiabatic approximation.

C. GW approximation within the ASC

In such calculation of the ZPR of diamond, Ref. [32], we have obtained a 40% change of direct gap ZPR when the quantities needed in the ASC approach were obtained from GW instead of DFT. This effect appears to be comparable to the difference between AHC and experimental values reported in the present manuscript, while (as for the non-quadratic effect) C is an IR-inactive material.

In Ref. [37], a similar change was reproduced for the direct gap, but the change for the indirect (true) gap of diamond was much smaller, about 23%. For the other materials studied in Ref. [37], using the ASC approach, the ratio between GW and DFT, systematically obtained, was usually closer to unity, but not always. One can wonder whether such difference between GW and DFT would not also be obtained within the non-adiabatic AHC approach.

The estimation of such an effect within the non-adiabatic treatment is at present out of reach. Actually, as already argued in the previous sections, one can hardly justify transposing the conclusions obtained in the ASC approach to the non-adiabatic AHC approach, because the ASC approach relies on the non-quadratic coupling to eliminate the divergence of the ZPR, while for most materials, we believe such elimination to originate from the coupled dynamics of electron and phonons. In view of the agreement obtained in Fig. 4.3 for non-adiabatic AHC+DFT, we can simply infer that the DFT to GW effect is apparently smaller in the non-adiabatic AHC case than in the ASC case.

Also, let us emphasize that such a DFT to GW effect can hardly equal the non-adiabatic effect reported in the present manuscript. Indeed, for IR-active materials, such effect suppresses an infinity and replaces it by a finite value, which is far more important than a 40% modification. Even if the non-quadratic effect is taken into account, Fig. 4.3 shows that the

difference between ASC and AHC is not merely 40% for many materials, but can be as large as a factor of 2 or 3.

So, we do not deny that other effects, like GW, could be important at the level of the 25% agreement that we show in Fig. 4.3. However, the non-adiabatic effect is not simply a 25% modification, it reduces a divergent quantity (possibly integrable) to a finite value.

Bibliography

- [1] P. B. Allen and M. Cardona. Theory of the temperature dependence of the direct gap of germanium. *Phys. Rev. B*, 23(4):1495–1505, 1981. doi: [10.1103/PhysRevB.23.1495](https://doi.org/10.1103/PhysRevB.23.1495).
- [2] P. Allen and M. Cardona. Temperature dependence of the direct gap of Si and Ge. *Phys. Rev. B*, 27(8):4760, 1983. doi: [10.1103/PhysRevB.23.1495](https://doi.org/10.1103/PhysRevB.23.1495).
- [3] S. Poncé, Y. Gillet, J. Laflamme Janssen, A. Marini, M. Verstraete, and X. Gonze. Temperature dependence of the electronic structure of semiconductors and insulators. *J. Chem. Phys.*, 143(10):102813, 2015. doi: [10.1063/1.4927081](https://doi.org/10.1063/1.4927081).
- [4] G. Antonius, S. Poncé, E. Lantagne-Hurtubise, G. Auclair, X. Gonze, and M. Côté. Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure. *Phys. Rev. B*, 92(8):085137, 2015. doi: [10.1103/PhysRevB.92.085137](https://doi.org/10.1103/PhysRevB.92.085137).
- [5] J. P. Nery and P. B. Allen. Influence of Fröhlich polaron coupling on renormalized electron bands in polar semiconductors: Results for zinc-blende GaN. *Phys. Rev. B*, 94(11):115135, 2016. doi: [10.1103/PhysRevB.94.115135](https://doi.org/10.1103/PhysRevB.94.115135).
- [6] M. Engel, H. Miranda, L. Chaput, A. Togo, C. Verdi, M. Marsman, and G. Kresse. Zero-point renormalization of the band gap of semiconductors and insulators using the projector augmented wave method. *Phys. Rev. B*, 106(9):094316, 2022. doi: [10.1103/PhysRevB.106.094316](https://doi.org/10.1103/PhysRevB.106.094316).
- [7] P. Y. Yu and M. Cardona. *Fundamentals of semiconductors*. Springer-Verlag, Berlin, 4th edition, 2010.
- [8] R. M. Martin. *Electronic Structure: Basic Theory and Methods*. Cambridge University Press, Cambridge, 2004.
- [9] J. P. Perdew and M. Levy. Physical content of the exact Kohn-Sham orbital energies: Band gaps and derivative discontinuities. *Phys. Rev. Lett.*, 51(20):1884–1887, 1983. doi: [10.1103/PhysRevLett.51.1884](https://doi.org/10.1103/PhysRevLett.51.1884).
- [10] L. J. Sham and M. Schlüter. Density-functional theory of the energy gap. *Phys. Rev. Lett.*, 51(20):1888–1891, 1983. doi: [10.1103/PhysRevLett.51.1888](https://doi.org/10.1103/PhysRevLett.51.1888).
- [11] R. Martin, L. Reining, and D. Ceperley. *Interacting Electrons. Theory and Computational Approaches*. Cambridge University Press, Cambridge, 2016.

- [12] M. S. Hybertsen and S. G. Louie. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B*, 34(8):5390–5413, 1986. doi: [10.1103/PhysRevB.34.5390](https://doi.org/10.1103/PhysRevB.34.5390).
- [13] M. van Schilfgaarde, T. Kotani, and S. Faleev. Quasiparticle self-consistent GW theory. *Phys. Rev. Lett.*, 96(22):226402, 2006. doi: [10.1103/PhysRevLett.96.226402](https://doi.org/10.1103/PhysRevLett.96.226402).
- [14] M. Shishkin, M. Marsman, and G. Kresse. Accurate quasiparticle spectra from self-consistent GW calculations with vertex corrections. *Phys. Rev. Lett.*, 99(24):246403, 2007. doi: [10.1103/PhysRevLett.99.246403](https://doi.org/10.1103/PhysRevLett.99.246403).
- [15] P. B. Allen and V. Heine. Theory of the temperature dependence of electronic band structures. *J. Phys. C*, 9(12):2305–2312, 1976. doi: [10.1088/0022-3719/9/12/013](https://doi.org/10.1088/0022-3719/9/12/013).
- [16] H. Y. Fan. Temperature dependence of the energy gap in semiconductors. *Phys. Rev.*, 82:900–905, 1951. doi: [10.1103/PhysRev.78.808.2](https://doi.org/10.1103/PhysRev.78.808.2).
- [17] E. Antoncik. On the theory of temperature shift of the absorption curve in non-polar crystals. *Czechoslovakij Fiziceskij Zurnal*, 5:449–461, 1955. doi: [10.1007/BF01687209](https://doi.org/10.1007/BF01687209).
- [18] C. Kim, P. Lautenschlager, and M. Cardona. Temperature dependence of the fundamental energy gap in GaAs. *Solid state communications*, 59(12):797–802, 1986. doi: [10.1016/0038-1098\(86\)90632-0](https://doi.org/10.1016/0038-1098(86)90632-0).
- [19] S. Zollner, S. Gopalan, and M. Cardona. The temperature dependence of the band gaps in InP, InAs, InSb, and GaSb. *Sol. State Comm.*, 77(7):485–488, 1991. doi: [10.1016/0038-1098\(91\)90725-B](https://doi.org/10.1016/0038-1098(91)90725-B).
- [20] A. Marini. Ab-initio finite temperature excitons. *Phys. Rev. Lett.*, 101(10):106405, 2008. doi: [10.1103/PhysRevLett.101.106405](https://doi.org/10.1103/PhysRevLett.101.106405).
- [21] F. Giustino, S. Louie, and M. Cohen. Electron-phonon renormalization of the direct band gap of diamond. *Phys. Rev. Lett.*, 105(26):265501, 2010. doi: [10.1103/PhysRevLett.105.265501](https://doi.org/10.1103/PhysRevLett.105.265501).
- [22] E. Cannuccia and A. Marini. Effect of the quantum zero-point atomic motion on the optical and electronic properties of diamond and trans-polyacetylene. *Phys. Rev. Lett.*, 107(25):255501, 2011. doi: [10.1103/PhysRevLett.107.255501](https://doi.org/10.1103/PhysRevLett.107.255501).
- [23] X. Gonze, P. Boulanger, and M. Côté. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Annalen der Physik (Leipzig)*, 523(1-2):168–178, 2011. doi: [10.1002/andp.201000100](https://doi.org/10.1002/andp.201000100).
- [24] E. Cannuccia and A. Marini. Zero point motion effect on the electronic properties of diamond, trans-polyacetylene and polyethylene. *Eur. Phys. J. B*, 85:320, 2012. doi: [10.1140/epjb/e2012-30105-4](https://doi.org/10.1140/epjb/e2012-30105-4).
- [25] H. Kawai, K. Yamashita, E. Cannuccia, and A. Marini. Electron-electron and electron-phonon correlation effects on the finite-temperature electronic and optical properties of zinc-blende GaN. *Phys. Rev. B*, 89(8):085202, 2014. doi: [10.1103/PhysRevB.89.085202](https://doi.org/10.1103/PhysRevB.89.085202).

- [26] S. Poncé, G. Antonius, P. Boulanger, E. Cannuccia, A. Marini, M. Côté, and X. Gonze. Verification of first-principles codes: Comparison of total energies, phonon frequencies, electron–phonon coupling and zero-point motion correction to the gap between ABINIT and QE/Yambo. *Comp. Mat. Sci.*, 83:341–348, 2014. doi: [10.1016/j.commatsci.2013.11.031](https://doi.org/10.1016/j.commatsci.2013.11.031).
- [27] S. Poncé, G. Antonius, Y. Gillet, P. Boulanger, J. Laflamme Janssen, A. Marini, M. Côté, and X. Gonze. Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation. *Phys. Rev. B*, 90(21):214304, 2014. doi: [10.1103/PhysRevB.90.214304](https://doi.org/10.1103/PhysRevB.90.214304).
- [28] J. Nery, P. Allen, G. Antonius, L. Reining, A. Miglio, and X. Gonze. Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron. *Phys. Rev. B*, 97(11):115145, 2018. doi: [10.1103/PhysRevB.97.115145](https://doi.org/10.1103/PhysRevB.97.115145).
- [29] R. B. Capaz, C. D. Spataru, P. Tangney, M. L. Cohen, and S. G. Louie. Temperature dependence of the band gap of semiconducting carbon nanotubes. *Phys. Rev. Lett.*, 94(3):036801, 2005. doi: [10.1103/PhysRevLett.94.036801](https://doi.org/10.1103/PhysRevLett.94.036801).
- [30] C. E. Patrick and F. Giustino. Quantum nuclear dynamics in the photophysics of diamondoids. *Nature Communications*, 4:2006, 2013. doi: [10.1038/ncomms3006](https://doi.org/10.1038/ncomms3006).
- [31] B. Monserrat, N. Drummond, and R. Needs. Anharmonic vibrational properties in periodic systems: energy, electron-phonon coupling, and stress. *Phys. Rev. B*, 87(14):144302, 2013. doi: [10.1103/PhysRevB.87.144302](https://doi.org/10.1103/PhysRevB.87.144302).
- [32] G. Antonius, S. Poncé, P. Boulanger, M. Côté, and X. Gonze. Many-body effects on the zero-point renormalization of the band structure. *Phys. Rev. Lett.*, 112(21):215501, 2014. doi: [10.1103/PhysRevLett.112.215501](https://doi.org/10.1103/PhysRevLett.112.215501).
- [33] B. Monserrat, N. Drummond, C. Pickard, and R. Needs. Electron-phonon Coupling and the metallization of solid helium at terapascal pressures. *Phys. Rev. Lett.*, 112(5):055504, 2014. doi: [10.1103/PhysRevLett.112.055504](https://doi.org/10.1103/PhysRevLett.112.055504).
- [34] B. Monserrat and R. Needs. Comparing electron-phonon coupling strength in diamond, silicon, and silicon carbide: First-principles study. *Phys. Rev. B*, 89(21):214304, 2014. doi: [10.1103/PhysRevB.89.214304](https://doi.org/10.1103/PhysRevB.89.214304).
- [35] B. Monserrat, E. Engel, and R. Needs. Giant electron-phonon interactions in molecular crystals and the importance of nonquadratic coupling. *Phys. Rev. B*, 92(14):140302(R), 2015. doi: [10.1103/PhysRevB.92.140302](https://doi.org/10.1103/PhysRevB.92.140302).
- [36] E. A. Engel, B. Monserrat, and R. J. Needs. Vibrational renormalisation of the electronic band gap in hexagonal and cubic ice. *J. Chem. Phys.*, 143(24):244708, 2015. doi: [10.1063/1.4938029](https://doi.org/10.1063/1.4938029).
- [37] F. Karsai, M. Engel, G. Kresse, and E. Flage-Larsen. Electron–phonon coupling in semiconductors within the GW approximation. *New Journal of Physics*, 20(12):123008,

2018. doi: [10.1088/1367-2630/aaf53f](https://doi.org/10.1088/1367-2630/aaf53f).
- [38] F. Giustino. Electron-phonon interactions from first principles. *Rev. Mod. Phys.*, 89(1):015003, 2017. doi: [10.1103/RevModPhys.89.015003](https://doi.org/10.1103/RevModPhys.89.015003).
- [39] H. Fröhlich. Electrons in lattice fields. *Adv. Phys.*, 3(11):325–361, 1954. doi: [10.1080/00018735400101213](https://doi.org/10.1080/00018735400101213).
- [40] R. P. Feynman. Slow electrons in a polar crystal. *Phys. Rev.*, 97(3):660–665, 1955. doi: [10.1103/PhysRev.97.660](https://doi.org/10.1103/PhysRev.97.660).
- [41] A. Mishchenko, N. Prokof'ev, A. Sakamoto, and B. Svistunov. Diagrammatic quantum Monte Carlo study of the Fröhlich polaron. *Phys. Rev. B*, 62(10):6317–6336, 2000. doi: [10.1103/PhysRevB.62.6317](https://doi.org/10.1103/PhysRevB.62.6317).
- [42] A. S. Mishchenko, N. Nagaosa, N. Prokof'ev, A. Sakamoto, and B. Svistunov. Optical conductivity of the Fröhlich polaron. *Phys. Rev. Lett.*, 91(23):236401, 2003. doi: [10.1103/PhysRevLett.91.236401](https://doi.org/10.1103/PhysRevLett.91.236401).
- [43] J. Devreese and A. Alexandrov. Fröhlich polaron and bipolaron: recent developments. *Rep. Prog. Phys.*, 72(6):066501, 2009. doi: [10.1088/0034-4885/72/6/066501](https://doi.org/10.1088/0034-4885/72/6/066501).
- [44] D. Emin. *Polarons*. Cambridge University Press, Cambridge, 2012.
- [45] T. Hahn, S. Klimin, J. Tempere, J. T. Devreese, and C. Franchini. Diagrammatic Monte Carlo study of Fröhlich polaron dispersion in two and three dimensions. *Phys. Rev. B*, 97(13):134305, 2018. doi: [10.1103/PhysRevB.97.134305](https://doi.org/10.1103/PhysRevB.97.134305).
- [46] KS-DFT ASC Calculations in Ref. [37] are based on GGA-PBE, except for Si, Ge, GaAs and CdSe, where the PBE0 hybrid functional has been used. Thus the better score of Ge and GaAs for ASC calculations than for AHC calculations might be partly explained by this different KS-DFT functional.
- [47] M. Cardona and M. Thewalt. Isotope effects on the optical spectra of semiconductors. *Rev. Mod. Phys.*, 77(4):1173–1224, 2005. doi: [10.1103/RevModPhys.77.1173](https://doi.org/10.1103/RevModPhys.77.1173).
- [48] T. Rangel *et al.* Reproducibility in G_0W_0 calculations for solids. *Comput. Phys. Commun.*, 255:107242, 2020. doi: [10.1016/j.cpc.2020.107242](https://doi.org/10.1016/j.cpc.2020.107242).
- [49] G. D. Mahan. Temperature dependence of the band gap in CdTe. *J. Phys. Chem. Solids*, 26(4):751–756, 1965. doi: [10.1016/0022-3697\(65\)90029-6](https://doi.org/10.1016/0022-3697(65)90029-6).
- [50] H. R. Trebin and U. Rössler. Polarons in the degenerate-band case. *Phys. Status Solidi B*, 70(2):717–726, 1975. doi: [10.1002/pssb.2220700232](https://doi.org/10.1002/pssb.2220700232).
- [51] J. T. Devreese, S. N. Klimin, J. L. M. van Mechelen, and D. van der Marel. Many-body large polaron optical conductivity in $\text{SrTi}_{(1-x)}\text{Nb}_x\text{O}_3$. *Phys. Rev. B*, 81(12):125119, 2010. doi: [10.1103/PhysRevB.81.125119](https://doi.org/10.1103/PhysRevB.81.125119).
- [52] M. Schlipf, S. Poncé, and F. Giustino. Carrier lifetimes and polaronic mass enhancement in the hybrid halide perovskite $\text{CH}_3\text{NH}_3\text{PbI}_3$ from multiphonon Fröhlich coupling. *Phys. Rev. Lett.*, 121(8):086402, 2018. doi: [10.1103/PhysRevLett.121.086402](https://doi.org/10.1103/PhysRevLett.121.086402).

- [53] P. Vogl. Microscopic theory of electron-phonon interaction in insulators or semiconductors. *Phys. Rev. B*, 13(2):694–704, 1976. doi: [10.1103/PhysRevB.13.694](https://doi.org/10.1103/PhysRevB.13.694).
- [54] C. Verdi and F. Giustino. Fröhlich electron-phonon vertex from first principles. *Phys. Rev. Lett.*, 115(17):176401, 2015. doi: [10.1103/PhysRevLett.115.176401](https://doi.org/10.1103/PhysRevLett.115.176401).
- [55] X. Gonze, F. Jollet, F. Abreu Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J. M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. Da Pieve, M. Delaveau, M. Di Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, A. Gerossier, M. Giantomassi, Y. Gillet, D. R. Hamann, L. He, G. Jomard, J. Laflamme Janssen, S. Le Roux, A. Levitt, A. Lherbier, F. Liu, I. Lukačević, A. Martin, C. Martins, M. J. T. Oliveira, S. Poncé, Y. Pouillon, T. Rangel, G. M. Rignanese, A. H. Romero, B. Rousseau, O. Rubel, A. A. Shukri, M. Stankovski, M. Torrent, M. J. Van Setten, B. Van Troeye, M. J. Verstraete, D. Waroquiers, J. Wiktoria, B. Xu, A. Zhou, and J. W. Zwanziger. Recent developments in the ABINIT software package. *Comput. Phys. Commun.*, 205:106–131, 2016. doi: [10.1016/j.cpc.2016.04.003](https://doi.org/10.1016/j.cpc.2016.04.003).
- [56] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, 1996. doi: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [57] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese. The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Comput. Phys. Commun.*, 226:39–54, 2018. doi: [10.1016/j.cpc.2018.01.012](https://doi.org/10.1016/j.cpc.2018.01.012).
- [58] B. Montanari and N. M. Harrison. Lattice dynamics of TiO₂ rutile: influence of gradient corrections in density functional calculations. *Chem. Phys. Lett.*, 364(5-6):528–534, 2002. doi: [10.1016/S0009-2614\(02\)01401-X](https://doi.org/10.1016/S0009-2614(02)01401-X).
- [59] S. Baroni, S. de Gironcoli, A. D. Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 73(2):515–562, 2001. doi: [10.1103/RevModPhys.73.515](https://doi.org/10.1103/RevModPhys.73.515).
- [60] X. Gonze and C. Lee. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B*, 55(16):10355–10368, 1997. doi: [10.1103/PhysRevB.55.10355](https://doi.org/10.1103/PhysRevB.55.10355).
- [61] J. Laflamme Janssen, Y. Gillet, S. Poncé, A. Martin, M. Torrent, and X. Gonze. Precise effective masses from density functional perturbation theory. *Phys. Rev. B*, 93(20):205147, 2016. doi: [10.1103/PhysRevB.93.205147](https://doi.org/10.1103/PhysRevB.93.205147).
- [62] M. Guennou, P. Bouvier, J. Kreisler, and D. Machon. Pressure-temperature phase diagram of SrTiO₃ up to 53 GPa. *Phys. Rev. B*, 81(5):054115, 2010. doi: [10.1103/PhysRevB.81.054115](https://doi.org/10.1103/PhysRevB.81.054115).
- [63] T. Tadano and S. Tsuneyuki. Self-consistent phonon calculations of lattice dynamical properties in cubic SrTiO₃ with first-principles anharmonic force constants. *Phys. Rev.*

- B*, 92(5):054301, 2015. doi: [10.1103/PhysRevB.92.054301](https://doi.org/10.1103/PhysRevB.92.054301).
- [64] K. A. Müller and H. Burkard. SrTiO₃: An intrinsic quantum paraelectric below 4 K. *Phys. Rev. B*, 19(7):3593–3602, 1979. doi: [10.1103/PhysRevB.19.3593](https://doi.org/10.1103/PhysRevB.19.3593).
- [65] W. Zhong and D. Vanderbilt. Effect of quantum fluctuations on structural phase transitions in SrTiO₃ and BaTiO₃. *Phys. Rev. B*, 53(9):5047–5050, 1996. doi: [10.1103/PhysRevB.53.5047](https://doi.org/10.1103/PhysRevB.53.5047).
- [66] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak. Temperature dependent effective potential method for accurate free energy calculations of solids. *Phys. Rev. B*, 87(10):104111, 2013. doi: [10.1103/PhysRevB.87.104111](https://doi.org/10.1103/PhysRevB.87.104111).
- [67] T. Tadano, Y. Gohda, and S. Tsuneyuki. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *J. Phys. Condens. Matter*, 26(22):225402, 2014. doi: [10.1088/0953-8984/26/22/225402](https://doi.org/10.1088/0953-8984/26/22/225402).
- [68] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino. Polarons from first principles, without supercells. *Phys. Rev. Lett.*, 122(24):246403, 2019. doi: [10.1103/PhysRevLett.122.246403](https://doi.org/10.1103/PhysRevLett.122.246403).
- [69] K. Persson *et al.* Materials project. <https://materialsproject.org>, 2018.
- [70] F. W. Lytle. X-Ray Diffractometry of Low-Temperature Phase Transformations in Strontium Titanate. *J. Appl. Phys.*, 35(7):2212–2215, 1964. doi: [10.1063/1.1702820](https://doi.org/10.1063/1.1702820).
- [71] M. Zacharias, M. Scheffler, and C. Carbogno. Fully anharmonic nonperturbative theory of vibronically renormalized electronic band structures. *Phys. Rev. B*, 102(4):045126, 2020. doi: [10.1103/PhysRevB.102.045126](https://doi.org/10.1103/PhysRevB.102.045126).
- [72] R. Pässler. Dispersion-related description of temperature dependencies of band gaps in semiconductors. *Phys. Rev. B*, 66(8):085201, 2002. doi: [10.1103/PhysRevB.66.085201](https://doi.org/10.1103/PhysRevB.66.085201).
- [73] R. Pässler. Semi-empirical descriptions of temperature dependences of band gaps in semiconductors. *Phys. Status Solidi B*, 236(3):710–728, 2003. doi: [10.1002/pssb.200301752](https://doi.org/10.1002/pssb.200301752).
- [74] Experimental gaps come from Refs. [83], [47], [84], [80], [25], [122], and [123].
- [75] C. Parks, A. K. Ramdas, S. Rodriguez, K. M. Itoh, and E. E. Hailer. Electronic band structure of isotopically pure germanium: Modulated transmission and reflectivity study. *Phys. Rev. B*, 49(20):14244–14250, 1994. doi: [10.1103/PhysRevB.49.14244](https://doi.org/10.1103/PhysRevB.49.14244).
- [76] D. Karaiskaj, M. Thewalt, T. Ruf, M. Cardona, and M. Konuma. Photoluminescence studies of isotopically enriched silicon: isotopic effects on the indirect electronic band gap and phonon energies. *Solid State Communications*, 123(5-6):87–92, 2002. doi: [10.1016/S0038-1098\(02\)00249-1](https://doi.org/10.1016/S0038-1098(02)00249-1).
- [77] J. M. Zhang, T. Ruf, R. Lauck, and M. Cardona. Isotope effects on exciton energies in CdS. *Phys. Rev. B*, 57(16):9716–9722, 1998. doi: [10.1103/PhysRevB.57.9716](https://doi.org/10.1103/PhysRevB.57.9716).
- [78] F. Manjon, M. Mollar, B. Mari, N. Garro, A. Cantarero, R. Lauck, and M. Cardona. Effect of isotopic mass on the photoluminescence spectra of β zinc sulfide. *Solid State*

- Communications*, 133(4):253–258, 2005. doi: [10.1016/j.ssc.2004.10.031](https://doi.org/10.1016/j.ssc.2004.10.031).
- [79] A. T. Collins, S. C. Lawson, G. Davies, and H. Kanda. Indirect energy gap of 13C diamond. *Phys. Rev. Lett.*, 65(7):891–894, 1990. doi: [10.1103/PhysRevLett.65.891](https://doi.org/10.1103/PhysRevLett.65.891).
- [80] A. Benrekia, N. Benkhattou, A. Nassour, M. Driz, M. Sahnoun, and S. Lebegue. Structural, electronic and optical properties of cubic SrTiO₃ and KTaO₃: Ab initio and GW calculations. *Physica B*, 407(13):2632–2636, 2012. doi: [10.1016/j.physb.2012.04.013](https://doi.org/10.1016/j.physb.2012.04.013).
- [81] D. J. Kok, K. Irmscher, M. Naumann, C. Guguschev, Z. Galazka, and R. Uecker. Temperature-dependent optical absorption of SrTiO₃. *Phys. Stat. Sol. (a)*, 212(9):1880–1887, 2015. doi: [10.1002/pssa.201431836](https://doi.org/10.1002/pssa.201431836).
- [82] F. Manjon, M. Mollar, M. Hernandez-Fenollosa, B. Mari, R. Lauck, and M. Cardona. Effect of isotopic mass on the photoluminescence spectra of zinc oxide. *Sol. State Comm.*, 128(1):35–39, 2003. doi: [10.1016/S0038-1098\(03\)00616-1](https://doi.org/10.1016/S0038-1098(03)00616-1).
- [83] M. Shishkin and G. Kresse. Self-consistent GW calculations for semiconductors and insulators. *Phys. Rev. B*, 75(23):235102, 2007. doi: [10.1103/PhysRevB.75.235102](https://doi.org/10.1103/PhysRevB.75.235102).
- [84] M. J. van Setten, M. Giantomassi, X. Gonze, G. M. Rignanese, and G. Hautier. Automation methodologies and large-scale validation for GW: Towards high-throughput GW calculations. *Phys. Rev. B*, 96(15):155207, 2017. doi: [10.1103/PhysRevB.96.155207](https://doi.org/10.1103/PhysRevB.96.155207).
- [85] M. Zacharias and F. Giustino. One-shot calculation of temperature-dependent optical spectra and phonon-induced band-gap renormalization. *Phys. Rev. B*, 94(7):075125, 2016. doi: [10.1103/PhysRevB.94.075125](https://doi.org/10.1103/PhysRevB.94.075125).
- [86] Y. Zhang, Z. Wang, J. Xi, and J. Yang. Temperature-dependent band gaps in several semiconductors: from the role of electron-phonon renormalization. *J. Phys. Condens. Matt.*, 32(47):475503, 2020. doi: [10.1088/1361-648X/aba45d](https://doi.org/10.1088/1361-648X/aba45d).
- [87] B. Monserrat. Correlation effects on electron-phonon coupling in semiconductors: Many-body theory along thermal lines. *Phys. Rev. B*, 93(10):100301(R), 2016. doi: [10.1103/PhysRevB.93.100301](https://doi.org/10.1103/PhysRevB.93.100301).
- [88] We refer to Fig. 8 in the Supplementary Material of Ref. [3] for AlN-w (note that in the figure caption, the unextrapolated values are mentioned), and to Fig. 15 in the Supplementary Material of Ref. [3] for BN-c (note that in the figure caption, the extrapolated values are mentioned). In these captions, there are errors concerning the name of the material corresponding to each figure.
- [89] H. Monkhorst and J. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188–5192, 1976. doi: [10.1103/PhysRevB.13.5188](https://doi.org/10.1103/PhysRevB.13.5188).
- [90] R. Pässler. Parameter sets due to fittings of the temperature dependencies of fundamental bandgaps in semiconductors. *Phys. Status Solidi B*, 216(2):975–1007, 1999. doi: [10.1002/\(SICI\)1521-3951\(199912\)216:2<975::AID-PSSB975>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1521-3951(199912)216:2<975::AID-PSSB975>3.0.CO;2-N).

- [91] N. Garro, A. Cantarero, M. Cardona, A. Göbel, T. Ruf, and K. Eberl. Dependence of the lattice parameters and the energy gap of zinc-blende-type semiconductors on isotopic masses. *Phys. Rev. B*, 54(7):4732–4740, 1996. doi: [10.1103/PhysRevB.54.4732](https://doi.org/10.1103/PhysRevB.54.4732).
- [92] C. Villegas, A. Rocha, and A. Marini. Anomalous thermal dependence of the band gap in black phosphorus. *Nanoletters*, 16(8):5095, 2016. doi: [10.1021/acs.nanolett.6b02035](https://doi.org/10.1021/acs.nanolett.6b02035).
- [93] G. M. Rignanese, J. P. Michenaud, and X. Gonze. Ab initio study of the volume dependence of dynamical and thermodynamical properties of silicon. *Phys. Rev. B*, 53(8):4488–4497, 1996. doi: [10.1103/PhysRevB.53.4488](https://doi.org/10.1103/PhysRevB.53.4488).
- [94] P. Carrier, R. Wentzcovitch, and J. Tsuchiya. First-principles prediction of crystal structures at high temperatures using the quasiharmonic approximation. *Phys. Rev. B*, 76(6):064116, 2007. doi: [10.1103/PhysRevB.76.064116](https://doi.org/10.1103/PhysRevB.76.064116).
- [95] M. Friedrich, A. Riefer, S. Sanna, W. Schmidt, and A. Schindlmayr. Phonon dispersion and zero-point renormalization of LiNbO_3 from density-functional perturbation theory. *J. Phys. Condens. Matter*, 27(38):385402, 2015. doi: [10.1088/0953-8984/27/38/385402](https://doi.org/10.1088/0953-8984/27/38/385402).
- [96] W. Saidi, S. Poncé, and B. Monserrat. Temperature dependence of the energy levels of methylammonium lead iodide perovskite from first principles. *J. Phys. Chem Lett.*, 7(24):5247, 2016. doi: [10.1021/acs.jpcllett.6b02560](https://doi.org/10.1021/acs.jpcllett.6b02560).
- [97] A. Molina-Sanchez, M. Palummo, A. Marini, and L. Wirtz. Temperature-dependent excitonic effects in the optical properties of single-layer MoS_2 . *Phys. Rev. B*, 93(16):155435, 2016. doi: [10.1103/PhysRevB.93.155435](https://doi.org/10.1103/PhysRevB.93.155435).
- [98] G. Antonius and S. G. Louie. Temperature-induced topological phase transitions: Promoted versus suppressed nontrivial topology. *Phys. Rev. Lett.*, 117(24):246401, 2016. doi: [10.1103/PhysRevLett.117.246401](https://doi.org/10.1103/PhysRevLett.117.246401).
- [99] C. E. P. Villegas, A. R. Rocha, and A. Marini. Electron-phonon scattering effects on electronic and optical properties of orthorhombic GeS . *Phys. Rev. B*, 94(13):134306, 2016. doi: [10.1103/PhysRevB.94.134306](https://doi.org/10.1103/PhysRevB.94.134306).
- [100] V. Ziaei and T. Bredow. Dynamical electron-phonon coupling, GW self-consistency, and vertex effect on the electronic band gap of ice and liquid water. *Phys. Rev. B*, 95(23):235105, 2017. doi: [10.1103/PhysRevB.95.235105](https://doi.org/10.1103/PhysRevB.95.235105).
- [101] R. Tutchton, C. Marchbanks, and Z. Wu. Structural impact on the eigenenergy renormalization for carbon and silicon allotropes and boron nitride polymorphs. *Phys. Rev. B*, 97(20):205104, 2018. doi: [10.1103/PhysRevB.97.205104](https://doi.org/10.1103/PhysRevB.97.205104).
- [102] J. Cao, J. Querales-Flores, S. Fahy, and I. Savić. Thermally induced band gap increase and high thermoelectric figure of merit of n-type PbTe . *Materials Today Physics*, 12:100172, 2020. doi: [10.1016/j.mtphys.2019.100172](https://doi.org/10.1016/j.mtphys.2019.100172).
- [103] J. Querales-Flores, J. Cao, S. Fahy, and I. Savić. Temperature effects on the electronic band structure of PbTe from first principles. *Phys. Rev. Mat.*, 3(5):055405, 2019. doi: [10.1103/PhysRevMaterials.3.055405](https://doi.org/10.1103/PhysRevMaterials.3.055405).

- [104] J. D. Querales-Flores, P. Aguado-Puente, D. Dangić, J. Cao, P. Chudzinski, T. N. Todorov, M. Grüning, S. Fahy, and I. Savić. Towards temperature-induced topological phase transition in SnTe: A first principles study. *Phys. Rev. B*, 101(23):235206, 2020. doi: [10.1103/PhysRevB.101.235206](https://doi.org/10.1103/PhysRevB.101.235206).
- [105] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12):121102, 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).
- [106] P. Han and G. Bester. Large nuclear zero-point motion effect in semiconductor nanoclusters. *Phys. Rev. B*, 88(16):165311, 2013. doi: [10.1103/PhysRevB.88.165311](https://doi.org/10.1103/PhysRevB.88.165311).
- [107] C. E. Patrick, K. W. Jacobsen, and K. S. Thygesen. Anharmonic stabilization and band gap renormalization in the perovskite CsSnI₃. *Phys. Rev. B*, 92(20):201205, 2015. doi: [10.1103/PhysRevB.92.201205](https://doi.org/10.1103/PhysRevB.92.201205).
- [108] M. Zacharias, C. E. Patrick, and F. Giustino. Stochastic approach to phonon-assisted optical absorption. *Phys. Rev. Lett.*, 115(17):177401, 2015. doi: [10.1103/PhysRevLett.115.177401](https://doi.org/10.1103/PhysRevLett.115.177401).
- [109] B. Monserrat. Electron-phonon coupling from finite differences. *J. Phys. Condens. Matter*, 30(8):083001, 2018. doi: [10.1088/1361-648X/aaa737](https://doi.org/10.1088/1361-648X/aaa737).
- [110] I. Bravic and B. Monserrat. Finite temperature optoelectronic properties of BAs from first principles. *Phys. Rev. Materials*, 3(6):065402, 2019. doi: [10.1103/PhysRevMaterials.3.065402](https://doi.org/10.1103/PhysRevMaterials.3.065402).
- [111] M. Engel, M. Marsman, C. Franchini, , and G. Kresse. Electron-phonon interactions using the projector augmented-wave method and Wannier functions. *Phys. Rev. B*, 101(18):184302, 2020. doi: [10.1103/PhysRevB.101.184302](https://doi.org/10.1103/PhysRevB.101.184302).
- [112] S. Botti and M. A. L. Marques. Strong renormalization of the electronic band gap due to lattice polarization in the GW formalism. *Phys. Rev. Lett.*, 110(22):226404, 2013. doi: [10.1103/PhysRevLett.110.226404](https://doi.org/10.1103/PhysRevLett.110.226404).
- [113] W. R. L. Lambrecht, C. Bhandari, and M. van Schilfgaarde. Lattice polarization effects on the screened coulomb interaction W of the GW approximation. *Phys. Rev. Mat.*, 1(4):043802, 2017. doi: [10.1103/PhysRevMaterials.1.043802](https://doi.org/10.1103/PhysRevMaterials.1.043802).
- [114] C. Bhandari, M. van Schilfgaarde, T. Kotani, and W. R. L. Lambrecht. All-electron quasi-particle self-consistent GW band structures for SrTiO₃ including lattice polarization corrections in different phases. *Phys. Rev. Mat.*, 2(1):013807, 2018. doi: [10.1103/PhysRevMaterials.2.013807](https://doi.org/10.1103/PhysRevMaterials.2.013807).
- [115] H. Fröhlich, H. Pelzer, and S. Zienau. Properties of slow electrons in polar materials. *Phil. Mag.*, 41(314):221–242, 1950. doi: [10.1080/14786445008521794](https://doi.org/10.1080/14786445008521794).
- [116] H. Fröhlich. Interaction of electrons with lattice vibrations. *Proc. R. Soc. Lond. A*, 215(1122):291–298, 1952. doi: [10.1098/rspa.1952.0212](https://doi.org/10.1098/rspa.1952.0212).
- [117] G. D. Mahan. *Many-Particle Physics*. Kluwer, New York, 2000.

- [118] N. Mecholsky, L. Resca, I. Pegg, and M. Fornari. Theory of band warping and its effects on thermoelectronic transport properties. *Phys. Rev. B*, 89(15):155131, 2014. doi: [10.1103/PhysRevB.89.155131](https://doi.org/10.1103/PhysRevB.89.155131).
- [119] M. Veithen, X. Gonze, and P. Ghosez. Nonlinear optical susceptibilities, raman efficiencies, and electro-optic tensors from first-principles density functional perturbation theory. *Phys. Rev. B*, 71(12):125107, 2005. doi: [10.1103/PhysRevB.71.125107](https://doi.org/10.1103/PhysRevB.71.125107).
- [120] J. M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Phys. Rev.*, 97(4):869–883, 1955. doi: [10.1103/PhysRev.97.869](https://doi.org/10.1103/PhysRev.97.869).
- [121] B. Monserrat. Vibrational averages along thermal lines. *Phys. Rev. B*, 93(1):014302, 2016. doi: [10.1103/PhysRevB.93.014302](https://doi.org/10.1103/PhysRevB.93.014302).
- [122] X. Yang, Y. Wang, H. Yan, and Y. Chen. Effects of epitaxial strains on spontaneous polarizations and band gaps of alkaline-earth-metal oxides MO (M = Mg, Ca, Sr, Ba). *Computational Materials Science*, 121:61–66, 2016. doi: [10.1016/j.commatsci.2016.04.021](https://doi.org/10.1016/j.commatsci.2016.04.021).
- [123] H. Singh, M. Singh, S. Kumar, and M. K. Kashyap. Full potential calculation of electronic properties of rutile RO₂ (R=Si, Ge, Sn and Pb) compounds via modified Becke Johnson potential. *Physica B*, 406(20):3825–3830, 2011. doi: [10.1016/j.physb.2011.07.004](https://doi.org/10.1016/j.physb.2011.07.004).

Chapitre 5

Dilatation du point zéro du réseau cristallin

5.0.1. Origine anharmonique de la dilatation thermique

Au chapitre 3, nous avons discuté de la contribution du couplage électron-phonon à la renormalisation de la bande interdite dans le cadre de l'approximation harmonique. Ce faisant, nous avons pu exprimer l'hamiltonien des noyaux comme un ensemble d'oscillateurs harmoniques indépendants, dont les quantas de vibrations forment les phonons qui interagissent avec les électrons à travers les mécanismes décrits par la self-énergie. Nous avons supposé que l'énergie totale du système varie quadratiquement avec les déplacements atomiques, ce qui nous a permis de construire la matrice dynamique et d'obtenir les fréquences des phonons. Or, puisque la courbure d'une fonction quadratique est constante, les fréquences des phonons sont indépendantes de l'amplitude des déplacements atomiques. De plus, rappelons-nous qu'à température finie l'occupation des différents niveaux vibrationnels est décrite par la distribution de Bose-Einstein. Pour un cristal parfaitement harmonique, le déplacement moyen des noyaux autour de leur position d'équilibre est donc toujours nul par symétrie, indépendamment de la température. Par conséquent, le volume ne peut pas changer avec la température.

La description de la dilatation thermique des matériaux nécessite donc d'introduire de l'anharmonicité dans notre modèle. Ce faisant, l'énergie libre devient asymétrique par rapport à la position d'équilibre statique, comme on peut observer à la figure 5.1a. La courbure de l'énergie va maintenant dépendre de l'amplitude des déplacements ou, de manière équivalente, de la séparation moyenne entre les atomes, que l'on peut relier au paramètre de maille et au volume pour les matériaux cubiques : la courbure est plus prononcée lorsque les distances interatomiques sont faibles et diminue quand elles sont plus grandes que la position d'équilibre. On peut interpréter cette asymétrie comme une conséquence de la répulsion coulombienne entre les noyaux, qui croît rapidement lorsque les ions se rapprochent les uns des autres. Puisque l'énergie n'est plus symétrique, le déplacement moyen des noyaux

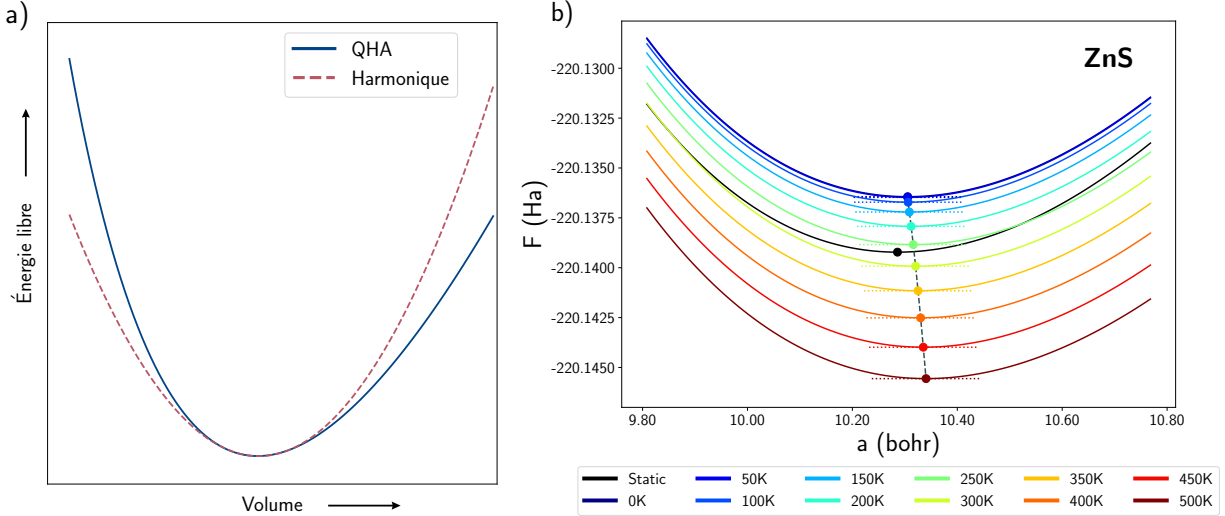


Figure 5.1 – Origine anharmonique de la dilatation thermique. a) Dans l’approximation quasi harmonique, l’énergie libre n’est plus symétrique par rapport au déplacement atomique moyen, ce qui permet au volume du cristal de changer avec la température. b) Dilatation thermique du ZnS. Le paramètre de maille à chaque température minimise l’énergie libre du système d’électrons et de phonons (cercles colorés reliés par les tirets gris). La dilatation du point zéro du réseau (ZPLE) correspond à la différence entre le paramètre de maille statique (noir) et celui à $T = 0$ K (indigo). Les lignes pointillées mettent en valeur que l’approche de Grüneisen cherche à identifier le volume pour lequel la pente correspond à la pression de référence (ici, la pression ambiante).

par rapport à leur position d’équilibre n’est plus nul et varie d’un niveau vibrationnel à l’autre. L’occupation thermique des niveaux vibrationnels plus élevés augmente ainsi progressivement la distance moyenne entre les noyaux. Cette dilatation du cristal se traduit par une modification du minimum global de l’énergie libre du système lorsqu’on y inclut la contribution des phonons [1, 2].

Pour caractériser la dilatation thermique d’un matériau, nous travaillerons dans l’approximation *quasi* harmonique (QHA). Bien que la QHA ne constitue pas un traitement explicite de l’anharmonicité telle que nous l’avons introduite à la section 3.2.3, Allen a démontré que ce formalisme permet d’évaluer correctement le terme de premier ordre de la dépendance en volume de l’énergie libre et, par conséquent, de la dilatation thermique [3], sans avoir besoin de recourir au formalisme à N -corps. Il a également montré que l’impact des interactions phonon-phonon sur la dilatation thermique est une correction d’ordre supérieur à celle fournie par la QHA.

Par simplicité, nous considérerons uniquement un système de symétrie cubique, pour lequel le volume est le seul paramètre libre. On négligera aussi la dépendance sur les coordonnées internes de noyaux dans la cellule unitaire. La généralisation des équations présentées dans cette section est développée à la section 5.2 de l’article.

Nous supposons que la contribution des phonons à l'énergie libre du cristal peut être décrite à l'aide de son expression analytique obtenue dans l'approximation harmonique, à l'exception que les fréquences des phonons dépendent maintenant explicitement du volume,

$$\begin{aligned} F^{\text{phon}}(V,T) &= -k_B T \ln Z^{\text{phon}} \\ &= \sum_{\mathbf{q}\nu} \frac{\omega_{\mathbf{q}\nu}(V)}{2} + k_B T \sum_{\mathbf{q}\nu} \ln \left(1 - e^{-\frac{\omega_{\mathbf{q}\nu}(V)}{k_B T}} \right), \end{aligned} \quad (5.0.1)$$

où k_B est la constante de Boltzmann et Z^{phon} est la fonction de partition de notre ensemble de phonons harmoniques. Le premier terme du membre de droite de la seconde ligne correspond à l'énergie du point zéro et nous indique que, tout comme les électrons interagissent avec le mouvement du point zéro des noyaux, l'énergie du point zéro des phonons modifie les paramètres de maille du cristal par rapport à leur valeur obtenue dans l'approximation des noyaux statiques. Cette correction porte le nom de dilatation du point zéro du réseau (*zero-point lattice expansion*, ZPLE).

On pourra ensuite évaluer l'énergie libre totale du système d'électrons et de noyaux,

$$F^{\text{tot}}(V,T) = E_{\text{stat}}^{\text{el}}(V) + F^{\text{phon}}(V,T), \quad (5.0.2)$$

où $E_{\text{stat}}^{\text{el}}$ est l'énergie totale du cristal dans l'approximation Born-Oppenheimer et où nous avons négligé la contribution entropique des électrons puisque nous traitons uniquement des semi-conducteurs et des isolants. En effet, dans la mesure où l'énergie de la bande interdite est beaucoup plus grande que $k_B T$, les électrons sont confinés dans les bandes de valence. N'ayant accès qu'à une seule configuration d'occupations, leur entropie est nulle. Cette approximation perd progressivement de sa validité lorsque les probabilités d'occupation des états excités deviennent statistiquement significatives. La dépendance en température du volume s'obtient finalement en minimisant l'énergie libre $F^{\text{tot}}(V,T = T')$ pour chaque température T' . Cette procédure est illustrée à la figure 5.1b pour le ZnS cubique. Notons que la courbe d'énergie libre à $T = 0$ K (indigo) a une énergie plus élevée que son équivalent statique (noir), puisque la contribution de l'énergie du point zéro est positive.

Une autre méthode couramment utilisée dans la littérature relie la dilatation thermique à la variation des fréquences phononiques par rapport au volume, qui découle du changement de courbure de l'énergie libre. Développée par Grüneisen au début du XX^e siècle [4], elle repose sur le développement de Taylor de l'énergie libre totale (éq. (5.0.2)) autour du volume à l'équilibre statique, V^0 . On obtient

$$\begin{aligned} F^{\text{tot}}(V,T) &\approx E_{\text{stat}}^{\text{el}}(V^0) + \frac{1}{2} \left(\frac{\partial^2 E_{\text{stat}}^{\text{el}}(V)}{\partial V^2} \right) \Big|_{V^0} (V - V^0)^2 \\ &\quad + F^{\text{phon}}(V^0,T) + \frac{\partial F^{\text{phon}}(V,T)}{\partial V} \Big|_{V^0} (V - V^0), \end{aligned} \quad (5.0.3)$$

où l'énergie électronique statique a été développée au 2^e ordre puisqu'elle est par définition minimale à $V = V^0$. À chaque température, ce développement devra vérifier la relation thermodynamique

$$P = -\frac{\partial F^{\text{tot}}}{\partial V}. \quad (5.0.4)$$

On peut réécrire la dérivée seconde de l'énergie électronique statique de l'équation (5.0.3) comme

$$\begin{aligned} \left. \frac{\partial^2 E_{\text{stat}}^{\text{el}}(V)}{\partial V^2} \right|_{V^0} &= \left(\frac{\partial}{\partial V} \frac{\partial E_{\text{stat}}^{\text{el}}}{\partial V} \right) \Big|_{V^0} \\ &= \left. \frac{\partial(-P)}{\partial V} \right|_{V^0} \\ &= \frac{B_0}{V^0}, \end{aligned} \quad (5.0.5)$$

où B_0 est le module l'élasticité volumique (*bulk modulus*).

La dérivée première de l'énergie libre des phonons s'écrit quant à elle

$$\begin{aligned} \left. \frac{\partial F^{\text{phon}}(V,T)}{\partial V} \right|_{V^0} &= \frac{\partial}{\partial V} \left[\sum_{\mathbf{q}\nu} \frac{\omega_{\mathbf{q}\nu}(V)}{2} + k_B T \sum_{\mathbf{q}\nu} \left(\ln(1 - e^{-\frac{\omega_{\mathbf{q}\nu}(V)}{k_B T}}) \right) \right] \Big|_{V^0} \\ &= \sum_{\mathbf{q}\nu} \frac{1}{2} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \right|_{V^0} + \sum_{\mathbf{q}\nu} \frac{e^{-\frac{\omega_{\mathbf{q}\nu}}{k_B T}}}{1 - e^{-\frac{\omega_{\mathbf{q}\nu}}{k_B T}}} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \right|_{V^0} \\ &= -\frac{1}{V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu} \omega_{\mathbf{q}\nu}(V^0) \left(n_{\mathbf{q}\nu}(V^0, T) + \frac{1}{2} \right), \end{aligned} \quad (5.0.6)$$

où $n_{\mathbf{q}\nu}(V^0, T)$ est la fonction d'occupation Bose-Einstein des phonons évaluée au volume à l'équilibre statique et où nous avons introduit les paramètres de Grüneisen volumiques,

$$\gamma_{\mathbf{q}\nu}^V = -\frac{V^0}{\omega_{\mathbf{q}\nu}(V^0)} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \right|_{V^0}. \quad (5.0.7)$$

Notons que nous avons multiplié le membre de droite de la seconde ligne de l'équation (5.0.6) par $-\omega_{\mathbf{q}\nu}(V^0)/V^0$ pour obtenir la ligne suivante.

En substituant les équations (5.0.5) et (5.0.6) dans les équations (5.0.3) et (5.0.4), on obtient

$$P \approx -\frac{B_0}{V^0} (V - V^0) + \frac{1}{V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu} \omega_{\mathbf{q}\nu}(V^0) \left(n_{\mathbf{q}\nu}(V^0, T) + \frac{1}{2} \right). \quad (5.0.8)$$

On pose ensuite $P = 0$ puisqu'on s'intéresse au résultat à pression ambiante, qui est négligeable pour les matériaux considérés. On peut dès lors exprimer le changement de volume relatif à une température donnée en termes des paramètres de Grüneisen :

$$\frac{V(T) - V^0}{V^0} = \frac{1}{B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V \omega_{\mathbf{q}\nu}(V^0) \left(n_{\mathbf{q}\nu}(V^0, T) + \frac{1}{2} \right). \quad (5.0.9)$$

Cette relation nous permet finalement d'écrire le coefficient de dilatation thermique volumique comme

$$\begin{aligned}\beta(T) &= \frac{d}{dT} \left(\frac{V(T) - V^0}{V^0} \right) \\ &= \frac{1}{B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu} \omega_{\mathbf{q}\nu}(V^0) \frac{dn_{\mathbf{q}\nu}(T)}{dT}.\end{aligned}\tag{5.0.10}$$

Le changement de volume par rapport à l'équilibre statique à une température T' s'obtient alors simplement en intégrant $\beta(T)$ de 0 à T' ¹.

On peut apprécier l'information physique contenue dans cette approche en observant la figure 5.1b. Sachant que la courbure de l'énergie anharmonique (et, par conséquent, les $\omega_{\mathbf{q}\nu}$) diminue lorsque le volume augmente, le signe des paramètres de Grüneisen indique si les différents modes phononiques favorisent la contraction ou la dilatation du cristal, alors que leur norme capture l'importance du changement de courbure liée à un mode donné. Puisque que la contribution de chaque mode au changement de volume du cristal à une température donnée est pondérée par son occupation thermique, les paramètres de Grüneisen permettent d'identifier les modes phononiques qui gouvernent la dilatation thermique dans les différents régimes de température.

Par ailleurs, l'équation (5.0.4) montre que l'approche de Grüneisen cherche à identifier le volume pour lequel la pente de l'énergie libre correspond à la pression ambiante, comme en témoignent les lignes pointillées horizontales de la figure 5.1b. On peut interpréter donc la contribution des phonons à l'énergie libre comme une « pression thermique » [5] qui est contrebalancée par la réponse élastique du matériau. Pour les systèmes non cubiques, nous verrons à la section 5.2.2.2 que la généralisation du formalisme de Grüneisen met en évidence la nature anisotrope de la dilatation thermique en reliant la variation des différents paramètres de maille aux constantes élastiques [5].

Tout comme l'interaction électron-phonon, la dilatation thermique va affecter l'énergie de la bande interdite par rapport à l'approximation des noyaux statiques. Nous supposons que l'on peut décorréler les deux contributions, c'est-à-dire que l'on peut évaluer séparément l'effet de l'électron-phonon au volume d'équilibre statique et celui de la dilatation thermique (TE),

$$\Delta E_g \simeq \Delta E_g^{\text{EPI}} + \Delta E_g^{\text{TE}}.\tag{5.0.11}$$

Pour de nombreux matériaux, l'interaction électron-phonon domine la renormalisation de la bande interdite pour une très grande plage de températures. Malgré tout, la contribution de la dilatation thermique est nécessaire pour reproduire les données expérimentales, qui capturent naturellement ces deux contributions. À titre d'exemple, la figure 5.2 compare la dépendance en température de la bande interdite directe du AlAs obtenue numériquement

1. Ou, de manière équivalente, avec l'équation (5.0.9).

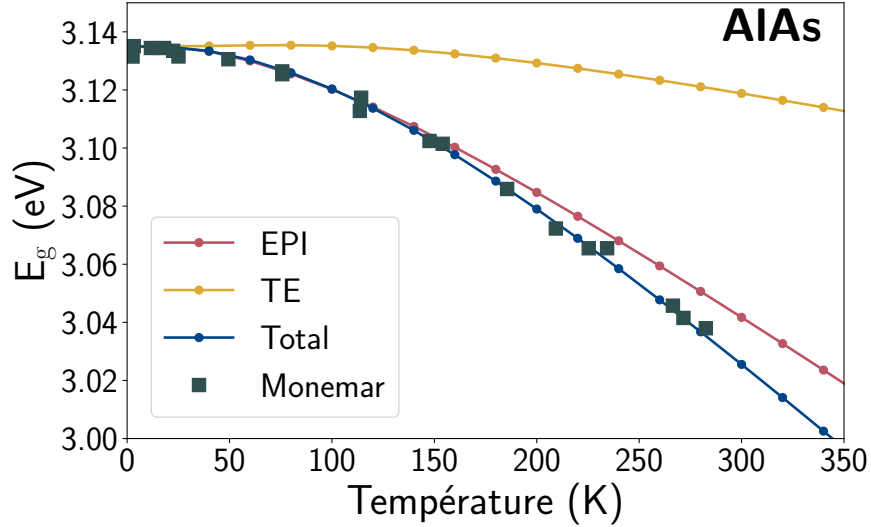


Figure 5.2 – Renormalisation de l'énergie de la bande interdite directe du AlAs. La renormalisation totale (bleu) somme des contributions de l'interaction électron-phonon (rouge) et de la dilatation thermique (jaune). Pour faciliter la comparaison de la dépendance en température, l'énergie de la bande interdite à $T = 0$ K a été ajustée aux données expérimentales de Monemar [6] (carrés gris). Dès $T \approx 200$ K, la contribution de la dilatation thermique est nécessaire pour reproduire correctement les données expérimentales. Les incertitudes expérimentales sont au plus de l'ordre de la grosseur des marqueurs.

aux valeurs mesurées par photoluminescence [6]. L'énergie de la bande interdite à $T = 0$ K a été ajustée aux données expérimentales. Bien qu'elle soit responsable de la plus grande partie de la renormalisation, l'EPI (rouge) sous-estime la variation de l'énergie de la bande interdite par rapport aux mesures (carrés gris) dès $T \simeq 200$ K. Lorsqu'on inclut également la dilatation thermique (jaune), la courbe totale (bleu) reproduit beaucoup mieux les données. Notons que cette température correspond environ à la moitié de la température de Debye du AlAs, $\theta_D \approx 414$ K [7], indiquant que cette contribution affecte la prédictibilité des calculs numériques bien en deçà du régime classique.

Pour conclure cette section, nous tenons à souligner que l'approximation quasi harmonique n'est pas équivalente à ajouter explicitement des termes cubiques ou quartiques avec les déplacements atomiques dans le développement en série de l'énergie totale que nous avons effectué à l'équation (3.2.9). De telles contributions impliqueraient d'ajouter à la self-énergie de l'interaction électron-phonon des diagrammes incluant des interactions multiphonons (voir, par exemple, la figure 3 de la référence [8]). De tels diagrammes correspondent aux interactions d'ordre supérieur qui sont capturées par la méthode des phonons gelés, auxquels nous avons référé dans le préambule du chapitre 4. En comparaison, dans l'approximation quasi harmonique, nous conférons uniquement une dépendance en volume aux fréquences phononiques pour être en mesure d'évaluer la dépendance en température de la structure cristalline. La contribution EPI, elle, demeure calculée dans l'approximation harmonique.

5.0.2. Mise en contexte

Ce projet met donc l'accent sur la contribution de la dilatation du point zéro du réseau cristallin en importance à la renormalisation des énergies électroniques à $T = 0$ K. Puisqu'elle est naturellement présente dans les mesures expérimentales, au même titre que l'interaction électron-phonon, une comparaison théoriquement justifiée entre des prédictions numériques et des données expérimentales requiert de considérer ces deux contributions sur un même pied d'égalité. Néanmoins, dans le régime des basses températures, la dilatation thermique est encore souvent perçue comme une petite correction à l'interaction électron-phonon, qui domine la dépendance en température. Ainsi, la contribution du ZPLE à la renormalisation de l'énergie de la bande interdite n'a pas reçu autant d'attention de la part de la communauté scientifique que l'EPI et mérite qu'on s'y attarde plus en profondeur.

L'angle d'approche de ce projet m'est venu en tête durant les premières années de mon doctorat, dans le cadre de mes travaux portant sur la transition de phase topologique dans le BiTeI, qui seront présentés au chapitre 8. Ce matériau possède une symétrie axiale et est formé de plans monoatomiques faiblement liés entre eux par des interactions Van der Waals, ce qui confère une importante anisotropie à ses propriétés élastiques. Des deux méthodes discutées plus haut, le formalisme de Grüneisen m'a paru le plus efficace pour évaluer la variation des deux paramètres de maille indépendants tout en prenant en compte correctement l'anisotropie élastique du matériau. Néanmoins, dans la plupart des documents de référence, cette méthode néglige la contribution de l'énergie du point zéro des phonons, vraisemblablement car son objectif d'origine était de modéliser le coefficient de dilatation thermique linéaire, $\alpha(T) = d(\Delta a(T)/a^0)/dT$, pour le comparer aux données expérimentales². À ce moment, nous n'avions trouvé aucune référence dans la littérature qui confirme l'équivalence de ces deux méthodes pour calculer la dilatation du point zéro du réseau. L'idée est donc demeurée en dormance.

Suite aux travaux présentés au chapitre 4, nous avons en main un ensemble de 22 matériaux pour lequel la contribution du ZPLE à la renormalisation du point zéro de la bande interdite avait été calculée par minimisation de l'énergie libre, incluant trois matériaux de structure wurtzite ayant deux paramètres de maille indépendants. Nous avons donc entrepris de vérifier numériquement l'équivalence entre la méthode de Grüneisen et la minimisation de l'énergie libre, tant pour le ZPLE lui-même que pour sa contribution au ZPR. Nous avons également quantifié l'importance relative de la contribution du ZPLE par rapport à celle de l'EPI et montré qu'elle pouvait dépasser 50% dans les matériaux lourds et atteindre plus de 20% dans certains matériaux ioniques comprenant des noyaux légers, mettant ainsi en garde contre le fait de négliger cette contribution dans ce type de matériau.

2. L'énergie du point zéro étant une constante, elle disparaît lorsqu'on évalue la dérivée par rapport à la température.

5.0.3. Contributions des auteurs

J'ai proposé l'angle d'approche de ce projet et effectué l'ensemble des analyses ainsi que la grande majorité des calculs. J'ai également assuré la rédaction de la première version du manuscrit. Émile Godbout a effectué les calculs du ZPLE pour 6 matériaux. Xavier Gonze a fourni un soutien théorique pour la rédaction de la section méthodologique. Michel Côté a supervisé le projet. Tous les auteurs ont contribué à la révision du manuscrit.

Deuxième article.

Zero-point lattice expansion and band gap renormalization : Grüneisen approach versus free energy minimization

par

Véronique Brousseau-Couture¹, Émile Godbout², Xavier Gonze³ et Michel Côté⁴

- (¹) Département de physique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7
- (²) Département de physique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7
- (³) Institute of Condensed Matter and Nanosciences, UCLouvain, B-1348 Louvain-la-Neuve, Belgium
- (⁴) Département de physique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7

Cet article a été publié dans Physical Review B.

RÉSUMÉ. La dilatation du point zéro du réseau (ZPLE) est une petite variation des paramètres du mailles du cristal induite par la présence des phonons dans un matériau, en comparaison au résultat obtenu dans l’approximation du réseau statique. Elle contribue de manière significative à la renormalisation du point zéro (ZPR) de l’énergie de la bande interdite, mais ses conséquences n’ont pas été étudiées de manière aussi approfondie que celles découlant de l’interaction électron-phonon. Dans l’approche *ab initio*, on évalue habituellement la ZPLE en minimisant l’énergie libre de Helmholtz à $T = 0$ K. Dans cet article, nous montrons que le formalisme basé sur les paramètres Grüneisen, qui néglige communément les effets du point zéro, peut être utilisé efficacement pour calculer la ZPLE de matériaux isotropes et anisotropes à un coût numérique beaucoup plus faible. Nous testons systématiquement ce formalisme sur 22 matériaux de structure cubique et wurtzite et obtenons un excellent accord avec les résultats provenant de la minimisation de l’énergie libre, tant pour le ZPLE que pour la renormalisation de la bande interdite qui en résulte. Nous utilisons nos résultats pour valider une expression empirique estimant le ZPR induit par la ZPLE et dévoiler sa sensibilité à la plage de température impliquée dans l’estimation de la ZPLE à partir de données expérimentales. Nos résultats révèlent enfin que la contribution de la ZPLE à la renormalisation de l’énergie de la bande interdite peut atteindre 20% à plus de 80% de la contribution de l’interaction électron-phonon pour les matériaux plus lourds ou plus ioniques, y compris des matériaux comportant des noyaux légers. Il est donc essentiel de considérer les deux contributions sur un pied d’égalité si l’on souhaite comparer les résultats théoriques du ZPR aux données expérimentales.

Mots clés : Dilatation thermique, dilatation du point zéro du réseau cristallin, renormalisation du point zéro de l’énergie de la bande interdite, théorie de la fonctionnelle de la densité perturbative, interaction électron-phonon.

ABSTRACT. The zero-point lattice expansion (ZPLE) is a small variation of the lattice parameters induced by the presence of phonons in a material compared to the static lattice picture. It contributes significantly to the zero-point renormalization (ZPR) of the band gap energy, but its consequences have not been investigated as thoroughly as those stemming from electron-phonon interactions. In the usual first-principles approach, one evaluates the ZPLE by minimizing the $T = 0$ K Helmholtz free energy. In this work, we show that the formalism based on the Grüneisen parameters, which commonly neglects zero-point effects, can be efficiently used to compute ZPLE for both isotropic and anisotropic materials at much lower computational cost. We systematically test this formalism on 22 cubic and wurtzite materials and obtain excellent agreement with free energy minimization results for both the ZPLE and the resulting band gap ZPR. We use our results to validate an empirical expression estimating the ZPLE-induced ZPR and unveil its sensitivity to the temperature range involved in estimating the ZPLE from experimental data. Our findings finally reveal that the ZPLE contribution to the band gap ZPR can reach 20% to more than 80% of the electron-phonon interaction contribution for heavier or more ionic materials, including materials containing light atoms. Considering both contributions on an equal footing is thus essential should one attempt to compare theoretical ZPR results with experimental data.

Keywords: Thermal expansion, zero-point lattice expansion, zero-point renormalization of the band gap, density-functional perturbation theory, electron-phonon interaction

5.1. Introduction

The electronic band gap and the lattice parameters are undoubtedly two of the most fundamental properties of materials. The former depends on the latter, as a modification of the lattice parameters modifies the ionic potential felt by the electrons, hence affecting the electronic eigenvalues. This is at the heart of band gap engineering through lattice mismatch [9]. While the standard Kohn-Sham DFT is not theoretically expected to deliver accurate excited-state properties such as band gap energies [10], it should in principle correctly predict the lattice parameters, as they are a ground state property.

The standard DFT framework assumes, however, a perfectly static lattice. This picture would be physically correct should the ions behave like classical particles as $T \rightarrow 0$ K. This is not the case in the quantum mechanical picture: quantum fluctuations of the nuclei positions remain present even at absolute zero temperature. This phenomenon, known as the zero-point (ZP) motion of the ions, is responsible for the zero-point renormalization (ZPR) of the band gap. The leading mechanism driving the ZPR is electron-phonon interaction (EPI), namely, the electronic response to the distorted ionic potential induced by the lattice vibrations. In addition, the phonon zero-point energy contributes to the total free energy of the crystal, which, in turn, affects the equilibrium lattice parameter at $T = 0$ K compared to the static value. This lattice parameter shift, known as the zero-point lattice expansion (ZPLE), also affects the band gap energy, thus contributing to the ZPR.

EPI has been widely investigated by the first-principles community throughout the last two decades [11]. In addition to the ZPR and the temperature dependence of the electronic structure, it is at the core of transport properties [12, 13], polaronic systems [14, 15] and superconductivity [16, 17]. ZPLE has been the subject of less attention by this community. It has been mostly considered in the broader context of thermal expansion and temperature-dependent thermodynamical quantities rather than as a zero-point correction itself [18–23]. It has also been addressed when investigating the accuracy of lattice parameters predictions by different exchange-correlation functionals [24–27], as this correction is essential to make a fair comparison to experimental values. In all cases, most approaches rely on minimizing the Helmholtz free energy (FE) with respect to volume. Some works have also proposed a more generalized framework which allows for treatment of anisotropic thermal expansion [28–31].

Another method, developed by Grüneisen in the early 1900’s [4], expresses the temperature-dependent thermal expansion coefficient in terms of the volume variation of the phonon frequencies. The main advantage of this method is its smaller numerical cost, as it is based on finite derivatives rather than the computation of the full temperature and volume-dependent free energy. However, in its typical form, the Grüneisen formalism neglects ZPLE, most likely for historical reasons, as this formalism was originally intended to model the thermal expansion coefficient, which does not depend on the zero-point energy.

Nevertheless, some authors recognized that, in principle, the ZPLE could be properly incorporated within this formalism [3, 18]. To the best of our knowledge, no systematic study of the predictive capabilities of the Grüneisen formalism compared to the standard FE minimization technique regarding ZPLE has been attempted.

Despite being a decades-old problem, developing a first-principles-based methodology that can accurately predict thermal-dependent material properties in a wide range of temperatures is still an active subject of investigation. Recent proposals include a fully anharmonic one-shot formalism that can account simultaneously for both EPI and thermal expansion [32], formalisms for temperature-dependent elastic constants [23, 33, 34], self-consistent volume change calculations [18], incorporation of higher-order phonon-phonon interactions within the Grüneisen formalism [35] and accounting for fully anharmonic temperature-dependent phonon frequencies [31, 36]. All these methodologies could, in principle, be used to assess the ZPLE. However, we can reasonably expect that the additional true anharmonicity captured by these methodologies will not significantly affect the $T = 0$ K ZPLE for weakly or moderately anharmonic materials.

On the other hand, the contribution of thermal expansion to the temperature-dependent band gap energy has been investigated in semiconductors [37–41], layered materials [42, 43], nanostructures [44, 45] and topological materials [46–48], mostly on a case-by-case basis. Outside this broader context, the explicit study of the importance of the ZPLE contribution to the $T = 0$ K band gap renormalization compared to the one stemming from EPI is more scarcely addressed throughout the literature, a noticeable exception being the work of Garro *et al.* [49]. Furthermore, it is typically thought of as negligible compared to the EPI contribution, without tangible statistical evidence. General trends allowing to qualitatively estimate the importance of ZPLE on the band gap ZPR in terms of simple bulk material properties are also still lacking, even for benchmark materials like diamond and zincblende structures.

In this work, we take a new look at the ZPLE and its contribution to the band gap ZPR from the first-principles perspective. We investigate 22 materials: 19 with cubic symmetry and 3 with wurtzite structure. We show that the well-established Grüneisen formalism can account for the ZPLE and its effect on the band gap as accurately as the typical FE minimization approach, for both isotropic and anisotropic materials, at a much lower computational cost. We also validate an empirical expression that estimates the ZPLE band gap correction from simple parameters for sufficiently symmetric materials and discuss its possible shortcomings when used in conjunction with experimental data. We finally demonstrate that, for most materials in our set, the ZPLE modifies the predicted band gap ZPR from EPI by more than 20%, even for materials containing light atoms, thus questioning the perception that ZPLE effects can be neglected in such materials.

The remainder of this paper is organized as follows. In Sec. 5.2, we briefly review the theoretical formalism used to compute the ZPLE and the EPI contributions to the band gap ZPR and demonstrate how to evaluate the ZPLE within the Grüneisen formalism. The relevant computational details are summarized in Sec. 5.3. Sections 5.4.1 and 5.4.2 respectively present our numerical results for the ZPLE and compare the free energy minimization and Grüneisen approaches, while Sec 5.4.3 focuses on the ZPLE contribution to the band gap renormalization. We finally discuss some implications of our results in Sec. 5.4.4 and summarize our findings.

5.2. Methodology

5.2.1. Zero-point renormalization of the band gap energy

Electronic eigenstates depend both on the unit cell structure and temperature. For a general material, the unit cell structure is fully defined by the lattice parameters, $\{a_i\}$, the angles between them, $\{b_i\}$, and the internal atomic coordinates, $\{z_i\}$. For lattice parameters and angles, the index $i = 1,2,3$ refers to the three independent lattice vectors, while for the internal atomic coordinates, i runs over the different atoms in the unit cell. We introduce the condensed notation

$$\{V\} \triangleq \{\{a_i\}, \{b_i\}, \{z_i\}\}, \quad (5.2.1)$$

to denote the full set of structure parameters.

We wish to evaluate the temperature dependence of an electronic eigenvalue $\varepsilon_{\mathbf{k}n}$, where \mathbf{k} is the electron wavevector and n is the branch index. From thermodynamic theory, we can express the variation of $\varepsilon_{\mathbf{k}n}$ with respect to temperature at constant stress $\boldsymbol{\sigma}$ (typically, at constant pressure) in terms of its variation at a fixed cell structure and of the structural variation at constant stress and temperature:

$$\begin{aligned} \left(\frac{\partial \varepsilon_{\mathbf{k}n}(\{V\}, T)}{\partial T}\right)_{\boldsymbol{\sigma}} &= \left(\frac{\partial \varepsilon_{\mathbf{k}n}(\{V\}, T)}{\partial T}\right)_{\{V^0\}} \\ &+ \left(\frac{\partial \varepsilon_{\mathbf{k}n}(\{V\}, T)}{\partial \{V\}}\right)_T \left(\frac{\partial \{V\}}{\partial T}\right)_{\boldsymbol{\sigma}}. \end{aligned} \quad (5.2.2)$$

The first term describes the effect of the electron interaction with a given phonon population at constant equilibrium configuration $\{V^0\} \triangleq \{\{a_i^0\}, \{b_i^0\}, \{z_i^0\}\}$, which minimizes total electronic energy in the static lattice approximation, that is, the EPI contribution that will be defined in Sec. 5.2.4. The second term captures the effect of the thermal modification of the lattice, including all variations of the lattice parameters, lattice angles and internal coordinates, and can be seen as a deformation potential of the band gap [50].

Working in a quasiparticle framework (e.g. DFT or GW approximations), one can obtain the temperature-dependent eigenvalue correction to the eigenvalue for static lattice, $\varepsilon_{\mathbf{kn}}^0$, by integrating Eq. (5.2.2):

$$\begin{aligned}
\Delta\varepsilon_{\mathbf{kn}}(T) &= \varepsilon_{\mathbf{kn}}(T) - \varepsilon_{\mathbf{kn}}^0(\{V^0\}) \\
&\approx \int_0^T dT' \left. \frac{\partial\varepsilon_{\mathbf{kn}}(\{V\}, T')}{\partial T'} \right|_{\{V^0\}} + \text{ZPR}_{\mathbf{kn}}^{\text{EPI}} \\
&+ \int_0^T dT' \left. \frac{\partial\varepsilon_{\mathbf{kn}}(\{V\}, T')}{\partial\{V\}} \right|_{T'} \left. \frac{\partial\{V\}}{\partial T'} \right|_{\sigma} + \text{ZPR}_{\mathbf{kn}}^{\text{ZPLE}} \\
&= \Delta\varepsilon_{\mathbf{kn}}^{\text{EPI}}(T) + \text{ZPR}_{\mathbf{kn}}^{\text{EPI}} + \Delta\varepsilon_{\mathbf{kn}}^{\text{TE}}(T) + \text{ZPR}_{\mathbf{kn}}^{\text{ZPLE}},
\end{aligned} \tag{5.2.3}$$

where $\Delta\varepsilon_{\mathbf{kn}}^{\text{EPI}}$ and $\Delta\varepsilon_{\mathbf{kn}}^{\text{TE}}$ respectively stand for the electron-phonon interaction and thermal expansion (TE) contribution to the renormalization. Note that the static equilibrium eigenvalues are implicitly computed at $T = 0$ K. The lower bound of both integrals refers to the *dynamical* $T = 0$ K, which considers both electrons and ions. The two integration constants, $\text{ZPR}_{\mathbf{kn}}^{\text{EPI}}$ and $\text{ZPR}_{\mathbf{kn}}^{\text{ZPLE}}$, capture the effect of the zero-point motion of the ions on the $T = 0$ K modification of the static electronic eigenenergies. We thus implicitly suppose that the true $\text{ZPR}_{\mathbf{kn}}$ correction, which corresponds to the EPI interaction computed at the $T = 0$ K renormalized lattice structure, can be evaluated in two distinct steps, hence the similar symbol on the second line. The band gap ZPR, which we define using the symbol ZPR_g , can thus be approximated at first order by the sum of the EPI and lattice deformation contributions [51]:

$$\text{ZPR}_g \simeq \text{ZPR}_g^{\text{EPI}} + \text{ZPR}_g^{\text{ZPLE}}. \tag{5.2.4}$$

We now focus on the contribution of lattice deformation. We can recast the second term of Eq. (5.2.2) by making explicit the derivative with respect to $\{V\}$ in terms of derivatives with respect to the lattice parameters, lattice angles and internal coordinates:

$$\begin{aligned}
&\left(\frac{\partial\varepsilon_{\mathbf{kn}}(\{V\}, T)}{\partial\{V\}} \right)_T \left(\frac{\partial\{V\}}{\partial T} \right)_{\{\sigma_i\}} \\
&= \sum_{j=1}^3 \left(\frac{\partial\varepsilon_{\mathbf{kn}}(\{V\}, T)}{\partial a_j} \right)_{\{a_{i \neq j}\}, \{b_i\}, \{z_i\}} \left(\frac{\partial a_j}{\partial T} \right)_{\{\sigma_i\}} \\
&+ \sum_{j=1}^3 \left(\frac{\partial\varepsilon_{\mathbf{kn}}(\{V\}, T)}{\partial b_j} \right)_{\{b_{i \neq j}\}, \{a_i\}, \{z_i\}} \left(\frac{\partial b_j}{\partial T} \right)_{\{\sigma_i\}} \\
&+ \sum_j \left(\frac{\partial\varepsilon_{\mathbf{kn}}(\{V\}, T)}{\partial z_j} \right)_{\{z_{i \neq j}\}, \{a_i\}, \{b_i\}} \left(\frac{\partial z_j}{\partial T} \right)_{\{\sigma_i\}}.
\end{aligned} \tag{5.2.5}$$

In this expression, all the structural parameters kept constant are set at their static equilibrium value, $\{V^0\}$.

We can, however, circumvent the explicit evaluation of the derivatives contained in Eq. (5.2.5) by supposing that the temperature-dependent eigenvalue correction induced by lattice deformation ($\Delta\varepsilon_{\mathbf{kn}}^{\text{TE}}$, third line of Eq. (5.2.3)) can be approximated from the static DFT eigenvalue evaluated at the temperature-dependent lattice structure, $\{V(T)\}$. As will be made explicit in the following section (Sec. 5.2.2), we furthermore suppose that the leading contribution to $\Delta\varepsilon_{\mathbf{kn}}^{\text{TE}}$ comes from the variation of the lattice parameters and lattice angles, and work with the set of structural parameters $\{\tilde{V}(T)\}$, in which the internal coordinates are chosen to minimize the static DFT total energy at fixed lattice vectors. We thus obtain

$$\Delta\varepsilon_{\mathbf{kn}}^{\text{TE}}(T) \approx \varepsilon_{\mathbf{kn}}^{\text{DFT}}(\{\tilde{V}(T)\}) - \varepsilon_{\mathbf{kn}}^{\text{DFT}}(\{V^0\}). \quad (5.2.6)$$

Within these approximations, $\text{ZPR}_g^{\text{ZPLE}}$ becomes the difference between the DFT band gap energy evaluated at the $T = 0$ dynamical lattice structure including zero-point motion effects and the one evaluated at the static equilibrium structure.

5.2.2. Zero-point lattice expansion in the quasiharmonic approximation

5.2.2.1. Free energy formalism. Throughout this work, we use the Hartree atomic unit system ($\hbar = m_e = c = |e| = 1$). At constant external stress $\{\boldsymbol{\sigma}\}$, the Helmholtz free energy (FE) of the crystal can be approximated by the sum of the electronic and vibrational contributions:

$$\begin{aligned} F(\{V\}, T) &= F^e(\{V\}) + F^{\text{vib}}(\{V\}, T) \\ &= E_{\text{stat}}^e(\{V\}) - k_B T \ln Z^{\text{ph}}(\{V\}, T), \end{aligned} \quad (5.2.7)$$

where the electronic part is further approximated by the total energy obtained at a given structural configuration $\{V\}$ in the static lattice approximation, since we are dealing with semiconductors and insulators, whose entropic contribution to the FE can be safely neglected. k_B is the Boltzmann constant.

From the phonon partition function in the harmonic approximation, Z^{ph} , the vibrational free energy can be decomposed into the zero-point (ZP) and thermal (th) contributions:

$$\begin{aligned} F^{\text{vib}} &= -k_B T \ln(Z^{\text{ZP}} Z^{\text{th}}) \\ &= \sum_{\mathbf{q}\nu} \frac{\omega_{\mathbf{q}\nu}(\{V\})}{2} + k_B T \sum_{\mathbf{q}\nu} \ln \left(1 - e^{-\frac{\omega_{\mathbf{q}\nu}(\{V\})}{k_B T}} \right) \\ &\triangleq F^{\text{ZP}}(\{V\}) + F^{\text{th}}(\{V\}, T), \end{aligned} \quad (5.2.8)$$

where the contributions of all phonon modes with wavevector \mathbf{q} , branch index ν and frequency $\omega_{\mathbf{q}\nu}$ are summed for all branch indices and all wavevectors in the Brillouin zone (BZ). The

first term of this expression is the well-known zero-point energy, which is responsible for the ZPLE. In this equation and for the remaining of this work, all summations over the phonon modes are implicitly normalized by the number of wavevectors in the BZ sampling.

In the quasiharmonic approximation (QHA), the main contribution to the temperature dependence of the phonon frequencies is represented by their variation with respect to the lattice parameters, which has the effect of causing thermal expansion at finite temperature and ZPLE at $T = 0$ K. It should be understood that the purpose of the QHA is not to capture the full anharmonicity of the interatomic force constants at a fixed structure, which is the true physical origin of thermal expansion and ZPLE, as would do non-perturbative methods or the TDEP method [36]. It rather provides us with an accurate leading-order expression for the temperature-dependent structural change of the crystal [3]. For most materials, the true anharmonicity correction to the ZPLE beyond QHA can be reasonably expected to be negligible, as we investigate weakly anharmonic materials and restrict ourselves to $T = 0$ K (see Ref. [3]).

The temperature-dependent lattice, $\{V(T)\}$, will therefore be the one which minimizes $F(\{V\}, T)$ at a given fixed temperature T . By comparison, the static equilibrium lattice, or bare lattice, $\{V^0\}$, minimizes the lattice-dependent static total energy, $E_{\text{stat}}^e(\{V\})$. The zero-point lattice modification can then be defined as the difference between the $T = 0$ K and bare lattices, $\{V(T = 0)\} - \{V^0\}$.

At this point, we emphasize that, for a general crystal, the temperature-dependent lattice structure, $\{V(T)\}$, includes, in principle, the internal coordinates. While this point is briefly addressed in Ref. [3], it is most customary in the literature to neglect this degree of freedom when evaluating Eq. (5.2.7). It is actually absorbed inside the sampled lattice structures: the internal coordinates are fully optimized in the static $T = 0$ K DFT framework, for the imposed T -dependent lattice parameters and angles. The phonon frequencies entering Eq. (5.2.8) are also evaluated at these static optimized configurations. In that case, the FE becomes

$$\begin{aligned} F(\{V\}, T) &\simeq F(\{a_i\}, \{b_i\}, T) \\ &\triangleq F(\{a_i\}, \{b_i\}, \{z_i | \min E_{\text{stat}}^e(\{a_i, b_i\})\}, T). \end{aligned} \quad (5.2.9)$$

We then define the temperature-dependent lattice structure $\{\tilde{V}(T)\}$ as the set of lattice parameters and angles, $\{\tilde{a}_i(T)\}$ and $\{\tilde{b}_i(T)\}$, which minimizes $F(\{a_i\}, \{b_i\}, T)$ at fixed temperature T , as well as the internal coordinates $\{\tilde{z}_i(T)\}$ which minimize the static DFT electronic energy at this lattice structure:

$$\begin{aligned} \{\tilde{V}(T)\} = \{ &\{\tilde{a}_i(T), \tilde{b}_i(T) | \min F(\{a_i\}, \{b_i\}, T), \\ &\{\tilde{z}_i(T) | \min E_{\text{stat}}^e(\{\tilde{a}_i(T), \tilde{b}_i(T)\})\} \}. \end{aligned} \quad (5.2.10)$$

For materials with axial symmetry, $F(\{V\}, T)$ thus reduces to a 2D surface, $F(a, c, T)$, such that the ZPLE is given by $\{a(T=0) - a^0, c(T=0) - c^0\}$. For cubic symmetry, one recovers a 1D FE curve, $F(V, T)$, with linear ZPLE = $a(T=0) - a^0$.

5.2.2.2. Grüneisen formalism. Another approach, known as the Grüneisen formalism, is frequently used in the literature to obtain the temperature-dependent lattice parameters. For simplicity, we first derive the expressions for the ZPLE and thermal expansion coefficients for cubic lattices. In that case, Eq. (5.2.7) reduces to a single dependency on the volume, $\{V\} \rightarrow V$.

We start from the equation of state relating pressure to the Helmholtz FE:

$$P(V, T) = - \left. \frac{\partial F}{\partial V} \right|_T. \quad (5.2.11)$$

F can be Taylor-expanded to lowest order around the static equilibrium volume, V^0 , yielding [52]

$$P(V, T) \approx - \frac{B_0}{V^0} (V - V^0) + \frac{1}{V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V \omega_{\mathbf{q}\nu}(V^0) \left(n_{\mathbf{q}\nu}(V^0, T) + \frac{1}{2} \right), \quad (5.2.12)$$

where B_0 is the isothermal bulk modulus, $n_{\mathbf{q}\nu}$, the Bose-Einstein occupation function of the phonon population, and where we have introduced the volumic mode Grüneisen parameters:

$$\gamma_{\mathbf{q}\nu}^V = - \frac{V^0}{\omega_{\mathbf{q}\nu}(V^0)} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \right|_{V^0}. \quad (5.2.13)$$

The average of these mode parameters weighted by the phonon contribution to the specific heat, $c_{\mathbf{q}\nu}(T)$, defines the temperature-dependent bulk Grüneisen constant,

$$\gamma(T) = \frac{\sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V c_{\mathbf{q}\nu}(T)}{\sum_{\mathbf{q}\nu} c_{\mathbf{q}\nu}(T)}, \quad (5.2.14)$$

often used in the experimental literature to parametrize the thermal expansion coefficient of cubic materials [52, 53].

The equilibrium configuration at any given temperature will verify $P = 0$. We can then express the relative volume change with respect to static equilibrium at temperature T as

$$\frac{\Delta V(T)}{V^0} = \frac{1}{B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V \omega_{\mathbf{q}\nu} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right), \quad (5.2.15)$$

where from now on we simplify the notation, $\omega_{\mathbf{q}\nu}(V^0) \rightarrow \omega_{\mathbf{q}\nu}$ and $n_{\mathbf{q}\nu}(V^0, T) \rightarrow n_{\mathbf{q}\nu}(T)$.

Evaluating Eq. (5.2.15) at $T = 0$ K yields the volumic ZPLE:

$$\Delta V(T=0) = \frac{1}{B_0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V \frac{\omega_{\mathbf{q}\nu}}{2}. \quad (5.2.16)$$

The Grüneisen formalism as presented in most textbooks neglects the zero-point energy in the evaluation of Eq. (5.2.12), as it focuses on the thermal expansion coefficient, hence the absence of zero-point correction to the lattice volume.

We now proceed to generalize the Grüneisen formalism to anisotropic materials, yet still retaining the previously defined approximation for the internal atomic coordinates (see Eq. (5.2.9) and the second line of Eq. (5.2.10), with $\{\tilde{a}_i(T)\}$ now referring to the temperature-dependent lattice parameters obtained from the Grüneisen approach). To simplify the notation, we will refer to $\{\tilde{a}_i(T)\}$ as $\{a_i(T)\}$. Until now, we have defined the lattice structure in terms of the $\{a_i\}$ and $\{b_i\}$, which are well defined regardless of any reference structure. This is a natural choice when working with the FE. As the Grüneisen approach is based on an expansion around the static equilibrium configuration, we will rather use the strain tensor, $\boldsymbol{\epsilon}$, whose components (here expressed in Voigt notation [52]) describe the relative deformations of the lattice parameters and lattice angles with respect to a given reference configuration, which we choose to be the static equilibrium structure:

$$\begin{aligned}\epsilon_{i=1,2,3} &= \frac{a_i - a_i^0}{a_i^0}, \\ \epsilon_{i=4,5,6} &= \frac{b_i - b_i^0}{b_i^0}.\end{aligned}\tag{5.2.17}$$

Note that other definitions of the reference configuration are commonly used in the literature, such as the room temperature and the dynamical $T = 0$ K (including zero-point effects) structures. These are natural choices when reporting experimental values of the thermal expansion coefficients.

We start from the differential form of the Helmholtz FE [52],

$$dF = -k_B \ln Z^{\text{ph}} dT + V^0 \sum_{i=1}^6 \sigma_i d\epsilon_i.\tag{5.2.18}$$

The stress tensor, $\boldsymbol{\sigma}$, is related to the strain tensor, $\boldsymbol{\epsilon}$, by the elastic stiffness tensor, \boldsymbol{c} , whose components c_{ij} are the isothermal elastic constants of the material,

$$\sigma_i = \sum_{j=1}^6 c_{ij} \epsilon_j.\tag{5.2.19}$$

From Eq. (5.2.18), we obtain the generalization of Eq. (5.2.11) to anisotropic materials by replacing the volume derivative by derivatives with respect to a given strain. Note that the sign of the derivative is now positive as the components σ_i describe the internal stresses:

$$\sigma_i(T) = \frac{1}{V^0} \frac{\partial F(T)}{\partial \epsilon_i}.\tag{5.2.20}$$

The generalization of Eq. (5.2.12) is again obtained by Taylor-expanding around the equilibrium configuration at zero strain:

$$\sigma_i \approx \frac{1}{V^0} \left[V^0 \sum_{j=1}^6 c_{ij} \epsilon_j - \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^i \omega_{\mathbf{q}\nu}(\boldsymbol{\epsilon} = 0) \left(n_{\mathbf{q}\nu}(\boldsymbol{\epsilon} = 0, T) + \frac{1}{2} \right) \right], \quad (5.2.21)$$

where we introduce the mode Grüneisen tensor

$$\gamma_{\mathbf{q}\nu}^i = - \frac{1}{\omega_{\mathbf{q}\nu}} \left. \frac{\partial \omega_{\mathbf{q}\nu}(\epsilon_i)}{\partial \epsilon_i} \right|_{\epsilon=0}, \quad (5.2.22)$$

in which the derivative are now taken with respect to the individual strains, ϵ_i .

From now on, we simplify the notation by assuming that the phonon frequencies and occupation function are evaluated at zero strain. Further assuming that the stress components are all zero at the equilibrium configuration of a given temperature, this reduces to

$$\sum_{j=1}^6 c_{ij} \epsilon_j(T) = \frac{1}{V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^i \omega_{\mathbf{q}\nu} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right). \quad (5.2.23)$$

In order to obtain the generalized form of Eq. (5.2.15), we take the tensor product with the elastic compliance tensor, \mathbf{s} :

$$\sum_{i=1}^6 s_{ki} \sum_{j=1}^6 c_{ij} \epsilon_j = \frac{1}{V^0} \sum_{i=1}^6 s_{ki} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^i \omega_{\mathbf{q}\nu} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right). \quad (5.2.24)$$

The left-hand side of this expression reduces to the individual strains as the elastic compliance tensor is, by definition, the inverse of the elastic stiffness tensor. At finite temperatures, one could further refine Eq. (5.2.23) and use temperature-dependent elastic constants, which can be obtained by differentiating the Helmholtz FE at temperature T with respect to strains [34].

For any anisotropic material, the temperature-dependent strains within the QHA can therefore be expressed as

$$\epsilon_k(T) = \frac{1}{V^0} \sum_{\mathbf{q}\nu} \sum_{i=1}^6 s_{ki} \gamma_{\mathbf{q}\nu}^i \omega_{\mathbf{q}\nu} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right). \quad (5.2.25)$$

For linear thermal expansion of the material along the crystallographic axis, the relevant strains are $\epsilon_k(T) = \Delta a_k(T)/a_k^0$, where $k = 1, 2, 3$ are Voigt indices. Recalling that the thermal expansion coefficients, α_k , are the temperature derivatives of the strains at constant stress,

$$\alpha_k = \left. \frac{d\epsilon_k}{dT} \right|_{\sigma}, \quad (5.2.26)$$

we obtain [54]

$$\begin{aligned}
\alpha_k(T) &= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \sum_{i=1}^6 s_{ki} \gamma_{\mathbf{q}\nu}^i \frac{d}{dT} (\omega_{\mathbf{q}\nu} n_{\mathbf{q}\nu}(T) + \frac{1}{2}) \\
&= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \sum_{i=1}^6 s_{ki} \gamma_{\mathbf{q}\nu}^i \frac{d}{dT} (\tilde{\varepsilon}_{\mathbf{q}\nu}(T)) \\
&= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \sum_{i=1}^6 s_{ki} \gamma_{\mathbf{q}\nu}^i c_{\mathbf{q}\nu}^V(T),
\end{aligned} \tag{5.2.27}$$

where $\tilde{\varepsilon}_{\mathbf{q}\nu}(T)$ is the average phonon energy at temperature T and $c_{\mathbf{q}\nu}^V(T)$ is the phonon specific heat at constant volume at this temperature.

Equation (5.2.27) thus generalizes the Grüneisen approach to any material, under our approximation for the internal coordinates. It has the exact same form as if one would have neglected the zero-point energy in Eq. (5.2.25), as is typically done within this formalism. Nevertheless, if we evaluate Eq. (5.2.25) at $T = 0$ K, we obtain a generalized expression for the ZPLE:

$$\Delta a_k(T = 0) = \frac{a_k^0}{V^0} \sum_{\mathbf{q}\nu} \sum_{i=1}^6 s_{ki} \gamma_{\mathbf{q}\nu}^i \frac{\omega_{\mathbf{q}\nu}}{2}. \tag{5.2.28}$$

Note that, for the remainder of this article, the ZPLE refers to the zero-point length shift of the lattice parameters only. In principle, one could also extract from Eq. (5.2.25) the zero-point variation of the cell angles, for materials in which they are not set to zero by symmetry.

By evaluating the temperature-dependent lattice constants as

$$a_k(T) = a_k^0 \int_0^T \alpha_k(T') dT' + a_k(T = 0), \tag{5.2.29}$$

where $a_k(T = 0) = a_k^0 + \Delta a_k(T = 0)$ is the ZPLE-renormalized lattice constant obtained from Eq. (5.2.28), the zero-point lattice effect can thus easily be taken into account within the anisotropic Grüneisen formalism. To the best of our knowledge, no ZPLE results were explicitly reported in the literature using this formula. We note that this expression can be seen as the anisotropic generalization of Eq. (4) of Ref. [18], in the $T = 0$ K, $P = 0$, one-shot limit. The isotropic Eq. (5.2.16) was also used by Querales-Flores *et al.* [37] to report the ZPLE contribution to the ZPR_g in cubic PbTe. See also Sec. IV of Ref. [3] and references therein.

Section S5.1 of the Supplemental Material [55] extends the Grüneisen formalism to a fixed external hydrostatic pressure, while Sec. S5.2 shows how the well-known literature results for the case of cubic and axial symmetries can be easily retrieved from Eq.(5.2.28). We state

their explicit ZPLE expressions and show that the isotropic Eq. (5.2.16) is contained in the anisotropic formulation of Eq. (5.2.25).

5.2.3. Estimation of ZPLE correction to the band gap energy

We now estimate the size of the ZPLE correction to the electronic eigenenergies, $\varepsilon_{\mathbf{k}n}$, induced by the zero-point motion in terms of simple parameters. In the following, we will first retain the temperature dependence of the electronic eigenenergies, starting from the volumic simplification of the TE contribution in Eq. (5.2.2) and replacing the partial derivatives with respect to volume by logarithmic derivatives:

$$\left(\frac{\partial\varepsilon_{\mathbf{k}n}(V,T)}{\partial T}\right)_P^{\text{TE}} = \left(\frac{\partial\varepsilon_{\mathbf{k}n}(V,T)}{\partial\ln V}\right)_T \left(\frac{\partial\ln V}{\partial T}\right)_P, \quad (5.2.30)$$

then take the $T \rightarrow 0$ K limit to obtain the final expression. We stress that, in principle, Eq. (5.2.30) should only be applied to cubic systems. Nevertheless, since we are seeking for an approximate expression, it can be expected to be reasonably accurate for sufficiently symmetric anisotropic materials, such as the wurtzite structure.

From the chain rule, we can recast the constant temperature term in terms of a derivative with respect to pressure and the bulk modulus, B_0 :

$$\begin{aligned} \frac{\partial\varepsilon_{\mathbf{k}n}}{\partial\ln V} &= \frac{\partial\varepsilon_{\mathbf{k}n}}{\partial P} \frac{\partial P}{\partial\ln V} \\ &= \frac{\partial\varepsilon_{\mathbf{k}n}}{\partial P} (-B_0). \end{aligned} \quad (5.2.31)$$

By definition, the constant pressure term is simply the volumic thermal expansion coefficient, $\beta(T)$:

$$\begin{aligned} \frac{1}{V} \frac{\partial V}{\partial T} &= \beta(T) \\ &= \frac{\partial}{\partial T} \left(\frac{\Delta V(T)}{V^0} \right). \end{aligned} \quad (5.2.32)$$

This expression can be adapted to any crystal symmetry by recalling that $\beta(T)$ is the trace of the thermal expansion tensor, whose coefficients describe the strain induced on the lattice by an infinitesimal temperature increment:

$$\begin{aligned} \beta(T) &= \sum_{k=1}^3 \alpha_k(T), \\ &= \sum_{k=1}^3 \left(\frac{\partial\epsilon_k}{\partial T} \right). \end{aligned} \quad (5.2.33)$$

In the case of cubic and axial symmetry, this reduces to

$$\beta(T) = \begin{cases} 3\alpha(T), & \text{(cubic),} \\ 2\alpha_a(T) + \alpha_c(T), & \text{(axial).} \end{cases} \quad (5.2.34)$$

Substituting Eq. (5.2.31) and (5.2.32) in the last term of Eq. (5.2.30), we can express the rate of change of the electronic eigenvalues with respect to temperature at any temperature T' as

$$\left(\frac{\partial \varepsilon_{\mathbf{k}n}(V, T)}{\partial T} \right)_P^{\text{TE}}(T') = -\frac{\partial \varepsilon_{\mathbf{k}n}}{\partial P} B_0(T') \beta(T'). \quad (5.2.35)$$

The thermal expansion-induced change to the electronic eigenenergies is obtained by integrating the previous expression from $T' = 0$ to $T' = T$, recalling that the lower bound is the dynamical $T = 0$ K lattice. Further neglecting the temperature dependence of the bulk modulus, which is typically very small at low temperatures, and taking advantage of the definition of $\beta(T')$ as a temperature derivative, we obtain

$$\begin{aligned} \Delta \varepsilon_{\mathbf{k}n}(T) &= \varepsilon_{\mathbf{k}n}(T) - \varepsilon_{\mathbf{k}n}^0 \\ &\approx -B_0 \frac{\partial \varepsilon_{\mathbf{k}n}}{\partial P} \left(\frac{\Delta V(T)}{V^0} + \frac{\Delta V^{\text{ZPLE}}}{V^0} \right), \end{aligned} \quad (5.2.36)$$

where the integration constant (second term in the large parentheses) captures the $T = 0$ K volume variation compared to the static lattice and is equivalent to the second integration constant in Eq. (5.2.3). Hence, one obtains the ZPLE-induced modification of an electronic eigenstate,

$$\Delta \varepsilon_{\mathbf{k}n}(T = 0) \approx -B_0 \frac{\partial \varepsilon_{\mathbf{k}n}}{\partial P} \frac{\Delta V^{\text{ZPLE}}}{V^0}, \quad (5.2.37)$$

where $\Delta V^{\text{ZPLE}}/V^0 = (V(T = 0) - V^0)/V^0$. In the following, we will refer to this term as $\Delta V/V^0$ for simplicity.

We finally note that, since the volume of the unit cell is allowed to change, the interatomic distances are modified, which in turn affects the ionic potential, hence the value of the Fermi energy. To remove any ambiguity regarding the reference energy of a given electronic band, we use Eq. (5.2.37) to compute directly the change to the band gap energy, which is defined irrespective of a reference Fermi energy.

The first two factors entering the right-hand side of Eq. (5.2.37), namely, the bulk modulus and the band gap derivative with respect to pressure (also known in the literature as the pressure coefficient), can easily be obtained either from experimental data or from static DFT calculations. The last one, $\Delta V/V^0$, requires a little more care and can be estimated in two ways. On the one hand, one can extrapolate the high-temperature limit slope of the curve for $\Delta a(T)/a^0$ and calculate its intercept at $T = 0$ K. The ZPLE is then simply the difference between the extrapolated value and the actual lattice parameter at $T = 0$ (see Fig. 2 of Ref. [50], and note that our definition of $\Delta a/a^0$ differs from theirs by a sign change). This

procedure is similar to the one frequently used in the literature to evaluate the full band gap ZPR from experimental values of $E_g(T)$ (see Fig. 21 of Ref. [50]). Note that the accuracy of such extrapolation depends on the measurements extending to sufficiently high temperature, i.e. much larger than the Debye temperature, as will be discussed in Sec. 5.4.3.2. One could also use theoretical results for $\Delta V/V^0$ from Eq. (5.2.25) to validate Eq. (5.2.37) a posteriori.

Lastly, one could also, in principle, make use of a simple semiempirical model proposed in the literature to express $\Delta V/V^0$ in terms of a few parameters [56]. This model relies on the Debye model for the expression for the zero-point energy and the Dugdale-MacDonald parametrization of the bulk Grüneisen constant in terms of the first derivative of the bulk modulus with respect to pressure, B_1 [57]. Other parametrizations of γ^V in terms of B_1 have been proposed (see Ref. [19] and references therein).

This simple model was widely used in the early 2000s to account for the ZPLE when comparing the accuracy of lattice parameter prediction from different exchange-correlation functionals (see Ref. [24] and references therein) and was originally derived for crystals with one atom in the unit cell. Ref. [26] showed that it strongly overestimates the ZPLE for diamond and zincblende structures. In that case, the poor predictive capabilities of the semiempirical model can be understood by recalling that it is based on bulk properties. Namely, it supposes that the average behavior of the whole phonon population can be approximated by the behavior of a single “average phonon”, whose frequency is lowered by increasing pressure. For most diamond and zincblende materials, the thermal expansion coefficient is negative at low temperatures. This behavior is driven by low-frequency modes, which are typically associated with bending modes that strengthen the atomic bonds during volumic expansion [58], hence a phonon hardening.

By looking at Eq. (5.2.14), we can understand that the proper definition of the bulk Grüneisen constant as a weighted sum of the individual Grüneisen mode parameters will correctly capture this unusual behavior, as it allows each mode parameter to be either positive or negative. Ignoring these alternating signs for low-frequency modes, as is done in the semiempirical model, naturally yields an overestimation of the total bulk Grüneisen constant. We will thus restrict our estimation of $\Delta a_k(T=0)/a_k^0$ to the high-temperature slope method explained earlier.

5.2.4. Electron-phonon interaction

The EPI contribution to the temperature dependence of the band gap energy, and, consequently, on ZPR_g , has been thoroughly discussed in the literature. Many theoretical formalisms have been used to address this question, among them ab initio molecular dynamics [59, 60], path integral molecular dynamics [61], importance sampling Monte Carlo [62], finite differences and distorted supercells [8, 63, 64], the special displacement method [65], the

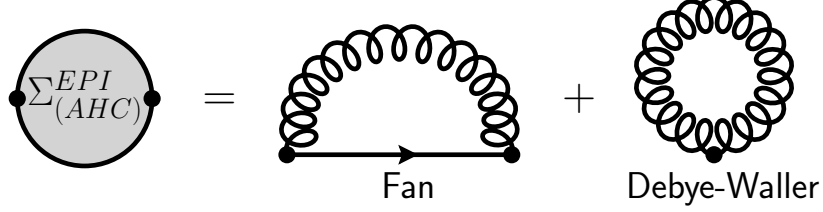


Figure 5.3 – Fan and Debye-Waller contributions entering the lowest order EPI self-energy in the AHC framework.

Bethe-Salpeter equation for excitonic gaps [66] as well as the Allen-Heine-Cardona (AHC) formalism [51, 67, 68], either adiabatic [69, 70] or non-adiabatic [71, 72]. In the following, we briefly review the key concepts of the non-adiabatic AHC framework. We invite our readers to consult Refs. [11, 70, 73] for a complete discussion of this methodology and of its underlying approximations.

In many-body perturbation theory [74], the zero-point renormalization of an electronic energy level $\varepsilon_{\mathbf{k}n}$ is captured by the real part of the electron-phonon self-energy, $\Sigma_{\mathbf{k}n}^{\text{EPI}}$, evaluated at the quasiparticle energy, $\varepsilon_{\mathbf{k}n}(T = 0)$:

$$\begin{aligned} \text{ZPR}_{\mathbf{k}n}^{\text{EPI}} &= \varepsilon_{\mathbf{k}n}(T = 0) - \varepsilon_{\mathbf{k}n}^0 \\ &= \Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}(T = 0), T = 0) \right]. \end{aligned} \quad (5.2.38)$$

In this expression, $\varepsilon_{\mathbf{k}n}^0$ is the unperturbed Kohn-Sham (KS) eigenvalue computed at the fixed lattice geometry which minimizes the total electronic energy of the system in the static approximation, $\{V^0\}$. For the remaining of this section, $T = 0$ is implied throughout.

The self-energy at second order in perturbation theory can be expressed as the sum of the Fan and Debye-Waller contributions, shown in Fig. 5.3. Assuming that the quasiparticle energy is close to the bare eigenenergy, one can evaluate the frequency-dependent Fan self-energy at the unperturbed electronic energy, in the so-called on-the-mass-shell approximation [75], yielding

$$\text{ZPR}_{\mathbf{k}n}^{\text{EPI}} \approx \Re \left[\Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega = \varepsilon_{\mathbf{k}n}^0) + \Sigma_{\mathbf{k}n}^{\text{DW}} \right]. \quad (5.2.39)$$

We finally approximate the fully-interacting wavefunction by the KS-DFT wavefunction and obtain the standard expression for the non-adiabatic Fan self-energy [11, 70]:

$$\begin{aligned} \Sigma_{\mathbf{k}n}^{\text{Fan}}(\varepsilon_{\mathbf{k}n}^0) &= \sum_{\mathbf{q}\nu} \sum_{n'}^{\text{BZ}} | \langle \mathbf{k} + \mathbf{q}n' | V_{\mathbf{q}\nu}^{(1)} | \mathbf{k}n \rangle |^2 \\ &\times \left[\frac{1 - f_{\mathbf{k}+\mathbf{q}n'}}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 - \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} + \frac{f_{\mathbf{k}+\mathbf{q}n'}}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 + \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} \right]. \end{aligned} \quad (5.2.40)$$

The electron-phonon matrix elements, $\langle \mathbf{k} + \mathbf{q}n' | V_{\mathbf{q}\nu}^{(1)} | \mathbf{k}n \rangle$, capture the probability amplitude of a coupling between two electronic eigenstates, given the first-order change in the KS potential induced by the collective atomic motion along the $\mathbf{q}\nu$ -phonon mode, $V_{\mathbf{q}\nu}^{(1)}$. Since

we work at $T = 0$ K, the two terms inside the square brackets can be interpreted as virtual phonon emission processes by respectively electrons and holes [76]. In the case of semiconductors, the Fermi-Dirac occupation functions at $T = 0$ K, $f_{\mathbf{k}+\mathbf{q}n'}$, are either 0 for conduction bands or 1 for occupied states. The small parameter $\eta_{\mathbf{k}}$ in the denominators correctly shifts the poles of the Green's function in the complex plane to maintain causality.

The first-order potential can be written as

$$V_{\mathbf{q}\nu}^{(1)} = \frac{1}{\sqrt{2\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha} U_{\nu,\kappa\alpha}(\mathbf{q}) \sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial V^{\text{KS}}}{\partial \mathbf{R}_{l\kappa\alpha}} \quad (5.2.41)$$

$$\triangleq \nabla_{\mathbf{q}\nu} V^{\text{KS}},$$

where $\mathbf{R}_{l\kappa\alpha}$ represents the displacement of atom κ , located in unit cell l , in cartesian direction α , and V^{KS} is the Kohn-Sham potential. The second line emphasizes the definition of the $\nabla_{\mathbf{q}\nu}$ operator. The phonon eigendisplacement vectors, $U_{\nu,\kappa\alpha}(\mathbf{q})$, are the solutions of the generalized eigenvalue equation

$$M_\kappa \omega_{\mathbf{q}\nu}^2 U_{\nu,\kappa\alpha}(\mathbf{q}) = \sum_{\kappa'\alpha'} \left(\sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial^2 E}{\partial \mathbf{R}_{l\kappa\alpha} \partial \mathbf{R}_{0\kappa'\alpha'}} \right) U_{\nu,\kappa'\alpha'}(\mathbf{q}), \quad (5.2.42)$$

where M_κ is the atomic mass of atom κ and E is the total energy of the static lattice. They verify the normalization condition

$$\sum_{\kappa\alpha} M_\kappa U_{\nu,\kappa\alpha}^*(\mathbf{q}) U_{\nu',\kappa\alpha}(\mathbf{q}) = \delta_{\nu\nu'}. \quad (5.2.43)$$

The static Debye-Waller self-energy formally involves second derivatives of the KS potential evaluated at first order in perturbation theory:

$$\Sigma_{\mathbf{k}n}^{\text{DW}} = \sum_{\mathbf{q}\nu} \frac{1}{2} \langle \mathbf{k}n | \nabla_{\mathbf{q}\nu} \nabla_{-\mathbf{q}\nu} V^{\text{KS}} | \mathbf{k}n \rangle. \quad (5.2.44)$$

However, in the rigid-ion approximation (RIA), the translational invariance of the lattice allows the second-order matrix elements to be recast in terms of the first derivatives entering the Fan self-energy [70]. By doing so, we suppose that the second derivatives of the Kohn-Sham potential with respect to the displacement of two different atoms within the unit cell vanish. The validity of this approximation was verified for crystals [70, 77], but strongly questioned for diatomic molecules [78].

5.3. Computational details

5.3.1. Ground state calculations and lattice optimization

All first-principles calculations were performed with the ABINIT software package [79]. We obtain the static ground state electronic structure and total energy at $T = 0$ K from density functional theory (DFT). Spin-orbit coupling is neglected throughout. For most

materials, we use norm-conserving pseudopotentials from the PseudoDojo project [80] and rely on the generalized gradient approximation of the Perdew-Burke-Ernzerhof functional (PBE-GGA) [81]. Lattice parameters and relevant internal atomic coordinates were optimized until all forces on the atoms were below 10^{-5} hartree/Bohr³. The optimized lattice parameters were used throughout, except for Ge, as the optimized structure led to an unphysical gapless band structure, and GaP, to ensure the correct relative energy ordering of the electron valleys in the conduction band. In those cases, we used the experimental lattice parameter extrapolated to $T = 0$ K. This was done to ensure a correct evaluation of the EPI contribution to ZPR_g , which can be very sensitive to such details.

For BN and AlN, the reported values for the EPI contribution to ZPR_g comes from previously published data [73]. Hence, for consistency, we used for these materials the Perdew and Wang parametrization of the local density approximation (LDA) [82], FHI98 pseudopotentials [83] and the reported lattice parameters for those materials. The energy cutoffs, optimized lattice parameters and \mathbf{k} -point sampling of the Brillouin zone (BZ) used for all materials can be found in Table S5.1 of the Supplemental Material [55].

5.3.2. Lattice dynamics and electron-phonon interaction

Response-function and electron-phonon coupling calculations were performed within the density functional perturbation theory (DFPT) framework [84]. The \mathbf{q} -point sampling used to compute the phonon frequencies and electron-phonon matrix elements for the different materials can be found in Table S5.1 of the Supplemental Material [55]. We used an imaginary broadening of $\eta = 0.01$ eV in the denominator of Eq. (5.2.40) and accelerated the convergence with respect to the \mathbf{q} -point sampling using the methodology based on the Fröhlich model presented in the Supplementary Methods of Ref. [71]. The Sternheimer equation method is used to reduce the explicit number of bands required in the active subspace of the self-energy [78]. The explicit number of bands was chosen such that the energy difference between the CBM and the highest band was at least 20 eV. Most of the ZPR_g^{EPI} results used in this paper were previously reported in Ref. [71]. For GaAs, the electron-phonon matrix elements for the maximal \mathbf{q} -point grid listed in Table S5.1 of the Supplemental Material [55] were obtained by interpolating the first-order potential computed self-consistently on a $16 \times 16 \times 16$ \mathbf{q} -point grid, using the methodology described in Ref. [79].

5.3.3. Zero-point lattice expansion and band gap correction

For both the FE minimization and Grüneisen methodologies, we evaluated the vibrational spectrum of all lattice configurations using an $8 \times 8 \times 8$ Γ -centered \mathbf{q} -point grid. We verified the convergence of our sampling by comparing with a $24 \times 24 \times 24$ Fourier-interpolated

\mathbf{q} -point grid for both approaches. Note that the \mathbf{q} -point grid used for ZPLE is coarser than the one used for EPI, since it only requires the phonon frequencies, which converge much faster with respect to the \mathbf{q} -point sampling.

For the Grüneisen approach, the relaxed-ion compliance tensor was computed from the strain-strain derivatives of the total energy within DFPT [85], taking into account the relaxation of the atomic coordinates within the unit cell. Note that properly converged elastic constants typically require a small smearing of the plane wave cutoff energy, hence, a slightly higher effective value compared to GS or phonon calculations, since shear strains modify the shape of the unit cell. The specific values used for the different materials are made explicit in Table S5.1 of the Supplemental Material [55]. The derivatives entering Eq. (5.2.22), which is used to evaluate Eq. (5.2.28), were computed from central finite differences using two volumes per degree of freedom and the static equilibrium configuration. To test the stability of the numerical differentiation, we used strained lattice parameters corresponding to a volume change of $\pm 0.5\%$ and $\pm 1\%$ for the unit cell for each independent crystallographic axis.

For the cubic materials, the FE curve was constructed using 9 configurations within a $\pm 5\%$ volume change with respect to the static equilibrium volume. The optimal lattice parameter was obtained by fitting a Murnaghan equation of state to the FE curve. For the wurtzite materials, the free energy surface was computed from 27 volumes for GaN, 42 for AlN and 55 for ZnO. The chosen lattice parameters were such that they individually yielded a volume change within $\pm 1\%$ of the equilibrium volume. The 2D surface was minimized by fitting a second-order polynomial. Most results from FE minimization were previously reported in Table II of the Supplementary Information of Ref. [71].

We obtain the ZPLE-induced band gap ZPR, ZPR_g^{ZPLE} , from static DFT at the renormalized lattice parameters. For materials whose internal coordinates are not fixed by symmetry, i.e. the three wurtzite materials, we use the internal coordinates which minimize the Born-Oppenheimer energy at such parameters.

In the case of Ge and GaP, as the static equilibrium configuration was set at the experimental lattice parameter instead of the theoretical one, we minimized the Gibbs free energy with an artificial pressure of respectively 7.88 GPa for Ge and 2.77 GPa for GaP, in order to shift the minimum of the static calculated Born-Oppenheimer energy curve at the experimental lattice parameter.

Table 5.1 – Zero-point lattice expansion for cubic materials. a^0 is the theoretical cubic lattice parameter in the static lattice approximation, except for Ge and GaP (see Sec. 5.3.1 for details). The variation of a^0 induced by the zero-point motion of the ions, $\Delta a^{\text{FE}}(T = 0)$, is obtained by minimizing the Helmholtz free energy at $T = 0$ K, while $\Delta a^{\text{Grun}}(T = 0)$ is obtained from the Grüneisen parameters, using Eq. (5.2.28). The indication ($T = 0$) is omitted in the table, for sake of simplicity. The central finite differences entering the mode Grüneisen parameters were evaluated with a volume variation of resp. $\pm 1\%$ and $\pm 0.5\%$. The three rightmost columns show the fractional ZPLE with respect to the bare lattice parameter.

Material	a^0 (Bohr)	Δa^{FE} (Bohr)	Δa^{Grun} (Bohr)		$\Delta a^{\text{FE}}/a^0$ (%)	$\Delta a^{\text{Grun}}/a^0$ (%)	
			1%	0.5%		1%	0.5%
C-dia	6.7534	0.0253	0.0250	0.0248	0.375	0.370	0.367
Si-dia	10.3348	0.0174	0.0172	0.0172	0.169	0.166	0.166
Ge-dia	10.5220	0.0102	0.0112	0.0112	0.097	0.106	0.106
SiC-zb	8.2772	0.0227	0.0226	0.0223	0.274	0.273	0.269
BN-zb	6.7460	0.0285	0.0271	0.0270	0.423	0.402	0.400
BAs	9.0880	0.0238	0.0237	0.0236	0.262	0.261	0.260
AlP-zb	10.4057	0.0180	0.0178	0.0178	0.173	0.171	0.171
AlAs-zb	10.8248	0.0166	0.0162	0.0163	0.153	0.150	0.151
AlSb-zb	11.7621	0.0163	0.0162	0.0161	0.139	0.138	0.137
GaN-zb	8.5984	0.0209	0.0208	0.0208	0.243	0.242	0.242
GaP-zb	10.2942	0.0152	0.0156	0.0156	0.148	0.152	0.152
GaAs-zb	10.8627	0.0157	0.0146	0.0145	0.145	0.134	0.134
ZnS-zb	10.2859	0.0194	0.0193	0.0192	0.189	0.188	0.187
ZnSe-zb	10.8330	0.0160	0.0157	0.0155	0.148	0.145	0.143
ZnTe-zb	11.6819	0.0154	0.0153	0.0153	0.132	0.131	0.131
CdS-zb	11.2021	0.0177	0.0170	0.0162	0.158	0.152	0.145
CdSe-zb	11.7114	0.0135	0.0133	0.0135	0.115	0.114	0.115
CdTe-zb	12.5133	0.0129	0.0126	0.0125	0.103	0.101	0.100
MgO-rs	8.0366	0.0333	0.0317	0.0318	0.414	0.394	0.396

5.4. Results and discussion

5.4.1. Zero-point lattice expansion

The numerical ZPLE results for the cubic and wurtzite materials are reported in Tables 5.1 and 5.2 respectively. Note that in most Tables, the material names are appended with their lattice structure, either diamond (dia), zincblende (zb), rocksalt (rs) or wurtzite (w). As expected, the corrections to the static lattice parameter are small, typically between 0.1-0.2% of the static value, except for materials containing light atoms (C, B, N, O), where

Table 5.2 – Zero-point lattice expansion for wurtzite materials. a^0 and c^0 are the theoretical lattice parameters in the static lattice approximation. The variation of a^0 and c^0 induced by the zero-point motion of the ions, respectively $\Delta a^{\text{FE}}(T = 0)$ and $\Delta c^{\text{FE}}(T = 0)$, are obtained by minimizing the 2D Helmholtz free energy surface, while $\Delta a^{\text{Grün}}(T = 0)$ and $\Delta c^{\text{Grün}}(T = 0)$ are obtained from the Grüneisen parameters, using Eq. (5.2.28). The indication ($T = 0$) is omitted in the table, for sake of simplicity. The central finite differences required to obtain the mode Grüneisen parameters were evaluated using a volume variation of resp. $\pm 1\%$ and $\pm 0.5\%$. The three rightmost columns show the fractional ZPLE with respect to the bare lattice parameters. The bottom section reports the resulting variation of the $c/a(T = 0)$ ratio.

Material	a^0 (Bohr)	Δa^{FE} (Bohr)	$\Delta a^{\text{Grün}}$ (Bohr)		$\Delta a^{\text{FE}}/a^0$ (%)	$\Delta a^{\text{Grün}}/a^0$ (%)	
			1%	0.5%		1%	0.5%
AlN	5.783	0.0170	0.0163	0.0162	0.294	0.282	0.280
GaN	6.0841	0.0151	0.0151	0.0151	0.248	0.248	0.248
ZnO	6.1910	0.0165	0.0165	0.0167	0.267	0.267	0.270
Material	c^0 (Bohr)	Δc^{FE} (Bohr)	$\Delta c^{\text{Grün}}$ (Bohr)		$\Delta c^{\text{FE}}/c^0$ (%)	$\Delta c^{\text{Grün}}/c^0$ (%)	
			1%	0.5%		1%	0.5%
AlN	9.255	0.0241	0.0242	0.0242	0.260	0.262	0.262
GaN	9.9110	0.0232	0.0228	0.0229	0.234	0.230	0.231
ZnO	9.9870	0.0213	0.0213	0.0207	0.213	0.214	0.207
Material	c^0/a^0	$(c/a)^{\text{FE}}$ ($T = 0$)	$(c/a)^{\text{Grün}}$ ($T = 0$)		$\Delta(c/a)^{\text{FE}}$	$\Delta(c/a)^{\text{Grün}}$	
			1%	0.5%		1%	0.5%
AlN	1.6004	1.5998	1.6001	1.6001	-0.0006	-0.0003	-0.0003
GaN	1.6290	1.6286	1.6286	1.6286	-0.0004	-0.0004	-0.0004
ZnO	1.6131	1.6122	1.6123	1.6121	-0.0010	-0.0009	-0.0010

it can reach up to 0.4%. This can intuitively be understood in terms of a simple mass and spring model, whose vibrational frequency varies as $\sqrt{k/M}$. Recall that the strength of ZPLE depends on how phonon frequencies vary with respect to volume. Lighter materials tend to have a more localized charge density. A small variation of the lattice parameter thus induces a larger charge reorganization, hence a greater variation of the spring constant k , compared to heavier materials. Note also that the typical size of the ZPLE is roughly one order of magnitude smaller than the typical error associated with the LDA and PBE functionals, the former usually underestimating the lattice parameter, the latter overestimating it. As ZPLE systematically increases the lattice parameter for our set of materials, it will bring the lattice constant closer to the experimental value for LDA and worsen the agreement for PBE.

In the case of wurtzite materials, our generalized formalism allows for anisotropic ZPLE along the independent crystallographic directions. For the three materials considered, we

find that the relative in-plane ZPLE tends to be slightly larger than the out-of-plane ZPLE, resulting in a small, almost negligible decrease of the c/a ratio. This could a posteriori justify using the volumic approach (Eq. (5.2.15)) for this class of materials, therefore approximating the c/a ratio to remain constant. Such a procedure would greatly improve numerical efficiency by reducing the dimensionality of the problem. This might not, however, be advisable for layered or weakly-bonded van der Waals materials, which can display significant elastic anisotropy [86]. In that case, as the independent crystallographic axes respond very differently to the zero-point “phonon pressure” [5], the c/a ratio could be more significantly affected even at very low temperatures.

We finally note that even a negligible modification of the c/a ratio at $T = 0$ K could lead to a noticeable effect as temperature is increased, should the in-plane and out-of-plane thermal expansion coefficients display sufficiently different behaviors, as can be seen for example in Fig. (3) of Ref. [87]. Hence, for non-cubic materials, the volumic and generalized methodologies can start to deviate even at moderate temperatures.

5.4.2. Free energy minimization vs Grüneisen approach

From Tables 5.1 and 5.2, we now focus on the predicted ZPLE from the Grüneisen approach (Eq. (5.2.28), $\Delta a^{\text{Grün}}(T = 0)$). First, comparing results obtained with the 1% and 0.5% volume variations for the finite differences, we find that the discrepancy is at most 0.2 mBohr for most materials, the few exceptions being SiC, CdS and the c parameter of ZnO. Thus, the materials in our dataset do not display great sensitivity to the configuration sampling used for the finite derivatives.

As the Grüneisen formalism is by construction an approximation of the FE approach within the QHA framework, one should naturally compare the accuracy of this method to the outcome of FE minimization. At this point, we stress that our goal is to provide a numerical “proof of concept” of the equivalent predictive capabilities of these two methodologies. Hence, we solely compare the ZPLE values obtained within our numerical calculations. For reasons that will be clarified further in this article (see Sec. 5.4.4), assessing the respective predictability of the two approaches with respect to experimental data would benefit from a full temperature-dependent comparative analysis, which is beyond the scope of the current study.

We observe a good general agreement between the methodologies, as is emphasized by Fig. 5.4. Note that the axes have a logarithmic scale. The relative error on the ZPLE with respect to FE minimization is under 5% for all but 6 of our 22 materials (see the shaded gray area) and below 3% for more than half of them. We stress that this error is relative to an already very small quantity; the relative error with respect to the renormalized lattice parameter, i.e. $a^0 + \Delta a(T = 0)$, would be significantly smaller. The absolute errors for

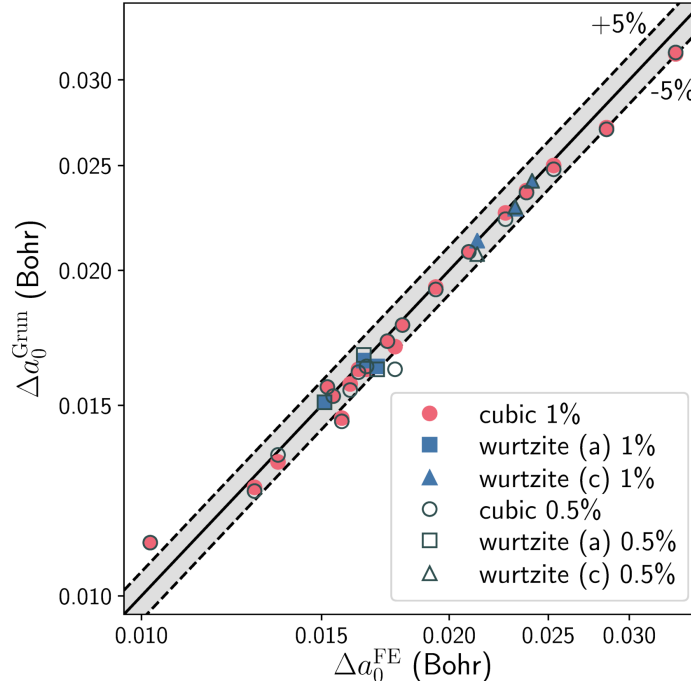


Figure 5.4 – Zero-point lattice expansion from FE minimization and Grüneisen formalism. The finite differences for mode Grüneisen parameters were computed with a volumic variation of 1% (solid colors) and 0.5% (empty gray markers). Note the use of logarithmic scales. The shaded gray area corresponds to the region where both approaches agree within 5% of each other. The ZPLE in the Grüneisen approach tends to be slightly smaller than with the FE minimization. Most materials show a better agreement between both methodologies when using a 1% variation.

all materials are detailed in Tables S5.2 and S5.3 of the Supplemental Material [55]. We also find that the agreement between both methodologies tends to decrease for materials with a larger bulk modulus, i.e. BN, MgO and C. The exceptions to this trend are CdS, as well as GaAs and Ge. For the two latter materials, we question the capability of the PBE functional to properly describe the electronic structure, as was suggested in Ref. [71]. We nevertheless restrain our analysis to this functional for consistency with the other materials in our dataset.

Two other observations can be drawn from Tables 5.1 and 5.2. On the one hand, we can see that the 1% volumic variation for the Grüneisen ZPLE improves the overall agreement with the FE ZPLE compared to the 0.5% variation. This can be counter-intuitive as one typically aims for smaller variable sampling when approximating derivatives by finite differences. On the other hand, we notice that the Grüneisen formalism tends to predict a slightly smaller ZPLE compared to the FE minimization. The only two exceptions to this trend are Ge and GaP. In that case, this behavior could be attributed to our choice of the experimental lattice parameter as a static reference, which required us to add an artificial pressure to the FE in order to compute the lattice expansion, as explained in Sec. 5.3.1.

Table 5.3 – Statistical errors on the ZPLE for the Grüneisen approach. The free energy minimization ZPLE results are taken as reference value. The mean absolute error (MAE), root-mean-square error (RMSE), and mean absolute relative error (MARE) for both the linear expansion and the resulting volumic expansion are displayed for the cubic materials, wurtzite materials and for the entire set.

Set	Grüneisen 1%			Grüneisen 0.5%		
	MAE (10^{-3} Bohr)	RMSE (10^{-3} Bohr)	MARE (%)	MAE (10^{-3} Bohr)	RMSE (10^{-3} Bohr)	MARE (%)
Cubic (a)	0.5	0.7	2.6	0.6	0.8	3.0
Wurtzite (a)	0.2	0.4	1.4	0.3	0.5	2.0
Wurtzite (c)	0.2	0.2	0.7	0.3	0.4	1.5
Wurtzite (a, c)	0.2	0.3	1.0	0.3	0.4	1.7
Full set (linear)	0.4	0.6	2.2	0.5	0.7	2.7

Set	Grüneisen 1%			Grüneisen 0.5%		
	MAE (Bohr ³)	RMSE (Bohr ³)	MARE (%)	MAE (Bohr ³)	RMSE (Bohr ³)	MARE (%)
Cubic (V)	0.134	0.176	2.6	0.157	0.214	3.0
Wurtzite (V)	0.030	0.041	1.3	0.032	0.044	1.4
Full set (V)	0.120	0.164	2.4	0.140	0.199	2.8

The mean absolute error (MAE), root mean square error (RMSE) and mean absolute relative error (MARE) for the Grüneisen approach are shown in Table 5.3, taking the FE results as a reference. Detailed results for the cubic and wurtzite materials are also provided. The statistical analysis for wurtzites (a, c) and the full set (linear) considers each independent lattice parameter as a single entry. Without surprise, the statistical errors decrease for the 1% variation compared to the 0.5% variation, in agreement with our previous observation. The cubic materials show the largest deviations in the set, followed respectively by the wurtzite a and c parameters. This could, however, be biased by the small proportion of wurtzite materials in our set compared to cubic ones. For the linear ZPLE ($\Delta a(T = 0)$), we obtain a MAE of 0.4 mBohr, a RMSE of 0.6 mBohr and a MARE of 2.2% for our entire dataset. This yields in a MRE of 0.120 Bohr³, RMSE of 0.164 Bohr³ and MARE of 2.4% for the zero-point volume expansion ($\Delta V(T = 0)$). Should we exclude Ge and GaAs from the statistical analysis for the reason discussed earlier, the MARE significantly drops by half a percent in each case, to 1.7% in the linear case and to 1.9% in the volumic case.

To fully appreciate these results, one has to recall that the derivation of the Grüneisen ZPLE relies on the lowest order approximation of the quasi-harmonic FE (see Sec. 5.2.2.2). In this approach, the ZPLE is described by the variation of the phonon frequencies with respect to the lattice parameters, truncated to first order. The absence of higher-order terms in the

FE expansion is most likely responsible for the small discrepancies between both methods. Nevertheless, our results clearly show that the first-order Grüneisen formalism can accurately account for the ZPLE, both for isotropic and anisotropic materials.

One should finally appreciate that, for low-symmetry materials, the FE can become up to a six-dimensional hypersurface if there are no symmetry constraints that prevent temperature-induced shear strains of the unit cell, not even counting the internal degrees of freedom, as detailed earlier. This will rapidly increase the minimal number of strained configurations that must be explicitly evaluated. While some promising methodologies have recently been developed to circumvent this difficulty [18, 31, 32, 36], a proper sampling of the FE surface can rapidly become a computational bottleneck. As pointed out for the wurtzite materials, one could try circumventing the explicit sampling of the multidimensional FE by minimizing a volume-dependent FE curve and relaxing both the internal coordinates and lattice parameters under a fixed volume constraint. Whether such an algorithm would correctly capture strong elastic anisotropy remains to be investigated. The Grüneisen formalism, on the other hand, requires only $n + 1$ configurations per degree of freedom, where n is the desired accuracy level chosen for the central finite difference. Hence, the lower the symmetry, the greater the reduction of the numerical cost. For example, computing the ZPLE for the wurtzite materials, which have only two degrees of freedom, required minimally more than 20 configurations for the 2D FE surface, while the Grüneisen approach provided the same result from only five configurations. Within our study we estimate the net gain to be about a factor of 3 for the cubic materials and a factor of at least 5 for the wurtzites.

5.4.3. ZPLE renormalization of the band gap energy

5.4.3.1. Free energy minimization vs Grüneisen approach. The ZPLE contribution to the band gap ZPR, ZPR_g^{ZPLE} , is reported in Table 5.4. To avoid any confusion with the full band gap ZPR or EPI-induced ZPR when we need to specify the calculation method, we denote it as ΔE_g . Note that for AlSb, the renormalization was computed for two conduction valleys: the L point is the theoretical CBM extracted from the static DFT calculation, and the minimum along the $\Delta = \Gamma - X$ high symmetry path corresponds to the experimental location of the fundamental band gap. From the two rightmost columns, we find that the choice of volume sampling for the finite differences used in the Grüneisen approach shows a negligible impact on the final value of ZPR_g^{ZPLE} . The 1% and 0.5% volumic variation results agree with each other within meV accuracy, which is of the same order as the precision typically reported for ZPR corrections. This is shown on Fig. 5.5, where the open and solid markers coincide almost perfectly. Note the axes' logarithmic scale again.

Making the comparison to FE results, we find again an agreement of about 1 meV, except for MgO (≈ 5 meV) and AlN (≈ 3 meV), as well as for GaAs (≈ 3 meV). In the former

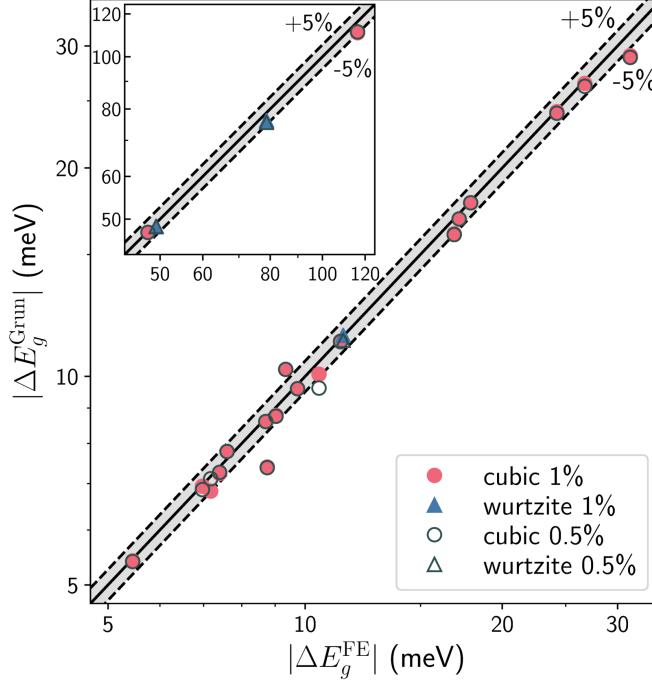


Figure 5.5 – Band gap ZPR induced by zero-point lattice expansion from FE minimisation and Grüneisen formalism. The finite differences for mode Grüneisen parameters were computed with volumic variation of 1% (solid colors) and 0.5% (empty gray markers). Our results show that for most materials, the difference between the two sampling sets is below meV accuracy, hence has no significant impact on ZPR_g^{ZPLE} . Note the use of logarithmic scales. The shaded gray area corresponds to the region where both approaches agree within 5% of each other. Materials whose ZPR_g^{ZPLE} is over 40 meV are displayed in the inset to enhance clarity.

case, these materials display by far the largest ZPR_g^{ZPLE} , such that the relative error remains coherent with the rest of the set. These three materials are also among those for which the absolute ZPLE discrepancy between FE and Grüneisen was among the largest. Just as for the ZPLE, the Grüneisen approach provides a very reliable value for ZPR_g^{ZPLE} , slightly smaller than the FE minimization. This can be understood as the numerical differences observed in Tables 5.1 and 5.2 between the two methods are in general on the order of the numerical precision typically reported in the literature for lattice parameters. Ge and GaP are again exceptions to this trend due to the small overestimation of the ZPLE by the Grüneisen method.

If we rather focus on the sign of ZPR_g^{ZPLE} , we can notice some general trends for the different band gap locations in the BZ. Indeed, all materials displaying a direct fundamental band gap find their gap energy decreased by ZPLE. In indirect band gap materials, the gap energy also decreases for CBM located at the L point, while it increases for gaps located along the $\Gamma - X$ high-symmetry path, diamond and BN being the only exceptions. To gain some insight into this trend, recall that increasing pressure, thus decreasing the lattice

Table 5.4 – Zero-point lattice expansion renormalization of the band gap. The VBM and CBM locations of the fundamental band gap in the BZ are specified in the second column. The band gap renormalization was computed from static DFT using the ZPLE-renormalized lattice parameters obtained from FE minimization (ΔE_g^{FE}) and with the Grüneisen formalism (ΔE_g^{Grun}), shown in Tables 5.1 and 5.2.

Material	Gap	ΔE_g^{FE} (meV)	$\Delta E_g^{\text{Grun 1\%}}$ (meV)	$\Delta E_g^{\text{Grun 0.5\%}}$ (meV)
C-dia	Γ - Δ	-26.74	-26.48	-26.23
Si-dia	Γ - Δ	8.71	8.60	8.61
Ge-dia	Γ -L	-9.34	-10.23	-10.24
SiC-zb	Γ -X	6.97	6.94	6.86
BN-zb	Γ -X	-16.89	-16.07	-16.02
BAs	Γ - Δ	8.75	7.40	7.38
AlP-zb	Γ -X	9.74	9.60	9.60
AlAs-zb	Γ -X	7.40	7.24	7.27
AlSb-zb	Γ -L	-11.33	-11.25	-11.22
AlSb-zb	Γ - Δ	5.45	5.41	5.40
GaN-zb	Γ - Γ	-47.49	-47.22	-47.25
GaP-zb	Γ - Δ	7.60	7.79	7.79
GaAs-zb	Γ - Γ	-31.37	-29.06	-28.88
ZnS-zb	Γ - Γ	-24.23	-24.11	-23.99
ZnSe-zb	Γ - Γ	-17.19	-16.88	-16.88
ZnTe-zb	Γ - Γ	-17.90	-17.81	-17.81
CdS-zb	Γ - Γ	-10.50	-10.07	-9.62
CdSe-zb	Γ - Γ	-7.17	-6.83	-7.12
CdTe-zb	Γ - Γ	-9.03	-8.78	-8.78
MgO-rs	Γ - Γ	-116.16	-110.75	-111.17
AlN-w	Γ - Γ	-78.77	-76.13	-75.55
GaN-w	Γ - Γ	-49.19	-48.42	-48.49
ZnO-w	Γ - Γ	-11.43	-11.46	-11.28

parameters, shifts the electronic bands to higher absolute energies due to the kinetic energy cost of bringing the ions closer to one another. Since ZPLE is linked to lattice expansion, hence a negative external pressure, the sign of the band gap correction will therefore be determined by how relatively fast the VBM and CBM rise in energy with increased external pressure. We thus recover the $-\partial E_g/\partial P$ dependency derived in Eq. (5.2.37), yielding a negative ZPR_g^{ZPLE} for direct band gaps at Γ . For indirect band gaps, this pressure-induced energy shift can be partly or fully counteracted by the phase factor entering the Bloch states at finite \mathbf{k} -points. The sign difference observed between diamond and silicon, which have the

Table 5.5 – Statistical errors on the zero-point lattice expansion renormalization of the band gap. The FE minimization results are taken as reference value (ΔE_g^{FE} from Table 5.4). The Grüneisen approximation with resp. $\pm 1\%$ and $\pm 0.5\%$ volumic variation were used. The mean absolute error (MAE), root-mean-square error (RMSE), and mean absolute relative error (MARE) are displayed for the cubic materials, wurtzite materials and for the entire set.

Set	Grüneisen 1%			Grüneisen 0.5%		
	MAE (meV)	RMSE (meV)	MARE (%)	MAE (meV)	RMSE (meV)	MARE (%)
Cubic	0.68	1.39	3.3	0.70	1.34	3.5
Wurtzite	1.14	1.59	1.7	1.36	1.91	2.3
Full set	0.74	1.42	3.1	0.79	1.43	3.4

same lattice structure and whose conduction band edge lie on similar high symmetry lines, has been explained[88] in terms of the presence or absence of d states with the same principal quantum number as the p -like valence states and the s -like conduction states, which drives a sign change in the pressure dependence of the band gap energy, hence, from Eq. (37), in the ZPR_g^{ZPLE} . In contrast with silicon (and with the zincblende materials whose CBM is at or close to the X point), whose band edges have principal quantum number $n = 3$, diamond (as well as zincblende BN) has $n = 2$, for which no unoccupied d states exist. See Table VI for the values of $\partial E_g/\partial P$.

The statistical errors for ZPR_g^{ZPLE} are displayed in Table 5.5. In agreement with the errors reported in Table 5.3 for the ZPLE, we find that the MARE is slightly larger for the cubic materials compared to the wurtzites. However, the larger MAE and RMSE observed for the wurtzites can be attributed to their larger ZPR values compared to the cubic set. For the entire dataset, we obtain a MAE of 0.74 meV, RMSE of 1.42 meV and MARE of 3.1%, thus confirming the reliability of the Grüneisen method for predicting ZPR_g^{ZPLE} within the QHA.

5.4.3.2. Comparison to the empirical expression. To assess the accuracy of the empirical expression of Eq. (5.2.37), we evaluate $\Delta V/V^0$ ($T = 0$ K) in 3 different ways. First, we make use of the numerical results presented in Tables 5.1 and 5.2 (ΔE_g^{data}) for 1% Grüneisen parameters. Second, we theoretically evaluate $\Delta a(T)/a^0$ from Eq. (5.2.27) and (5.2.29) using 1% Grüneisen parameters, and compute the intersect at $T = 0$ K of a linear fit in the high temperature regime. We favor the Grüneisen approach for this purpose as it delivers a smooth temperature dependence for the lattice parameters up to arbitrary high temperatures without the need for additional calculations, in contrast with the FE minimization method, for which our current volume sampling does not explicitly capture the FE minima at very high temperatures. The third method evaluates the $T = 0$ K intersect of

Table 5.6 – Estimation of the zero-point lattice expansion renormalization of the band gap, using Eq. (5.2.37). The band gap edges, bulk modulus extracted from elastic constants and pressure derivative of the band gap used to compute the estimates are provided for the different materials. The $\Delta V(T=0)/V^0$ term was evaluated in three different ways: from the numerical data presented in Tables 5.1 and 5.2 for 1% Grüneisen parameters (ΔE_g^{data}), from the extrapolation to $T=0$ K of the high- T slope of our theoretical calculation of $\Delta a(T)/a^0$ obtained from Eq. (5.2.27) and (5.2.29) using 1% Grüneisen parameters, and from the same extrapolation technique applied to available experimental data (ΔE_g^{exp} , see text for details). Two estimations are presented using the theoretical high- T slope method: the first takes the slope between $T \approx 2\theta_D$ and $T \approx 3\theta_D$ (except for diamond, SiC and BN, see text) ($\Delta E_g^{T>2\theta_D}$), while the second takes the slope between 700 and 1000 K ($\Delta E_g^{700-1000\text{K}}$), regardless of the Debye temperature. This emphasizes that the accuracy of the estimation strongly relies on the choice of the “high- T regime”. $ZPR_g^{\text{Grun } 1\%}$ (from Table 5.4) is taken as reference to assess the accuracy of the estimation. The numerical values of $\Delta a_k(T=0)/a_k^0$ used to obtain the different estimates are made explicit in Table S5.4 of the Supplemental Material [55].

Material	Gap	θ_D [7] (K)	B_0^{elast} (GPa)	$\partial E_g/\partial P$ (meV/GPa)	$\Delta E_g^{\text{Grun } 1\%}$ (meV)	ΔE_g^{data} (meV)	$\Delta E_g^{T>2\theta_D}$ (meV)	$\Delta E_g^{700-1000\text{K}}$ (meV)	ΔE_g^{exp} (meV)
C-dia	Γ - Δ	2220	431	5.5	-26.48	-26.15	-22.39	-11.78	-27.06 [50]
Si-dia	Γ - Δ	636	88	-19.7	8.61	8.67	7.43	6.43	9.72 [50]
Ge-dia	Γ -L	374	87	34.1	-10.24	-9.44	-8.04	-7.88	-12.76 [50]
SiC-zb	Γ -X	1270	211	-4.0	6.94	7.01	6.14	4.24	6.95 [89]
BN-zb	Γ -X	1730	403	3.3	-16.07	-15.81	-13.23	-7.84	
BAs	Γ - Δ	800	131	-7.2	7.40	7.35	6.26	4.95	3.56 [90]
AlP-zb	Γ -X	588	82	-22.9	9.60	9.69	8.36	7.39	
AlAs-zb	Γ -X	414	67	-24.2	7.24	7.33	6.20	5.95	
AlSb-zb	Γ -L	292	49	55.5	-11.25	-11.29	-9.14	-9.42	
AlSb-zb	Γ - Δ	292	49	-26.8	5.42	5.44	4.41	4.55	
GaN-zb	Γ - Γ	608	172	37.9	-47.22	-47.34	-38.90	-32.91	
GaP-zb	Γ - Δ	445	86	-25.5	7.79	9.98	8.46	8.00	11.40 [89]
GaAs-zb	Γ - Γ	344	61	118.8	-29.06	-29.16	-25.48	-25.30	
ZnS-zb	Γ - Γ	352	70	61.7	-24.11	-24.22	-20.48	-20.31	
ZnSe-zb	Γ - Γ	339	57	68.7	-16.88	-16.93	15.04	-14.98	
ZnTe-zb	Γ - Γ	180	43	105.0	-17.81	-17.78	-14.75	-16.08	
CdS-zb	Γ - Γ	219	53	41.6	-10.07	-10.06	-7.89	-8.57	
CdSe-zb	Γ - Γ	182	45	46.3	-6.83	-7.05	-5.86	-6.35	
CdTe-zb	Γ - Γ	158	35	83.5	-8.78	-8.83	-7.36	-8.12	
MgO-rs	Γ - Γ	948	149	63.5	-110.75	-112.11	-102.50	-86.49	-110.60 [91, 92]
AlN-w	Γ - Γ	950	208	42.4	-76.13	-72.87	-63.02	-46.34	-70.44 [93]
GaN-w	Γ - Γ	600	172	39.2	-48.42	-49.04	-37.77	-34.11	-35.73 [94]
ZnO-w	Γ - Γ	440	130	12.6	-11.46	-12.25	-9.88	-9.22	-15.80 [87, 95]

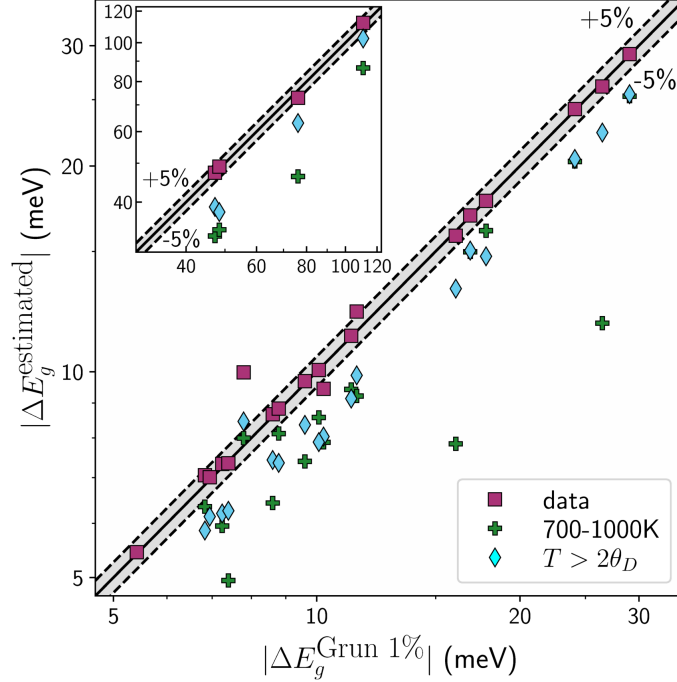


Figure 5.6 – Estimation of the ZPLE-induced band gap renormalization from the empirical expression, Eq. (5.2.37), using $\Delta a(T=0)/a^0$ as computed from 1% Grüneisen parameters (purple squares), and extracted from the $T=0$ K intercept of a linear fit $\Delta a(T)/a^0$ in the high temperature regime with $T \in [700, 1000]$ K (green plus) and $T \in [2\theta_D, 3\theta_D]$ (cyan diamonds). Note that the scales are logarithmic. Numerical results from Tables 5.1 and 5.2 are taken as reference values.

Table 5.7 – Statistical errors on the estimation of the ZPLE-induced band gap ZPR from Eq. (5.2.37), using the three different methods explained in the caption of Table 5.6. For consistency, the ZPLE-induced band gap ZPR obtained with 1% Grüneisen parameters are taken as reference value ($\Delta E_g^{\text{Grün } 1\%}$ from Table 5.4, see text). The mean absolute error (MAE), root-mean-square error (RMSE), and mean absolute relative error (MARE) are displayed for the cubic materials, wurtzite materials and for the entire set. ΔE_g^{data} quantifies the accuracy of the estimation obtained from Eq. (5.2.37), while $\Delta E_g^{T > 2\theta_D}$ and $\Delta E_g^{700-1000\text{K}}$ reveal its sensitivity to the choice of temperature range used for the extrapolation of $\Delta a(T=0)/a^0$.

Set	ΔE_g^{data}			$\Delta E_g^{T > 2\theta_D}$			$\Delta E_g^{700-1000\text{K}}$		
	MAE (meV)	RMSE (meV)	MARE (%)	MAE (meV)	RMSE (meV)	MARE (%)	MAE (meV)	RMSE (meV)	MARE (%)
Cubic	0.30	0.62	2.5	2.58	3.36	15.1	4.57	7.58	21.7
Wurtzite	1.56	1.97	4.2	8.45	9.80	16.7	15.45	19.13	29.4
Full set	0.47	0.91	2.7	3.34	4.72	15.4	5.99	9.88	22.7

the high- T slope from available experimental data (ΔE_g^{exp}). Details about our treatment of experimental data can be found in Sec. S5.4 of the Supplemental Material [55]. We finally investigate the sensitivity of the estimation to the choice of temperature regime for the linear fit by considering two different ranges for the second method. The first one is material-specific and covers $T \in [2\theta_D, 3\theta_D]$ ($\Delta E_g^{T>2\theta_D}$), while the second covers $T \in [700, 1000]$ K regardless of the Debye temperature ($\Delta E_g^{700-1000\text{K}}$). For diamond, Si and BN, we set the first range to be respectively $T \in [1.5\theta_D, 1.8\theta_D]$, $T \in [2\theta_D, 2.4\theta_D]$ and $T \in [1.5\theta_D, 1.8\theta_D]$, as otherwise, it would go beyond the melting temperature.

The estimation of $\text{ZPR}_g^{\text{ZPLE}}$ from the different methods, as well as relevant parameters used to obtain them, are presented in Table 5.6 and Fig. 5.6. The estimates are compared with the DFT-computed $\Delta E_g^{\text{Grun 1\%}}$ previously reported in Table 5.4. Statistical errors are presented in Table 5.7. Note that all estimations rely on the same values of B_0 and $\partial E_g/\partial P$; hence, the numerical differences are solely explained by variations of $\Delta a/a^0$. Without surprise, ΔE_g^{data} (purple squares) provides the most reliable estimate, as it makes use of the same value of $\Delta a/a^0$ computed with the first-principles approach as the reference $\Delta E_g^{\text{Grun 1\%}}$, without relying on any extrapolation procedure. These results are thus a direct assessment of the accuracy of Eq. (5.2.37). For the cubic materials, most materials display a remarkable numerical agreement below 0.5 meV (except, again, GaP, Ge and MgO, for which the agreement is still nevertheless excellent), yielding a MARE of 2.5%. The wurtzites materials show a slightly higher MARE of 4.2%, which can most likely be attributed to AlN. In all cases, we obtain a global MARE of 2.7% for the whole dataset, thus validating a posteriori Eq. (5.2.37).

We now shift our focus on $\Delta E_g^{T>2\theta_D}$ (cyan diamonds) and $\Delta E_g^{700-1000\text{K}}$ (green plus), which both rely on the extrapolation of the linearized high- T solution to $T = 0$ K from our first-principles Grüneisen results. In contrary to the previous assessment, both results display important discrepancies compared to the reference values. The statistical errors increase accordingly, yielding a MARE of resp. 15.4% and 22.7%. We can group the general trends into three categories. First, for materials with a high Debye temperature ($\theta_D \gtrsim 500$ K), choosing a fixed, arbitrary temperature range ($\Delta E_g^{700-1000\text{K}}$) strongly underestimates the ZPLE-induced band gap ZPR, while the material-dependent approach ($\Delta E_g^{T>2\theta_D}$) visibly improves the accuracy of the estimation. The difference is even more striking for materials with $\theta_D > 1000$ K. In that case, the chosen fixed range does not cover a region where the classical regime approximation is physically valid, hence its incapacity to deliver a reliable estimate. A second category regroups materials whose Debye temperature is below room temperature. In that case, the fixed 700-1000 K range delivers a better estimate than the material-specific approach. This observation may seem surprising at first, but can be understood by recalling that the linearized behavior in the high- T regime originates by construction from a first-order approximation of the classical Maxwell-Boltzmann distribution function,

which may not be entirely accurate at intermediate temperatures (below 700 K), despite the material being in principle properly described by the classical picture. The third category re-groups the few materials with $\theta_D \in [300, 500]$ K, for which the apparent agreement between the two methods is simply a consequence of their respective temperature range being almost identical. The rightmost column of Table 5.6 computes the estimation of ZPR_g^{ZPLE} from available experimental data. From our previous observations, we stress that such an analysis requires the lattice parameter to be measured in a broad range of temperature, namely, both sufficiently above the Debye temperature, to ensure that the linear fit is indeed in the true high- T regime, and at low enough temperatures, to evaluate the experimental $T = 0$ K value correctly. Our results nevertheless allow us to conclude that, when the experimental data range is appropriate for this type of analysis, Eq. (5.2.37) provides a more than reliable reference for ZPR_g^{ZPLE} . In most cases, our numerical results are in very good agreement with the experimental estimation. They always minimally show the correct order of magnitude, thus validating our first-principles approach for computing the ZPLE correction to the band gap.

5.4.3.3. Comparison of the ZPLE and EPI contributions to the band gap ZPR.

We finally compare the respective contributions of ZPLE and EPI to the total band gap ZPR in Table 5.8. Throughout the literature, the EPI contribution is commonly recognized as the leading mechanism driving the zero-point renormalization of the electronic bands, especially for materials containing light atoms, like B, C, N and O. To assess the importance of the ZPLE contribution, we also evaluate the relative variation of ZPR_g in the full QHA picture, with respect to EPI contribution only: $R^{\text{ZPLE/EPI}} = (ZPR_g^{\text{tot}} - ZPR_g^{\text{EPI}})/ZPR_g^{\text{EPI}} = ZPR_g^{\text{ZPLE}}/ZPR_g^{\text{EPI}}$. In this section, we only consider the FE methodology for the ZPLE contribution. As the results presented in Sec. 5.4.3 showed an excellent agreement between the FE and Grüneisen approaches, the following observations will remain qualitatively valid should one rather consider the Grüneisen ZPLE.

From the second leftmost column, we can see that EPI interaction reduces the band gap energy for all the materials considered. This is a typical behavior for semiconductors, known as the Varshni effect [96]. As discussed earlier, the ZPLE contribution either closes or opens the band gap, depending on the band edge location in the BZ. As the magnitude of the ZPR_g^{ZPLE} is smaller than its EPI counterpart, the resulting total band gap ZPR remains negative for all materials. The rightmost column shows the relative variation of the band gap ZPR induced by ZPLE. Note that since the total band gap ZPR is negative, a positive variation indicates that the EPI and ZPLE contributions reinforce each other; namely, the ZPR becomes more negative. A negative variation consequently signifies that the ZPLE contribution competes with EPI and thus reduces its impact on the total band gap ZPR.

Table 5.8 – Zero-point renormalization of the band gap. The ZPLE contribution to the band gap from FE minimization (Table 5.4) is rounded to 1 meV. The rightmost column gives the relative variation of the band gap ZPR when considering the combined effects of ZPLE and EPI compared to EPI contribution only: $R^{\text{ZPLE/EPI}} = \text{ZPR}_g^{\text{ZPLE}}/\text{ZPR}_g^{\text{EPI}} = (\text{ZPR}_g^{\text{tot}} - \text{ZPR}_g^{\text{EPI}})/\text{ZPR}_g^{\text{EPI}}$. As $\text{ZPR}_g^{\text{EPI}}$ is negative, i.e. decrease of the band gap energy, a positive variation indicates that the total ZPR has a larger magnitude, hence is “more negative”, while a negative variation indicates a decrease in magnitude, as it becomes “less negative”.

Material	$\text{ZPR}_g^{\text{EPI}}$ (meV) [71]	ZPR_g^{FE} (meV)	$\text{ZPR}_g^{\text{tot}}$ (meV)	$R^{\text{ZPLE/EPI}}$ (%)
C-dia	-330	-27	-357	+ 8.2
Si-dia	-56	9	-47	-16.1
Ge-dia	-33	-9	-42	+ 27.3
SiC-zb	-179	7	-172	-3.9
BN-zb	-406	-17	-423	+ 4.2
BAs	-97	9	-88	-9.3
AlP-zb	-93	10	-83	-10.8
AlAs-zb	-74	7	-67	-9.5
AlSb-zb (L)	-43	-11	-54	+ 25.6
AlSb-zb (Δ)	-51	5	-46	-9.8
GaN-zb	-176	-47	-223	+ 26.7
GaP-zb	-65	8	-57	-12.3
GaAs-zb	-21	-31	-52	+147.6
ZnS-zb	-88	-24	-112	+ 27.3
ZnSe-zb	-44	-17	-61	+ 38.6
ZnTe-zb	-22	-18	-40	+ 81.8
CdS-zb	-70	-10	-80	+ 14.3
CdSe-zb	-34	-7	-41	+ 20.6
CdTe-zb	-20	-9	-29	+ 45.0
MgO-rs	-524	-116	-640	+ 22.1
AlN-w	-399	-79	-478	+ 19.8
GaN-w	-189	-49	-238	+ 25.9
ZnO-w	-157	-11	-168	+ 7.0

If we rather consider the magnitude of the relative variation, $R^{\text{ZPLE/EPI}}$, we find that, out of our 22 materials, 7 of them see their total ZPR_g modified by up to 10% by the inclusion of ZPLE (C, AlSb, AlAs, SiC, BN, BAs and ZnO), 5 are modified by 10-20% (Si, GaP, AlP, CdS and AlN), and 10 of them, hence almost half of our set, see their ZPR affected by 20% or more (Ge, GaAs, CdTe, CdSe, ZnTe, ZnSe, ZnS, cubic and wurtzite GaN and MgO). For the two tellurides, which are the heaviest materials in the set, $\text{ZPR}_g^{\text{ZPLE}}$ is about half the size

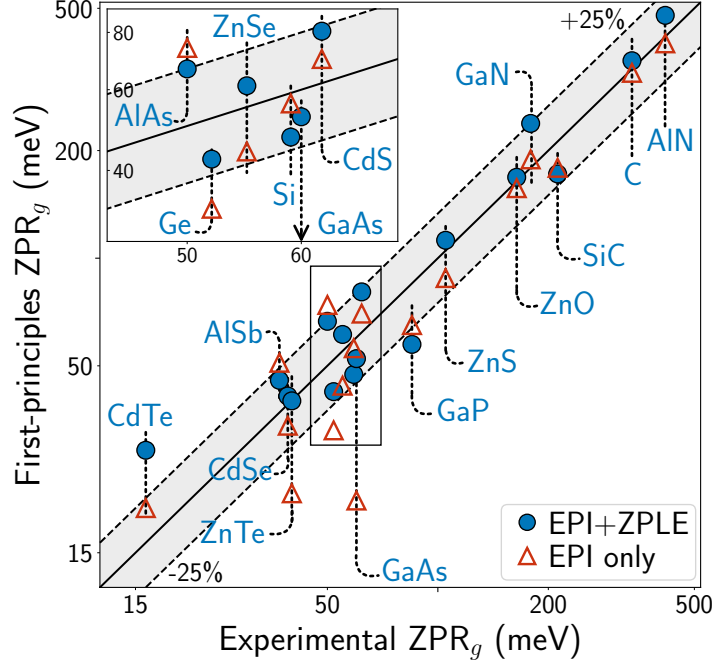


Figure 5.7 – Comparison of the predicted first-principles band gap ZPR to experimental values, from EPI only (empty red triangles) and from the total EPI+ZPLE contributions (blue circles). The dashed lines and shaded gray area highlight the region where both values agree within 25% of each other (see text). Note that the scales are logarithmic. When the ZPLE contribution is properly accounted for, more materials lie in this target region compared to when it is neglected. Detailed numerical values can be found in Table S5.5 of the Supplemental Material [55].

of ZPR_g^{EPI} or more, which is all but negligible. Moreover, three materials containing N and O see their ZPR_g increased by 20%-25%, which cautions about disregarding ZPLE effects even in materials containing light atoms.

We now briefly analyze the data in Table 5.8 in terms of the elastic response of the materials (B_0) and their degree of ionicity. For this purpose, the three topmost sections of Table 5.8 respectively separate group IV, III-V and II-VI cubic materials. While no obvious quantitative trends can be extracted, we can nevertheless gain some qualitative insight into the relative strength of ZPR_g^{ZPLE} for the different types of materials. First, despite the three exceptions mentioned earlier, materials with light atoms and high B_0 tend to be less affected by the inclusion of ZPLE, most likely thanks to their very large ZPR_g^{EPI} (C, BN, SiC). As a corollary, heavier materials typically show smaller values of ZPR_g^{EPI} , hence are more strongly affected by ZPR_g^{ZPLE} (see, for example, the ZnX and CdX families, where $X=\{\text{S, Se, Te}\}$). Second, a higher degree of ionicity seems to be associated with a larger impact of ZPR_g^{ZPLE} on the total band gap ZPR (i.e. II-VI materials compared to III-V materials). There are, however, outliers to both of these observations. Hence, a more thorough analysis of the effect of ZPLE on the EPI might prove useful to explain our results in a more meaningful way.

Table 5.9 – Effect of exchange-correlation functional on the zero-point lattice expansion and zero-point renormalization of the band gap. The fractional ZPLE, as well as the EPI and ZPLE contributions to the band gap ZPR, are compared for three semilocal XC functionals. The ratios $R^{\text{ZPLE/EPI}}$ give the relative variation of the band gap ZPR when considering the combined effects of ZPLE and EPI compared to EPI contribution only. See the caption of Table 5.8 for the significance of the positive or negative sign. Despite their respective static equilibrium lattice parameters varying by more than the typical ZPLE correction, all three XC functionals yield ratios in good agreement with each other, thus validating the ratios presented in Table 5.8 for all materials.

Material	a^0 (Bohr)			$\Delta a^{\text{FE}}/a^0$ (%)			ZPR_g^{FE} (meV)		
	PBE	PBESOL	LDA	PBE	PBESOL	LDA	PBE	PBESOL	LDA
GaAs-zb	10.8627	10.6994	10.5861	0.145	0.127	0.128	-31	-28	-29
ZnS-zb	10.2859	10.1196	10.0064	0.189	0.180	0.176	-24	-23	-24
ZnSe-zb	10.8330	10.6437	10.5192	0.148	0.136	0.137	-17	-16	-17
ZnTe-zb	11.6819	11.4675	11.3316	0.132	0.122	0.121	-18	-17	-18
MgO-rs	8.0366	7.9629	7.8615	0.414	0.412	0.394	-116	-118	-120
Material	$\text{ZPR}_g^{\text{EPI}}$ (meV)			$\text{ZPR}_g^{\text{tot}}$ (meV)			$R^{\text{ZPLE/EPI}}$ (%)		
	PBE	PBESOL	LDA	PBE	PBESOL	LDA	PBE	PBESOL	LDA
GaAs-zb	-21	-24	-24	-52	-52	-53	+ 147.6	+ 116.7	+ 120.8
ZnS-zb	-88	-82	-78	-112	-105	-102	+ 27.3	+ 28.0	+ 30.1
ZnSe-zb	-44	-41	-40	-61	-57	-57	+ 38.6	+ 39.0	+ 42.5
ZnTe-zb	-22	-21	-22	-40	-38	-40	+ 81.8	+ 81.0	+ 81.8
MgO-rs	-524	-511	-526	-640	-629	-646	+ 22.1	+ 23.1	+ 22.8

In Fig. 5.7, we compare the accuracy of the predicted ZPR_g to experimental values for 17 of our 22 materials, when either including (blue circles) or excluding (empty red triangles) the ZPLE contribution. All numerical values, including the reference experimental values, can be found in Table S5.5 of the Supplemental Material [55]. The dashed lines and shaded gray area highlight the region where both values agree within 25% of each other (namely, when the smallest of the two values is 25% smaller than the largest one, such that the gray area is symmetrical). As pointed out in the previous section for the fractional ZPLE, zero-point corrections are usually extracted from experimental data from extrapolation procedures, which yield an uncertainty which can be quite substantial. Hence, we believe that a 25% agreement is acceptable.

From Fig. 5.7, we note that the effect of the ZPLE contribution on the experimental agreement is somewhat mitigated, as about half of the materials (9 materials) show an improved accuracy, while it lessens for the other half (8 materials). Nevertheless, the ZPLE contribution has an overall positive effect on the global agreement of the entire dataset, as more materials enter (Ge, GaAs, AlAs, ZnTe) than leave (CdTe, GaP) the shaded area

when it is properly accounted for. For CdTe, the lack of spin-orbit coupling could explain the observed overestimation. Moreover, the relative agreement improves for most materials for which the ZPLE contribution is significant. Despite this, we stress that a decrease in accuracy for a given material does not justify purposely disregarding a real physical effect.

As the typical ZPLE correction is on the order of or smaller than the variation of the static lattice parameters between different exchange-correlation (XC) functionals, one can also wonder if the large ratios reported in Table 5.8, hence the experimental agreement discussed in Fig. 5.7, are simply an artifact of our choice of the PBE-GGA. See Ref. [97] for an investigation of the sensitivity of phonon frequencies to the XC functionals, including the effect of differing lattice constant. To verify the numerical stability of our results to the choice of functional, we computed the ZPR_g^{EPI} and ZPR_g^{ZPLE} contributions with the Perdew-Wang parametrization of the local density approximation (LDA) [82] and the PBE approximation for solids (PBESOL) [98] for five representative materials: the ZnX family (X=S, Se, Te), GaAs and MgO. These results are presented in Table 5.9. Despite their different respective static equilibrium lattice parameters, the fractional ZPLE remains on the same order for all three functionals. The slightly larger fractional ZPLE obtained from PBE can be understood in terms of the typical lattice parameter overestimation associated with this functional: as the material is slightly less rigid, the lattice response to the zero-point energy is marginally larger. Nevertheless, this difference has no noticeable effect on the resulting ZPR_g^{FE} . We also find that the ratio $R^{\text{ZPLE/EPI}}$ is not significantly affected by the choice of XC functional. It should be understood that these ratios are to be interpreted semi-quantitatively, as an indicator of the importance of the ZPR_g^{ZPLE} compared to the ZPR_g^{EPI} contribution. Hence, despite small variations (less than 5% except for GaAs) between the different XC functionals, Table 5.9 clearly establishes that the ratios presented in Table 5.8 are robust within semilocal DFT. In the case of GaAs, it indicates that the ZPLE contribution to the total ZPR_g is necessary to bridge the gap with experimental data, as it steadily contributes more than half of the total ZPR_g . A similar conclusion can also be drawn for ZnTe.

However, whether the relative strength of the ZPLE contribution compared to EPI reported in Table 5.8 maintains the same order when using higher-level XC functionals remains to be investigated. As a more exact treatment of electronic correlations has been shown to increase the strength of the ZPR_g^{EPI} in some materials [77, 99, 100], one could reasonably expect the ratios to decrease slightly. However, as Ref. [77] reported a 10 meV increase of the magnitude of the ZPR_g^{EPI} of GaAs when computing EPI with single-shot GW (from -23 to -33 meV), this suggests that electronic correlations alone are not sufficient to obtain experimental agreement for this material. The ZPLE contribution is thus most likely the missing ingredient.

Lastly, one fundamental assumption made throughout this work is that, following Eq. (5.2.4), the total band gap ZPR can be approximated by the individual contributions of

EPI and ZPLE. This supposes that the output of $\Sigma_{\mathbf{kn}}^{\text{Fan}}$ (Eq. (5.2.40)) and $\Sigma_{\mathbf{kn}}^{\text{DW}}$, which are computed at the static equilibrium lattice parameters, is not significantly affected by this small change of the lattice parameters, beyond what corresponds to the ZPLE effect on the band gap in the static DFT picture. In other words, it assumes that the QHA picture, which decorrelates both mechanisms and neglects the effect of anharmonicity in its description of EPI, is valid. We stress that, in principle, one should evaluate the temperature-dependent EPI self-energy at the temperature-dependent lattice parameter. While anharmonic effects can reasonably be expected to be small at absolute zero temperature, it remains to be investigated if computing the EPI self-energy at the ZPLE-renormalized lattice parameter instead of the bare lattice parameter, as was done for example in Ref. [40], would yield a total ZPR_g that differs from the EPI-only picture by a similar ratio, even when the ZPLE contribution is large compared to EPI.

5.4.4. Further discussion

At this point, we would like to bring up the fact that zero-point corrections, whether on the lattice parameter, on the band gap energy or on any other physical quantity, cannot be obtained from a direct nor an indirect measurement. By essence, they represent the difference between a true quantum behavior of the lattice at $T = 0$ K and a hypothetical lattice behavior, should the ions behave classically as $T \rightarrow 0$ K and remain at rest. Hence, one cannot escape applying extrapolation procedures to measured data. Assessing the accuracy of the current numerical methods for computing zero-point corrections is thus no straightforward task.

In this work, we addressed the ZPLE question from a pure first-principles point of view, taking the FE minimization within the QHA as a reference value. We have not attempted to assess the accuracy of this framework with respect to experiments. Many studies have, however, partially addressed this question by investigating the temperature dependence of the lattice constants or the thermal expansion coefficients (see, for example, Refs. [18, 22, 34, 101]). As these works focus on reproducing the correct shape of the T -dependence, the numerical results are conveniently rescaled to the $T = 0$ K measured values (such that, for example, the fractional ZPLE is null). In this picture, one can draw conclusions about the accuracy of numerical methods in different temperature regimes, but not about zero-point corrections.

Along the same line, another point brought to light by the results of Table 5.6 is the great sensitivity of Eq. (5.2.37) to the actual value of the fractional ZPLE, $\Delta a/a^0$. Naturally, they also depend on B_0 and $\partial E_g/\partial P$; still, it is quite unlikely that these well-studied quantities vary by more than a few percent (drastically) between experiments or numerical studies from static DFT. Hence, one must be particularly careful when using experimental

data to obtain an experimental value of the fractional ZPLE. Again, we do not question the accuracy of published experimental data. We rather want to point out that not all available data are suitable for extracting experimental fractional ZPLE. On this subject, we also refer to the Supplementary Material of Ref. [71] and to Ref. [102], which discuss experimental evaluations of the total band gap ZPR from temperature-dependent band gap energies. As pointed out by our results, it should be understood that, unless the experimental temperature range extends well above the Debye temperature *and* at sufficiently high temperatures, a fractional ZPLE extracted from a linearization and extrapolation procedure as described in Sec. 5.4.3.2 should be thought of as a lower bound for the true ZPLE. While experimental results for temperature-dependent lattice parameters are abundant in the literature, datasets that fit both of these requirements are not.

5.5. Conclusion

In this work, we investigated the zero-point lattice expansion within the quasi-harmonic approximation and its contribution to the total zero-point renormalization of the band gap using first-principles methodologies for a set of 22 benchmark semiconductors and insulators. We demonstrated that the formalism based on Grüneisen parameters, which commonly neglects zero-point corrections, can efficiently predict the ZPLE both for isotropic and anisotropic materials, at a lower computational cost than the traditional Helmholtz free energy minimization method.

The Grüneisen approach reproduces the results obtained from free energy minimization with a mean absolute relative error of 2.2% for the linear lattice expansion, which is typically of the order of 0.1-0.3% of the bare lattice parameter. The resulting ZPLE contribution to the band gap ZPR from both methodologies agrees within meV accuracy. Using our numerical results, we validated a posteriori an empirical expression that estimates the ZPLE contribution to the band gap ZPR from few simple parameters. The quality of this estimation is strongly sensitive to the temperature range chosen to extrapolate the fractional ZPLE, either from numerical or experimental data.

We finally assess the importance of the ZPLE contribution to the total band gap ZPR compared to the contribution of electron-phonon interaction. For heavier or more ionic materials, the inclusion of the ZPLE contribution typically modifies the predicted band gap ZPR by 20-80%, including some materials containing light atoms. These proportions are shown to be stable within semilocal DFT. For GaAs, ZPLE is shown to be the missing ingredient to obtain agreement with the experimental band gap ZPR value. Both renormalization mechanisms must therefore be considered on an equal footing to deliver predictive theoretical ZPR values, and should be systematically incorporated in the different methodologies used to evaluate the band gap ZPR.

Overall, our findings bring to light that the decades-old Grüneisen formalism can provide the community with a simple and accessible framework to incorporate zero-point lattice corrections in thermal expansion studies. Our results also indicate that a general understanding of the relative strength of the electron-phonon interaction and ZPLE contributions to the band gap ZPR in terms of simple material properties is still lacking. Investigating such a relationship from a broader perspective would further our fundamental understanding of zero-point quantum effects.

Acknowledgments

This research was financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), under the Discovery Grants program grant No. RGPIN-2016-06666, and by the Fonds de la Recherche Scientifique (FRS-FNRS Belgium) through the PdR Grant No. T.0103.19 - ALPS. This research was enabled in part by support provided by Calcul Québec (www.calculquebec.ca) and the Digital Research Alliance of Canada (www.alliancecan.ca). The operation of the supercomputers used for this research is funded by the Canada Foundation for Innovation (CFI), the Ministère de la Science, de l'Économie et de l'Innovation du Québec (MESI), and the Fonds de recherche du Québec – Nature et technologies (FRQ-NT). V.B.-C. acknowledges support by the NSERC Alexander Graham Bell Canada Graduate Scholarship doctoral program, the FRQ-NT B2 Doctoral Scholarship and the Hydro-Québec Excellence Scholarship. V.B.-C, É.G. and M.C. are members of the Regroupement québécois sur les matériaux de pointe (RQMP).

Supplemental Material

S5.1. Grüneisen formalism at finite external pressure

At finite external pressure P , one has to work with the Gibbs free energy (FE) of the combined electronic and lattice system which is approximated in the same fashion as described in the main text,

$$G(P, T) = \min_V \left(E_{\text{stat}}^e(V) + F^{\text{vib}}(V, T) + PV \right). \quad (\text{S5.1})$$

For fixed pressure P and temperature T , the equilibrium lattice configuration, $V_P \triangleq V_P(P, T)$, will verify

$$\left. \frac{\partial E_{\text{stat}}^e}{\partial V} \right|_{V=V_P} + \left. \frac{\partial F^{\text{vib}}}{\partial V} \right|_{V=V_P, T} + P = 0. \quad (\text{S5.2})$$

In the following, we generalize the results of Sec. 5.2.2.2 of the main text to finite external pressure for the volumic case only. Assuming that the internal coordinates can safely be approximated by those minimizing the Born-Oppenheimer static electronic energy at volume $V_P(P, T)$, the extension to anisotropic expansion is straightforward.

We first define the static equilibrium lattice configuration, V_P^0 , which minimizes the electronic Gibbs FE, G^e , at fixed external pressure P ,

$$G^e(P, T) = \min_V \left(E_{\text{stat}}^e(V) + PV \right). \quad (\text{S5.3})$$

giving

$$\left. \frac{\partial E_{\text{stat}}^e}{\partial V} \right|_{V=V_P^0} + P = 0. \quad (\text{S5.4})$$

In all these equations, we neglected the entropic electronic contribution to the free energy as we are dealing with semiconductors and insulators. From a computational point of view, P is the external constraint applied on the static lattice during the relaxation process. It can be expressed as minus the slope of the static electronic energy evaluated at the static equilibrium volume V_P^0 .

Coming back to the temperature-dependent case, in order to emphasize that P is a target pressure, we will now denote it as P_{target} . The equilibrium condition on the total Gibbs FE (Eq. (S5.2)) yields the relationship

$$P_{\text{target}} = - \left. \frac{\partial E_{\text{stat}}^e}{\partial V} \right|_V - \left. \frac{\partial F^{\text{vib}}(V, T)}{\partial V} \right|_{V, T}, \quad (\text{S5.5})$$

where P_{target} now contains the derivatives of both the electronic and phononic contributions. We can understand this equation as a search for the volume $V_P(P, T)$ for which the total Helmholtz FE at temperature T has the same slope as in the static case.

Since G^e is minimal at $V = V_P^0$, we Taylor-expand around this reference volume. The second-order expansion of the static electronic energy gives

$$E_{\text{stat}}^e(V) \approx E_{\text{stat}}^e(V_P^0) + \left. \frac{\partial E_{\text{stat}}^e}{\partial V} \right|_{V_P^0} (V - V_P^0) + \frac{1}{2} \left. \frac{\partial^2 E_{\text{stat}}^e}{\partial V^2} \right|_{V_P^0} (V - V_P^0)^2, \quad (\text{S5.6})$$

while the expansion of F^{vib} yields

$$F^{\text{vib}}(V, T) \approx F^{\text{vib}}(V_P^0, T) + \left. \frac{\partial F^{\text{vib}}}{\partial V} \right|_{V=V_P^0, T} (V - V_P^0), \quad (\text{S5.7})$$

where the partial derivative is identical to the second line of Eq. (5.2.12) of the main text, with V^0 replaced by the new reference volume V_P^0 . Substituting these expansions in Eq. (S5.5), we obtain

$$\begin{aligned} P_{\text{target}} &\approx - \frac{\partial}{\partial V} \left(- P_{\text{target}}(V - V_P^0) + \frac{1}{2} \left. \frac{\partial^2 E_{\text{stat}}^e}{\partial V^2} \right|_{V_P^0} (V - V_P^0)^2 - \left. \frac{\partial F^{\text{vib}}}{\partial V} \right|_{V=V_P^0, T} (V - V_P^0) \right) \\ &\approx P_{\text{target}} - \left. \frac{\partial^2 E^e}{\partial V^2} \right|_{V_P^0} (V - V_P^0) - \left. \frac{\partial F^{\text{vib}}}{\partial V} \right|_{V=V_P^0, T}. \end{aligned} \quad (\text{S5.8})$$

The second derivative in the first term on the right-hand side can be recast in terms of the static bulk modulus evaluated at constant pressure P_{target} :

$$\begin{aligned} \left. \frac{\partial^2 E_{\text{stat}}^e}{\partial V^2} \right|_{V_P^0} &= - \left. \frac{\partial P}{\partial V} \right|_{V_P^0} \\ &= \frac{B_0(P_{\text{target}})}{V_P^0}. \end{aligned} \quad (\text{S5.9})$$

Substituting Eq. (S5.9) in Eq. (S5.8) and cancelling out the P_{target} terms, we recover an expression analogous to Eq. (5.2.12) of the main text, in which the reference static configuration used to evaluate the equilibrium volume, bulk modulus, and phonon frequencies already captures the effect of finite external pressure on the lattice. The resulting expression for the ZPLE defined in the main text can therefore be applied to finite external pressure by using a strained reference configuration.

S5.2. Generalized Grüneisen ZPLE limit for cubic and axial crystals

In the following, we show that Eq. (5.2.27) of the main text reduces to the well-known literature results for cubic and axial crystals and provide explicit expressions for the Grüneisen ZPLE (Eq. (5.2.28) of the main text). In the case of cubic symmetry, the only non-zero coefficients in the compliance tensor are $s_{11} = s_{22} = s_{33}$, $s_{12} = s_{13} = s_{23}$ and $s_{44} = s_{55} = s_{66}$. Furthermore, the infinitesimal strains ϵ_i , $i = 4, 5, 6$ are null, as the crystal symmetry is preserved.

There is only one independent lattice constant, hence $\epsilon_1 = \epsilon_2 = \epsilon_3$ and $\gamma_{\mathbf{q}\nu}^1 = \gamma_{\mathbf{q}\nu}^2 = \gamma_{\mathbf{q}\nu}^3$. Note that $\epsilon_1 = \Delta a/a^0$ by definition of the stress tensor components. Eq. (5.2.27) of the main text thus reduces to [52]

$$\begin{aligned}\alpha(T) &= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \left(s_{11} \gamma_{\mathbf{q}\nu}^1 + s_{12} \gamma_{\mathbf{q}\nu}^2 + s_{13} \gamma_{\mathbf{q}\nu}^3 \right) \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}^V(T) \\ &= \frac{1}{V^0} (s_{11} + 2s_{12}) \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^1 \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}^V(T) \\ &= \frac{1}{3B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^1 \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}^V(T),\end{aligned}\tag{S5.10}$$

since $s_{11} + 2s_{12} = (3B_0)^{-1}$. In this expression, $\gamma_{\mathbf{q}\nu}^1$ refers to the derivative taken by straining only *one* of the three lattice parameters, the other two remaining at their fixed static equilibrium value. One could recast this term to compute the derivative with respect to a simultaneous change of length of *all three* lattice vectors at the cost of a factor 1/3 :

$$\begin{aligned}\gamma_{\mathbf{q}\nu}^1 &= -\frac{a_i}{\omega_{\mathbf{q}\nu}} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial a_i} \right|_{a_{j \neq i} = a^0} \\ &= -\frac{1}{3} \frac{a^0}{\omega_{\mathbf{q}\nu}} \left(\frac{\partial \omega_{\mathbf{q}\nu}}{\partial a} \right) \\ &= \frac{1}{3} \gamma_{\mathbf{q}\nu}^a\end{aligned}\tag{S5.11}$$

The ZPLE expression for cubic symmetry (Eq. (5.2.28) of the main text) then reads

$$\begin{aligned}\Delta a(T=0) &= \frac{a^0}{3B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^1 \frac{\omega_{\mathbf{q}\nu}}{2} \\ &= \frac{a^0}{9B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^a \frac{\omega_{\mathbf{q}\nu}}{2}\end{aligned}\tag{S5.12}$$

One can also show that the anisotropic case contains the volumic result, Eq. (5.2.15) of the main text. First, from Eq. (5.2.34) of the main text, we know that $\beta(T) = 3\alpha(T)$ for cubic materials. Hence,

$$\begin{aligned}\frac{\Delta V(T)}{V^0} &= 3 \frac{\Delta a(T)}{a^0} \\ &= \frac{1}{3B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^a \omega_{\mathbf{q}\nu} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right),\end{aligned}\tag{S5.13}$$

where we reintroduced the temperature dependence using the relations shown in Eq. (5.2.25) to (5.2.27) of the main text. We now find the relation between the linear mode Grüneisen

parameters $\gamma_{\mathbf{q}\nu}^a$ and their volumic counterpart,

$$\begin{aligned}
\gamma_{\mathbf{q}\nu}^a &= -\frac{a^0}{\omega_{\mathbf{q}\nu}} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial a} \right|_{a^0} \\
&= -\frac{a^0}{\omega_{\mathbf{q}\nu}} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \frac{\partial V}{\partial a} \right|_{a^0} \\
&= -\frac{a^0}{\omega_{\mathbf{q}\nu}} (3(a^0)^2) \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \right|_{V^0} \\
&= 3 \left(-\frac{V^0}{\omega_{\mathbf{q}\nu}} \left. \frac{\partial \omega_{\mathbf{q}\nu}}{\partial V} \right|_{V^0} \right) \\
&= 3\gamma_{\mathbf{q}\nu}^V.
\end{aligned} \tag{S5.14}$$

Substituting this result in Eq. (S5.13), we recover the volumic expression, Eq. (5.2.15) of the main text,

$$\frac{\Delta V(T)}{V^0} = \frac{1}{B_0 V^0} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V \omega_{\mathbf{q}\nu} \left(n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right). \tag{S5.15}$$

For axial symmetry (i.e. hexagonal, trigonal and tetragonal lattices), there are only two non-zero components of the mode Grüneisen tensor, $\gamma_{\mathbf{q}\nu}^1 = \gamma_{\mathbf{q}\nu}^2$ and $\gamma_{\mathbf{q}\nu}^3$. Hence, the relevant independent compliance coefficients are $s_{11} = s_{22}$, s_{33} , s_{12} and $s_{13} = s_{23}$. We obtain

$$\begin{aligned}
\alpha^1(T) &= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \left((s_{11} + s_{12}) \gamma_{\mathbf{q}\nu}^1 + s_{13} \gamma_{\mathbf{q}\nu}^3 \right) \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}(T) \\
&= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \left(\frac{1}{2} (s_{11} + s_{12}) \gamma_{\mathbf{q}\nu}^a + s_{13} \gamma_{\mathbf{q}\nu}^c \right) \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}(T), \\
\alpha^3(T) &= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \left(2s_{13} \gamma_{\mathbf{q}\nu}^1 + s_{33} \gamma_{\mathbf{q}\nu}^3 \right) \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}(T) \\
&= \frac{1}{V^0} \sum_{\mathbf{q}\nu} \left(s_{13} \gamma_{\mathbf{q}\nu}^a + s_{33} \gamma_{\mathbf{q}\nu}^c \right) \omega_{\mathbf{q}\nu} c_{\mathbf{q}\nu}(T),
\end{aligned} \tag{S5.16}$$

which are the expressions tabulated in the literature [103]. The ZPLE thus takes the form

$$\begin{aligned}
\Delta a(T=0) &= \frac{a^0}{V^0} \sum_{\mathbf{q}\nu} \left(\frac{1}{2} (s_{11} + s_{12}) \gamma_{\mathbf{q}\nu}^a + s_{13} \gamma_{\mathbf{q}\nu}^c \right) \frac{\omega_{\mathbf{q}\nu}}{2}, \\
\Delta c(T=0) &= \frac{c^0}{V^0} \sum_{\mathbf{q}\nu} \left(s_{13} \gamma_{\mathbf{q}\nu}^a + s_{33} \gamma_{\mathbf{q}\nu}^c \right) \frac{\omega_{\mathbf{q}\nu}}{2}.
\end{aligned} \tag{S5.17}$$

S5.3. Calculation parameters for the different materials

Table S5.1 specifies the Materials Project ID number, space group, cut-off energy for the plane-wave basis set for the GS, phonon and elastic constants calculations, \mathbf{k} -point sampling of the BZ, as well as the \mathbf{q} -point sampling used for the ZPLE and EPI calculations,

for the 22 materials investigated for this work. For elastic constant calculations, the aforementioned cutoff energies were used in conjunction with a cutoff energy smearing of 0.5 hartree, to prevent discontinuities in the total energy curve when computing shear strains. The larger increase of the cutoff energy for AlN can be attributed to older pseudopotentials. For more complete details about the EPI calculation for ZnO, see the METHODS section and Supplementary Material of Ref. [71].

Table S5.1 – Calculations parameters for the twenty-two materials in our set. From left to right, material ID on the Materials Project, cut-off energy for the plane-wave basis set for GS/phonons (E_{cut}) and elastic constants (E_{cut}, c_{ij}) calculations, \mathbf{k} -point sampling for the BZ, and \mathbf{q} -point sampling for ZPLE and ZPR_{*g*}^{EPI} ($N \times N \times N (\times S)$ shifts) Monkhorst-Pack grids). All \mathbf{q} -point grids are Γ -centered. The elastic constants cutoff energy was used in conjunction with a cutoff energy smearing of 0.5 hartree. See Sec. 5.3.2 for more details regarding the \mathbf{q} -point sampling for GaAs.

Material	MP-ID	Space group	E_{cut} (ha)	E_{cut}, c_{ij} (ha)	\mathbf{k} -points	\mathbf{q} -points (ZPLE)	\mathbf{q} -points (EPI) [71]
C	66	Fd $\bar{3}m$ [227]	35	45	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$125 \times 125 \times 125$ [73]
Si	149	Fd $\bar{3}m$ [227]	25	25	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$100 \times 100 \times 100$
Ge	32	Fd $\bar{3}m$ [227]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
SiC	8062	F $\bar{4}3m$ [216]	35	45	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
BN	1639	F $\bar{4}3m$ [216]	35	40	$8 \times 8 \times 8 (\times 4)$	$8 \times 8 \times 8$	$100 \times 100 \times 100$ [73]
BA _s	10044	F $\bar{4}3m$ [216]	40	45	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
AlP	1550	F $\bar{4}3m$ [216]	25	30	$8 \times 8 \times 8 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
AlAs	2172	F $\bar{4}3m$ [216]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
AlSb	2624	F $\bar{4}3m$ [216]	40	45	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
GaN	830	F $\bar{4}3m$ [216]	40	45	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
GaP	2690	F $\bar{4}3m$ [216]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
GaAs	2534	F $\bar{4}3m$ [216]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$64 \times 64 \times 64$
ZnS	10695	F $\bar{4}3m$ [216]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
ZnSe	1190	F $\bar{4}3m$ [216]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
ZnTe	2176	F $\bar{4}3m$ [216]	40	40	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
CdS	2469	F $\bar{4}3m$ [216]	45	50	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
CdSe	2691	F $\bar{4}3m$ [216]	50	50	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
CdTe	406	F $\bar{4}3m$ [216]	50	55	$6 \times 6 \times 6 (\times 4)$	$8 \times 8 \times 8$	$48 \times 48 \times 48$
MgO	1265	Fm $\bar{3}m$ [225]	50	50	$8 \times 8 \times 8 (\times 4)$	$8 \times 8 \times 8$	$96 \times 96 \times 96$
AlN	661	P6 ₃ mc [186]	35	60	$6 \times 6 \times 6$	$8 \times 8 \times 8$	$34 \times 34 \times 34$ [73]
GaN	804	P6 ₃ mc [186]	40	45	$8 \times 8 \times 8$	$8 \times 8 \times 8$	$64 \times 64 \times 64$
ZnO	2133	P6 ₃ mc [186]	50	50	$6 \times 6 \times 6$	$8 \times 8 \times 8$	$48 \times 48 \times 48$

S5.4. Additional numerical tables

Tables S5.2 and S5.3 reproduce the ZPLE results of Tables 5.1 and 5.2 of the main text, making explicit the absolute relative error with respect to FE minimization results for each material.

Table S5.4 displays the numerical values of the fractional ZPLE ($\Delta a(T=0)/a^0$ and $\Delta c(T=0)/c^0$) used in Eq. (5.2.37) of the main text to obtain the different estimations of the ZPR_g^{ZPLE} reported in Table 5.6 of the main text. Note that, since we use $\Delta E_g^{\text{Grun 1\%}}$ as a reference, $\Delta a^{\text{Grun 1\%}}/a^0$ and $\Delta a^{\text{data}}/a^0$ are by construction identical. For $\Delta a^{700-1000\text{K}}/a^0$ and $\Delta a^{T>2\theta_D}/a^0$, we used the ZP-renormalized lattice constants, $a(T=0)$, as the reference configuration to define the temperature-dependent strain:

$$\frac{\Delta a(T)}{a^0} \rightarrow \frac{a(T) - a(T=0)}{a(T=0)}. \quad (\text{S5.18})$$

This definition is slightly different from Eq. (5.2.17) of the main text. Since the purpose of this analysis is to evaluate the reliability of Eq. (5.2.37) of the main text when applied to experimental data, we chose to treat our theoretical data as if it were experimental results. As $a(T=0) = a^0 + \Delta a(T=0)$, this induces an error of the order of $(\Delta a(T=0)/a^0)^2$, which will not affect our conclusions. Note finally that such an error is naturally present in the evaluation of $\Delta a^{\text{exp}}/a^0$.

The reported data in the rightmost column was obtained from the following procedures. For Ge, Si and diamond, we average the $\Delta a(T=0)/a^0$ data reported in Sec. III.B.2 of Ref. [50] from 3 different techniques, our definition of $\Delta a(T=0)/a^0$ differing by a sign change from theirs. For AlN and GaN, we use the experimental data displayed in Figs. (1) and (3) of Ref. [93] from references therein. For GaN, this value of $\Delta a(T=0)/a^0$ is averaged with the one extracted from Fig. (1) of Ref. [94], from markers labelled “this work” and “Ref. 10”, to obtain the ZPR_g estimation reported in the main text. For ZnO, we extract $\Delta a(T=0)/a^0$ by combining high-temperature data from Ref. [87] and low-temperature data from Ref. [95]. Note that this method may be biased by a small lattice parameter mismatch between the two combined datasets, as they have been obtained from different samples and measuring devices.

Table S5.5 finally reports the experimental values shown in Fig. 5.7 of the main text and compares them to our numerical results. When available, we favored values extracted from isotopic substitution data, which is less sensitive to the extrapolation procedure as the high-temperature asymptote method described in Sec. 5.4.4 of the main text (see Supplementary Note 1 of the Supplementary Information of Ref. [71] for a more detailed analysis of the available experimental ZPR_g data). For the latter method, we rely on the works of Pässler [104, 105]. For AlN, we set our reference value as the average of the reported values from those two articles, resp. -350 meV and -483 meV.

Table S5.2 – Zero-point lattice expansion for cubic materials: absolute relative error (ARE) of the Grüneisen approach with respect to FE minimization. a^0 is the theoretical lattice parameter in the static lattice approximation, except for Ge and GaP (see Sec. 5.3.1 of the main text for details). The variation of a^0 induced by the zero-point motion of the ions, $\Delta a^{\text{FE}}(T = 0)$, is obtained by minimizing the Helmholtz free energy, while $\Delta a^{\text{Grün}}(T = 0)$ is obtained from the Grüneisen parameters, using Eq. (5.2.28) of the main text.

Material	Δa^{FE} (Bohr)	$\Delta a^{\text{Grün}}$ (Bohr)		ARE vs FE (%)	
		1%	0.5%	1%	0.5%
C-dia	0.0253	0.0250	0.0248	1.19	1.98
Si-dia	0.0174	0.0172	0.0172	1.15	1.15
Ge-dia	0.0102	0.0112	0.0112	9.80	9.80
SiC-zb	0.0227	0.0226	0.0223	0.44	1.76
BN-zb	0.0285	0.0271	0.0270	5.01	5.26
BAs	0.0238	0.0237	0.0236	0.42	0.84
AlP-zb	0.0180	0.0178	0.0178	1.11	1.11
AlAs-zb	0.0166	0.0162	0.0163	2.41	1.81
AlSb-zb	0.0163	0.0162	0.0161	0.61	1.23
GaN-zb	0.0209	0.0208	0.0208	0.48	0.48
GaP-zb	0.0152	0.0156	0.0156	2.63	2.63
GaAs-zb	0.0157	0.0146	0.0145	7.01	7.64
ZnS-zb	0.0194	0.0193	0.0192	0.52	1.03
ZnSe-zb	0.0160	0.0157	0.0155	1.88	3.13
ZnTe-zb	0.0154	0.0153	0.0153	0.65	0.65
CdS-zb	0.0177	0.0170	0.0162	3.95	8.47
CdSe-zb	0.0136	0.0133	0.0135	2.21	0.74
CdTe-zb	0.0129	0.0126	0.0125	2.33	3.10
MgO-rs	0.0333	0.0317	0.0318	5.09	4.79

Table S5.3 – Zero-point lattice expansion for wurtzite materials: absolute relative error (ARE) of the Grüneisen approach with respect to FE minimization. a^0 and c^0 are the theoretical lattice parameter in the static lattice approximation. The variation of a^0 and c^0 induced by the zero-point motion of the ions, resp. $\Delta a^{\text{FE}}(T = 0)$ and $\Delta c^{\text{FE}}(T = 0)$, are obtained by minimizing the Helmholtz FE, while $\Delta a^{\text{Grün}}(T = 0)$ and $\Delta c^{\text{Grün}}(T = 0)$ are calculated from the Grüneisen parameters, using Eq. 5.2.28 of the main text.

Material	Δa^{FE} (Bohr)	$\Delta a^{\text{Grün}}$ (Bohr)		ARE vs FE (%)	
		1%	0.5%	1%	0.5%
AlN	0.0170	0.0163	0.0162	4.12	4.71
GaN	0.0151	0.0151	0.0151	0.00	0.00
ZnO	0.0165	0.0165	0.0167	0.00	1.21
Material	Δc^{FE} (Bohr)	$\Delta c^{\text{Grün}}$ (Bohr)		ARE vs FE (%)	
		1%	0.5%	1%	0.5%
AlN	0.0241	0.0242	0.0242	0.42	0.42
GaN	0.0232	0.0228	0.0229	1.73	1.29
ZnO	0.0213	0.0213	0.0207	0.00	2.82

Table S5.4 – Numerical values of $\Delta a(T = 0)/a^0$ and $\Delta c(T = 0)/c^0$ used to obtain estimations of the ZPLE band gap ZPR in Table 5.6 of the main text. At linear order, we can express the $\Delta V(T = 0)/V^0$ term entering Eq. (5.2.37) of the main text as $3\Delta a(T = 0)/a^0$ for cubic materials, and as $2\Delta a(T = 0)/a^0 + \Delta c(T = 0)/c^0$ for wurtzites. The different methodologies used to evaluate the fractional ZPLE are described in Sec. 5.4.3.2 of the main text. Note that, since the Grüneisen 1% results as a reference value in Table 5.6, $\Delta a^{\text{Grün 1\%}}(T = 0)/a^0 = \Delta a^{\text{data}}(T = 0)/a^0$.

Material	$\frac{\Delta a^{\text{data}}}{a^0}$	$\frac{\Delta a^{T>2\theta_D}}{a^0}$	$\frac{\Delta a^{700-1000\text{K}}}{a^0}$	$\frac{\Delta a^{\text{exp}}}{a^0}$
	(10^{-3})	(10^{-3})	(10^{-3})	(10^{-3})
C-dia	3.70	3.17	1.67	3.9, 3.7, 3.9 [50]
Si-dia	1.66	1.43	1.23	1.7, 1.9, 2.0 [50]
Ge-dia	1.06	0.90	0.89	1.3, 1.2, 1.8 [50]
SiC-zb	2.73	2.39	1.65	2.71 [89]
BN-zb	4.02	3.36	1.99	
BAs	2.60	2.22	1.75	1.30 [90]
AlP-zb	1.71	1.48	1.30	
AlAs-zb	1.50	1.27	1.22	
AlSb-zb	1.38	1.11	1.15	
GaN-zb	2.42	1.99	1.68	
GaP-zb	1.52	1.28	1.21	1.73 [89]
GaAs-zb	1.34	1.17	1.16	
ZnS-zb	1.88	1.59	1.58	
ZnSe-zb	1.45	1.29	1.28	
ZnTe-zb	1.31	1.09	1.18	
CdS-zb	1.52	1.19	1.29	
CdSe-zb	1.14	0.94	1.02	
CdTe-zb	1.01	0.84	0.92	
MgO-rs	3.95	3.61	3.07	4.01 [91], 3.77 [92]
AlN-w	2.81	2.44	1.80	2.64 [93]
GaN-w	2.48	1.91	1.72	1.13 [93], 2.50 [94]
ZnO-w	2.67	2.15	2.00	4.23 [87, 95]
Material	$\frac{\Delta c^{\text{data}}}{c^0}$	$\frac{\Delta c^{T>2\theta_D}}{c^0}$	$\frac{\Delta c^{700-1000\text{K}}}{c^0}$	$\frac{\Delta c^{\text{exp}}}{c^0}$
	(10^{-3})	(10^{-3})	(10^{-3})	(10^{-3})
AlN-w	2.62	2.26	1.64	2.71 [93]
GaN-w	2.30	1.78	1.60	1.41 [93], 1.91 [94]
ZnO-w	2.14	1.73	1.62	1.19 [87, 95]

Table S5.5 – Comparison of the predicted band gap ZPR to experimental values from the EPI contribution only and from the combined EPI+ZPLE contributions (see Table 5.8 of the main text). The unitless ratio R^{tot} is defined as $R = \text{ZPR}_g^{\text{tot}}/\text{ZPR}_g^{\text{exp}}$, and reported as $R^{\text{tot}} = 1/(\text{ZPR}_g^{\text{exp}}/\text{ZPR}_g^{\text{tot}})$ when $|\text{ZPR}_g^{\text{tot}}| > |\text{ZPR}_g^{\text{exp}}|$, in order to emphasize the symmetric boundaries of the shaded gray area in Fig. 5.6 of the main text (see Sec. 5.4.3.3 of the main text for details). The ratio R^{EPI} is defined in a similar manner, using $\text{ZPR}_g^{\text{EPI}}$. †The reference value for AlN is the average value of the reported values from Ref. [104] and [105] (resp. -350 meV and -483 meV).

Material	$\text{ZPR}_g^{\text{EPI}}$ (meV)	$\text{ZPR}_g^{\text{tot}}$ (meV)	$\text{ZPR}_g^{\text{exp}}$ (meV)	R^{EPI}	R^{tot}
C-dia	-330	-357	-338 [106]	0.98	1/0.95
Si-dia	-56	-47	-59 [107]	0.95	0.80
Ge-dia	-33	-42	-52 [108]	0.63	0.81
SiC-zb	-179	-172	-212 [105]	0.84	0.81
AlAs-zb	-74	-67	-50 [104]	1/0.68	1/0.75
GaP-zb	-65	-57	-85 [104]	0.76	0.68
GaAs-zb	-21	-52	-60 [104]	0.35	0.87
ZnS-zb	-88	-112	-105 [109]	0.84	1/0.94
ZnSe-zb	-44	-61	-55 [104]	0.80	1/0.90
ZnTe-zb	-22	-40	-40 [104]	0.55	1.00
CdS-zb	-70	-80	-62 [110]	1/0.89	1/0.77
CdSe-zb	-34	-41	-39 [105]	0.87	1/0.95
CdTe-zb	-20	-29	-16 [104]	1/0.80	1/0.55
AlN-w	-399	-478	-417 [†]	0.96	1/0.87
GaN-w	-189	-238	-180 [104]	1/0.95	1/0.76
ZnO-w	-157	-168	-164 [111]	0.96	1/0.97

Bibliography

- [1] D. I. Khomskii. *Basic Aspects of The Quantum Theory of Solids: Order and Elementary Excitations*. Cambridge University Press, Cambridge, 2010. doi: [10.1017/CBO9780511780271](https://doi.org/10.1017/CBO9780511780271).
- [2] O. Madelung. Phonon-Phonon Interaction: Thermal Properties. In M. Cardona, P. Fulde, K. von Klitzing, H. J. Queisser, and H. K. V. Lotsch, editors, *Introduction to Solid-State Theory*, volume 2, pages 314–326. Springer Berlin Heidelberg, Berlin, Heidelberg, 1978. doi: [10.1007/978-3-642-61885-7_7](https://doi.org/10.1007/978-3-642-61885-7_7).
- [3] P. B. Allen. Theory of thermal expansion: Quasi-harmonic approximation and corrections from quasiparticle renormalization. *Modern Physics Letters B*, 34(02):2050025, 2020. doi: [10.1142/S0217984920500256](https://doi.org/10.1142/S0217984920500256).
- [4] E. Grüneisen. Theorie des festen Zustandes einatomiger Elemente. *Annalen der Physik*, 344(12):257–306, 1912. doi: [10.1002/andp.19123441202](https://doi.org/10.1002/andp.19123441202).
- [5] R. W. Munn. Role of the elastic constants in negative thermal expansion of axial solids. *Journal of Physics C: Solid State Physics*, 5(5):535–542, 1972. doi: [10.1088/0022-3719/5/5/005](https://doi.org/10.1088/0022-3719/5/5/005).
- [6] B. Monemar. Fundamental Energy Gaps of AlAs and AlP from Photoluminescence Excitation Spectra. *Physical Review B*, 8(12):5711–5718, 1973. doi: [10.1103/PhysRevB.8.5711](https://doi.org/10.1103/PhysRevB.8.5711).
- [7] O. Madelung. *Semiconductors: Data Handbook*. Springer-Verlag, Berlin Heidelberg, 3rd edition, 2004. doi: [10.1007/978-3-642-18865-7](https://doi.org/10.1007/978-3-642-18865-7).
- [8] G. Antonius, S. Poncé, E. Lantagne-Hurtubise, G. Auclair, X. Gonze, and M. Côté. Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure. *Phys. Rev. B*, 92(8):085137, 2015. doi: [10.1103/PhysRevB.92.085137](https://doi.org/10.1103/PhysRevB.92.085137).
- [9] C. P. Kuo, S. K. Vong, R. M. Cohen, and G. B. Stringfellow. Effect of mismatch strain on band gap in III-V semiconductors. *Journal of Applied Physics*, 57(12):5428–5432, 1985. doi: [10.1063/1.334817](https://doi.org/10.1063/1.334817).
- [10] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, Cambridge, 2004. doi: [10.1017/CBO9780511805769](https://doi.org/10.1017/CBO9780511805769).
- [11] F. Giustino. Electron-phonon interactions from first principles. *Reviews of Modern Physics*, 89(1):015003, 2017. doi: [10.1103/RevModPhys.89.015003](https://doi.org/10.1103/RevModPhys.89.015003).
- [12] S. Poncé, W. Li, S. Reichardt, and F. Giustino. First-principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two-dimensional materials. *Reports on Progress in Physics*, 83(3):036501, 2020. doi: [10.1088/1361-6633/ab6a43](https://doi.org/10.1088/1361-6633/ab6a43).

- [13] G. Brunin, H. P. C. Miranda, M. Giantomassi, M. Royo, M. Stengel, M. J. Verstraete, X. Gonze, G. M. Rignanese, and G. Hautier. Phonon-limited electron mobility in Si, GaAs, and GaP with exact treatment of dynamical quadrupoles. *Phys. Rev. B*, 102(9):094308, 2020. doi: [10.1103/PhysRevB.102.094308](https://doi.org/10.1103/PhysRevB.102.094308).
- [14] C. Verdi, F. Caruso, and F. Giustino. Origin of the crossover from polarons to Fermi liquids in transition metal oxides. *Nature Communications*, 8:15769, 2017. doi: [10.1038/ncomms15769](https://doi.org/10.1038/ncomms15769).
- [15] J. P. Nery, P. B. Allen, G. Antonius, L. Reining, A. Miglio, and X. Gonze. Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron. *Phys. Rev. B*, 97(11):115145, 2018. doi: [10.1103/PhysRevB.97.115145](https://doi.org/10.1103/PhysRevB.97.115145).
- [16] F. Giustino, J. R. Yates, I. Souza, M. L. Cohen, and S. G. Louie. Electron-Phonon Interaction via Electronic and Lattice Wannier Functions: Superconductivity in Boron-Doped Diamond Reexamined. *Phys. Rev. Lett.*, 98(4):047005, 2007. doi: [10.1103/PhysRevLett.98.047005](https://doi.org/10.1103/PhysRevLett.98.047005).
- [17] X. Blase, C. Adessi, and D. Connétable. Role of the Dopant in the Superconductivity of Diamond. *Physical Review Letters*, 93(23):237004, 2004. doi: [10.1103/PhysRevLett.93.237004](https://doi.org/10.1103/PhysRevLett.93.237004).
- [18] L. F. Huang, X. Z. Lu, E. Tennesen, and J. M. Rondinelli. An efficient ab-initio quasiharmonic approach for the thermodynamics of solids. *Computational Materials Science*, 120:84–93, 2016. doi: [10.1016/j.commatsci.2016.04.012](https://doi.org/10.1016/j.commatsci.2016.04.012).
- [19] K. Lejaeghere, J. Jaeken, V. Van Speybroeck, and S. Cottenier. Ab initio based thermal property predictions at a low cost: An error analysis. *Physical Review B*, 89(1):014304, 2014. doi: [10.1103/PhysRevB.89.014304](https://doi.org/10.1103/PhysRevB.89.014304).
- [20] A. Erba. On combining temperature and pressure effects on structural properties of crystals with standard ab initio techniques. *The Journal of Chemical Physics*, 141(12):124115, 2014. doi: [10.1063/1.4896228](https://doi.org/10.1063/1.4896228).
- [21] A. Otero-de-la Roza and V. Luaña. Treatment of first-principles data for predictive quasiharmonic thermodynamics of solids: The case of MgO. *Phys. Rev. B*, 84(2):024109, 2011. doi: [10.1103/PhysRevB.84.024109](https://doi.org/10.1103/PhysRevB.84.024109).
- [22] N. Mounet and N. Marzari. First-principles determination of the structural, vibrational and thermodynamic properties of diamond, graphite, and derivatives. *Phys. Rev. B*, 71(20):205214, 2005. doi: [10.1103/PhysRevB.71.205214](https://doi.org/10.1103/PhysRevB.71.205214).
- [23] G. M. Rignanese, J. P. Michenaud, and X. Gonze. Ab initio study of the volume dependence of dynamical and thermodynamical properties of silicon. *Phys. Rev. B*, 53(8):4488–4497, 1996. doi: [10.1103/PhysRevB.53.4488](https://doi.org/10.1103/PhysRevB.53.4488).
- [24] P. Hao, Y. Fang, J. Sun, G. I. Csonka, P. H. T. Philipsen, and J. P. Perdew. Lattice constants from semilocal density functionals with zero-point phonon correction.

- Physical Review B*, 85(1):014111, 2012. doi: [10.1103/PhysRevB.85.014111](https://doi.org/10.1103/PhysRevB.85.014111).
- [25] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán. Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B*, 79(15):155107, 2009. doi: [10.1103/PhysRevB.79.155107](https://doi.org/10.1103/PhysRevB.79.155107).
- [26] P. Haas, F. Tran, and P. Blaha. Calculation of the lattice constant of solids with semilocal functionals. *Physical Review B*, 79(8):085104, 2009. doi: [10.1103/PhysRevB.79.085104](https://doi.org/10.1103/PhysRevB.79.085104).
- [27] L. Schimka, J. Harl, and G. Kresse. Improved hybrid functional for solids: The HSEsol functional. *The Journal of Chemical Physics*, 134(2):024116, 2011. doi: [10.1063/1.3524336](https://doi.org/10.1063/1.3524336).
- [28] A. I. Lichtenstein, R. O. Jones, S. de Gironcoli, and S. Baroni. Anisotropic thermal expansion in silicates: A density functional study of β -eucryptite and related materials. *Physical Review B*, 62(17):11487–11493, 2000. doi: [10.1103/PhysRevB.62.11487](https://doi.org/10.1103/PhysRevB.62.11487).
- [29] P. Carrier, R. Wentzcovitch, and J. Tsuchiya. First-principles prediction of crystal structures at high temperatures using the quasiharmonic approximation. *Physical Review B*, 76(6):064116, 2007. doi: [10.1103/PhysRevB.76.064116](https://doi.org/10.1103/PhysRevB.76.064116).
- [30] M. Råsander and M. A. Moram. On the thermal expansion in MgSiN_2 . *Journal of Physics D: Applied Physics*, 51(10):105108, 2018. doi: [10.1088/1361-6463/aaa932](https://doi.org/10.1088/1361-6463/aaa932).
- [31] N. A. Pike and O. M. Løvvik. Calculation of the anisotropic coefficients of thermal expansion: A first-principles approach. *Computational Materials Science*, 167:257–263, 2019. doi: [10.1016/j.commatsci.2019.05.045](https://doi.org/10.1016/j.commatsci.2019.05.045).
- [32] M. Zacharias, M. Scheffler, and C. Carbogno. Fully anharmonic nonperturbative theory of vibronically renormalized electronic band structures. *Physical Review B*, 102(4):045126, 2020. doi: [10.1103/PhysRevB.102.045126](https://doi.org/10.1103/PhysRevB.102.045126).
- [33] Z. Wu and R. M. Wentzcovitch. Quasiharmonic thermal elasticity of crystals: An analytical approach. *Physical Review B*, 83(18):184115, 2011. doi: [10.1103/PhysRevB.83.184115](https://doi.org/10.1103/PhysRevB.83.184115).
- [34] C. Malica and A. D. Corso. Quasi-harmonic temperature dependent elastic constants: applications to silicon, aluminum, and silver. *Journal of Physics: Condensed Matter*, 32(31):315902, 2020. doi: [10.1088/1361-648X/ab8426](https://doi.org/10.1088/1361-648X/ab8426).
- [35] R. Masuki, T. Nomoto, R. Arita, and T. Tadano. Anharmonic Grüneisen theory based on self-consistent phonon theory: Impact of phonon-phonon interactions neglected in the quasiharmonic theory. *Phys. Rev. B*, 105(6):064112, 2022. doi: [10.1103/PhysRevB.105.064112](https://doi.org/10.1103/PhysRevB.105.064112).
- [36] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak. Temperature dependent effective potential method for accurate free energy calculations of solids. *Phys. Rev. B*, 87(10):104111, 2013. doi: [10.1103/PhysRevB.87.104111](https://doi.org/10.1103/PhysRevB.87.104111).

- [37] J. D. Querales-Flores, J. Cao, S. Fahy, and I. Savić. Temperature effects on the electronic band structure of PbTe from first principles. *Phys. Rev. Materials*, 3(5):055405, 2019. doi: [10.1103/PhysRevMaterials.3.055405](https://doi.org/10.1103/PhysRevMaterials.3.055405).
- [38] B. Monserrat and D. Vanderbilt. Temperature Effects in the Band Structure of Topological Insulators. *Physical Review Letters*, 117(22):226801, 2016. doi: [10.1103/PhysRevLett.117.226801](https://doi.org/10.1103/PhysRevLett.117.226801).
- [39] I. Bravić and B. Monserrat. Finite temperature optoelectronic properties of BAs from first principles. *Phys. Rev. Materials*, 3(6):065402, 2019. doi: [10.1103/PhysRevMaterials.3.065402](https://doi.org/10.1103/PhysRevMaterials.3.065402).
- [40] Y. Zhang, Z. Wang, J. Xi, and J. Yang. Temperature-dependent band gaps in several semiconductors: from the role of electron-phonon renormalization. *Journal of Physics: Condensed Matter*, 32:475503, 2020. doi: [10.1088/1361-648X/aba45d](https://doi.org/10.1088/1361-648X/aba45d).
- [41] Y. N. Wu, W. A. Saidi, P. Ohodnicki, B. Chorpening, and Y. Duan. First-Principles Investigations of the Temperature Dependence of Electronic Structure and Optical Properties of Rutile TiO₂. *The Journal of Physical Chemistry C*, 122(39):22642–22649, 2018. doi: [10.1021/acs.jpcc.8b06941](https://doi.org/10.1021/acs.jpcc.8b06941).
- [42] C. E. P. Villegas, A. R. Rocha, and A. Marini. Anomalous Temperature Dependence of the Band Gap in Black Phosphorus. *Nano Letters*, 16(8):5095–5101, 2016. doi: [10.1021/acs.nanolett.6b02035](https://doi.org/10.1021/acs.nanolett.6b02035).
- [43] B. Monserrat and D. Vanderbilt. Temperature dependence of the bulk Rashba splitting in the bismuth tellurohalides. *Physical Review Materials*, 1(5):054201, 2017. doi: [10.1103/PhysRevMaterials.1.054201](https://doi.org/10.1103/PhysRevMaterials.1.054201).
- [44] H. Kamisaka, S. V. Kilina, K. Yamashita, and O. V. Prezhdo. Ab Initio Study of Temperature and Pressure Dependence of Energy and Phonon-Induced Dephasing of Electronic Excitations in CdSe and PbSe Quantum Dots. *ACS Publications*, 112(21):7800–7808, 2008. doi: [10.1021/jp710435q](https://doi.org/10.1021/jp710435q).
- [45] M. A. Abdulsattar, B. B. Kadhim, and A. M. Ali. Effect of temperature and pressure on electronic and spectroscopic properties of CdS wurtzoid molecule: A DFT study. *Optik*, 188:126–132, 2019. doi: <https://doi.org/10.1016/j.ijleo.2019.05.030>.
- [46] J. D. Querales-Flores, P. Aguado-Puente, D. Dangić, J. Cao, P. Chudzinski, T. N. Todorov, M. Grüning, S. Fahy, and I. Savić. Towards temperature-induced topological phase transition in SnTe: A first-principles study. *Physical Review B*, 101(23):235206, 2020. doi: [10.1103/PhysRevB.101.235206](https://doi.org/10.1103/PhysRevB.101.235206).
- [47] V. Brousseau-Couture, G. Antonius, and M. Côté. Temperature dependence of the topological phase transition of BiTeI from first principles. *Phys. Rev. Research*, 2(2):023185, 2020. doi: [10.1103/PhysRevResearch.2.023185](https://doi.org/10.1103/PhysRevResearch.2.023185).

- [48] B. Monserrat and A. Narayan. Unraveling the topology of ZrTe_5 by changing temperature. *Physical Review Research*, 1(3):033181, 2019. doi: [10.1103/PhysRevResearch.1.033181](https://doi.org/10.1103/PhysRevResearch.1.033181).
- [49] N. Garro, A. Cantarero, M. Cardona, A. Göbel, T. Ruf, and K. Eberl. Dependence of the lattice parameters and the energy gap of zinc-blende-type semiconductors on isotopic masses. *Phys. Rev. B*, 54(7):4732–4740, 1996. doi: [10.1103/PhysRevB.54.4732](https://doi.org/10.1103/PhysRevB.54.4732).
- [50] M. Cardona and M. L. W. Thewalt. Isotope effects on the optical spectra of semiconductors. *Reviews of Modern Physics*, 77(4):1173–1224, 2005. doi: [10.1103/RevModPhys.77.1173](https://doi.org/10.1103/RevModPhys.77.1173).
- [51] P. B. Allen and M. Cardona. Temperature dependence of the direct gap of Si and Ge. *Physical Review B*, 27(8):4760–4769, 1983. doi: [10.1103/PhysRevB.27.4760](https://doi.org/10.1103/PhysRevB.27.4760).
- [52] G. Grimvall. *Thermophysical properties of materials*. Number XVIII in Selected topics in solid state physics. North-Holland, Amsterdam, 1986.
- [53] H. M. Kagaya and T. Soma. Mode Grüneisen parameters and thermal expansion of GaP and InP. *Solid State Communications*, 58(7):479–482, 1986. doi: [10.1016/0038-1098\(86\)90037-2](https://doi.org/10.1016/0038-1098(86)90037-2).
- [54] T. H. K. Barron, J. F. Collins, T. W. Smith, and G. K. White. Thermal expansion, Gruneisen functions and static lattice properties of quartz. *Journal of Physics C: Solid State Physics*, 15(20):4311–4326, 1982. doi: [10.1088/0022-3719/15/20/016](https://doi.org/10.1088/0022-3719/15/20/016).
- [55] See Supplemental Material for an extension of the Grüneisen formalism to finite external pressure, explicit ZPLE equations in the Grüneisen approach for cubic and axial symmetries, detailed calculation parameters for all materials, as well as additional numerical tables, which include Ref. [50, 52, 71, 73, 87, 89–95, 103–111].
- [56] A. B. Alchagirov, J. P. Perdew, J. C. Boettger, R. C. Albers, and C. Fiolhais. Energy and pressure versus volume: Equations of state motivated by the stabilized jellium model. *Physical Review B*, 63(22):224115, 2001. doi: [10.1103/PhysRevB.63.224115](https://doi.org/10.1103/PhysRevB.63.224115).
- [57] J. S. Dugdale and D. K. C. MacDonald. The Thermal Expansion of Solids. *Physical Review*, 89(4):832–834, 1953. doi: [10.1103/PhysRev.89.832](https://doi.org/10.1103/PhysRev.89.832).
- [58] C. H. Xu, C. Z. Wang, C. T. Chan, and K. M. Ho. Theory of the thermal expansion of Si and diamond. *Physical Review B*, 43(6):5024–5027, 1991. doi: [10.1103/PhysRevB.43.5024](https://doi.org/10.1103/PhysRevB.43.5024).
- [59] A. Franceschetti. First-principles calculations of the temperature dependence of the band gap of Si nanocrystals. *Phys. Rev. B*, 76(16):161301(R), 2007. doi: [10.1103/PhysRevB.76.161301](https://doi.org/10.1103/PhysRevB.76.161301).
- [60] R. Ramírez, C. P. Herrero, E. R. Hernández, and M. Cardona. Path-integral molecular dynamics simulation of 3C-SiC . *Phys. Rev. B*, 77(4):045210, 2008. doi: [10.1103/PhysRevB.77.045210](https://doi.org/10.1103/PhysRevB.77.045210).

- [61] A. Kundu, M. Govoni, H. Yang, M. Ceriotti, F. Gygi, and G. Galli. Quantum vibronic effects on the electronic properties of solid and molecular carbon. *Phys. Rev. Materials*, 5(7):L070801, 2021. doi: [10.1103/PhysRevMaterials.5.L070801](https://doi.org/10.1103/PhysRevMaterials.5.L070801).
- [62] B. Monserrat, R. J. Needs, and C. J. Pickard. Temperature effects in first-principles solid state calculations of the chemical shielding tensor made simple. *The Journal of Chemical Physics*, 141(13):134113, 2014. doi: [10.1063/1.4897261](https://doi.org/10.1063/1.4897261).
- [63] B. Monserrat and R. J. Needs. Comparing electron-phonon coupling strength in diamond, silicon, and silicon carbide: First-principles study. *Phys. Rev. B*, 89(21):214304, 2014. doi: [10.1103/PhysRevB.89.214304](https://doi.org/10.1103/PhysRevB.89.214304).
- [64] B. Monserrat. Electron-phonon coupling from finite differences. *J. Phys.: Condens. Matter*, 30(8):083001, 2018. doi: [10.1088/1361-648X/aaa737](https://doi.org/10.1088/1361-648X/aaa737).
- [65] M. Zacharias and F. Giustino. Theory of the special displacement method for electronic structure calculations at finite temperature. *Phys. Rev. Research*, 2(1):013357, 2020. doi: [10.1103/PhysRevResearch.2.013357](https://doi.org/10.1103/PhysRevResearch.2.013357).
- [66] A. Marini. Ab initio finite-temperature excitons. *Phys. Rev. Lett.*, 101(10):106405, 2008. doi: [10.1103/PhysRevLett.101.106405](https://doi.org/10.1103/PhysRevLett.101.106405).
- [67] P. B. Allen and V. Heine. Theory of the temperature dependence of electronic band structures. *Journal of Physics C: Solid State Physics*, 9(12):2305, 1976. doi: [10.1088/0022-3719/9/12/013](https://doi.org/10.1088/0022-3719/9/12/013).
- [68] P. B. Allen and M. Cardona. Theory of the temperature dependence of the direct gap of germanium. *Physical Review B*, 23(4):1495–1505, 1981. doi: [10.1103/PhysRevB.23.1495](https://doi.org/10.1103/PhysRevB.23.1495).
- [69] F. Giustino, S. G. Louie, and M. L. Cohen. Electron-phonon renormalization of the direct band gap of diamond. *Phys. Rev. Lett.*, 105(26):265501, 2010. doi: [10.1103/PhysRevLett.105.265501](https://doi.org/10.1103/PhysRevLett.105.265501).
- [70] S. Ponc e, G. Antonius, Y. Gillet, P. Boulanger, J. Laffamme Janssen, A. Marini, M. C ot e, and X. Gonze. Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation. *Phys. Rev. B*, 90(21):214304, 2014. doi: [10.1103/PhysRevB.90.214304](https://doi.org/10.1103/PhysRevB.90.214304).
- [71] A. Miglio, V. Brousseau-Couture, E. Godbout, G. Antonius, Y. H. Chan, S. G. Louie, M. C ot e, M. Giantomassi, and X. Gonze. Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap. *npj Computational Materials*, 6:167, 2020. doi: [10.1038/s41524-020-00434-z](https://doi.org/10.1038/s41524-020-00434-z).
- [72] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12):121102(R), 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).
- [73] S. Ponc e, Y. Gillet, J. L. Janssen, A. Marini, M. Verstraete, and X. Gonze. Temperature dependence of the electronic structure of semiconductors and insulators. *The Journal of Chemical Physics*, 143:102813, 2015. doi: [10.1063/1.4927081](https://doi.org/10.1063/1.4927081).

- [74] G. Mahan. *Many-Particle Physics*. Plenum, New York, 2nd edition, 1990.
- [75] A. Marini, S. Poncé, and X. Gonze. Many-body perturbation theory approach to the electron-phonon interaction with density-functional theory as a starting point. *Phys. Rev. B*, 91(22):224310, 2015. doi: [10.1103/PhysRevB.91.224310](https://doi.org/10.1103/PhysRevB.91.224310).
- [76] P. Coleman. *Introduction to Many-Body Physics*. Cambridge University Press, Cambridge, 2015. doi: [10.1017/CBO9781139020916](https://doi.org/10.1017/CBO9781139020916).
- [77] G. Antonius, S. Poncé, P. Boulanger, M. Côté, and X. Gonze. Many-Body Effects on the Zero-Point Renormalization of the Band Structure. *Phys. Rev. Lett.*, 112(21):215501, 2014. doi: [10.1103/PhysRevLett.112.215501](https://doi.org/10.1103/PhysRevLett.112.215501).
- [78] X. Gonze, P. Boulanger, and M. Côté. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Annalen der Physik*, 523(1-2):168–178, 2011. doi: [10.1002/andp.201000100](https://doi.org/10.1002/andp.201000100).
- [79] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J. M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J. B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. L. Marques, A. Martin, C. Martins, H. P. C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G. M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. Van Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger. The ABINIT project: Impact, environment and recent developments. *Computer Physics Communications*, 248:107042, 2020. doi: [10.1016/j.cpc.2019.107042](https://doi.org/10.1016/j.cpc.2019.107042).
- [80] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese. The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Computer Physics Communications*, 226:39–54, 2018. doi: [10.1016/j.cpc.2018.01.012](https://doi.org/10.1016/j.cpc.2018.01.012).
- [81] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, 1996. doi: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [82] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244–13249, 1992. doi: [10.1103/PhysRevB.45.13244](https://doi.org/10.1103/PhysRevB.45.13244).
- [83] M. Fuchs and M. Scheffler. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory. *Computer Physics Communications*, 119(1):67–98, 1999. doi: [https://doi.org/10.1016/S0010-4655\(98\)00201-X](https://doi.org/10.1016/S0010-4655(98)00201-X).

- [84] X. Gonze and C. Lee. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Physical Review B*, 55(16):10355–10368, 1997. doi: [10.1103/PhysRevB.55.10355](https://doi.org/10.1103/PhysRevB.55.10355).
- [85] D. R. Hamann, X. Wu, K. M. Rabe, and D. Vanderbilt. Metric tensor formulation of strain in density-functional perturbation theory. *Phys. Rev. B*, 71(3):035117, 2005. doi: [10.1103/PhysRevB.71.035117](https://doi.org/10.1103/PhysRevB.71.035117).
- [86] R. McKinney, P. Gorai, S. Manna, E. Toberer, and V. Stevanović. Ionic vs. van der Waals Layered Materials: Identification and Comparison of Elastic Anisotropy. *Journal of Materials Chemistry A*, 6(32):15828–15838, 2018. doi: [10.1039/C8TA04933E](https://doi.org/10.1039/C8TA04933E).
- [87] H. Iwanaga, A. Kunishige, and S. Takeuchi. Anisotropic thermal expansion in wurtzite-type crystals. *Journal of Materials Science*, 35:2451–2452, 2000. doi: <https://doi.org/10.1023/A:1004709500331>.
- [88] S. Fahy, K. J. Chang, S. G. Louie, and M. L. Cohen. Pressure coefficients of band gaps of diamond. *Phys. Rev. B*, 35(11):5856–5859, 1987. doi: [10.1103/PhysRevB.35.5856](https://doi.org/10.1103/PhysRevB.35.5856).
- [89] R. R. Reeber and K. Wang. Thermal Expansion of β -SiC, GaP and InP. *MRS Proceedings*, 410:211, 1995. doi: [10.1557/PROC-410-211](https://doi.org/10.1557/PROC-410-211).
- [90] J. S. Kang, M. Li, H. Wu, H. Nguyen, and Y. Hu. Basic physical properties of cubic boron arsenide. *Applied Physics Letters*, 115(12):122103, 2019. doi: [10.1063/1.5116025](https://doi.org/10.1063/1.5116025).
- [91] L. S. Dubrovinsky and S. K. Saxena. Thermal Expansion of Periclase (MgO) and Tungsten (W) to Melting Temperatures. *Physics and Chemistry of Minerals*, 24(8):547–550, 1997. doi: [10.1007/s002690050070](https://doi.org/10.1007/s002690050070).
- [92] G. Fiquet, P. Richet, and G. Montagnac. High-temperature thermal expansion of lime, periclase, corundum and spinel. *Physics and Chemistry of Minerals*, 27(2):103–111, 1999. doi: [10.1007/s002690050246](https://doi.org/10.1007/s002690050246).
- [93] K. Wang and R. R. Reeber. Thermal Expansion Of GaN And AlN. *MRS Proceedings*, 482:863, 1997. doi: [10.1557/PROC-482-863](https://doi.org/10.1557/PROC-482-863).
- [94] C. Roder, S. Einfeldt, S. Figge, and D. Hommel. Temperature dependence of the thermal expansion of GaN. *Phys. Rev. B*, 72(8):085218, 2005. doi: [10.1103/PhysRevB.72.085218](https://doi.org/10.1103/PhysRevB.72.085218).
- [95] R. R. Reeber. Lattice parameters of ZnO from 4.2° to 296°K. *Journal of Applied Physics*, 41(13):5063–5066, 1970. doi: [10.1063/1.1658600](https://doi.org/10.1063/1.1658600).
- [96] Y. P. Varshni. Temperature Dependence of the Energy Gap in Semiconductors. *Physica*, 34(1):149–154, 1967. doi: [https://doi.org/10.1016/0031-8914\(67\)90062-6](https://doi.org/10.1016/0031-8914(67)90062-6).
- [97] L. He, F. Liu, G. Hautier, M. J. T. Oliveira, M. A. L. Marques, F. D. Vila, J. J. Rehr, G. M. Rignanese, and A. Zhou. Accuracy of generalized gradient approximation functionals for density-functional perturbation theory calculations. *Phys. Rev. B*, 89(6):064305, 2014. doi: [10.1103/PhysRevB.89.064305](https://doi.org/10.1103/PhysRevB.89.064305).

- [98] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Physical Review Letters*, 100(13):136406, 2008. doi: [10.1103/PhysRevLett.100.136406](https://doi.org/10.1103/PhysRevLett.100.136406).
- [99] Z. Li, G. Antonius, M. Wu, F. H. da Jornada, and S. G. Louie. Electron-phonon coupling from ab initio linear-response theory within the *GW* method: Correlation-enhanced interactions and superconductivity in $\text{Ba}_{1-x}\text{K}_x\text{BiO}_3$. *Phys. Rev. Lett.*, 122(18):186402, 2019. doi: [10.1103/PhysRevLett.122.186402](https://doi.org/10.1103/PhysRevLett.122.186402).
- [100] Z. Li, M. Wu, Y. H. Chan, and S. G. Louie. Unmasking the origin of kinks in the photoemission spectra of cuprate superconductors. *Phys. Rev. Lett.*, 126(14):146401, 2021. doi: [10.1103/PhysRevLett.126.146401](https://doi.org/10.1103/PhysRevLett.126.146401).
- [101] A. Otero-de-la Roza and V. Luaña. Equations of state and thermodynamics of solids using empirical corrections in the quasiharmonic approximation. *Phys. Rev. B*, 84(18):184103, 2011. doi: [10.1103/PhysRevB.84.184103](https://doi.org/10.1103/PhysRevB.84.184103).
- [102] B. Monserrat, G. J. Conduit, and R. J. Needs. Extracting semiconductor band gap zero-point corrections from experimental data. *Phys. Rev. B*, 90(18):184302, 2014. doi: [10.1103/PhysRevB.90.184302](https://doi.org/10.1103/PhysRevB.90.184302).
- [103] R. W. Munn. Grüneisen functions for axial crystals. *Philosophical Magazine*, 17(146):433–435, 1968. doi: [10.1080/14786436808226173](https://doi.org/10.1080/14786436808226173).
- [104] R. Pässler. Dispersion-related description of temperature dependencies of band gaps in semiconductors. *Phys. Rev. B*, 66(8):085201, 2002. doi: [10.1103/PhysRevB.66.085201](https://doi.org/10.1103/PhysRevB.66.085201).
- [105] R. Pässler. Semi-empirical descriptions of temperature dependences of band gaps in semiconductors. *Physica Status Solidi (B)*, 236(3):710–728, 2003. doi: <https://doi.org/10.1002/pssb.200301752>.
- [106] A. T. Collins, S. C. Lawson, G. Davies, and H. Kanda. Indirect energy gap of ^{13}C diamond. *Phys. Rev. Lett.*, 65(7):891–894, 1990. doi: [10.1103/PhysRevLett.65.891](https://doi.org/10.1103/PhysRevLett.65.891).
- [107] D. Karaiskaj, M. Thewalt, T. Ruf, M. Cardona, and M. Konuma. Photoluminescence studies of isotopically enriched silicon: isotopic effects on the indirect electronic band gap and phonon energies. *Solid State Communications*, 123(3):87–92, 2002. doi: [https://doi.org/10.1016/S0038-1098\(02\)00249-1](https://doi.org/10.1016/S0038-1098(02)00249-1).
- [108] C. Parks, A. K. Ramdas, S. Rodriguez, K. M. Itoh, and E. E. Haller. Electronic band structure of isotopically pure germanium: Modulated transmission and reflectivity study. *Phys. Rev. B*, 49(20):14244–14250, 1994. doi: [10.1103/PhysRevB.49.14244](https://doi.org/10.1103/PhysRevB.49.14244).
- [109] F. Manjón, M. Mollar, B. Marí, N. Garro, A. Cantarero, R. Lauck, and M. Cardona. Effect of isotopic mass on the photoluminescence spectra of β zinc sulfide. *Solid State Communications*, 133(4):253–258, 2005. doi: <https://doi.org/10.1016/j.ssc.2004.10.031>.

- [110] J. M. Zhang, T. Ruf, R. Lauck, and M. Cardona. Isotope effects on exciton energies in CdS. *Phys. Rev. B*, 57(16):9716–9722, 1998. doi: [10.1103/PhysRevB.57.9716](https://doi.org/10.1103/PhysRevB.57.9716).
- [111] F. Manjón, M. Mollar, M. Hernández-Fenollosa, B. Marí, R. Lauck, and M. Cardona. Effect of isotopic mass on the photoluminescence spectra of zinc oxide. *Solid State Communications*, 128(1):35–39, 2003. doi: [https://doi.org/10.1016/S0038-1098\(03\)00616-1](https://doi.org/10.1016/S0038-1098(03)00616-1).

Chapitre 6

Effet du couplage spin-orbite sur la renormalisation du point zéro de l'énergie de la bande interdite des matériaux cubiques

6.0.1. Mise en contexte

Ce projet s'inscrit dans la continuité de l'article présenté au chapitre 4. Nous cherchions à expliquer la surestimation du ZPR obtenu pour le CdTe par rapport aux données expérimentales, que l'on a pu observer dans les figures 4.3 et 5.7. Puisque ce matériau est composé de deux éléments lourds, l'hypothèse de l'absence du couplage spin-orbite dans nos calculs s'est naturellement imposée. Si cette hypothèse s'avérait justifiée, il devenait alors primordial de vérifier si l'inclusion du SOC allait altérer l'accord global entre les résultats du ZPR prédits par la méthode AHC non adiabatique sans SOC et les données expérimentales.

En parcourant la littérature, nous avons de plus constaté que la communauté du calcul numérique *ab initio* omet fréquemment le couplage spin-orbite pour le calcul de l'interaction électron-phonon, sauf s'il est nécessaire pour décrire adéquatement la structure électronique du matériau dans l'approximation des noyaux statiques ou la dispersion des phonons. On peut penser, par exemple, à la famille des matériaux Rashba discutés à la section 2.5.4, tels que les tellurohalogénures de bismuth [1]. Ainsi, encore à ce jour, la dépendance en température de la structure électronique de matériaux de référence comme le Ge et le GaAs est le plus souvent étudiée et comparée aux données expérimentales en négligeant l'effet du SOC. Il est pourtant bien connu que l'inclusion du SOC dans le cadre de l'approximation des noyaux statiques lève les dégénérescences des états électroniques. Ainsi, sans même faire de calculs, on peut raisonnablement soupçonner que la séparation en énergie entre le VBM et les bandes *split-off* induite par le SOC va nécessairement affecter la self-énergie du couplage électron-phonon et, corollairement, la renormalisation des énergies électroniques.

Un des objectifs de ce projet était donc d'effectuer une analyse comparative du ZPR, afin d'identifier les mécanismes au travers desquels le SOC affecte l'interaction électron-phonon et de déceler des tendances générales qui pourraient ensuite permettre d'estimer d'une manière plus fiable l'erreur induite dans les résultats de calculs où on choisit d'omettre le SOC. Pour des matériaux monoatomiques, on peut raisonnablement s'attendre à ce que l'effet du SOC augmente avec la masse des noyaux¹. En ce sens, l'ensemble des composés binaires cubiques analysés dans l'article précédent s'avère particulièrement intéressant à étudier, puisqu'il contient différentes combinaisons de matériaux isoélectroniques. Ce faisant, nous pouvons extraire l'effet du SOC sur le cation et l'anion.

Dans un autre ordre d'idées, nous nous questionnions sur les conséquences de l'inclusion du SOC dans le modèle de Fröhlich généralisé développé au chapitre précédent. Puisque ce dernier implique une moyenne sur le sous-espace des bandes dégénérées du VBM, on peut d'ores et déjà déduire que la présence du SOC va réduire le nombre de bandes contribuant au ZPR. De même, il est bien connu que l'inclusion du SOC modifie la masse effective des électrons et des trous². On peut donc à juste titre se demander si la perte de dégénérescence et le changement des masses effectives au centre de la zone de Brillouin sont suffisants pour capturer la variation du ZPR obtenue à l'aide des principes premiers. Si cette hypothèse est vérifiée, nous serions dès lors en mesure d'estimer l'erreur sur le ZPR induite par l'absence de SOC par le biais d'un calcul simple.

Nos travaux ont permis de démontrer que la présence du SOC réduit le ZPR des matériaux cubiques et que cette diminution peut être presque entièrement attribuée à la perte de dégénérescence du VBM. De plus, l'inclusion du SOC au modèle de Fröhlich généralisé a mené à une discussion particulièrement intéressante sur la représentation de la structure électronique en termes de bandes paraboliques situées au centre de la zone de Brillouin pour le VBM des matériaux cubiques sans symétrie d'inversion.

6.0.2. Contributions des auteurs

J'ai mis sur pied ce projet et effectué l'ensemble des calculs et analyses qui y sont présentés. J'ai également assuré la rédaction de la première version du manuscrit. Xavier Gonze a fourni un soutien théorique et proposé l'utilisation du modèle analytique de Dresselhaus pour évaluer le modèle de Fröhlich généralisé du VBM des matériaux sans symétrie d'inversion. Michel Côté a suggéré d'analyser l'effet du SOC en termes de la norme des vecteurs d'onde des phonons et a assuré la supervision le projet.

1. Nous avons vu à la section 2.5.1 que l'hamiltonien du couplage spin-orbite est proportionnel à la charge des noyaux, qui augmente avec la masse des ions.

2. L'effet du SOC sur la masse effective des électrons et des trous a été largement étudié du point de vue théorique dans les années 1950, notamment par les modèles de Kane, Dresselhaus et Luttinger-Kohn.

Troisième article.

Effect of spin-orbit coupling on the zero-point renormalization of the electronic band gap in cubic materials : first principles and generalized Fröhlich model

par

Véronique Brousseau-Couture¹, Xavier Gonze² et Michel Côté³

- (¹) Département de physique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7
- (²) Institute of Condensed Matter and Nanosciences, UCLouvain, B-1348 Louvain-la-Neuve, Belgium
- (³) Département de physique, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7

Cet article a été soumis à Physical Review B.

RÉSUMÉ. La structure électronique des semi-conducteurs et des isolants est affectée par le mouvement des noyaux via l'interaction électron-phonon, qui confère une dépendance en température à l'énergie de la bande interdite et induit une renormalisation du point zéro (ZPR) à température nulle. Pour les matériaux polaires, la contribution la plus significative au ZPR de l'énergie de la bande interdite ZPR peut être interprétée en termes du modèle de Fröhlich, qui se concentre sur l'interaction non adiabatique entre un électron et la polarisation électrique macroscopique créée par un mode phononique longitudinal optique de grande longueur d'onde. D'autre part, le couplage spin-orbite (SOC) modifie la structure électronique dans l'approximation des noyaux statiques, ce qui, en retour, affectera l'interaction électron-phonon et le ZPR. Nous présentons une étude comparative de l'effet du SOC sur le ZPR de l'énergie de la bande interdite de vingt semi-conducteurs et isolants de symétrie cubique à l'aide de calculs basés sur les principes premiers. Nous observons une diminution du ZPR induite par le SOC pouvant atteindre 30%, laquelle est dominée par l'extrema de la bande de valence et provient presque entièrement de la modification des énergies propres électroniques statiques et de la diminution des masses effectives des trous au voisinage du point Γ . Nous incorporons également le SOC dans un modèle Fröhlich généralisé, en considérant la séparation Dresselhaus qui se produit dans les matériaux non centrosymétriques, et confirmons que la prédominance des effets non adiabatiques sur la renormalisation du point zéro de l'énergie de la bande interdite des matériaux polaires est inchangée lorsque le SOC est inclus. Notre modèle de Fröhlich généralisé incluant le SOC fournit une estimation fiable de la diminution de l'énergie de formation de polaron induite par le SOC obtenue à partir des principes premiers et met en lumière certaines subtilités fondamentales dans l'évaluation numérique des masses effectives en présence du SOC pour les matériaux non centrosymétriques. Nous mettons enfin en garde contre une éventuelle perte de validité de l'approximation parabolique, l'une des hypothèses centrales du modèle Fröhlich, à l'intérieur de la plage d'énergie où l'interaction Fröhlich est dominante, pour les matériaux possédant des fréquences phononiques élevées qui seraient traités avec SOC.

Mots clés : Théorie de la fonctionnelle de la densité perturbative, renormalisation du point zéro de l'énergie de la bande interdite, semi-conducteurs, modèle de Fröhlich, couplage spin-orbite, interaction électron-phonon

ABSTRACT. The electronic structure of semiconductors and insulators is affected by ionic motion through electron-phonon interaction, yielding temperature-dependent band gap energies and zero-point renormalization (ZPR) at absolute zero temperature. For polar materials, the most significant contribution to the band gap ZPR can be understood in terms of the Fröhlich model, which focuses on the non-adiabatic interaction between an electron and the macroscopic electrical polarization created by a long-wavelength optical longitudinal phonon mode. On the other hand, spin-orbit interaction (SOC) modifies the bare electronic structure, which will, in turn, affect the electron-phonon interaction and the ZPR. We present a comparative investigation of the effect of SOC on the band gap ZPR of twenty semiconductors and insulators with cubic symmetry using first-principles calculations. We observe a SOC-induced decrease of the ZPR, up to 30%, driven by the valence band edge, which almost entirely originates from the modification of the bare electronic eigenenergies and the decrease of the hole effective masses near the Γ point. We also incorporate SOC into a generalized Fröhlich model, addressing the Dresselhaus splitting which occurs in non-centrosymmetric materials, and confirm that the predominance of non-adiabatic effects on the band gap ZPR of polar materials is unchanged when including SOC. Our generalized Fröhlich model with SOC provides a reliable estimate of the SOC-induced decrease of the polaron binding energy obtained from first principles and brings to light some fundamental subtleties in the numerical evaluation of the effective masses with SOC for non-centrosymmetric materials. We finally warn about a possible breakdown of the parabolic approximation, one of the most fundamental assumptions of the Fröhlich model, within the physically relevant energy range of the Fröhlich interaction for materials with high phonon frequencies treated with SOC.

Keywords: Density-functional perturbation theory, zero-point renormalization of the band gap, semiconductors, Fröhlich model, spin-orbit coupling, electron-phonon interaction

6.1. Introduction

Electron-phonon interaction (EPI) has been widely investigated from a theoretical point of view since the late 1940s, through pioneering works of Pekar [2], Landau and Pekar [3], Fröhlich [4], and Feynman [5], and numerous subsequent works [6, 7]. Using model Hamiltonians, those first theories essentially addressed the interaction of an electron in an isotropic, continuous medium with the macroscopic polarization induced by longitudinal optical (LO) long-range lattice vibrations, which yields in a correlated state called a polaron. The Fröhlich model since became the foundation stone of modern large-polaron studies. From a first-principles point of view, the works of Allen, Heine and Cardona [8–10] (AHC) in the 1980s clarified earlier theories by Fan [11] and Antončík [12]. They provided a unified formalism for the EPI self-energy, rooted in many-body perturbation theory, which addresses all types of lattice vibrations.

Despite their fundamentally different perspectives, the model Hamiltonian and first-principles approaches address the same problem, namely, the consequences of EPI on the electronic structure. Amongst numerous effects on transport and optical properties of materials [7], EPI modifies the quasiparticle energy and introduces finite quasiparticle lifetimes,

which depend on the phonon population at a given temperature. As a consequence, the electronic structure does not only acquire a temperature dependence: it is affected even at absolute zero temperature, through the zero-point motion of the ions. This $T = 0$ K correction is known as the zero-point renormalization (ZPR). From the Fröhlich model perspective, the band edge ZPR corresponds to the polaron formation energy.

In recent years, considerable efforts have been directed towards tackling the Fröhlich interaction within the full complexity of real materials as captured by first-principles methods (see Ref. [13, 14] and references therein). Among others, Sio *et al.* [15, 16] developed a first-principles theory of polarons, later reformulated using a variational principle [17]. More recently, Lafuente-Bartolome *et al.* proposed a self-consistent many-body Green’s function theory which simultaneously addresses phonon-induced band structure renormalization and small polaron formation [18, 19]. From another perspective, Houtput and Tempere [20] derived anharmonic corrections to the Fröhlich Hamiltonian, and Kandolf *et al.* [21] and Macheda *et al.* [22] investigated the Fröhlich interaction in doped solids. Other works proposed models retaining certain fundamental assumptions of the original Fröhlich model while lifting some of its hypotheses. Schlipf *et al.* [23] addressed the case of multiple phonon branches, relying on the first-principles Fröhlich vertex proposed by Verdi and Giustino [24]. Miglio *et al.* [13] introduced a generalized Fröhlich model (gFr), based on a simplified electron-phonon vertex, that allows for multiple phonon branches, degenerate band extrema and anisotropic band warping. The authors used this model to reveal the predominance of non-adiabatic effects in the ZPR of semiconductors and insulators and explain why including such effects in calculations is essential to obtain an agreement between the first-principles band gap ZPR (ZPR_g) and experimental data. Their gFr model was recently used to obtain polaron effective masses and localization lengths in cubic materials [14], as well as to investigate the domain of applicability of the Fröhlich model using a high-throughput computational framework [25].

One question which remains unaddressed in Ref. [13] is the effect spin-orbit coupling (SOC). It is well known that SOC lifts the spin degeneracy of the Bloch states throughout the Brillouin zone, except at time-reversal invariant \mathbf{k} -points. For the valence band maximum (VBM) of cubic materials, which is triply-degenerate when neglecting SOC, this leads to a $4 + 2$ degeneracy: the two split-off bands are moved to lower energies compared to the heavy hole and light hole bands, which remain degenerate at the Γ point. This loss of degeneracy would directly affect the ZPR_g predicted by the gFr model. In addition to the electronic eigenvalues, the inclusion of a SOC term in the external potential of the first-principles Hamiltonian will also have repercussions on the first-order Hamiltonian perturbed by atomic displacements, which is a key quantity for computing the ZPR.

SOC has often been neglected throughout literature when investigating the Fröhlich interaction since strong polaronic effects are most likely to occur in materials where the

LO phonon frequency is large. Such systems typically contain light atoms (e.g. oxides), for which SOC can reasonably be expected to be weak. Some theoretical studies have addressed the consequences of SOC on EPI in 2D materials [26–28], mainly through Rashba-Holstein [29, 30] and Rashba-Fröhlich [31, 32] model Hamiltonians. To the best of our knowledge, only Trebin and Rössler [33] explicitly investigated the effect of SOC on the Fröhlich polaron for triply-degenerate band extrema in 3D materials. However, they relied on an isotropic model Hamiltonian, thus neglecting the effect of band warping.

From the first-principles perspective, density-functional perturbation theory calculations including SOC have been available for about 15 years [34, 35]. Other formalisms relying on finite differences and distorted supercells [36], as well as the recent special displacement method [37], have also been used to investigate this question. Nevertheless, SOC remains commonly neglected in ZPR calculations to this day. Full first-principles EPI calculations with SOC are typically done on a case-by-case basis [23, 38, 39]. Rashba materials [40], for which SOC is known to have a profound impact on either the electronic structure or the phonon frequencies, and topological materials [41, 42], in which SOC is necessary to induce the band inversion, have naturally been investigated by including SOC in first-principles EPI calculations. Some compound-specific comparative studies have been made, for example, in PbTe [43, 44], $\text{CH}_3\text{NH}_3\text{PbI}_3$ [45] and BAs [46], as well as when investigating the superconducting coupling constant [47–50]. Yet, even in the most simple case of cubic materials, the effect of SOC on EPI and the ZPR has not received the thorough investigation it deserves.

In this article, we investigate the effect of SOC on the ZPR of twenty semiconductors using the non-adiabatic AHC framework. We focus on representative cubic materials, as their well-characterized electronic structure provides a simple framework to investigate the mechanisms at play. Their triply-degenerate VBM also proves ideal to investigate the effect of SOC on the polaron formation energy of degenerate extrema within the gFr model. We evaluate the first-principles ZPR with the AHC methodology and extend the generalized Fröhlich model of Miglio *et al.* [13] to include SOC. First-principles calculations show that spin-orbit coupling reduces the zero-point renormalization of the valence band edge by 15%–30% for the heavier materials, e.g. the tellurides. We address the SOC-induced Dresselhaus splitting [51] occurring in non-centrosymmetric materials, which shifts the band extrema slightly away from its location without SOC in reciprocal space. The leading mechanism driving the observed SOC-induced decrease of the ZPR_g is found to be the variation of the electronic eigenenergies of the occupied bands and the loss of degeneracy at the valence band edge. We also confirm the claims of Miglio *et al.* [13] regarding the predominance of non-adiabatic effects in the ZPR_g of polar materials. We relate the results from the two approaches and bring to light some limitations of the approximations inherent to the gFr model when SOC is considered.

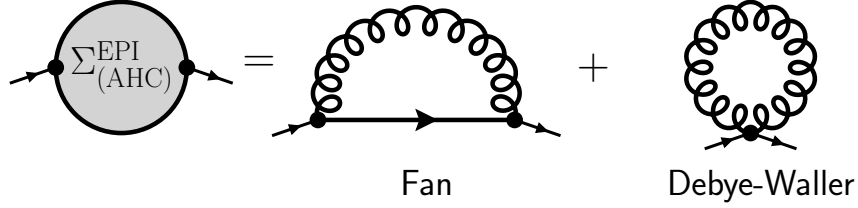


Figure 6.1 – Fan and Debye-Waller diagrams contributing to the AHC self-energy.

Section 6.2 presents an overview of the theoretical concepts used throughout this work. We first review the AHC formalism for EPI (Sec. 6.2.1), then briefly discuss some key consequences of SOC in the first-principles perspective (Sec. 6.2.2) before demonstrating how to incorporate SOC into the gFr model of Ref. [13] (Sec. 6.2.3) and investigating the consequences of Dresselhaus splitting on our results (Sec. 6.2.4). Section 6.3 provides the relevant technical details regarding our calculations. We respectively analyze our first-principles and gFr model results in Sec. 6.4.1 and 6.4.2, then summarize our findings in Sec. 6.5.

6.2. Methodology

6.2.1. AHC formalism

In the following, we briefly summarize the key concepts of the non-adiabatic Allen-Heine-Cardona (AHC) framework [8–10]. We work with the Hartree atomic unit system, such that $\hbar = m_e = c = |e| = 1$.

Within the many-body perturbation theory formalism, the electron-phonon interaction at temperature T affects the electronic Green’s function through a frequency-dependent electron-phonon self-energy, $\Sigma_{\mathbf{k}n}(\omega, T)$, where \mathbf{k} and n are respectively the electron wavevector and band index. At the lowest order of perturbation, known as AHC theory [10], the self-energy contains two terms, called the Fan and Debye-Waller (DW) contributions:

$$\Sigma_{\mathbf{k}n}^{\text{AHC}}(\omega, T) = \Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega, T) + \Sigma_{\mathbf{k}n}^{\text{DW}}(T). \quad (6.2.1)$$

The dynamical Fan self-energy contains two first-order vertices treated at second order in perturbation theory, while the static Debye-Waller self-energy has one second-order vertex treated at first order in perturbation theory. The Feynman diagrams corresponding to these contributions are shown in Fig. 6.1.

Note that we implicitly suppose that the full self-energy matrix can be approximated by its diagonal counterpart, i.e. $\Sigma_{\mathbf{k}nn'} \propto \delta_{nn'}$. The non-diagonal contributions hybridize the unperturbed electronic eigenstates within the interacting Green’s function [52] and become important when the band gap nearly vanishes. These can be safely neglected here as we work with semiconductors and insulators.

Within this framework, the temperature dependence of an electronic eigenstate with eigenvalue $\varepsilon_{\mathbf{k}n}$ then reads

$$\varepsilon_{\mathbf{k}n}(T) = \Re \left[\Sigma_{\mathbf{k}n}^{\text{AHC}}(\omega = \varepsilon_{\mathbf{k}n}(T), T) \right] + \varepsilon_{\mathbf{k}n}^0. \quad (6.2.2)$$

From this point, we work exclusively at $T = 0$ K. The ZPR of an electronic eigenstate $|\mathbf{k}n\rangle$ is obtained from Eq. (6.2.2),

$$\text{ZPR}_{\mathbf{k}n} = \varepsilon_{\mathbf{k}n}(T = 0) - \varepsilon_{\mathbf{k}n}^0, \quad (6.2.3)$$

while the band gap ZPR is the difference between the ZPR of the conduction and valence band edges (resp. ZPR_c and ZPR_v),

$$\text{ZPR}_g = \text{ZPR}_c - \text{ZPR}_v. \quad (6.2.4)$$

We apply the on-the-mass-shell approximation to Eq. (6.2.1), thus evaluating the Fan self-energy at the poles of the Green's function, namely, at the bare electronic eigenvalue, $\varepsilon_{\mathbf{k}n}^0$,

$$\Sigma_{\mathbf{k}n}^{\text{Fan}}(\varepsilon_{\mathbf{k}n}(T = 0), T = 0) \approx \Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega = \varepsilon_{\mathbf{k}n}^0, T = 0). \quad (6.2.5)$$

Furthermore approximating the interacting electronic Green's function by the noninteracting Kohn-Sham wavefunction obtained from density-functional theory (DFT), one obtains the standard expression for the non-adiabatic Fan self-energy [7],

$$\begin{aligned} \Sigma_{\mathbf{k}n}^{\text{Fan}}(\varepsilon_{\mathbf{k}n}^0, T = 0) &= \sum_{\mathbf{q}\nu} \sum_{n'}^{\text{BZ}} |\langle \mathbf{k} + \mathbf{q}n' | \nabla_{\mathbf{q}\nu} V^{\text{KS}} | \mathbf{k}n \rangle|^2 \\ &\times \left[\frac{1 - f_{\mathbf{k}+\mathbf{q}n'}}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 - \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} + \frac{f_{\mathbf{k}+\mathbf{q}n'}}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 + \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} \right]. \end{aligned} \quad (6.2.6)$$

The contributions of all phonon modes with frequency $\omega_{\mathbf{q}\nu}$ are summed for all wavevector \mathbf{q} and branch index ν in the Brillouin zone (BZ). In Eq. (6.2.6) and throughout this work, all phonon modes summations are implicitly normalized by the number of phonon wavevectors used to sample the Brillouin zone. Since we work at $T = 0$ K, the Fermi-Dirac occupation functions, $f_{\mathbf{k}n}$, are either 1 for the occupied states or 0 for the conduction bands. The small imaginary parameter $\eta_{\mathbf{k}} = \eta \text{sgn}(\varepsilon_{\mathbf{k}n}^0 - \mu)$, with μ the chemical potential and η real and positive, shifts the poles of the Green's function in the complex plane to maintain causality. Without SOC, the electronic bands are implicitly spin degenerate.

The electron-phonon matrix elements squared,

$$\left| g_{\mathbf{k}n n'}^{\text{Fan}}(\mathbf{q}\nu) \right|^2 \triangleq |\langle \mathbf{k} + \mathbf{q}n' | \nabla_{\mathbf{q}\nu} V^{\text{KS}} | \mathbf{k}n \rangle|^2, \quad (6.2.7)$$

capture the probability that an electron in eigenstate $\varepsilon_{\mathbf{k}n}^0$ interacts with a $\mathbf{q}\nu$ -phonon, given the self-consistent first-order variation of the Kohn-Sham potential (labelled with superscript ‘‘KS’’) induced by the collective atomic motion along this phonon mode [7, 53]. The operator

$\nabla_{\mathbf{q}\nu}$ expressed in the position basis can be written as

$$\begin{aligned}\nabla_{\mathbf{q}\nu} &= \frac{1}{\sqrt{2\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha} U_{\nu,\kappa\alpha}(\mathbf{q}) \sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial}{\partial \mathbf{R}_{l\kappa\alpha}} \\ &= \frac{1}{\sqrt{2\omega_{\mathbf{q}\nu}}} \sum_{\kappa\alpha} U_{\nu,\kappa\alpha}(\mathbf{q}) \partial_{\kappa\alpha}(\mathbf{q}),\end{aligned}\quad (6.2.8)$$

where $\mathbf{R}_{l\kappa\alpha}$ denotes the displacement of atom κ , located in unit cell l , in cartesian direction α . The phonon eigendisplacement vector, $U_{\nu,\kappa\alpha}(\mathbf{q})$, verifies the generalized eigenvalue equation

$$M_\kappa \omega_{\mathbf{q}\nu}^2 U_{\nu,\kappa\alpha}(\mathbf{q}) = \sum_{\kappa'\alpha'} \Phi_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q}) U_{\nu,\kappa'\alpha'}(\mathbf{q}) \quad (6.2.9)$$

and the normalization condition

$$\sum_{\kappa\alpha} M_\kappa U_{\nu,\kappa\alpha}^*(\mathbf{q}) U_{\nu',\kappa\alpha}(\mathbf{q}) = \delta_{\nu\nu'}, \quad (6.2.10)$$

where M_κ is the atomic mass of atom κ . The dynamical matrix, $\Phi_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q})$, is the Fourier transform of the second derivative of the total energy with respect to two atomic displacements,

$$\Phi_{\kappa\kappa'}^{\alpha\alpha'}(\mathbf{q}) = \sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial^2 E}{\partial \mathbf{R}_{l\kappa\alpha} \partial \mathbf{R}_{0\kappa'\alpha'}}. \quad (6.2.11)$$

For its part, the Debye-Waller self-energy is formally defined as [7]

$$\Sigma_{\mathbf{k}n}^{\text{DW}} = \sum_{\mathbf{q}\nu} \frac{1}{2} \langle \mathbf{k}n | \nabla_{\mathbf{q}\nu} \nabla_{-\mathbf{q}\nu} V^{\text{KS}} | \mathbf{k}n \rangle. \quad (6.2.12)$$

The direct evaluation of the second-order derivative of the Kohn-Sham potential with respect to atomic displacements entering Eq. (6.2.12) is a computational bottleneck in the density-functional perturbation theory approach. By applying the rigid-ion approximation, i.e. assuming that the potentials created by each nucleus are independent of each other, one can replace the second-order derivatives by the same first-order derivatives entering $\Sigma_{\mathbf{k}n}^{\text{Fan}}$ [53], yielding

$$\Sigma_{\mathbf{k}n}^{\text{DW,RIA}}(T=0) = \sum_{\mathbf{q}\nu} \sum_{n' \neq n}^{\text{BZ}} -\frac{1}{4\omega_{\mathbf{q}\nu}} \frac{|g_{\mathbf{k}nn'}^{\text{DW}}(\mathbf{q}\nu)|^2}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}n'}^0 + i\eta}, \quad (6.2.13)$$

where RIA stands for rigid ion approximation and where

$$\begin{aligned}\left| g_{\mathbf{k}nn'}^{\text{DW}}(\mathbf{q}\nu) \right|^2 &= \sum_{\kappa\kappa'} \sum_{\alpha\alpha'} [U_{\nu,\kappa\alpha}(\mathbf{q}) U_{\nu,\kappa\alpha'}(\mathbf{q})^* + U_{\nu,\kappa'\alpha}(\mathbf{q}) U_{\nu,\kappa'\alpha'}(\mathbf{q})^*] \\ &\times \langle \mathbf{k}n | V_{\kappa\alpha}^{(1)}(0)^* | \mathbf{k} + \mathbf{q}n' \rangle \langle \mathbf{k} + \mathbf{q}n' | V_{\kappa'\alpha'}^{(1)}(0) | \mathbf{k}n \rangle,\end{aligned}\quad (6.2.14)$$

with

$$V_{\kappa\alpha}^{(1)}(0) = \partial_{\kappa\alpha}(\mathbf{q} = \mathbf{0}) V^{\text{KS}}, \quad (6.2.15)$$

following the definition of the operator $\partial_{\kappa\alpha}(\mathbf{q})$ in the second line Eq. (6.2.8). The consequences of the rigid-ion approximation on the ZPR have been discussed in Ref. [53] for crystals and in Ref. [54] for molecules.

6.2.2. Spin-orbit interaction

We now examine how SOC can affect Eq. (6.2.6) and (6.2.13). Upon inclusion of SOC, the electronic wavefunction becomes a spinor,

$$|\mathbf{k}n\rangle = \begin{pmatrix} |\mathbf{k}n \uparrow\rangle \\ |\mathbf{k}n \downarrow\rangle \end{pmatrix}, \quad (6.2.16)$$

and the Hamiltonian, a 2×2 matrix,

$$\hat{H}_{\mathbf{k}} = \begin{pmatrix} H_{\mathbf{k}\uparrow\uparrow} & H_{\mathbf{k}\uparrow\downarrow} \\ H_{\mathbf{k}\downarrow\uparrow} & H_{\mathbf{k}\downarrow\downarrow} \end{pmatrix}. \quad (6.2.17)$$

In real space, the general form of the SOC contribution to the electronic Hamiltonian writes [51]

$$\hat{H}^{\text{SOC}}(\mathbf{r}) = \frac{1}{4} (\nabla V(\mathbf{r}) \times \hat{\mathbf{P}}) \cdot \boldsymbol{\sigma}, \quad (6.2.18)$$

where $\hat{\mathbf{P}}$ is the momentum operator and $\boldsymbol{\sigma}$ are the Pauli matrices.

For a plane wave basis set and norm-conserving pseudopotentials, SOC only enters the Hamiltonian through the electron-ion term. Assuming that the pseudopotentials are fully separable and substituting the Coulomb potential in Eq. (6.2.18), one recovers the typical $\mathbf{L} \cdot \mathbf{S}$ term from introductory quantum mechanics. For a single atom, one gets [34]

$$V^{\text{e-ion}}(\mathbf{r}, \mathbf{r}') = \sum_l V_l^{\text{SR}}(r, r') |ls\rangle \langle ls| + \sum_l V_l^{\text{SOC}}(r, r') \mathbf{L} \cdot \mathbf{S} |ls\rangle \langle ls|, \quad (6.2.19)$$

where $V_l^{\text{SR}}(r, r')$ and $V_l^{\text{SOC}}(r, r')$ follow the Kleinman-Bylander construction [55],

$$V_l^x = f_l^x(r) E_l^{\text{KB},x} f_l^x(r'), \quad x \in \{\text{SR}, \text{SOC}\}, \quad (6.2.20)$$

where $E_l^{\text{KB},x}$ is the Kleinman-Bylander energy [56]. SR stands for the scalar-relativistic contribution to the electron-ion potential (hence, without SOC), and $|ls\rangle \langle ls|$ is the projector on the tensor product subspace of angular momentum L and spin S , which has dimension $2(2l+1)$. Detailed expressions for V_l^{SR} and V_l^{SOC} can be found in Ref. [57]. No magnetism is considered, such that the electronic density is given by a single scalar function, $\rho(\mathbf{r})$.

The consequences of SOC on the explicit density-functional perturbation theory equations have been derived in Ref. [34] and [35] for norm-conserving pseudopotentials. In our case, the general form of the equations presented in Sec. 6.2.1 remain unchanged, but all the relevant physical quantities, i.e. $\omega_{\mathbf{q}\nu}$, $\varepsilon_{\mathbf{k}n}^0$ and the electron-phonon matrix elements squared, Eqs. (6.2.7) and (6.2.14), now capture the effect of SOC. There is no implicit sum on the spin

degree of freedom, as the spinorial electronic wavefunctions mix the spin-up and spin-down components.

6.2.3. Generalized Fröhlich model

In the following, we discuss how to incorporate SOC into the generalized Fröhlich model developed in Ref. [13]. For completeness, we start by reviewing the key elements of this model. First neglecting SOC, the Hamiltonian at the first order of interaction writes:

$$H = \sum_{\mathbf{k}n\sigma} \frac{\theta k^2}{2m_n^*(\hat{\mathbf{k}})} c_{\mathbf{k}n\sigma}^\dagger c_{\mathbf{k}n\sigma} + \sum_{\mathbf{q}j} \omega_{\mathbf{0}j}(\hat{\mathbf{q}}) \left(a_{\mathbf{q}j}^\dagger a_{\mathbf{q}j} + \frac{1}{2} \right) + \sum_{\mathbf{k}n n' \sigma} \sum_{\mathbf{q}j} g_{\mathbf{k}n n'}^{\text{gFr}}(\mathbf{q}j) c_{\mathbf{k}+\mathbf{q}n'\sigma}^\dagger c_{\mathbf{k}n\sigma} \left(a_{\mathbf{q}j} + a_{-\mathbf{q}j}^\dagger \right). \quad (6.2.21)$$

The first term corresponds to parabolic bare electronic eigenenergies $\varepsilon_{\mathbf{k}n}$ with direction-dependent effective mass $m_n^*(\hat{\mathbf{k}})$, while the second term allows for multiple phonon branches j with direction-dependent Einstein frequency $\omega_{\mathbf{0}j}(\hat{\mathbf{q}})$, evaluated at the zone center Γ . The $c_{\mathbf{k}n\sigma}^\dagger$, $c_{\mathbf{k}n\sigma}$, $a_{\mathbf{q}j}^\dagger$, $a_{\mathbf{q}j}$ are respectively the creation and annihilation operators for electrons and phonons, while $\hat{\mathbf{k}}$ and $\hat{\mathbf{q}}$ are unit vectors. The parameter θ gives the sign of the effective mass: $\theta = -1$ for the hole-like bands and $\theta = 1$ for the electron-like bands. The sum on spin index σ implies that all electronic states are doubly degenerate.

The last term couples the electron and phonon subsystems through the Fröhlich interaction, with matrix element

$$g_{\mathbf{k}n n'}^{\text{gFr}}(\mathbf{q}j) = \frac{1}{q} \frac{4\pi}{\Omega_0} \left(\frac{1}{2\omega_{\mathbf{0}j}(\hat{\mathbf{q}}) V_{\text{BvK}}} \right)^{1/2} \frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \sum_m s_{n'm}(\hat{\mathbf{k}}') s_{nm}^*(\hat{\mathbf{k}}), \quad (6.2.22)$$

where $\mathbf{k}' = \mathbf{k} + \mathbf{q}$, Ω_0 is the primitive unit cell volume, V_{BvK} is the Born-von Karman normalization volume associated with the \mathbf{k} and \mathbf{q} samplings, and ϵ^∞ is the macroscopic optic dielectric constant, obtained from the dielectric tensor,

$$\epsilon^\infty(\hat{\mathbf{q}}) = \sum_{\alpha\beta} \hat{q}_\alpha \epsilon_{\alpha\beta}^\infty \hat{q}_\beta. \quad (6.2.23)$$

Here, $\mathbf{p}_j(\hat{\mathbf{q}})$ is the mode-polarity vector of the j -phonon mode [13], constructed from the Born effective charges, $Z_{\kappa\alpha,\alpha'}^*$, and the phonon eigendisplacement vectors, $U_{j,\kappa\alpha}(\mathbf{q})$, summing over all Cartesian directions α and all atoms κ in the unit cell,

$$p_{j,\alpha'}(\hat{\mathbf{q}}) = \lim_{q \rightarrow 0} \sum_{\kappa\alpha} Z_{\kappa\alpha,\alpha'}^* U_{j,\kappa\alpha}(q\hat{\mathbf{q}}). \quad (6.2.24)$$

Note also that our formulation of the Fröhlich matrix element, Eq. (6.2.22), relies on the Born and Huang convention for the phonon eigenvectors [58], which implies the following relation:

$$U_{j,\kappa\alpha}(-\mathbf{q}) = U_{j,\kappa\alpha}^*(\mathbf{q}), \quad (6.2.25)$$

such that Eq. (6.2.21) is hermitian. See Ref. [59] for a thorough discussion of the different phase conventions in the literature.

The unitary matrix $s_{nm}(\hat{\mathbf{k}})$ describes the direction-dependent overlap between the electronic states at the band extrema located at Γ and states along the \mathbf{k} direction, in the $k \rightarrow 0$ limit, computed from the periodic part of the wavefunction (indicated by the subscript P):

$$s_{nm}(\hat{\mathbf{k}}) = \lim_{k \rightarrow 0} \langle k\hat{\mathbf{k}}n | \Gamma m \rangle_P. \quad (6.2.26)$$

While we set the band extrema at Γ for convenience, the previous definition allows for a band extrema located at any wavevector in the Brillouin zone.

In all previous expressions, the sums on electronic bands index n , n' and m run on the degenerate subset of bands connected to extrema, thus allowing interband couplings within this subset. The sum on phonon branches j is restricted to LO modes, as $\mathbf{p}_j(\hat{\mathbf{q}})$ is zero otherwise.

When SOC is considered, the Hamiltonian is no longer diagonal in spin space. We can, however, define new electronic creation and annihilation operators, $\tilde{c}_{\mathbf{k}n}^\dagger$ and $\tilde{c}_{\mathbf{k}n}$, such that the electronic part of the Hamiltonian can be written as

$$H^{\text{el}} = \sum_{\mathbf{k}n} \varepsilon_{\mathbf{k}n}^{\text{SOC}} \tilde{c}_{\mathbf{k}n}^\dagger \tilde{c}_{\mathbf{k}n}. \quad (6.2.27)$$

In order to formulate a Fröhlich Hamiltonian for the SOC case, we will suppose that the relevant part of the electronic structure is at a band extremum, with quadratic departure from the extremal eigenenergy as a function of the wavevector. This is the same hypothesis as for the generalized Fröhlich model without SOC. Generally speaking, this hypothesis is correct when the band extrema are non-degenerate (except for the spin degeneracy) when the SOC is not present. It will hold also when the starting band extremum is degenerate, provided the typically spin-orbit coupling energy is much bigger than the phonon energy, so that, after applying the SOC, one is left with a new band extremum with quadratic departure of the eigenenergy in a sufficiently large zone, where the phonon energy is relevant.

Supposing this hypothesis to be valid, we take the extremum eigenvalue as zero of energy, and expand the eigenvalue as

$$\varepsilon_{\mathbf{k}n}^{\text{SOC}} = \frac{\theta k^2}{2\tilde{m}_n^*(\hat{\mathbf{k}})}, \quad (6.2.28)$$

which captures the modification of the electronic effective masses near the band extrema induced by SOC, $\tilde{m}_n^*(\hat{\mathbf{k}})$ replacing $m_n^*(\hat{\mathbf{k}})$. We also neglect spin-phonon interaction, thus assuming that SOC affects the vibrational properties through electronic properties only. Within these assumptions, we recover an expression identical to Eq. (6.2.21), in which the sum on σ has been absorbed inside the new electronic operators. The starting point of Ref. [13] can therefore be taken as implicitly incorporating the effects of SOC on the electronic

and vibrational properties. From now on, we simplify the notation by dropping all tilde on electronic quantities which include SOC, i.e. $\widetilde{m}_n^*(\hat{\mathbf{k}}) \rightarrow m_n^*(\hat{\mathbf{k}})$.

We now follow the same procedure as described in Section 5 of the Supplementary Notes of Ref. [13]: we substitute $g_{\mathbf{k}n'n}^{\text{gFr}}(\mathbf{q}_j)$ for the matrix elements in the general expression for $\Sigma_{\mathbf{k}n}^{\text{Fan}}$ (Eq. (6.2.6)) and, as per the original Fröhlich model, take the continuum macroscopic limit, replacing the discrete sum over \mathbf{q} by an integral over the \mathbf{q} coordinate ($\sum_{\mathbf{q}} f(\mathbf{q})/V_{\text{BvK}} \rightarrow \Omega_0/(2\pi)^3 \int d^3\mathbf{q} f(\mathbf{q})$), thus extending the Brillouin zone boundaries to infinity. Since we only consider interband contributions within the degenerate subset of bands connected to the extrema, only the second (first) term inside the brackets of Eq. (6.2.6) contribute for the hole-like (electron-like) bands. Taking the $q \rightarrow 0$ limit of the denominator for purely parabolic electronic bands, we are left with

$$\begin{aligned} \text{ZPR}_{n_\theta}^{\text{gFr}} = & -\frac{\theta}{\pi\Omega_0} \int d^3\mathbf{q} \sum_{jn} \frac{1}{q^2} \frac{|s_{n,n_\theta}(\hat{\mathbf{q}})|^2}{\omega_{\mathbf{0}j}(\hat{\mathbf{q}})} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2 \\ & \times \frac{1}{\frac{q^2}{2m_n^*(\hat{\mathbf{q}})} + \omega_{\mathbf{0}j}(\hat{\mathbf{q}})}, \end{aligned} \quad (6.2.29)$$

where $n_{\theta=-1}$ is the band index of the VBM and $n_{\theta=1}$, that of the conduction band minimum (CBM). Using spherical coordinates, the radial part of this three-dimensional integral has an analytic solution of the form

$$\int_0^\infty dq \frac{1}{C_1 q^2 + C_2} = \frac{1}{\sqrt{C_1 C_2}} \frac{\pi}{2}, \quad (6.2.30)$$

where the parameters $C_1 = (2m_n^*(\hat{\mathbf{q}}))^{-1}$ and $C_2 = \omega_{\mathbf{0}j}(\hat{\mathbf{q}})$ are positive. Recall that, for the VBM, the negative curvature of the electronic bands is parametrized by θ . This yields

$$\text{ZPR}_{n_\theta}^{\text{gFr}} = -\frac{\theta}{\sqrt{2}\Omega_0} \oint_{4\pi} d\hat{\mathbf{q}} \sum_{jn} |s_{n,n_\theta}(\hat{\mathbf{q}})|^2 \frac{(m_n^*(\hat{\mathbf{q}}))^{1/2}}{\omega_{\mathbf{0}j}(\hat{\mathbf{q}})^{3/2}} \times \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2. \quad (6.2.31)$$

With the previous definition of θ , we thus obtain a positive (negative) ZPR for the VBM (CBM).

Up to this point, no special treatment was made to consider SOC in the treatment of the Fröhlich interaction outside incorporating it implicitly in the static electronic and vibrational properties, i.e. $m_n^*(\hat{\mathbf{q}})$, $\omega_{\mathbf{0}j}(\hat{\mathbf{q}})$, $\mathbf{p}_j(\hat{\mathbf{q}})$ and $\epsilon^\infty(\hat{\mathbf{q}})$ are computed with SOC.

We finally argue that the treatment of $|s_{n,n_\theta}(\hat{\mathbf{q}})|^2$ based on the point group symmetry argument detailed in the Supplementary Information of Ref. [13] remains valid in the presence of SOC. For this paper, we will treat the $3 \times 2 \rightarrow 4 + 2$ degeneracy arising from a cubic space group, taking the VBM of cubic materials as a typical example. The argument could be generalized to any space group symmetry using group theory.

As the degeneracy arises from symmetry, i.e. it is not accidental, the degenerate electronic wavefunctions at the extrema can be decomposed in a basis of orthonormal eigenfunctions

that form an irreducible representation of the symmetry group, $\mathcal{G} = T_d$ for the zincblende structure and $\mathcal{G} = O_h$ for the diamond structure. Without SOC, this basis contains three eigenfunctions, denoted $\{|X\rangle, |Y\rangle, |Z\rangle\}$, which each are doubly degenerate in the spin space. When considering SOC, the basis functions $\{|X \uparrow\rangle, |Y \uparrow\rangle, |Z \uparrow\rangle, |X \downarrow\rangle, |Y \downarrow\rangle, |Z \downarrow\rangle\}$ no longer form a good basis choice as they do not form an irreducible representation of the double group, $\mathcal{G} \otimes D_{1/2}$. One rather has to use linear combinations of those states, namely, the 4-fold $\{|j = 3/2\rangle\}$ states for the degenerate heavy hole and light hole bands, which form the VBM, and the 2-fold $\{|j = 1/2\rangle\}$ states should one wish to evaluate the ZPR for the split-off bands.

We now express the eigenstates entering the $s_{nm}(\hat{\mathbf{q}})$ overlap integrals (Eq. (6.2.26)) in this basis. The $|\Gamma v\rangle$ state, where $v = n_{\theta=-1}$ is the band index of the VBM, can be written as

$$|\Gamma v\rangle = \sum_{m \in \{\pm 3/2, \pm 1/2\}} u_{vm} \left| \frac{3}{2} m \right\rangle, \quad (6.2.32)$$

where u_{vm} is the coefficient of the basis function $|3/2 m\rangle$, and the $|q\hat{\mathbf{q}}n\rangle$ state becomes

$$\begin{aligned} |q\hat{\mathbf{q}}n\rangle &= \sum_{m \in \{\pm 3/2, \pm 1/2\}} \left| \frac{3}{2} m \right\rangle \langle \frac{3}{2} m | q\hat{\mathbf{q}}n \rangle, \\ &= \sum_{m \in \{\pm 3/2, \pm 1/2\}} s_{nm}(\hat{\mathbf{q}}) \left| \frac{3}{2} m \right\rangle, \end{aligned} \quad (6.2.33)$$

where $s_{nm}(\hat{\mathbf{q}})$ is the overlap integral with the basis function $|\frac{3}{2} m\rangle$. The $q \rightarrow 0$ limit is implied.

Substituting the last two equations in Eq. (6.2.31), we obtain an expression which is identical to the Supplementary Eq. (24) and (25) of Ref. [13], with the double sum on $m, m' \in \{X, Y, Z\}$ replaced by a double sum on $m, m' \in \{\pm 3/2, \pm 1/2\}$. The remaining of the argument thus holds, yielding a final expression for the ZPR^{gFr} which has the same form as their Eq. (6):

$$\text{ZPR}_{n_\theta}^{\text{gFr}} = - \sum_{jn} \frac{\theta}{\sqrt{2}\Omega_0 n_{\text{deg}}} \oint_{4\pi} d\hat{\mathbf{q}} \frac{(m_n^*(\hat{\mathbf{q}}))^{1/2}}{\omega_{0j}(\hat{\mathbf{q}})^{3/2}} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2, \quad (6.2.34)$$

in which n_{deg} is the degree of degeneracy of the band extrema in presence of SOC. As the n summation is made over degenerate states, the division by n_{deg} yields an average over degenerate states. Note that this last expression can be further simplified when applied to cubic systems, as in that case the phonon frequencies and mode-polarity vector do not depend on the wavevector orientation, and the dielectric tensor is isotropic (see Eq. (45) of Ref. [14]).

For materials whose vibrational properties are not significantly affected by SOC, such as cubic semiconductors and insulators, the main consequence of SOC on the ZPR^{gFr} is therefore the modification of the electronic effective masses. The reduced dimensionality of the irreducible representation at the extrema should not play a significant role on the

ZPR, as the overlap between the electronic wavefunctions belonging to the $\{|j = 1/2\rangle\}$ and $\{|j = 3/2\rangle\}$ states in the vicinity of the extrema can be expected to be small. Nevertheless, the consequences on the ZPR of this smaller number of states contributing to the average on the degenerate subspace remain to be investigated.

6.2.4. Generalized Fröhlich model in the presence of Dresselhaus splitting

In Sec. 6.2.3, we implicitly assumed that the location of the band extrema in reciprocal space is unchanged by the inclusion of SOC, i.e. it remains at the high-symmetry, degenerate \mathbf{k} -point. However, for non-centrosymmetric materials such as those of zincblende structure, SOC acts as an effective magnetic field which splits the previously spin-degenerate states. As a consequence, the band extrema are slightly displaced from their location without SOC, both in energy and momentum (see Fig. 6.9 of Appendix B), thus breaking, in principle, one of our underlying hypotheses. This effect was originally discussed by Dresselhaus [51] in 1955. In the following, we analyze the consequences of Dresselhaus splitting on Eq. (6.2.34) for the degenerate \mathbf{k} -point (i.e. the Γ point in the current work). See Appendix B for more details about the Dresselhaus Hamiltonian and its consequences on the band structure of zincblende materials.

In the presence of Dresselhaus splitting, the electronic dispersion of band n at $\mathbf{k} = \Gamma + \mathbf{q}$ is no longer $\theta q^2/2m_n^*$, following Eq. (6.2.28), but rather

$$\varepsilon_{\mathbf{q}n}^{\text{SOC}} = \frac{\theta(q - k_n^0(\hat{\mathbf{q}}))^2}{2m_n^*(\hat{\mathbf{q}})} - \theta\Delta E_n(\hat{\mathbf{q}}), \quad (6.2.35)$$

where $k_n^0(\hat{\mathbf{q}})$ and $\Delta E_n(\hat{\mathbf{q}})$ are respectively the momentum and energy offsets characterizing the Dresselhaus splitting of band n along direction $\hat{\mathbf{q}}$. As for the effective masses, we define $\Delta E_n(\hat{\mathbf{q}})$ as positive and let θ parametrize the sign of the energy offset. Eq. (6.2.29) therefore becomes

$$\text{ZPR}_{n\theta}^{\text{gFr}} = -\frac{\theta}{\pi\Omega_0} \int d^3\mathbf{q} \sum_{jn} \frac{1}{q^2} \frac{f(\hat{\mathbf{q}})}{\frac{(q - k_n^0(\hat{\mathbf{q}}))^2}{2m_n^*(\hat{\mathbf{q}})} + \omega_{0j}(\hat{\mathbf{q}}) - \Delta E_n(\hat{\mathbf{q}})}, \quad (6.2.36)$$

where $f(\hat{\mathbf{q}})$ is a purely angular function which includes the two rightmost fractions of the first line of Eq. (6.2.29).

As in Sec. 6.2.3, we express the integral in spherical coordinates. However, instead of the usual integral boundaries, we rather integrate on half the sphere (note the lower bound of the $\cos\theta$ integral), while simultaneously extending the lower bound of the radial integral

to $-\infty$:

$$\begin{aligned}
I &= \int d^3\mathbf{q} \sum_{jn} \frac{1}{q^2} \frac{f(\hat{\mathbf{q}})}{\frac{(q-k_n^0(\hat{\mathbf{q}}))^2}{2m_n^*(\hat{\mathbf{q}})} + \omega_{0j}(\hat{\mathbf{q}}) - \Delta E_n(\hat{\mathbf{q}})} \\
&= \int_0^1 d(\cos\theta) \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} dq \sum_{jn} \frac{f(\hat{\mathbf{q}})}{\frac{(q-k_n^0(\hat{\mathbf{q}}))^2}{2m_n^*(\hat{\mathbf{q}})} + \omega_{0j}(\hat{\mathbf{q}}) - \Delta E_n(\hat{\mathbf{q}})}.
\end{aligned} \tag{6.2.37}$$

With the change of variable $q'(\hat{\mathbf{q}}) = q - k_0(\hat{\mathbf{q}})$, this integral becomes

$$I = \int_0^1 d(\cos\theta) \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} dq' \sum_{jn} \frac{f(\hat{\mathbf{q}})}{\frac{(q'(\hat{\mathbf{q}}))^2}{2m_n^*(\hat{\mathbf{q}})} + \omega_{0j}(\hat{\mathbf{q}}) - \Delta E_n(\hat{\mathbf{q}})}. \tag{6.2.38}$$

Note that, by construction, the unit vectors $\hat{\mathbf{q}}$ and $\hat{\mathbf{k}}_n^0$ point in the same direction, hence $\hat{\mathbf{q}}' = \hat{\mathbf{q}}$. Transforming back with $q' = q$, we note that the contribution of the momentum offset has been exactly eliminated. One can then return the radial and angular integral boundaries to their usual values and perform the radial integral, which takes a form similar to Eq. (6.2.30):

$$\int_0^{\infty} dq \frac{1}{C_1 q^2 + C_2'}, \tag{6.2.39}$$

where the parameter C_2 has been replaced by $C_2' = \omega_{0j}(\hat{\mathbf{q}}) - \Delta E_n(\hat{\mathbf{q}})$. The parameter C_1 and the angular function $f(\hat{\mathbf{q}})$, respectively defined below Eq. (6.2.30) and in Eq. (6.2.36), therefore remain unchanged by Dresselhaus splitting. As a consequence, the treatment of the overlap matrices which enter $f(\hat{\mathbf{q}})$, discussed in Sec. 6.2.3, is also unaltered.

In light of this analysis, one finds that, in presence of Dresselhaus splitting, Eq. (6.2.34) generalizes exactly to

$$\text{ZPR}_{n\theta}^{\text{gFr}} = - \sum_{jn} \frac{\theta}{\sqrt{2}\Omega_0 n_{\text{deg}}}_{4\pi} \oint d\hat{\mathbf{q}} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2 \frac{(m_n^*(\hat{\mathbf{q}}))^{1/2}}{\omega_{0j}(\hat{\mathbf{q}}) \sqrt{\omega_{0j}(\hat{\mathbf{q}}) - \Delta E_n(\hat{\mathbf{q}})}}. \tag{6.2.40}$$

Supposing that $\Delta E_n(\hat{\mathbf{q}})$ is small compared to $\omega_{0j}(\hat{\mathbf{q}})$, one can further simplify the last expression by Taylor expanding the inverse square root, yielding

$$\text{ZPR}_{n\theta}^{\text{gFr}} = - \sum_{jn} \frac{\theta}{\sqrt{2}\Omega_0 n_{\text{deg}}}_{4\pi} \oint d\hat{\mathbf{q}} \left(\frac{\hat{\mathbf{q}} \cdot \mathbf{p}_j(\hat{\mathbf{q}})}{\epsilon^\infty(\hat{\mathbf{q}})} \right)^2 \frac{(m_n^*(\hat{\mathbf{q}}))^{1/2}}{\omega_{0j}(\hat{\mathbf{q}})^{3/2}} \left(1 + \frac{\Delta E_n(\hat{\mathbf{q}})}{2\omega_{0j}(\hat{\mathbf{q}})} \right). \tag{6.2.41}$$

Comparing this expression with Eq. (6.2.34), one finds that the energy offset stemming from the Dresselhaus splitting slightly enhances the ZPR at the Γ point (recall that $\Delta E_n(\hat{\mathbf{q}})$ is defined as positive). The angular-dependent energy offset can therefore be interpreted as direction-dependent modulation of the integrand. Should the contribution of the energy offsets be neglected, one recovers Eq. (6.2.34), which can now be seen as a lower bound to the true value of ZPR^{gFr} . An upper bound could also be obtained by estimating the largest value of $\Delta E_n(\hat{\mathbf{q}})$ for a given physical system. Eq. (6.2.40) could also, in principle, be used

to investigate the Fröhlich-induced ZPR of the degenerate band crossing points in Rashba systems [60], provided that the LO frequency is larger than the largest energy offset. Greater care would be required otherwise, as Eq. (6.2.40) would have poles.

Lastly, if we consider the Dresselhaus-split bands to be independent of each other, the known polaron effective mass enhancement induced by the isotropic Fröhlich interaction [61] should remain valid even in presence of anisotropic bands (see the numerical results of Ref. [14]). This suggests that EPI would attenuate the (already small) magnitude of the energy and momentum offsets stemming from SOC. However, the effect of the possible couplings between the almost-degenerate bands, as well as continuity conditions between the single band picture at the true extrema and the degeneracies at the Γ point remain to be investigated. Besides, as the momentum offset plays no role in Eq. (6.2.40) and (6.2.41), one can conclude that non-centrosymmetric cubic materials retain the Fröhlich physical picture of Γ -centered parabolic bands, which has been well corroborated by the experimental literature. See, for example, Fig. 2b) of Ref. [62].

6.3. Computational details

6.3.1. First-principles calculations

All first-principles calculations were performed with the ABINIT software package [63]. The bulk ground state properties were obtained from density functional theory, while vibrational properties and electron-phonon coupling were computed within density-functional perturbation theory [64, 65]. When SOC is taken into account, it is included both in the ground state and the density-functional perturbation theory calculations. We use norm-conserving pseudopotentials from the Pseudo-Dojo project [66] and rely on the generalized gradient approximation of the Perdew-Burke-Ernzerhof functional (PBE-GGA) [67]. The lattice parameters were optimized in the absence of SOC until all forces on the atoms were below 10^{-7} hartree/bohr³, except for Ge, where we used the experimental lattice parameter, as otherwise the obtained optimized lattice parameter for the PBE-GGA functional predicts a metallic ground state.

In order to isolate the effect of SOC on the EPI, we kept the lattice parameter fixed to the theoretical value without SOC. The electron-phonon self-energy was computed with the `ElectronPhononCoupling` Python code [68]. All relevant calculation parameters, including the relaxed lattice parameters, the maximal plane wave energy, the Monkhorst-Pack sampling of the Brillouin zone for \mathbf{k} -points and \mathbf{q} -points, and the broadening parameter η for the self-energy, can be found in Table S6.2 of the Supplemental Material [69].

We evaluate the sum on band index n' in the self-energy (Eq. (6.2.6) and (6.2.13)) using a semi-static approach [53, 70]: we replace the explicit evaluation of the non-adiabatic

contribution of the high energy bands, namely, bands where the phonon frequencies are negligible compared to the difference between the electronic eigenenergies, by the solution of a Sternheimer equation [54] for the subspace orthonormal to the active subspace. We chose the explicit number of bands in the active subspace such that the energy difference between the CBM and the highest band was at least 20 eV. We finally obtain the converged ZPR values in the $N_q \rightarrow \infty$ limit using the linear extrapolation method described in Ref. [71].

6.3.2. Generalized Fröhlich model

The generalized Fröhlich model (Eq. (6.2.34)) relies on the evaluation of angular-averaged square-root effective masses. From the first-principles perspective, electronic effective masses are typically computed either from finite differences or from density-functional perturbation theory. In the absence of SOC, we use the latter to evaluate the effective mass tensor, or, in the case of degenerate states, the transport-equivalent effective mass tensor defined in Ref. [72].

When SOC is taken into account, the calculation of the effective mass tensor from density-functional perturbation theory is not currently implemented in the ABINIT code for norm-conserving pseudopotentials. Hence, we evaluate the effective masses with SOC using order-4 central finite differences from the first-principles electronic eigenvalues. The VBM of zincblende materials constitutes a special case, as the electronic dispersion displays Dresselhaus splitting due to the lack of inversion symmetry [51]. Therefore, we model the dispersion with the Dresselhaus Hamiltonian [51] and obtain the effective masses from quadratic fits. For comparison, we also compute the angular-averaged effective masses for the VBM using the electronic dispersion obtained from the Luttinger-Kohn [73] Hamiltonian in the presence of SOC. See Appendix A and B for more details about our treatment of the VBM. When Dresselhaus splitting is noticeable near the CBM of zincblende materials, we evaluate the effective masses from quadratic fits using the first-principles electronic dispersion.

We finally note that, despite not being very accurate, the electronic effective masses computed with GGA-PBE are sufficient for the purpose of this work, which focuses on EPI. Lastly, as the effective masses computed from PBE for GaAs at the theoretically relaxed lattice parameter are particularly small, we also provide results computed at the experimental lattice parameter [74].

6.4. Results and discussion

6.4.1. First-principles

6.4.1.1. Effect of SOC on the VBM and CBM ZPR. The effect of SOC on the ZPR computed from first principles for our twenty materials is shown in Fig. 6.2, for both the VBM

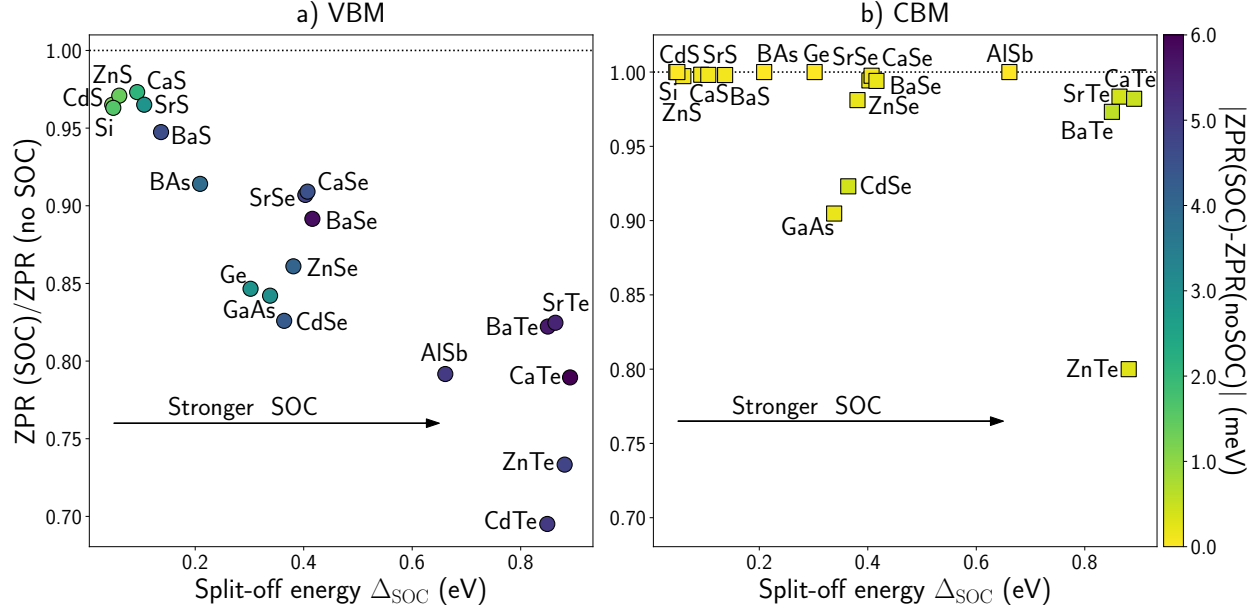


Figure 6.2 – Effect of SOC on the ZPR from first principles, for a) the VBM (circle markers) and b) the CBM (square markers). The ZPR reduction ratio, $ZPR(\text{SOC})/ZPR(\text{no SOC})$, is presented with respect to the split-off energy, Δ_{SOC} , which is an indicator of the SOC’s strength. The color scale shows the absolute value of the SOC-induced correction to the ZPR, $|ZPR(\text{SOC}) - ZPR(\text{no SOC})|$, in meV. For the VBM, the absolute difference remains under 6 meV but increases with Δ_{SOC} , yielding a relative decrease that reaches up to 30% for the heavier materials (AlSb and the tellurides). For the CBM, most materials show a relative decrease below 5%, with the exception of ZnTe, CdSe and GaAs. The absolute value of the difference remains negligible, below 1 meV. CdTe is absent from the CBM figure (see text). Numerical values are provided in Table S6.3 of the Supplemental Material [69].

(left, circle markers) and the CBM (right, square markers). Both subfigures display the ZPR reduction ratio, $ZPR(\text{SOC})/ZPR(\text{no SOC})$, with respect to the split-off energy, Δ_{SOC} , which is a direct indicator of the SOC’s strength. The color scale indicates the absolute value of the SOC-induced correction to the ZPR for each band extrema, $|ZPR(\text{SOC}) - ZPR(\text{no SOC})|$. Numerical values can be found in Table S6.3 of the Supplemental Material [69].

On the one hand, one can observe that the relative decrease of the valence band edge ZPR, ZPR_v , qualitatively increases with Δ_{SOC} . Recall that, for zincblende materials, the leading orbital character of the VBM is p states from the anion. Hence, the less affected materials regroup the lighter anions, namely all sulfides and Si, for which the relative decrease is below 5%. The selenides, arsenides and Ge display a relative decrease ranging from ~ 10 – 15% , while the heavier materials in our set, AlSb and the tellurides, see their ZPR_v reduced by ~ 15 – 30% . However, the numerical value of $|ZPR_v(\text{SOC}) - ZPR_v(\text{no SOC})|$ remains small, under 6 meV for all materials. Nevertheless, the color scale clearly indicates that the absolute value of the correction increases with Δ_{SOC} . Small absolute differences were to be expected, as stronger SOC is naturally present in heavier materials, which typically display smaller ZPR.

However, we note that the effect of SOC we observe on the ZPR_v , hence on the real part of the self-energy, is not nearly as significant as the relative impact reported in the literature for the hole mobility [75–77], which can be over 10% in weak-SOC materials like Si and reach more than 50% in heavier materials. In this context, taking SOC into account reduces the number of scattering channels, hence increasing the mobility. In contrast, despite their respective contributions being reduced by SOC, all phonon wavevectors still contribute to the ZPR. Recall that the mobility depends on the electron-phonon self-energy through the relaxation time of the electronic states, which goes as the inverse of the imaginary part of $\Sigma_{\mathbf{k}n}^{\text{AHC}}$ [7]. Thus, one could expect the inverse relaxation time of a given electronic state when including SOC to decrease by a similar ratio as the mobility when SOC is neglected. Nevertheless, it is not entirely clear how those two ratios should correlate, as the inverse relaxation time is defined for each electronic state, while the mobility is a global quantity integrated on the BZ, hence all the neighboring states around the Γ point contribute to the hole mobility. As the real and imaginary parts of $\Sigma_{\mathbf{k}n}^{\text{AHC}}$ are related to one another by the Kramers-Kronig relations, further investigation will be required to fully understand the effect of SOC on the full electron-phonon self-energy.

The CBM, on the other hand, displays very little modification of the ZPR from SOC. The relative decrease remains under 5% for all materials, including AlSb and some tellurides. These results are in line with the atomic picture, in which a s -like band such as the CBM of zincblende materials is not affected by SOC since $l = 0$. This argument does not hold for the d -like CBM of the rocksalt alkaline earth chalcogenides; in that case, the conduction band edge ZPR (ZPR_c) decrease remains negligible as the CBM is well-isolated in energy from the other bands. The only three exceptions to this trend are ZnTe (20%), GaAs (9%) and CdSe (8%). CdTe is absent from the CBM figure as its ZPR_c changes sign when including SOC, going from -0.4 meV to 0.4 meV. Nevertheless, these seemingly large relative corrections are not numerically significant: the absolute correction remains under 1 meV for all materials (see the color scale), which is within the typical numerical accuracy of this type of calculations.

To understand these different tendencies from a qualitative point of view, one can picture a typical zincblende band structure. In Fig. 6.3, the energy bands of CdTe with SOC (solid red lines) have been shifted such that the VBM coincides with the VBM without SOC (dashed blue lines). The Fermi energy without SOC has been set to zero. When comparing both sets of bands, one can see that the general shape of the unoccupied subset of bands is scarcely affected by SOC, other than a relatively small energy shift between spin-split bands, located mostly close to the Brillouin zone boundaries. Recall that the Fan and Debye-Waller contributions to the self-energy have opposite signs (see Eq. (3.3.13) and (3.3.29)) and almost equal magnitudes. Nevertheless, when considering the contribution from either the occupied or unoccupied subset of bands to the AHC self-energy, the Fan term typically governs the net sign of the renormalization in semiconductors. Hence, as the CBM is mostly repelled by

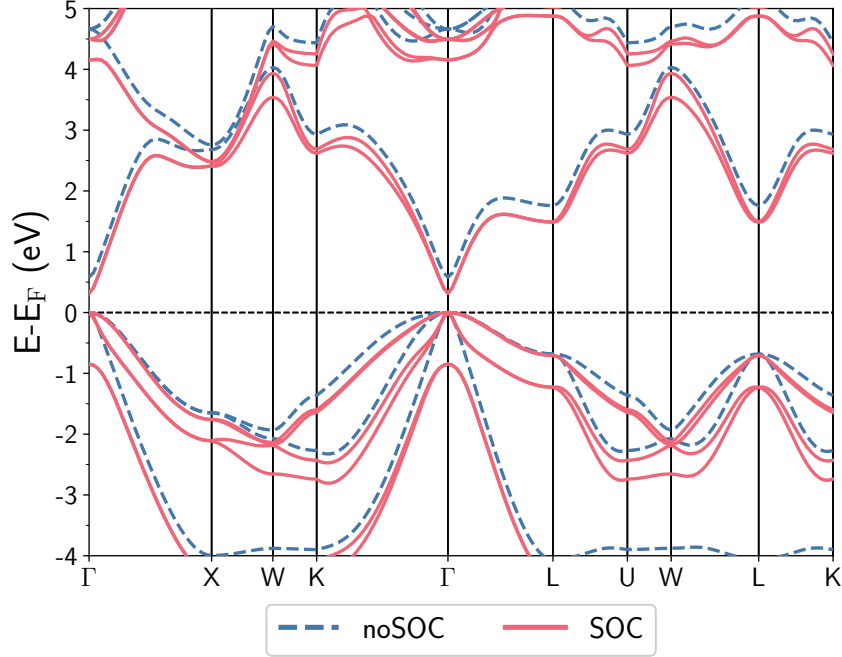


Figure 6.3 – Band structure of zincblende CdTe, with SOC (solid red lines) and without SOC (dashed blue lines). As a guide to the eye, the SOC energy bands have been shifted such that the VBM coincides with its equivalent without SOC. The Fermi energy without SOC has been set to zero.

couplings with nearby conduction states of higher energy, one can deduce that the ZPR of the CBM will barely be affected by SOC, since the effective mass does not change significantly and the band extrema is well isolated in energy. The three apparent outliers, CdSe, ZnTe and GaAs, all feature a direct fundamental gap at the zone center, small CBM effective masses and a relatively small band gap energy (albeit always much larger than the highest phonon frequency, GaAs with SOC being the exception). In the light of the Kane model [78], recall that the CBM and VBM at the Γ point are linked through an avoided crossing. The CBM is thus indirectly affected by SOC, through its interplay with the heavy hole and light hole p -like bands. A small band gap reinforces this interaction, resulting in a greater relative decrease of the ZPR_c for these three materials. We do not observe such an effect in indirect band gap materials like AlSb.

When rather considering the VBM, one can observe the loss of degeneracy predicted by group theory, as the two split-off bands have been shifted by Δ_{SOC} below the heavy hole and light hole bands. Fig. 6.3 also reveals two main consequences of SOC on the occupied bands: on the one hand, the effective masses in the vicinity of the zone center are reduced, and on the other hand, the energy shift between the spin-split bands occurs throughout the Brillouin zone, thus globally lowering the eigenenergies of the occupied states with respect to the VBM energy.

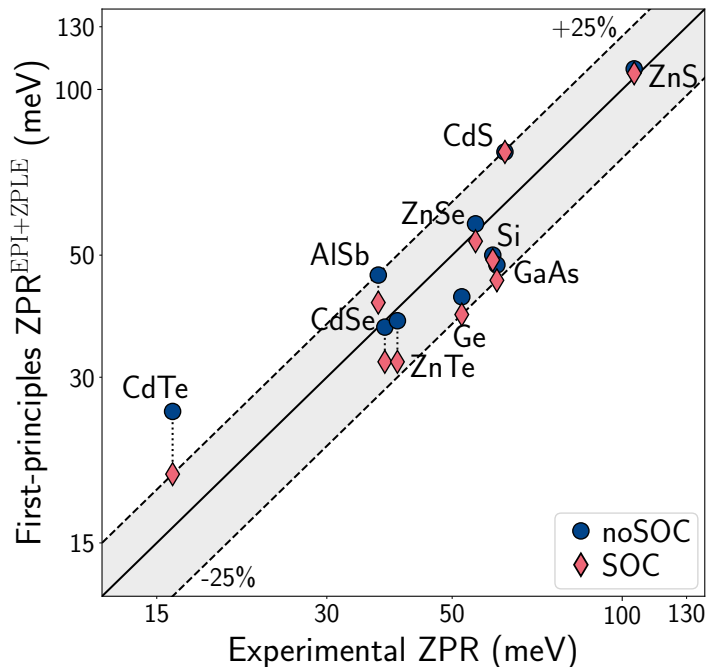


Figure 6.4 – Comparison between first-principles AHC band gap ZPR and experimental values, without SOC (blue circles) and with SOC (red diamonds). The dashed lines, emphasized by the shaded gray area, indicate the theoretical ZPR_g agrees with the experimental one within 25%. Note that the scales are logarithmic. When SOC is included, all materials lie within the target area. Numerical experimental values are provided in Table S6.5 of the Supplemental Material [69].

With this simple picture in mind, our results suggest that the effect of SOC can be safely neglected for band extrema which are well-isolated in energy, should the modification of the effective mass remain small. In contrast, degenerate extrema or densely entangled bands must be treated more carefully. We finally emphasize that one must not systematically overlook the apparent small magnitude of the SOC corrections to the ZPR_v . What may at first be perceived as “only a few meV” nevertheless captures a significant relative decrease for the heavier materials, that reaches 15–30% of the predicted ZPR_v without SOC. This effect cannot be neglected when aiming for predictive results, especially if one seeks to validate numerical predictions with experimental data.

6.4.1.2. Experiment vs first principles. We now examine how does the inclusion of SOC affects the global agreement between ZPR_g and available experimental data. To make a fair comparison with experimental data, the first-principles data shown in Fig. 6.4 includes the theoretical contribution of the zero-point expansion of the lattice [79, 80]. This term originates from the phonon contribution to the total free energy of the crystal, which increases the $T = 0$ K lattice parameter compared to the static equilibrium value and, in turn, affects the band gap energy (see Appendix D for more details). Note that the scales

are logarithmic. The shaded gray area highlights the region where both ZPR (first-principles and experimental) agree within 25% of each other. As experimental values of the ZPR are obtained from extrapolation procedures rather than from direct measurements, we have to keep in mind when analyzing the accuracy of theoretical results that there is an experimental uncertainty which can be quite substantial, especially when few experimental datasets are available for a given material. See the Supplementary Note 1 of Ref. [13] for a detailed discussion about the uncertainties associated with the experimental values of the ZPR.

The results without SOC (blue circles) are equivalent to the non-adiabatic AHC data shown in Fig. 2 of Ref. [13] for the 10 materials considered. When SOC is taken into account (red diamonds), all materials now lie within the tolerance criterion, including CdTe, which is largely overestimated without SOC. Thus, SOC does not alter the quantitative agreement between first principles and experiment, although Ge and GaAs reach the lower limit of the tolerance criterion.

One can also wonder if the greater predictability of the non-adiabatic AHC approach compared to the adiabatic supercell method claimed in Fig. 2 of Ref. [13] (see empty red triangles, labelled ASC-DFT) would remain upon inclusion of SOC. Although we have not attempted any adiabatic supercell calculation with SOC, our results suggest that the inclusion of SOC would not reduce the significant underestimation of the ZPR by adiabatic supercells for the lighter, more ionic materials. For intermediate to strong SOC, our data show a reduction of the total band gap ZPR ranging between 8%–34% (see the rightmost column of Table S6.3 of the Supplemental Material [69]). Should we infer that a similar effect would be observed in adiabatic supercell calculations, one could expect the result obtained from adiabatic supercells based on DFT calculations for CdTe to lie inside the tolerance criterion. At the same time, the underestimation would worsen for CdSe, and the ZnSe adiabatic supercell result would likely exit the target. Our results, therefore, support the general conclusion of Miglio *et al.* [13] regarding the adiabatic supercell method being outperformed by the non-adiabatic AHC approach.

6.4.1.3. Origin of the SOC-induced ZPR decrease. In Sec. 6.4.1.1, we discussed the effect of SOC on the ZPR in terms of the variation of the electronic eigenvalues. We now refine this analysis by constructing histograms of the different \mathbf{q} -points contribution to ZPR_v with respect to the norm of the phonon wavevector, for a $48 \times 48 \times 48$ \mathbf{q} -point grid. Fig. 6.5 displays such decomposition for two polar materials, CdS (left) where SOC has little effect on the calculated ZPR, and CdTe (right) where, on the contrary, the calculated ZPR is more strongly reduced by SOC. The bottom panels show the distribution of the Brillouin zone weight for the different wavevector bins. The solid and shaded hatched histograms refer to the ZPR contributions computed respectively with SOC and without SOC.

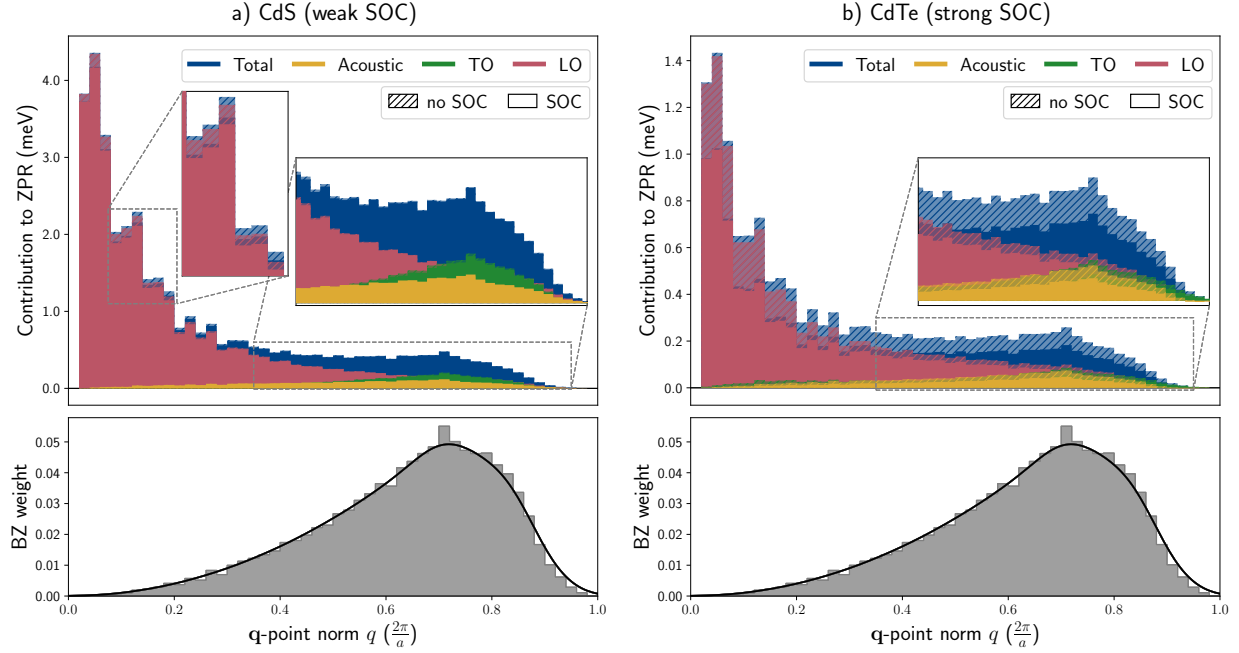


Figure 6.5 – Total and mode contribution to ZPR_v for different q -point norms, for polar materials, for a $48 \times 48 \times 48$ Γ -centered q -point grid. a) CdS (weak SOC), b) CdTe (strong SOC). Solid histograms include SOC while the shaded hatched ones do not. Lower panels show the Brillouin zone weight distribution for the different q -points. Both materials display a clear signature of the Fröhlich interaction, as the LO mode (red) at small q accounts for the greatest part of the total ZPR. Notice the different y-axis scaling for both materials. The total contribution at large q (blue) is more equally splitted between the LO, acoustic (yellow) and TO (green) modes. For CdTe (right), the SOC-induced decrease of the ZPR originates from the whole Brillouin zone, and can be heuristically understood in terms of the variation of the effective masses (small q) and of a global decrease of the eigenenergies of the occupied bands (large q). Only the small q behavior is observed in CdS (left), as emphasized by the insets.

A first observation that emerges from this figure is the extremely similar shape of the mode histograms, apart from their respective energy scale. For both materials, the vast majority of the ZPR originates from the LO phonon mode (red) in a very small portion of the Brillouin zone, near the zone center (small q). This behavior is a clear signature of the Fröhlich interaction. The contribution of the large q modes, which cover most of the Brillouin zone, is significantly smaller and split more equally between the acoustic (yellow), transverse optical (TO, green) and LO modes.

We now compare the solid and shaded histograms to identify how the different q regimes are affected. The total contribution to ZPR (blue) for CdTe shows that strong SOC reduces the contribution of all phonon modes throughout the Brillouin zone. This suggests that, in the large q regime, the ZPR decrease can be associated with the global reduction of the electronic eigenvalues in the occupied subset of bands. In contrast, in the small q regime, it

can be linked to the decrease of the effective masses. On the contrary, for CdS (weak SOC), the large q regime is unaffected by SOC; thus, only the variation of the band curvatures in the vicinity of the VBM seems to be responsible for reducing the ZPR.

At this point, we insist that our interpretation of the SOC-induced ZPR reduction in terms of the variation of electronic eigenvalues remains a heuristic analysis, since the full ZPR expressions for a given \mathbf{q} -point (Eq. (6.2.6) and (6.2.13)) include other physical quantities which are, in principle, affected by SOC: the phonon frequencies, $\omega_{\mathbf{q}\nu}$, and the squared EPI matrix elements, $|g_{\mathbf{k}nn'}^{\text{Fan}}(\mathbf{q}\nu)|^2$ and $|g_{\mathbf{k}nn'}^{\text{DW}}(\mathbf{q}\nu)|^2$. To test our hypothesis, we computed all the physical quantities entering $\Sigma_{\mathbf{k}n}^{\text{Fan}}$ and $\Sigma_{\mathbf{k}n}^{\text{DW}}$ for CdTe while artificially reducing SOC to 1% of its full strength. This allows us to decompose the electronic states correctly in terms of the double group irreducible representations while still reproducing the electronic and phononic dispersions without SOC adequately (see Sec. S6.1 of the Supplemental Material [69] for more details). With these in hand, we can precisely control which ingredients of the self-energy are affected by SOC. While such arbitrary combinations of SOC and lowSOC quantities have no physical meaning per se, they will prove insightful for understanding our previous results.

Fig. 6.6 show the histogram decomposition of the total ZPR_v (the equivalent of the blue histograms of Fig. 6.5) for different combinations of contributions. Table 6.1 contains the numerical results for ZPR_v , ZPR_c and ZPR_g for the same combinations. The data labelled “only $\varepsilon_{\mathbf{k}n}^{\text{low}}$ ” (olive dashes) refers to $\Sigma_{\mathbf{k}n}^{\text{EPI}}$ being computed with the full SOC except for the electronic eigenenergies, which are taken at 1% SOC. In contrast, “only $\varepsilon_{\mathbf{k}n}^{\text{SOC}}$ ” (cyan dashes) is computed with 1% SOC except for the electronic eigenenergies, which are taken at full SOC.

Table 6.1 – VBM, CBM and band gap ZPR of CdTe for different combinations of SOC strength. See Sec. 6.4.1.3 and the caption of Fig. 6.6 for details.

Combination	ZPR (meV)		
	VBM	CBM	ZPR_g
noSOC	16.4	-0.4	-16.8
SOC	11.4	0.4	-11.0
1% SOC	16.2	-0.5	-16.7
only $\varepsilon_{\mathbf{k}n}^{\text{low}}$	17.4	-0.8	-18.3
only $\varepsilon_{\mathbf{k}n}^{\text{SOC}}$	11.1	-0.8	-11.9

From Fig. 6.6 and Table 6.1, the histograms can be grouped in two categories. On the one hand, the data computed using low SOC eigenvalues, “only $\varepsilon_{\mathbf{k}n}^{\text{low}}$ ”, yields both histograms and total ZPR values which qualitatively reproduce the calculation without SOC (dark indigo). On the other hand, the data computed by including SOC only through the electronic eigenvalues, “only $\varepsilon_{\mathbf{k}n}^{\text{SOC}}$ ”, captures almost all the ZPR decrease of the SOC data throughout

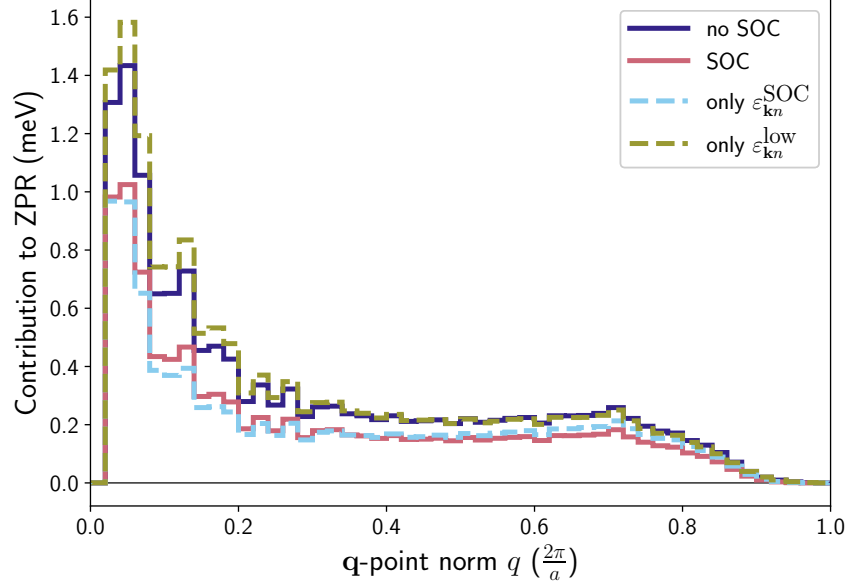


Figure 6.6 – Contribution to the total ZPR_v of CdTe with varying SOC strength for different q -point norms, for a $48 \times 48 \times 48$ Γ -centered q -point grid. The different physical quantities entering Σ^{EPI} (Eq. (6.2.6) and (6.2.13)) are computed without SOC (dark indigo), with SOC (light red), and from mixed combinations of full SOC and 1% SOC quantities (cyan and olive dashes, see text for details). The four histograms can be qualitatively grouped into two classes according to the strength of SOC used to evaluate the electronic eigenvalues, in agreement with the heuristic explanation developed in Fig. 6.5.

the Brillouin zone (see the light red reference line). These results validate our heuristic explanation and confirm that the modification of the electronic eigenvalues dominates the SOC-induced decrease of ZPR_v in the different q regimes. While the phonon frequencies and the EPI matrix elements undoubtedly influence the quantitative results, we argue here that the SOC-induced variation of the ϵ_{kn} are sufficient to estimate the effect of SOC on the VBM ZPR and, by extension, on the total band gap ZPR for this class of cubic materials (see also Sec. S6.1 of the Supplemental Material [69]).

We note that our conclusions differ significantly from those of Ref. [45], who reported that SOC enhances the electron-phonon coupling strength in perovskite methylammonium lead iodine (MAPbI_3), thus *increasing* the temperature-dependent band gap renormalization compared to a scalar relativistic calculation. The electronic structure of MAPbI_3 is, however, very different from the zincblende, diamond and rocksalt structures we investigated; the VBM is non-degenerate and reasonably well isolated in energy, while the CBM, of Pb character, can couple to more electronic states within a small energy window (see Fig. 4 of their Supplemental Material), resulting in the band gap opening with increasing temperature, in contrast with our set of materials. While we do observe some variations of the ZPR_v and in the small q regime of its histogram decomposition when including SOC in the

EPI matrix elements and excluding it in the eigenvalues (see “only $\varepsilon_{\mathbf{k}n}^{\text{low}}$ ” data in Table 6.1 and Fig. 6.6), the effect is too small to allow us to draw conclusions. In fact, for our set of materials, any effect of SOC on the EPI matrix elements is entirely washed out by the variation of the static electronic eigenvalues. We also verified that the band gap correction at $T = 300$ K decreased by a similar ratio as the ZPR_g when including SOC (see Table S6.4 of the Supplemental Material), thus confirming that our conclusions hold beyond $T = 0$ K.

6.4.2. Generalized Fröhlich model with SOC

We now present the results from our gFr model with SOC. As mentioned in Sec. 6.2.4 and 6.3, and discussed thoroughly in Appendix B, the VBM of zincblende materials with SOC requires a special treatment, as it displays Dresselhaus splitting. As a consequence, the VBM in a generic direction $\hat{\mathbf{k}}$ is slightly shifted from Γ , both in momentum and in energy. We found an energy offset $\Delta E_n(\hat{\mathbf{q}})$ smaller than 0.5 meV for all materials, and an average momentum offset $k_n^0(\hat{\mathbf{q}})$ of $\sim 5 \times 10^{-3} \text{\AA}^{-1}$, at most 10^{-2}\AA^{-1} for CdS. For comparison, the change in momentum expected from a photon doing a vertical optical transition in a semiconductor with a band gap energy of $E_g = 1$ eV would be $\Delta k \sim E_g/hc \sim 10^{-4} \text{\AA}^{-1}$, with h the Planck constant and c the speed of light. As we argue in Sec. 6.2.4, the momentum offset has no consequence on the ZPR^{gFr} . Furthermore, for our set of materials, the largest value of $\Delta E_n(\hat{\mathbf{q}})$ is at least two orders of magnitude smaller than the LO frequency. Hence, we can safely neglect the energy offset and use Eq. (6.2.34).

Fig. 6.7 compares ZPR_g computed with the gFr model to the first-principles AHC result, both with SOC (solid markers) and without SOC (shaded dotted markers). Note that the scales are logarithmic. The shaded gray area delimits the region where the gFr value deviates from the first-principles ZPR by at most 25%. In this Figure, the VBM contribution to the ZPR_g with SOC was computed with the Dresselhaus Hamiltonian (see Appendix B for details). The Luttinger-Kohn Hamiltonian yields qualitatively equivalent results. The materials are grouped into three sets: the alkaline earth chalcogenides (purple diamonds), the Zn/Cd chalcogenides (yellow squares), both fairly ionic, and the less ionic III-V materials (blue triangles). Si and Ge are not considered in this Section, as their vanishing Born effective charges yield a null ZPR_g in the Fröhlich picture.

The alkaline earth chalcogenides, which display a larger ZPR_g ranging from 50 to 150 meV, are very well described by the gFr model, which captures more than 75% of the AHC ZPR_g both without and with SOC. The Zn and Cd chalcogenides ZPR_g is also reasonably well captured by the model, which accounts for more than two-thirds of the AHC value. The absolute value of their ZPR_g is smaller compared to the isoelectronic alkaline earth chalcogenides, which can be attributed to a smaller band gap at the DFT level that strengthens the contribution of interband couplings between valence and conduction states,

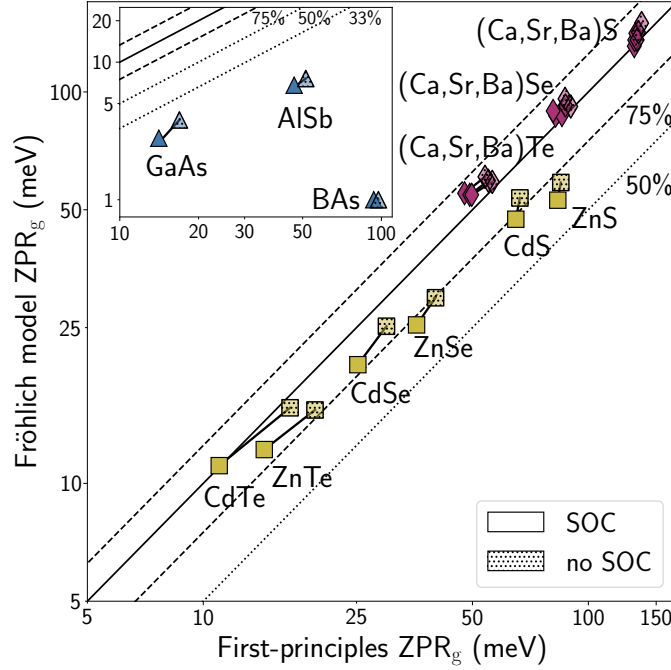


Figure 6.7 – Comparison between the generalized Fröhlich model ZPR_g and the first principles AHC ZPR_g , with and without SOC. The different markers indicate materials with similar ionicity and band orbital characters: Zn and Cd chalcogenides (yellow squares), alkaline earth chalcogenides (purple diamonds) and group III-V materials. Solid (dotted) markers include (exclude) SOC. Note that the scales are logarithmic. The dashed lines and shaded gray area delimits the region where they agree within 25% of each other, while the dotted lines indicate where the ZPR_g^{gFr} captures 50% (also 33% for the inset) of the AHC value. Tables S6.3, S6.6 and S6.7 of the Supplemental Material [69] respectively report the numerical values for the ZPR_g obtained from AHC and the gFr model, as well as the physical parameters entering Eq. (6.2.34)

thus reducing ZPR_g . In contrast, the gFr model captures less than one third of the ZPR_g for III-V materials.

Upon inclusion of SOC (going from the dotted to solid markers), the gFr model qualitatively retains the same proportion of the AHC ZPR_g value for all three material families. However, we observe a very small decrease for the sulfides (upper rightmost groups of yellow and purple markers). These lighter materials have higher LO frequencies than selenides and tellurides. Hence, the Fröhlich interaction has a greater energy window to feel the SOC-induced decrease of the effective masses.

These results confirm that the claim of Ref. [13] is robust to the inclusion of SOC: the ZPR of ionic materials, here both chalcogenide families, is dominated by the physical picture of a large polaron, where the movement of the slow electron is correlated to the dynamically adjusting phonon cloud, as emphasized by their Fig. 5(b) [13].

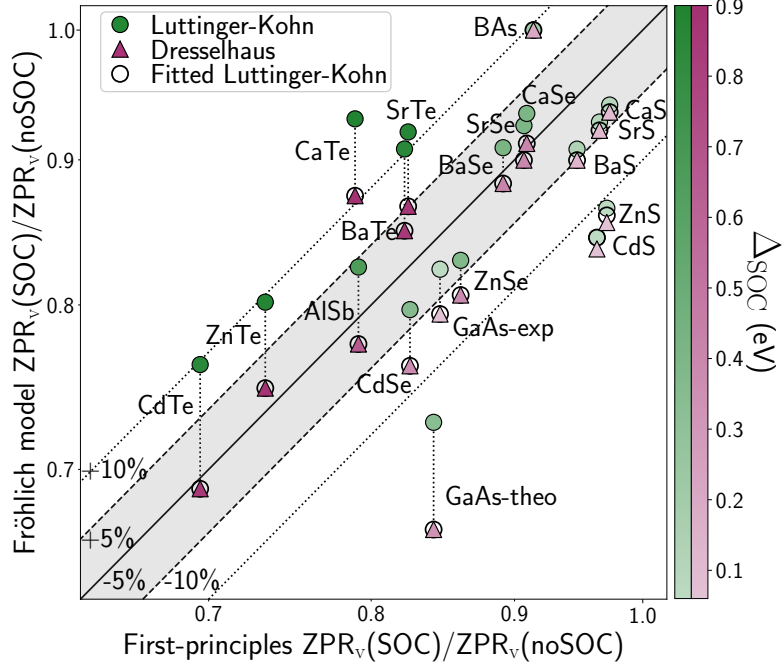


Figure 6.8 – Comparison of the SOC-induced reduction of the ZPR_v obtained with the generalized Fröhlich model and with first-principles AHC methodology. The electronic dispersion used to compute the angular-averaged effective masses entering the ZPR_v^{gFr} (see Eq. (6.2.34)) was obtained either from the Luttinger-Kohn (green circles) Hamiltonian or from the Dresselhaus (purple triangles) Hamiltonian. The color intensity of the markers is proportional to the split-off energy. Note that the scales are logarithmic. The dashed (dotted) lines indicate where the ratio computed with the gFr model deviates from the first-principles ratio by 5% (10%). For most materials, the gFr model provides a fairly good estimate of the first-principles ZPR_v decrease, although using effective masses from the Dresselhaus Hamiltonian delivers superior result. The discrepancy between results from both Hamiltonians increases with SOC, as the Luttinger-Kohn Hamiltonian tends to overestimate the light hole effective masses (see text). The Luttinger-Kohn Hamiltonian can be reparameterized with values extracted from the Dresselhaus model which reproduce more accurately the curvature of light holes around Γ . This delivers the SOC-induced reduction represented by empty circle, that agree better with each Dresselhaus-based results. Numerical values are provided in Table S6.6 of the Supplemental Material [69].

As in Ref. [13], we stress that a perfect agreement between the gFr and AHC methodologies was not expected, hence the labelling of the reference lines on Fig. 6.7 emphasize the fraction of the total first-principles ZPR_g captured by the gFr model rather than the level of agreement between the two values. By construction, the gFr model does not include the contribution from acoustic and TO modes, nor the Debye-Waller contribution, as its purpose is to solely capture the contribution of the non-adiabatic Fröhlich interaction to the ZPR. Moreover, all interband couplings outside the 4-fold degenerate subset at the VBM are neglected. Discrepancies could also arise from non-parabolic behavior of the electronic bands,

which will naturally occur at some point since the Brillouin zone is finite and periodic, and from LO phonon dispersion.

On a different note, one can wonder if our gFr model with SOC would provide a reliable estimate of the SOC-induced decrease of ZPR_v , without resorting to a full AHC calculation with SOC. In the spirit of the discussion presented in Sec. 6.4.1.3, the gFr model would capture the decrease of the hole effective masses near the Γ point. To answer this question, Fig. 6.8 compares the ratio $ZPR(\text{SOC})/ZPR(\text{noSOC})$ for the VBM obtained from the gFr model, using both the Luttinger-Kohn (green circles) and Dresselhaus (purple triangles) Hamiltonians, to the AHC value displayed in Fig. 6.2. The color intensity of the markers is proportional to the split-off energy, Δ_{SOC} . Note that the scales are logarithmic. The shaded gray area shows the region where both ratios agree within 5% of each other.

Two observations can be drawn from Fig. 6.8. On the one hand, the two ratios agree within 5% of each other for about half of the 19 materials (dashed lines): 11 for the Dresselhaus model and 9 for the Luttinger-Kohn model. The discrepancy with the first-principles ratio is below 10% for most materials (dotted lines), thus providing a reasonable estimate of the SOC-induced decrease of the VBM observed in the full AHC calculation, from a single phonon calculation at the Γ point. The main exceptions are CaTe with the Luttinger-Kohn model, GaAs at the theoretical lattice parameter, ZnS and CdS.

On the other hand, the difference between both models increases with SOC, the Dresselhaus model being more accurate for heavier materials. This discrepancy can be attributed to the construction of the models: as it couples the 4-fold heavy hole and light hole bands to the 2-fold split-off bands, the Luttinger-Kohn model intends to describe the effect of SOC in a broader region of the Brillouin zone, thus capturing the warping of the light hole bands around the split-off energy observed in many cubic materials (see, for example, Fig. 2 of Ref. [72]). At stronger SOC, we observed that the cost of qualitatively capturing the correct band warping is a less accurate representation of the curvature of the light hole bands, yielding an overestimated m^* compared to the first-principles dispersion. In contrast, the Dresselhaus Hamiltonian treats only the 4-fold VBM and provides an accurate description of the bands in the vicinity of the Γ point, at the cost of not describing the band warping further away in the Brillouin zone (see Fig. 6.10 of Appendix C). Moreover, while our Dresselhaus model parameters are fitted to the first-principles dispersion with SOC, the Luttinger-Kohn Hamiltonian relies on the theoretical Luttinger parameters obtained without SOC, Δ_{SOC} being the only parameter related to SOC. The original work by Luttinger and Kohn already cautioned about a decreased accuracy of their model for strong SOC [73].

To validate this interpretation, we also compute the Luttinger-Kohn model using fitted Luttinger parameters extracted from the Dresselhaus model (empty circles, see Appendix C for details). The purpose of these parameters is to reproduce more accurately the curvature of the light hole bands around Γ , at the cost of deviating more severely from the first-principles

dispersion around the split-off energy and underestimating the m^* of the split-off bands compared to the original Luttinger parameters. In that case, both models agree very well. The fundamental differences between the two models suggest that the modified Luttinger-Kohn Hamiltonian constructed from fitted Luttinger parameters would be a more suitable choice to include SOC in the Fröhlich polaron effective mass theory developed in Guster *et al.* [14], as it simultaneously captures the correct curvature for the light hole bands in presence of SOC and lacks the numerical complications arising from the Dresselhaus splitting. One should nevertheless make sure that the agreement region between the fitted Luttinger-Kohn model and the first-principles dispersion covers at least a few ω_{LO} for the predicted polaron effective mass to be reliable.

Lastly, we come back to the underestimation of the ratio observed for ZnS and CdS, despite both models agreeing quite well with each other. To understand this result, recall the radial integral that occurred during the derivation of the gFr model with SOC, Eq. (6.2.29) and (6.2.30). At this point, we suppose that the parabolic behavior of the electronic bands can be extended to infinity to replace the integral with its asymptotic solution. By doing so, we assume that the Brillouin zone region where the effective mass approximation holds is sufficiently large such that one has reached a significant fraction of the asymptotic value once the bands start to deviate from parabolicity. Such fraction can be estimated by looking at the analytical solution of Eq. (6.2.30) at a finite upper bound q_c :

$$\int_0^{q_c} dq \frac{1}{\frac{q^2}{2m_n^*(\hat{\mathbf{q}})} + \omega_{0j}(\hat{\mathbf{q}})} = \sqrt{\frac{2m_n^*(\hat{\mathbf{q}})}{\omega_{0j}(\hat{\mathbf{q}})}} \arctan \left(\frac{q_c}{\sqrt{2m_n^*(\hat{\mathbf{q}})\omega_{0j}(\hat{\mathbf{q}})}} \right). \quad (6.4.1)$$

Assuming that q_c corresponds to the wavevector where the electronic bands start to deviate from parabolicity, the argument of the arctan function can be recast as

$$\sqrt{\frac{q_c^2/2m_n^*(\hat{\mathbf{q}})}{\omega_{0j}(\hat{\mathbf{q}})}} = \sqrt{\frac{E_c}{\omega_{0j}(\hat{\mathbf{q}})}}, \quad (6.4.2)$$

namely the square-root of the ratio of the eigenenergy of the electronic bands where they stop being parabolic and the LO frequency. Should the departure from parabolicity occur at an energy smaller than $\omega_{0j}(\hat{\mathbf{q}}) = \omega_{\text{LO}}$ with respect to the VBM energy, like for ZnS and CdS, the radial integral would have reached at most $\arctan(1) = 0.5$, less than half of the asymptotic value, and one can reasonably question the validity of such an approximation. In physical terms, we warn against a possible breakdown of the parabolic approximation within the energy window that is physically relevant to the Fröhlich interaction. As the effective mass approximation is a cornerstone assumption of the original Fröhlich model, one should therefore be particularly careful when introducing SOC in the treatment of polarons for materials with high phonon frequencies.

6.5. Conclusion

In the present study, we investigate the consequences of spin-orbit coupling on the electron-phonon interaction contribution to the zero-point renormalization of cubic materials. Our first-principles calculations show that spin-orbit coupling reduces the zero-point renormalization of the valence band edge by 15%–30% for the heavier materials, while the conduction band edge is scarcely affected. The leading mechanism behind this behavior, brought to light by an Allen-Heine-Cardona calculation where the strength of spin-orbit coupling is artificially reduced to 1%, is revealed to be the variation of the electronic eigenvalues entering the electron-phonon self-energy and the decrease of the hole effective masses near the valence band maximum.

We also extend the generalized Fröhlich model presented in Miglio *et al.* [13] to include the spin-orbit coupling, revealing some numerical subtleties in the treatment of the valence band maximum of zincblende materials due to Dresselhaus splitting. We show that the predominance of non-adiabatic effects on the zero-point renormalization of ionic materials is robust to the inclusion of spin-orbit coupling and that the generalized Fröhlich model can be used to estimate the magnitude of the SOC-induced ZPR decrease with reasonable accuracy. We finally warn about the accuracy of the Luttinger-Kohn model with spin-orbit coupling for heavier materials and propose a method relying on the Dresselhaus model to extract fitted Luttinger parameters more suitable for the purpose of our generalized Fröhlich model with spin-orbit coupling, as well as highlight a possible breakdown of the parabolic approximation on which the original Fröhlich model with spin-orbit coupling is built for materials with high phonon frequencies.

Acknowledgments

The authors acknowledge fruitful discussions with M. J. Verstraete and M. Giantomassi. This research was financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), under the Discovery Grants program grant No. RGPIN-2016-06666, by the Fonds de la Recherche Scientifique (FRS-FNRS Belgium) through the PdR Grant No. T.0103.19 - ALPS and by the European Union’s Horizon 2020 research and innovation program under grant agreement No. 951786 - NOMAD CoE. This work is part of the SHAPeMe project (EOS ID 400077525) that has received funding from the FWO and F.R.S.-FNRS under the Excellence of Science (EOS) program. This research was enabled in part by support provided by Calcul Québec (www.calculquebec.ca) and the Digital Research Alliance of Canada (www.alliancecan.ca). The operation of the supercomputers used for this research is funded by the Canada Foundation for Innovation (CFI), the Ministère de la Science, de l’Économie et de l’Innovation du Québec (MESI), and the Fonds de recherche du Québec – Nature et technologies (FRQ-NT). V.B.-C. acknowledges support by the NSERC

Alexander Graham Bell Canada Graduate Scholarship doctoral program, the FRQ-NT B2 Doctoral Scholarship and the Hydro-Québec Excellence Scholarship. V.B.-C and M.C. are members of the Regroupement québécois sur les matériaux de pointe (RQMP).

Appendix

A. Luttinger-Kohn Hamiltonian

The effective mass theory derived in 1955 by Luttinger and Kohn [73] describes the behavior of the electronic bands in the vicinity of band extrema, using 2nd order $\mathbf{k} \cdot \mathbf{p}$ theory. For the 3-fold degenerate VBM of cubic materials, the Hamiltonian without SOC writes

$$H_{n,n'}(\mathbf{k}) = \begin{bmatrix} Ak_x^2 + B(k_y^2 + k_z^2) & Ck_xk_y & Ck_xk_z \\ Ck_xk_y & Ak_y^2 + B(k_x^2 + k_z^2) & Ck_yk_z \\ Ck_xk_y & Ck_yk_z & Ak_z^2 + B(k_x^2 + k_y^2) \end{bmatrix}, \quad (\text{A1})$$

where \mathbf{k} is the electronic wavevector, n, n' are band indices, and $H_{n,n'}(\mathbf{k})$ is expressed in the $\{|X\rangle, |Y\rangle, |Z\rangle\}$ basis, which forms an irreducible representation of the symmetry group of the wavevector at the VBM.

The three parameters A , B and C , known in the literature as the Luttinger parameters, can be deduced from the effective masses, m_n^* , along the [100], [110] and [111] cartesian directions in reciprocal space,

$$\begin{aligned} m_n^{*-1}[100] &= \begin{cases} 2A, \\ 2B \text{ (2-fold)}, \end{cases} \\ m_n^{*-1}[110] &= \begin{cases} (A + B \pm C), \\ 2B, \end{cases} \\ m_n^{*-1}[111] &= \begin{cases} \frac{2}{3}(A + 2B + 2C), \\ \frac{2}{3}(A + 2B - C) \text{ (2-fold)}. \end{cases} \end{aligned} \quad (\text{A2})$$

In the presence of SOC, the Hamiltonian contains an additional term, H^{SOC} (see Eq. (6.2.18)), which is treated as a perturbation in the $\mathbf{k} \cdot \mathbf{p}$ expansion. The Hamiltonian is expressed in the basis of the zeroth-order wavefunctions, which are the 4-fold $\{|3/2, m_j\rangle\}$

(heavy holes and light holes) and the 2-fold $\{|1/2, m_j\rangle\}$ (split-off),

$$H_{j,m_j}(\mathbf{k}) = \begin{bmatrix} \frac{P}{2} & L & M & 0 & i\frac{L}{\sqrt{2}} & -i\sqrt{2}M \\ L^* & \frac{P}{6} + \frac{2Q}{3} & 0 & M & -i\frac{(P-2Q)}{3\sqrt{2}} & i\frac{\sqrt{3}L}{\sqrt{2}} \\ M^* & 0 & \frac{P}{6} + \frac{2Q}{3} & -L & -i\frac{\sqrt{3}L^*}{\sqrt{2}} & -i\frac{(P-2Q)}{3\sqrt{2}} \\ 0 & M^* & -L^* & \frac{P}{2} & -i\sqrt{2}M^* & -i\frac{L^*}{\sqrt{2}} \\ -i\frac{L^*}{\sqrt{2}} & i\frac{(P-2Q)}{3\sqrt{2}} & i\frac{\sqrt{3}L}{\sqrt{2}} & i\sqrt{2}M & \frac{(P+Q)}{3} - \Delta_{\text{SOC}} & 0 \\ i\sqrt{2}M^* & -i\frac{\sqrt{3}L^*}{\sqrt{2}} & i\frac{(P-2Q)}{3\sqrt{2}} & i\frac{L}{\sqrt{2}} & 0 & \frac{(P+Q)}{3} - \Delta_{\text{SOC}} \end{bmatrix}, \quad (\text{A3})$$

where the \mathbf{k} -dependent parameters P , Q , L and M are constructed from the Luttinger parameters,

$$\begin{aligned} P(\mathbf{k}) &= (A + B)(k_x^2 + k_y^2) + 2Bk_z^2, \\ Q(\mathbf{k}) &= B(k_x^2 + k_y^2) + Ak_z^2, \\ L(\mathbf{k}) &= -\frac{iC}{\sqrt{3}}(k_x k_z - ik_y k_z), \\ M(\mathbf{k}) &= \frac{1}{\sqrt{12}} \left[(A - B)(k_x^2 - k_y^2) - 2iCk_x k_y \right]. \end{aligned} \quad (\text{A4})$$

The split-off energy, Δ_{SOC} , makes the Luttinger-Kohn Hamiltonian with SOC non-homogenous in (k_x, k_y, k_z) , resulting in the typical band warping observed around the split-off energy in light hole bands of cubic materials (see, for example, the Γ -L direction of Fig. 2 of Ref. [72]). Around this energy level, the light hole bands typically depart from the quadratic behavior characterized by m^* (SOC) to recover a curvature that resembles m^* (noSOC). Δ_{SOC} can be extracted either from experiments or from a DFT ground state calculations with SOC. Note that Eq. (A3) inherently relies on the approximation that SOC is sufficiently small, such that the zeroth-order wavefunctions form a good basis set for the H matrix. Luttinger and Kohn explicitly warned that this approximation would be less accurate in the presence of very strong SOC. In practice, we observe that the light hole bands predicted by the Luttinger-Kohn model with SOC for tellurides meander around the first-principles dispersion and predict larger light hole effective masses compared to the first-principles results (see the blue curve in Fig. 6.10, which will be further discussed in Appendix C).

In the present work, we extract the three Luttinger parameters from the effective mass tensor without SOC using density-functional perturbation theory and evaluate Δ_{SOC} from our ground state calculation with SOC. The angular-averaged effective masses for the heavy

hole and light hole bands are computed with order-4 finite differences, using the electronic dispersion calculated by diagonalizing Eq. (A3).

We note that, by construction, the Luttinger-Kohn Hamiltonian with SOC assumes inversion symmetry (without SOC, time-reversal symmetry can be used to show that the Hamiltonian matrix elements retain the same form). Hence, it should not, in principle, be applied to zincblende materials when SOC is considered. Nevertheless, we found that, despite missing some microscopic features in the very close vicinity of Γ , which will be described in the following Appendix, the Luttinger-Kohn model with SOC provides a qualitatively reliable description of the electronic bands, which does not invalidate its use for zincblende materials.

B. Dresselhaus Hamiltonian

A correct treatment of the lack of inversion symmetry in presence of SOC was made by Dresselhaus [51]. As time-reversal symmetry is preserved, the electronic dispersion without SOC still verifies $\varepsilon_{-\mathbf{k}n} = \varepsilon_{\mathbf{k}n}$, as per Kramer's theorem, but the Bloch functions no longer have to verify $u_{-\mathbf{k}n}(\mathbf{r}) = u_{\mathbf{k}n}(-\mathbf{r})$ up to a phase factor. The inclusion of SOC acts as an effective magnetic field which splits the spin-degenerate states at finite crystal momentum, with condition $\varepsilon_{-\mathbf{k}n\uparrow} = \varepsilon_{\mathbf{k}n\downarrow}$. This effect, known as Dresselhaus splitting, creates a well-defined spin texture in reciprocal space, locking the spin orientation to the crystal momentum [60]. Dresselhaus splitting has been recently observed in GaAs and InSb by circular dichroic photoemission [62]. The underlying physical mechanism is similar to the one driving Rashba splitting [81, 82], in which an asymmetry in the crystal potential along a preferred axial direction acts as an effective electric field to break inversion symmetry. The resulting spin texture, however, is different from the helical polarization associated with Rashba systems [1, 60]. See Ref. [60] for more details about the difference between the Rashba and Dresselhaus effects.

For a generic direction in \mathbf{k} -space, the eigenvalues for the heavy hole and light hole bands in the Dresselhaus model are given by

$$E_{\mathbf{k}} = \frac{k^2}{2} + \left(\frac{L + 2M}{3} \right) k^2 + y, \quad (\text{B1})$$

where the bare electron mass $m = 1$ in atomic units. The variables $y = y(\mathbf{k}, L, M, N, W)$ are the roots of a 4th-order polynomial,

$$\begin{aligned}
& y^4 - 2y^2 \left[\frac{\alpha^2}{9} k^4 + \beta (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2) + W^2 k^2 \right] + 4y W^2 N (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2) \\
& + \left[\frac{\alpha^2}{9} k^4 + \beta (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2) \right]^2 + W^4 [k^4 - 3 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)] \\
& + 2 \frac{\alpha^2}{9} W^2 (k_x^6 + k_y^6 + k_z^6) - \left[3 \frac{\alpha^2}{9} + \frac{2N^2}{3} \right] W^2 k^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2) + 21 \frac{\alpha^2}{9} W^2 k_x^2 k_y^2 k_z^2 = 0,
\end{aligned} \tag{B2}$$

in which we have defined the shorthands

$$\begin{aligned}
\alpha^2 &= (L - M)^2, \\
\beta &= \frac{N^2 - \alpha^2}{3}.
\end{aligned} \tag{B3}$$

Eq. (B2) results from the secular determinant of the $\mathbf{k} \cdot \mathbf{p}$ expansion in the 4-fold degenerate subspace at the VBM. Note that our W parameter is labelled C in the original work from Dresselhaus [51]. We renamed it to avoid confusion with the 3rd Luttinger parameter.

The Dresselhaus model is thus parametrized by four real numbers; three of them, L , M and N , play a similar role to the Luttinger parameters [83], while the last one, W , captures information about SOC and the breaking of inversion symmetry in the crystal potential [51]. Rewriting Eq. (B1) as

$$E_{\mathbf{k}} = \lambda k^2 + y, \tag{B4}$$

hence defining

$$\lambda = \frac{1}{2} + \left(\frac{L + 2M}{3} \right), \tag{B5}$$

one can conveniently parametrize the Dresselhaus model by α^2 , λ , W and $|N|$. These four parameters can be extracted from the electronic dispersion in the close vicinity of the Γ point along two high-symmetry paths: in cartesian direction $\hat{\mathbf{k}} = [1,0,0]$, Eq. (B1) reduces to

$$E_{k[1,0,0]} = \lambda k^2 \pm \sqrt{\frac{\alpha^2}{9} k^4 + W^2 k^2}, \tag{B6}$$

in which both solutions are doubly degenerate, while for $\hat{\mathbf{k}} = [1,1,1]$, one obtains

$$E_{k[1,1,1]} = \lambda k^2 + \begin{cases} \frac{|N|}{3} k^2, & (2\text{-fold}), \\ -\frac{|N|}{3} k^2 \pm \sqrt{2} W k. \end{cases} \tag{B7}$$

Eq. (B6) and (B7) reveal two peculiarities of the generic electronic dispersion with SOC near the VBM of zincblende crystals which complicate the numerical evaluation of electronic effective masses. These features are sketched in Fig. 2 and 5 of Ref. [51] and can be observed

in Fig. 6.9 for CdS. On the one hand, the Dresselhaus field shifts the heavy hole band extrema away from the Γ point; this can be read from the second case of Eq. (B7), which contains a linear term in k . This momentum offset is, however, much smaller in magnitude than what is typically observed in Rashba materials [60]. In the present work, we found it to be of the order of $5 \times 10^{-3} \text{\AA}^{-1}$ on average and at most 10^{-2}\AA^{-1} , about half a percent of the length of the reciprocal lattice vectors or less in all cases. The energy difference between the true VBM and the Γ point was at most 0.5 meV. Such a small effect will be barely visible on the band structure but will nevertheless render the finite central differences around Γ unreliable, regardless of the fact that the momentum offset does not contribute to the gFr model (see Eq.(6.2.40)).

On the other hand, if we take the higher energy solution of Eq. (B6) to be the heavy hole bands and the lower energy solutions to be the light hole bands, their slopes will be discontinuous across Γ . As a consequence, the effective mass computed from finite differences for these bands will diverge as the \mathbf{k} -point sampling gets denser. A similar feature occurs in for the light hole bands in the $\hat{\mathbf{k}} = [1,1,0]$ cartesian direction (see the upper panel of Fig. 2 of Ref. [51]), as well as for generic $\hat{\mathbf{k}}$.

Nevertheless, in line with Dresselhaus [83] and as addressed in Sec. 6.2.4, we argue that these features have little to no impact on our current application since the energy offset between Γ and the true extrema is about two orders of magnitude smaller than the LO frequency.

With these precautions in mind, we obtain the model parameters from the calculated dispersion along the [100] and [111] directions through the following relations:

- (1) The difference between the dispersion of the two non-degenerate bands given by Eq. (B7) is linear in k , with a slope proportional to W ,
- (2) the average value of these two bands is quadratic in k . The difference between this average and the dispersion of the degenerate bands in the same direction is also quadratic in k , with a curvature proportional to $|N|$,
- (3) averaging the difference between the non-degenerate bands of Eq.(B7) with the degenerate bands in the same direction results in a quadratic function in k , with curvature λ ,
- (4) the difference between the dispersion of the two pairs of degenerate states in Eq. (B6) allows extracting the α^2 factor.

To circumvent any numerical instabilities arising from the band peculiarities associated with Dresselhaus splitting, we computed the effective masses by fitting quadratic functions to the Dresselhaus model dispersion rather than relying on finite differences. The different parameters used for the Dresselhaus model are tabulated in Table S6.8 of the Supplemental Material [69]. Lastly, we emphasize that our Dresselhaus parameters should be perceived as

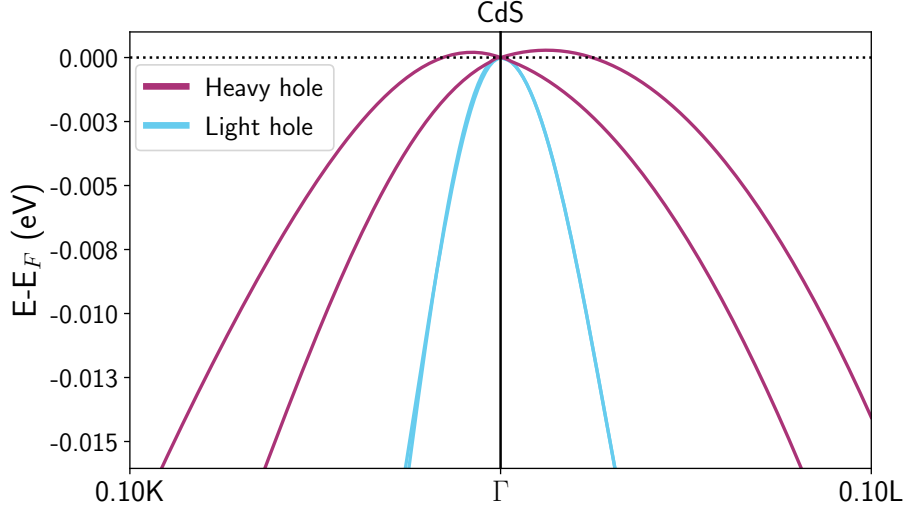


Figure 6.9 – Dresselhaus splitting in CdS. In non-centrosymmetric materials, SOC shifts the band extrema away from the Γ point, both in momentum and energy. As a consequence, numerical difficulties arise in the evaluation of the angular-averaged effective masses required by the gFr model with SOC. Zincblende CdS displays the largest effect detected in our dataset, with a momentum offset of $\sim 0.01 \text{ \AA}^{-1}$ and an energy offset of $\sim 0.5 \text{ meV}$. As these features occur very close to Γ and since $\omega_{LO} = 34.4 \text{ meV}$, we retain the physical picture of Γ -centered parabolic bands inherent to the Fröhlich model (see text).

fitting parameters rather than physical parameters. While they retain some physical essence from their original formulation (see Eq. (47) of Ref. [83]), we did not evaluate them through the $\mathbf{k} \cdot \mathbf{p}$ framework but rather fitted them to the first-principles dispersion with SOC.

C. Fitted Luttinger parameters from Dresselhaus Hamiltonian

As discussed in Appendix A, the Luttinger-Kohn Hamiltonian becomes less accurate in predicting the light holes effective masses as SOC increases, resulting in an overestimation of the SOC-induced ZPR_v decrease compared to the first-principles result, as emphasized in Sec. 6.4.2 of the main text. Nevertheless, it lacks all the numerical difficulties arising from Dresselhaus splitting. In the spirit of the original Fröhlich model, we aim for reliable effective masses with respect to the first-principles dispersion, while remaining in the physical picture of Γ -centered parabolic bands. The Luttinger-Kohn Hamiltonian thus provides a more efficient framework to evaluate the angular-averaged effective masses required by our gFr model. In this context, we propose a very simple solution: to extract fitted Luttinger parameters from the Dresselhaus model that yield more predictive light hole effective masses compared to the first-principles results in the presence of SOC.

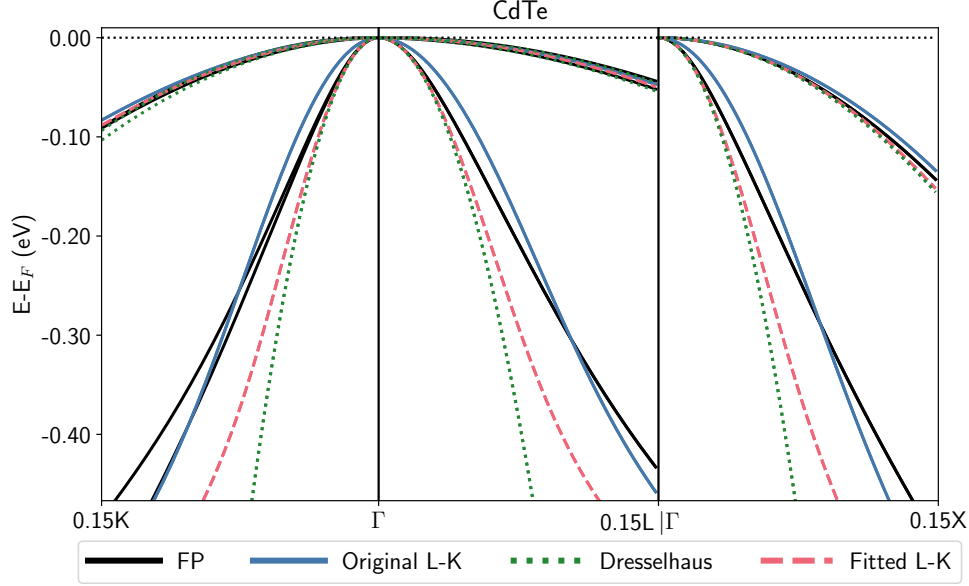


Figure 6.10 – Electronic dispersion from the Luttinger-Kohn and Dresselhaus models for CdTe, computed with the original Luttinger parameters $((A, B, C))$, solid blue lines), the fitted Dresselhaus model (green dotted lines) and the fitted Luttinger parameters extracted from the Dresselhaus model $((\tilde{A}, \tilde{B}, \tilde{C}))$, dashed red lines) as described in Appendix. C. The first-principles dispersion is shown in black. While the original model qualitatively captures the correct band warping, it overestimates the light holes’ effective masses. The modified parameters reproduce the first-principles curvature of the light holes with greater accuracy and produce the same ZPR^{gFr} as the Dresselhaus model, without the numerical difficulties.

First, we note that the momentum offset in the heavy hole bands of zincblende materials prevents us from directly fitting the Luttinger-Kohn model on the first-principles bands. The Dresselhaus model, in contrast, parametrizes this offset correctly through the W parameter. From our fitted values of λ and α^2 , we can extract fitted values for L and M , through Eq. (B3) and (B4). To do so, one nevertheless requires prior knowledge of the relative strength of L and M , as $\alpha = |L - M|$.

One can then go back to the original works of Luttinger-Kohn and Dresselhaus (see Eq. (V.10) of Ref. [73] and Eq. (47) of Ref. [83]) and note that the definitions of the (A, B) and (L, M) parameters differ only by the bare electron mass term, $\hbar^2/2m_e$, which is equal to 1/2 in atomic units. The definitions of the C and N parameters are equivalent. As L maps to A and M maps to B up to a constant shift, we can first deduce that $\text{sgn}(L - M) = \text{sgn}(A - B)$. Then, we can use this mapping to extract fitted Luttinger parameters that are optimized to reproduce the first-principles dispersion with SOC along the [100] and [111] cartesian directions in reciprocal space using the Dresselhaus model. Note that the only purpose of these parameters is to provide an accurate description of the first-principles bands in the close

vicinity of the Γ point. Hence, the predicted dispersion will rapidly become less accurate than the original Luttinger parameters for larger $|\mathbf{k}|$ values, as well as for the split-off bands.

Fig. 6.10 compares the first-principles dispersion (black) to the Luttinger-Kohn model dispersion obtained with the original set of Luttinger parameters (blue) and the fitted parameters (red) for CdTe. The dispersion obtained from the Dresselhaus model (green dotted lines) is also shown for comparison. While the original Luttinger-Kohn model closely follows the band warping for a broader portion of the Brillouin zone, the fitted parameters reproduce the light holes' curvature around Γ more accurately, similar to the one predicted by the Dresselhaus model. In turn, they predict a SOC-induced decrease of the ZPR^{gFr} in better agreement with the first-principles result, as shown by the open circle markers in Fig 6.8. The numerical values of the fitted Luttinger parameters are tabulated in Table S6.8 of the Supplemental Material [69].

D. Zero-point expansion of the lattice

The zero-point motion of the ions affects the band gap energy in two distinct ways: on the one hand, the electrons interact with the ionic motion through EPI, which is computed at the static lattice geometry, a^0 (typically, the lattice parameter obtained from a standard DFT relaxation). On the other hand, the ionic vibrations contribute to the total Helmholtz free energy of the crystal, yielding a small variation of the lattice parameter, called zero-point lattice expansion. In addition to the EPI contribution discussed in Sec. 6.2.1, the zero-point lattice expansion induces a $T = 0$ K modification of the band gap energy.

We work within the quasi-harmonic approximation, in which the main contribution to the temperature dependence of the phonon frequencies is expressed as their variation with respect to the lattice parameter, which causes the crystal to expand [84]. The Helmholtz free energy then reads:

$$\begin{aligned} F^{\text{tot}}(V,T) &\simeq F^{\text{e}}(V) + F^{\text{vib}}(V,T), \\ &= E_{\text{stat}}^{\text{e}}(V) - k_B T \ln Z^{\text{ph}}(V, T), \end{aligned} \tag{D1}$$

in which we approximate that the electronic and vibrational degrees of freedom can be separated. Since we are dealing with semiconductors and insulators, the entropic contribution of the electrons to the free energy is neglected. $E_{\text{stat}}^{\text{e}}(V)$ is the Born-Oppenheimer energy obtained from DFT and k_B is the Boltzmann constant.

At $T = 0$, the contribution of the phonon partition function is simply the zero-point energy, such that

$$F^{\text{tot}}(V,T = 0) = E_{\text{stat}}^{\text{e}}(V) + \sum_{\mathbf{q}\nu} \frac{\omega_{\mathbf{q}\nu}(V)}{2} \tag{D2}$$

For cubic materials, the lattice parameter including zero-point motion, $a(T = 0)$, is obtained from the $V(T = 0)$ volume by minimizing the $T = 0$ K Helmholtz free energy. The zero-point lattice expansion is simply the difference between the static and dynamical lattice parameters:

$$\Delta a(T = 0) = a(T = 0) - a^0 \quad (\text{D3})$$

Once $a(T = 0)$ is known, we approximate the zero-point lattice expansion contribution to the band gap ZPR, labelled $\text{ZPR}_g^{\text{ZPLE}}$, by computing the difference between the DFT band gap energy, E_g^{DFT} , evaluated at the static and $T = 0$ K geometries:

$$\text{ZPR}_g^{\text{ZPLE}} \simeq E_g^{\text{DFT}}(a(T = 0)) - E_g^{\text{DFT}}(a^0). \quad (\text{D4})$$

For all materials, we computed the zero-point lattice expansion using a an $8 \times 8 \times 8$ Γ -centered \mathbf{q} -point grid. As the construction of the Helmholtz free energy only requires the vibrational spectrum of the crystal, it converges faster with respect to the \mathbf{q} -point sampling compared to EPI. The values of $\text{ZPR}_g^{\text{ZPLE}}$ used to obtain Fig. 6.4 can be found in Table S6.5 of the Supplemental Material [69], along with experimental values of ZPR_g . Note that we did not attempt to evaluate the effect of SOC on $\text{ZPR}_g^{\text{ZPLE}}$; the same value was used to correct $\text{ZPR}_g^{\text{EPI}}$ both with and without SOC. As the zero-point lattice expansion depends mostly on the phonon frequencies (see Eq. (D2)), which are not significantly affected by SOC in the cubic materials investigated (see Table S6.7 of the Supplemental Material [69]), we expect this approximation to be fairly accurate for the purpose of this work. For a more thorough discussion of the zero-point lattice expansion and its effect on the band gap ZPR, see Ref. [80].

Supplemental Material

S6.1. Low SOC ZPR of CdTe

In this Section, we present complementary information regarding the 1% SOC calculation used to validate our heuristic interpretation of the origin of the SOC-induced ZPR decrease (Sec. 6.4.1.3 of the main text). Fig. S6.1a) and b) respectively show the electronic and phononic dispersions of CdTe without SOC (blue) and with 1% SOC (yellow dashes). The inset of Fig. S6.1a) shows a small residual split-off energy of ~ 8 meV, namely 1% of the full Δ_{SOC} for CdTe. Both band structures are virtually identical. The phonon dispersions also show excellent agreement. The largest energy difference between equivalent phonon modes in our denser $48 \times 48 \times 48$ \mathbf{q} -point grid is 0.1 meV, occurring for the LO mode at $\mathbf{q} = \Gamma$. The average energy difference per mode is at most 0.01 meV. Fig. S6.1c) confirms that the histogram decomposition of the ZPR_ν computed with 1% SOC (yellow dashes) is almost identical to the reference histogram without SOC (dark indigo) shown in Fig. 6.6 of the main text. A negligible discrepancy remains in the small q bins, most likely due to the inevitable non-zero split-off energy in the low SOC framework. The histogram with SOC (light red) is shown for comparison purposes.

Table S6.1 extends Table 6.1 of the main text to all six possible combinations of the self-energy ingredients: $\varepsilon_{\mathbf{kn}}$, $\omega_{\mathbf{q}\nu}$ and $|g_{\mathbf{knn}'}|^2$, are computed with either full SOC (“SOC” superscripts) or artificially reduced 1% SOC (“low” superscripts). The corresponding histograms are shown in Fig. S6.2. Note that the data labelled “only $\varepsilon_{\mathbf{kn}}^{\text{low}}$ ” and “only $\varepsilon_{\mathbf{kn}}^{\text{SOC}}$ ” is respectively equivalent to “ $\varepsilon_{\mathbf{kn}}^{\text{low}}, \omega_{\mathbf{q}\nu}^{\text{SOC}}, |g_{\mathbf{knn}'}^{\text{SOC}}|^2$ ” and “ $\varepsilon_{\mathbf{kn}}^{\text{SOC}}, \omega_{\mathbf{q}\nu}^{\text{low}}, |g_{\mathbf{knn}'}^{\text{low}}|^2$ ”.

As discussed in the main text, the different datasets can be grouped according to whether SOC is included or not in the bare electronic eigenvalues (see the caption of Fig. S6.2 and notice the difference between the upper and lower panels). When either including or excluding SOC only in $\varepsilon_{\mathbf{kn}}^0$ (resp. Fig. S6.2a) and d)), some small discrepancies yet remain when compared to the reference lines (SOC (dark indigo) and noSOC (light red)), which can be attributed to small variations of the EPI matrix elements. Including or excluding SOC only in the phonon frequencies has virtually no effect on the final results for this material, as their corresponding histograms (resp. Fig. S6.2b) and e)) reproduce almost perfectly the corresponding reference lines.

Table S6.1 – VBM, CBM and band gap ZPR for different combinations of SOC strength. In extension to Table 6.1 of the main text, all possible combinations of $\varepsilon_{\mathbf{kn}}$, $\omega_{\mathbf{q}\nu}$ and $|g_{\mathbf{knn}'}|^2$ computed with either full SOC (“SOC” superscripts) or artificially reduced 1% SOC (“low” superscripts) are displayed.

Combination	ZPR (meV)		
	VBM	CBM	ZPR _g
noSOC	16.4	-0.4	-16.8
SOC	11.4	0.4	-11.0
1% SOC	16.2	-0.5	-16.7
only $\varepsilon_{\mathbf{kn}}^{\text{low}}$	17.4	-0.8	-18.3
only $\varepsilon_{\mathbf{kn}}^{\text{SOC}}$	11.1	-0.8	-11.9
$\varepsilon_{\mathbf{kn}}^{\text{low}}, \omega_{\mathbf{q}\nu}^{\text{SOC}}, g_{\mathbf{knn}'}^{\text{low}} ^2$	16.3	-0.6	-16.9
$\varepsilon_{\mathbf{kn}}^{\text{SOC}}, \omega_{\mathbf{q}\nu}^{\text{low}}, g_{\mathbf{knn}'}^{\text{SOC}} ^2$	11.3	0.4	-10.9
$\varepsilon_{\mathbf{kn}}^{\text{low}}, \omega_{\mathbf{q}\nu}^{\text{low}}, g_{\mathbf{knn}'}^{\text{SOC}} ^2$	17.3	-0.8	-18.1
$\varepsilon_{\mathbf{kn}}^{\text{SOC}}, \omega_{\mathbf{q}\nu}^{\text{SOC}}, g_{\mathbf{knn}'}^{\text{low}} ^2$	11.1	-0.9	-12.0

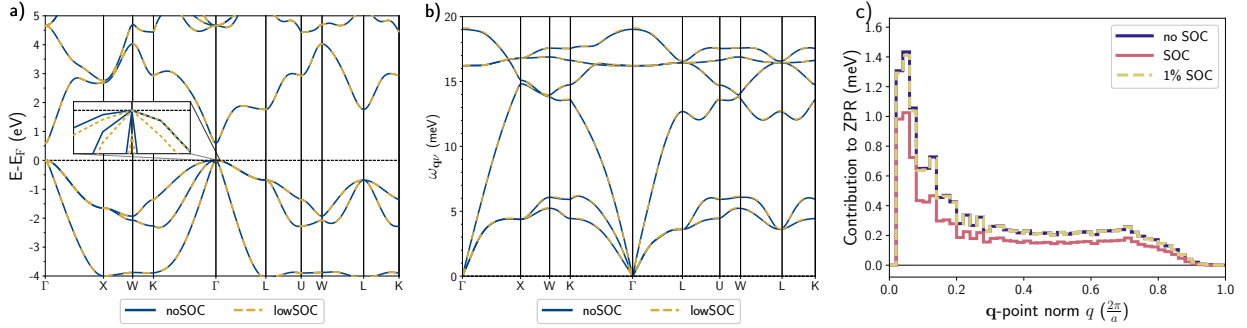


Figure S6.1 – Band structure, phonon dispersion and ZPR decomposition of zincblende CdTe, without SOC (solid blue or indigo lines) and with SOC artificially reduced to 1% (yellow dashed lines). a) Both band structures are virtually identical, except in the close vicinity of the Γ point (see inset), where a small split-off energy of $\Delta_{\text{SOC}}^{\text{low}} \simeq 8 \text{ meV} \approx 0.01 \Delta_{\text{SOC}}$ can be observed. b) The low SOC data adequately reproduces the phonon dispersion without SOC. The maximal difference of 0.1 meV occurs for the LO mode at $\mathbf{q} = \Gamma$. c) The contribution of the different phonon modes with wavevector norm $|q|$ to the ZPR_v computed with low SOC (yellow dashes) reproduces the histogram without SOC (dark indigo) shown in Fig. 6.6 of the main text. The ZPR decomposition with SOC (light red) is also shown for comparison.

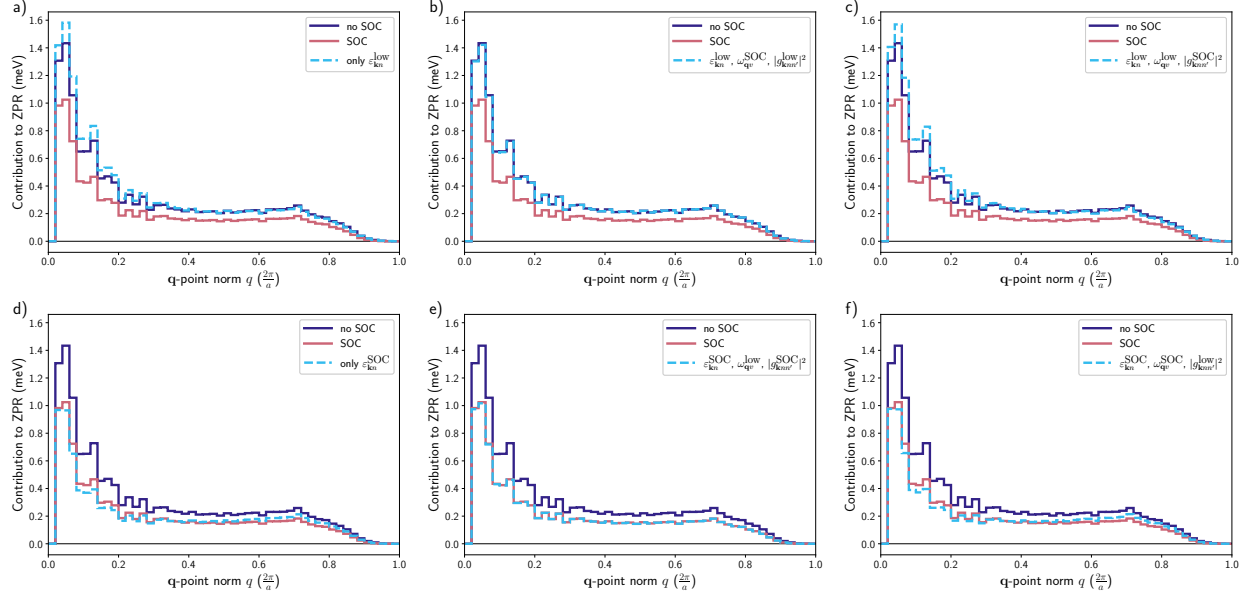


Figure S6.2 – Contribution to the total ZPR_v of CdTe for different q -point norms, for a $48 \times 48 \times 48$ Γ -centered q -point grid. In extension to Fig. 6.6 of the main text, all possible combinations of $\varepsilon_{\mathbf{k}n}^0$, $\omega_{\mathbf{q}\nu}$ and $|g_{\mathbf{k}n n'}|^2$ computed with either full SOC (“SOC” superscripts) or artificially reduced 1% SOC (“low” superscripts) are displayed with cyan dashes in the different subfigures (see subfigure legends for details). The histograms with SOC (light red) and without SOC (dark indigo) serve as references. The strength of SOC used to evaluate the electronic eigenvalues is sufficient to group the different histograms into two classes: one with 1% SOC (upper panels a), b) and c)), and one with full SOC (lower panels, d), e) and f)). The small variations with respect to the reference lines can most likely be attributed to small discrepancies in EPI matrix elements.

S6.2. Numerical tables

Table S6.2 regroups the relevant calculation parameters for all materials investigated in the present work, as well as the band gap location and band gap energy with and without SOC. Table S6.3 provides the numerical values of the ZPR for the VBM, CBM and band gap with and without SOC computed from first principles, and the ratio $R = \text{ZPR}(\text{SOC}) / \text{ZPR}(\text{noSOC})$ which is shown in Fig. 6.2 of the main text. Table S6.4 compares the band edge and band gap renormalization ratio R computed at $T = 300$ K and $T = 0$ K, and confirms that the main conclusions of this article, obtained at $T = 0$ K, remain valid at finite temperature. Table S6.5 reports the contribution of the zero-point lattice expansion to the band gap ZPR, which was added to the EPI contribution to produce Fig. 6.4 of the main text, as well as the experimental values used for comparison. Table S6.6 contains numerical results for the VBM, CBM and band gap ZPR from the gFr model (Eq. (6.2.34) of the main text), evaluated using the physical parameters reported in Table S6.7. Table S6.8 finally reports the model parameters used to construct the Luttinger-Kohn and Dresselhaus Hamiltonians, enabling the computation of the angular-averaged effective masses for the VBM of zincblende materials. The fitted Luttinger parameters obtained using the procedure described in Appendix C of the main text are also reported.

Table S6.2 – Calculations parameters for twenty materials. From left to right, material ID on the Materials Project, cubic lattice parameter (a), cut-off energy for the plane-wave basis set (E_{cut}), band gap edges (VBM-CBM), band gap energy (E_g) with and without SOC, and \mathbf{q} -point sampling for the BZ ($N \times N \times N$ Γ -centered Monkhorst-Pack grids). All materials were computed with $(6 \times 6 \times 6) \times 4$ shifts Monkhorst-Pack \mathbf{k} -point grids. The broadening parameter η for the self-energy was set to 0.01 eV for all materials.

Material	MP-ID	Space group	a (bohr)	E_{cut} (Ha)	Gap	E_g^{noSOC} (eV)	E_g^{SOC} (eV)	\mathbf{q} -point sampling
Si	149	Fd $\bar{3}m$ [227]	10.3348	25	$\Gamma - \Delta$	0.607	0.591	$24 \times 24 \times 24$
ZnS	10695	F $\bar{4}3m$ [216]	10.2859	40	$\Gamma - \Gamma$	2.020	2.000	$48 \times 48 \times 48$
CdS	2469	F $\bar{4}3m$ [216]	11.2021	45	$\Gamma - \Gamma$	1.040	1.024	$48 \times 48 \times 48$
CaS	1672	Fm $\bar{3}m$ [225]	10.7900	45	$\Gamma - X$	2.382	2.350	$48 \times 48 \times 48$
SrS	1087	Fm $\bar{3}m$ [225]	11.4434	45	$\Gamma - X$	2.520	2.484	$48 \times 48 \times 48$
BaS	1500	Fm $\bar{3}m$ [225]	12.1734	45	$\Gamma - X$	2.185	2.140	$48 \times 48 \times 48$
BAs	10044	F $\bar{4}3m$ [216]	9.0880	40	$\Gamma - \Delta$	1.216	1.149	$48 \times 48 \times 48$
Ge	32	Fd $\bar{3}m$ [227]	10.5220	40	$\Gamma - L$	0.333	0.234	$48 \times 48 \times 48$
GaAs	2534	F $\bar{4}3m$ [216]	10.8627	40	$\Gamma - \Gamma$	0.149	0.044	$64 \times 64 \times 64$
GaAs-exp	2534	F $\bar{4}3m$ [216]	10.6770	40	$\Gamma - \Gamma$	0.531	0.424	$64 \times 64 \times 64$
ZnSe	1190	F $\bar{4}3m$ [216]	10.8330	40	$\Gamma - \Gamma$	1.155	1.031	$48 \times 48 \times 48$
CdSe	2691	F $\bar{4}3m$ [216]	11.7114	50	$\Gamma - \Gamma$	0.493	0.374	$48 \times 48 \times 48$
SrSe	2758	Fm $\bar{3}m$ [225]	11.8947	45	$\Gamma - X$	2.252	2.118	$48 \times 48 \times 48$
CaSe	1415	Fm $\bar{3}m$ [225]	11.2563	45	$\Gamma - X$	2.082	1.946	$48 \times 48 \times 48$
BaSe	1253	Fm $\bar{3}m$ [225]	12.6054	45	$\Gamma - X$	1.979	1.842	$48 \times 48 \times 48$
AlSb	2624	F $\bar{4}3m$ [216]	11.7621	40	$\Gamma - \Delta$	1.268	1.059	$48 \times 48 \times 48$
ZnTe	2176	F $\bar{4}3m$ [216]	11.6819	40	$\Gamma - \Gamma$	1.071	0.797	$48 \times 48 \times 48$
CdTe	406	F $\bar{4}3m$ [216]	12.5133	50	$\Gamma - \Gamma$	0.580	0.315	$48 \times 48 \times 48$
BaTe	1000	Fm $\bar{3}m$ [225]	13.3727	45	$\Gamma - X$	1.620	1.347	$48 \times 48 \times 48$
SrTe	1958	Fm $\bar{3}m$ [225]	12.6950	45	$\Gamma - X$	1.785	1.506	$48 \times 48 \times 48$
CaTe	1519	Fm $\bar{3}m$ [225]	12.0793	45	$\Gamma - X$	1.563	1.275	$48 \times 48 \times 48$

Table S6.3 – First-principles ZPR for the VBM, CBM and band gap, split-off energy Δ_{SOC} and LO phonon frequency ω_{LO} for the twenty materials in our set. $R = \text{ZPR}(\text{SOC})/\text{ZPR}(\text{noSOC})$ is unitless. The ratio for the CBM of CdTe is not meaningful as its ZPR changes sign upon inclusion of SOC. For GaAs, results obtained with both the theoretically relaxed lattice parameter (“GaAs-theo”) and experimental lattice parameter [74] (“GaAs-exp”) are provided. Note also that the values without SOC are slightly different but very close to those reported in Ref. [13], as we use a different method to extrapolate the converged value to an infinite number of \mathbf{q} -points.

Material	Δ_{SOC} (meV)	ω_{LO} (meV)	VBM (meV)			CBM (meV)			Total ZPR (meV)		
			noSOC	SOC	R	noSOC	SOC	R	noSOC	SOC	R
Si	47	62.4	34.3	33.1	0.965	-24.9	-24.9	1.000	-59.2	-58.0	0.980
ZnS	60	40.6	48.1	46.7	0.971	-36.7	-36.6	0.997	-84.8	-83.3	0.982
CdS	49	34.4	43.3	41.7	0.963	-23.1	-23.1	1.000	-66.4	-64.8	0.976
CaS	93	41.6	78.3	76.2	0.973	-55.5	-55.4	0.998	-133.8	-131.7	0.984
SrS	106	34.3	83.0	80.1	0.965	-51.9	-51.8	0.998	-134.9	-132.0	0.978
BaS	137	30.1	87.5	82.9	0.947	-50.0	-49.9	0.998	-137.5	-132.7	0.966
BAs	209	84.4	45.4	41.5	0.914	-52.0	-52.0	1.000	-97.4	-93.5	0.960
Ge	302	38.7	16.4	14.0	0.854	-15.9	-15.9	1.000	-32.5	-30.1	0.926
GaAs	328	33.5	19.0	16.0	0.842	2.1	1.9	0.905	-16.9	-14.1	0.834
GaAs-exp	335	35.4	19.0	16.0	0.842	-4.1	-4.0	0.976	-23.1	-20.0	0.866
ZnSe	381	29.3	29.5	25.4	0.861	-10.6	-10.4	0.981	-40.1	-35.8	0.893
CdSe	364	23.6	24.7	20.4	0.826	-5.2	-4.8	0.923	-29.9	-25.2	0.843
SrSe	403	24.0	51.6	46.8	0.907	-35.6	-35.4	0.994	-87.2	-82.2	0.943
CaSe	407	31.3	49.5	45.0	0.909	-40.5	-40.4	0.998	-90.0	-85.4	0.949
BaSe	416	20.3	53.5	47.7	0.892	-33.5	-33.3	0.994	-87.0	-81.0	0.931
AlSb	661	39.8	24.0	19.0	0.792	-27.4	-27.4	1.000	-50.8	-46.4	0.903
ZnTe	881	24.1	18.0	13.2	0.733	-1.5	-1.2	0.800	-19.5	-14.4	0.738
CdTe	849	19.1	16.4	11.4	0.695	-0.4	0.4	n/a	-16.8	-11.0	0.655
BaTe	850	16.3	31.5	25.9	0.822	-22.4	-21.8	0.973	-53.9	-47.6	0.885
SrTe	864	19.5	30.8	25.4	0.825	-24.2	-23.8	0.983	-55.0	-49.2	0.895
CaTe	891	26.1	28.5	22.5	0.789	-27.8	-27.3	0.982	-56.3	-49.8	0.885

Table S6.4 – Ratio of eigenenergy renormalization with and without SOC, $\Delta\varepsilon(\text{SOC})/\Delta\varepsilon(\text{noSOC})$ evaluated at $T = 300\text{ K}$ and $T = 0\text{ K}$, for the VBM, CBM and band gap. All ratios are unitless. Note that $\Delta\varepsilon(T = 0\text{ K})$ corresponds by definition to the ZPR reported in Table S6.3; hence, the ratios at $T = 0\text{ K}$ are identical. We include them here for ease of comparison. The CBM ratios for CdTe are absent from this Table for the same reason as in Table S6.3.

Material	$\Delta\varepsilon(\text{SOC})/\Delta\varepsilon(\text{noSOC})$					
	VBM		CBM		Total ZPR	
	0 K	300 K	0 K	300 K	0 K	300 K
Si	0.965	0.970	1.000	1.000	0.980	0.982
ZnS	0.971	0.948	0.997	0.996	0.982	0.973
CdS	0.963	0.930	1.000	1.021	0.976	0.970
CaS	0.973	0.944	0.998	0.998	0.984	0.963
SrS	0.965	0.917	0.998	1.000	0.978	0.942
BaS	0.947	0.906	0.998	0.997	0.966	0.930
BAs	0.914	0.907	1.000	1.000	0.960	0.958
Ge	0.847	0.837	1.000	1.002	0.917	0.937
GaAs	0.842	0.801	0.905	1.112	0.834	0.641
ZnSe	0.861	0.875	0.981	1.000	0.893	0.921
CdSe	0.826	0.828	0.923	0.918	0.843	0.845
SrSe	0.907	0.874	0.994	0.995	0.943	0.916
CaSe	0.909	0.827	0.998	1.002	0.949	0.889
BaSe	0.892	0.878	0.994	0.995	0.931	0.916
AlSb	0.792	0.799	1.000	1.007	0.903	0.909
ZnTe	0.733	0.704	0.800	0.954	0.738	0.747
CdTe	0.695	0.651	n/a	n/a	0.655	0.555
BaTe	0.822	0.824	0.973	0.981	0.885	0.881
SrTe	0.825	0.822	0.983	0.982	0.895	0.885
CaTe	0.789	0.736	0.982	0.982	0.885	0.838

Table S6.5 – Comparison between the total band gap ZPR and experimental values. The contribution of the zero-point lattice expansion to the band gap ZPR, ZPR_g^{ZPLE} , is added to the EPI contribution, ZPR_g^{EPI} (see Table S6.3), to obtain the total band gap ZPR, ZPR_g^{tot} , which is compared to experimental data shown in Fig. 6.4 of the main text, ZPR_g^{exp} . Note that the same value of ZPR_g^{ZPLE} , computed without SOC, was applied to our ZPR_g^{EPI} results, both with and without SOC. See Appendix D of the main text for more details. All ZPR values are in meV. $R=ZPR^{tot}/ZPR^{exp}$ is unitless. ZPR^{tot} is rounded to 1 meV before computing R.

Material	ZPR_g^{ZPLE}	ZPR_g^{exp}	ZPR_g^{tot} (noSOC)	R (noSOC)	ZPR_g^{tot} (SOC)	R (SOC)
Si	8.7	-59 [85]	-51	0.86	-49	0.83
ZnS	-24.2	-105 [86]	-109	1.04	-108	1.02
CdS	-10.5	-62 [87]	-77	1.24	-76	1.24
Ge	-9.3	-52 [88]	-42	0.81	-39	0.75
GaAs	-31.4	-60 [89]	-48	0.80	-46	0.77
ZnSe	-17.2	-55 [89]	-57	1.04	-53	0.96
CdSe	-7.2	-38 [89]	-37	0.97	-32	0.84
AlSb	5.5	-35 [89]	-45	1.24	-40	1.11
ZnTe	-17.9	-40 [89]	-37	0.93	-32	0.80
CdTe	-9.0	-16 [89]	-26	1.63	-20	1.25

Table S6.6 – Generalized Fröhlich model ZPR. $R=ZPR(\text{SOC})/ZPR(\text{noSOC})$ is unitless. When SOC is included, the angular-averaged effective masses entering Eq. (6.2.34) of the main text were evaluated from the electronic dispersion obtained either with the Luttinger-Kohn Hamiltonian (SOC-LK), from DFT (SOC-DFT), or from the Luttinger-Kohn Hamiltonian with fitted Luttinger parameters (SOC- $\widetilde{\text{LK}}$, see Appendix C of the main text). For the VBM, the special treatment of the SOC-DFT data based on the Dresselhaus Hamiltonian is described in Appendix B of the main text. The different ratios (R-LK, R-DFT, R- $\widetilde{\text{LK}}$) refer to the calculation method with SOC. All effective masses without SOC were computed with density-functional perturbation theory. For GaAs, results obtained with both theoretically relaxed lattice parameter (“GaAs”) and experimental lattice parameter (“GaAs-exp”) are provided.

Material	VBM (meV)							CBM (meV)			Total ZPR (meV)		
	no SOC	SOC-LK	SOC-DFT	SOC- $\widetilde{\text{LK}}$	R-LK	R-DFT	R- $\widetilde{\text{LK}}$	no SOC	SOC-DFT	R-DFT	no SOC	SOC-DFT	R-DFT
ZnS	40.1	34.7	34.3	34.5	0.865	0.855	0.860	-18.6	-18.6	1.000	-58.7	-52.9	0.901
CdS	38.7	32.9	32.4	32.7	0.850	0.837	0.845	-14.9	-14.9	1.000	-53.6	-47.3	0.882
CaS	74.5	70.1	69.7	69.7	0.941	0.936	0.936	-61.3	-61.2	0.998	-135.8	-130.9	0.964
SrS	80.6	74.8	74.3	74.3	0.928	0.922	0.922	-61.5	-61.4	0.998	-142.1	-135.7	0.955
BaS	86.8	78.8	78.1	78.1	0.908	0.900	0.900	-63.1	-63.1	1.000	-149.9	-141.2	0.942
BAs	0.5	0.5	0.5	0.5	1.000	1.000	1.000	-0.5	-0.5	1.000	-1.0	-1.0	1.000
GaAs	3.3	2.4	2.2	2.2	0.727	0.667	0.667	-0.5	-0.4	0.800	-3.8	-2.6	0.690
GaAs-exp	3.4	2.9	2.7	2.7	0.853	0.794	0.794	-1.0	-1.0	1.000	-4.4	-3.7	0.841
ZnSe	21.7	18.0	17.5	17.5	0.829	0.806	0.806	-8.1	-7.9	0.975	-29.8	-25.4	0.852
CdSe	19.7	15.7	15.0	15.0	0.797	0.761	0.761	-5.5	-5.1	0.927	-25.2	-20.1	0.798
SrSe	50.9	47.1	45.8	45.8	0.925	0.900	0.900	-41.4	-41.2	0.995	-92.3	-87.0	0.943
CaSe	48.9	45.5	44.6	44.6	0.935	0.912	0.912	-42.9	-42.7	0.995	91.8	-87.3	0.951
BaSe	53.8	48.9	47.5	47.5	0.909	0.883	0.883	-42.2	-41.9	0.993	-96.0	-89.4	0.931
AlSb	4.0	3.3	3.1	3.1	0.825	0.775	0.775	-3.6	-3.5	0.972	-7.6	-6.6	0.868
ZnTe	11.1	8.9	8.3	8.3	0.802	0.748	0.748	-4.3	-3.9	0.907	-15.4	-12.2	0.792
CdTe	11.9	9.1	8.2	8.2	0.765	0.689	0.689	-3.7	-2.9	0.784	-15.6	-11.1	0.712
BaTe	32.6	29.6	27.7	27.7	0.908	0.850	0.850	-28.1	-27.4	0.975	-60.7	-55.1	0.908
SrTe	31.5	29.0	27.3	27.3	0.921	0.867	0.867	-27.8	-27.3	0.982	-59.3	-54.6	0.921
CaTe	30.2	28.1	26.4	26.4	0.930	0.874	0.874	-28.7	-28.2	0.983	-58.9	-54.6	0.927

Table S6.7 – Generalized Fröhlich model parameters used in the evaluation of Eq. (6.2.34) of the main text: LO phonon frequency ω_{LO} (in meV), optical dielectric constant ϵ^∞ , Born effective charges Z^* and absolute value of the angular and band averaged square root effective masses $\langle\langle(m^*)^{\frac{1}{2}}\rangle\rangle$ for the VBM and CBM, with and without SOC. Square root effective masses with SOC were computed with the Dresselhaus model.

Material	noSOC					SOC				
	ω_{LO}	ϵ^∞	Z^*	$\langle\langle(m^*)^{\frac{1}{2}}\rangle\rangle_{\text{v}}$	$\langle\langle(m^*)^{\frac{1}{2}}\rangle\rangle_{\text{c}}$	ω_{LO}	ϵ^∞	Z^*	$\langle\langle(m^*)^{\frac{1}{2}}\rangle\rangle_{\text{v}}$	$\langle\langle(m^*)^{\frac{1}{2}}\rangle\rangle_{\text{c}}$
ZnS	40.6	5.97	2.03	0.882	0.409	40.6	5.97	2.03	0.756	0.409
CdS	34.4	6.21	2.23	0.893	0.343	34.4	6.22	2.23	0.749	0.343
CaS	41.6	4.99	2.37	0.836	0.688	41.6	4.99	2.37	0.784	0.688
SrS	34.2	4.63	2.40	0.893	0.681	34.3	4.63	2.40	0.824	0.681
BaS	30.1	4.78	2.60	0.950	0.691	30.0	4.78	2.60	0.855	0.690
BAs	84.4	9.81	0.45	0.596	0.576	84.4	9.81	0.45	0.527	0.575
GaAs	33.5	15.31	2.23	0.581	0.095	33.5	15.42	2.23	0.395	0.063
GaAs-exp	35.4	13.78	2.12	0.594	0.175	35.5	13.85	2.12	0.454	0.169
ZnSe	29.3	7.35	2.10	0.802	0.299	29.3	7.38	2.10	0.651	0.296
CdSe	23.6	7.83	2.31	0.812	0.226	23.5	7.91	2.31	0.630	0.215
SrSe	24.0	5.15	2.39	0.812	0.660	24.0	5.17	2.40	0.734	0.660
CaSe	31.3	5.61	2.38	0.761	0.668	31.4	5.62	2.38	0.696	0.667
BaSe	20.3	5.26	2.59	0.857	0.672	20.3	5.28	2.59	0.760	0.671
AlSb	39.8	12.02	1.83	0.656	0.584	39.8	12.13	1.84	0.514	0.585
ZnTe	24.1	9.05	2.09	0.713	0.276	24.1	9.25	2.09	0.554	0.261
CdTe	19.1	8.89	2.29	0.738	0.228	19.0	9.27	2.29	0.545	0.191
BaTe	16.3	5.93	2.59	0.749	0.645	16.3	6.02	2.59	0.651	0.642
SrTe	19.5	5.90	2.42	0.719	0.633	19.4	5.96	2.42	0.633	0.632
CaTe	26.1	6.57	2.42	0.673	0.640	26.1	6.63	2.42	0.598	0.638

Table S6.8 – Luttinger-Kohn and Dresselhaus model parameters. This table re-groups the parameters used in Eq. (A3) and (B1)–(B3) of the main text to evaluate the electronic dispersion from which one can extract the angular-averaged effective masses entering our gFr model with SOC (Eq. (6.2.34) of the main text). Note that the W parameter, labeled C in the original work from Dresselhaus [83], was renamed to avoid confusion with the 3rd Luttinger parameter. To extract these parameters from our fitting parameters, λ and α^2 (see Eq. (B3) and (B5)), keep note that Eq. (B1) is always negative for hole-like bands, and that we suppose that $|L| > |M|$, as observed for the Luttinger parameters. Note that $W = 0$ indicates a centrosymmetric materials. See Appendix C for more details. We also provide the fitted Luttinger parameters (\tilde{A} , \tilde{B} , \tilde{C}) derived from the Dresselhaus model. All parameters are in atomic units. The split-off energy used in the Luttinger-Kohn Hamiltonian can be found in Table S6.3. Note that the Luttinger parameters (A , B , C) listed here for the zincblende materials were originally published by us in Ref. [14]. We include them here for completeness.

Material	Dresselhaus				Luttinger-Kohn			Fitted Luttinger-Kohn		
	L	M	N	$10^{-3}W$	A	B	C	\tilde{A}	\tilde{B}	\tilde{C}
ZnS	-3.293	-1.206	-3.234	1.352	-2.751	-0.694	-3.170	-2.793	-0.706	-3.234
CdS	-4.575	-1.124	-4.443	2.495	-3.999	-0.605	-4.321	-4.075	-0.624	-4.443
CaS	-2.715	-0.912	-1.290	0.000	-2.186	-0.409	-1.276	-2.215	-0.412	-1.290
SrS	-2.774	-0.799	-1.139	0.000	-2.220	-0.300	-1.122	-2.274	-0.299	-1.139
BaS	-2.859	-0.694	-0.753	0.000	-2.249	-0.204	-0.736	-2.359	-0.194	-0.753
BA _s	-2.853	-2.598	-3.925	0.183	-2.337	-2.104	-3.912	-2.353	-2.098	-3.925
GaAs	-189.915	-1.957	-190.432	0.323	-54.896	-1.362	-55.859	-189.415	-1.457	-190.432
GaAs-exp	-21.263	-2.049	-21.968	0.360	-16.371	-1.450	-17.412	-20.763	-1.549	-21.968
ZnSe	-6.589	-1.327	-6.630	1.074	-5.340	-0.791	-5.834	-6.089	-0.827	-6.630
CdSe	-13.236	-1.231	-13.202	2.120	-9.504	-0.684	-9.881	-12.736	-0.731	-13.202
SrSe	-3.440	-0.894	-1.703	0.000	-2.752	-0.380	-1.592	-2.940	-0.394	-1.703
CaSe	-3.428	-1.032	-1.924	0.000	-2.769	-0.508	-1.817	-2.928	-0.532	-1.924
BaSe	-3.416	-0.773	-1.131	0.000	-2.680	-0.270	-1.047	-2.916	-0.273	-1.131
AlSb	-8.382	-1.992	-9.040	0.109	-6.473	-1.372	-7.520	-7.882	-1.492	-9.040
ZnTe	-9.533	-1.641	-9.832	0.769	-6.495	-1.032	-7.174	-9.033	-1.141	-9.832
CdTe	-18.628	-1.486	-18.819	1.711	-9.517	-0.867	-10.033	-18.128	-0.986	-18.819
BaTe	-4.176	-0.941	-1.543	0.000	-3.164	-0.399	-1.329	-3.676	-0.441	-1.543
SrTe	-4.198	-1.094	-2.291	0.000	-3.226	-0.534	-1.984	-3.698	-0.594	-2.291
CaTe	-4.209	-1.293	-2.587	0.000	-3.261	-0.707	-2.273	-3.709	-0.793	-2.587

Bibliography

- [1] M. S. Bahramy and N. Ogawa. Bulk Rashba Semiconductors and Related Quantum Phenomena. *Advanced Materials*, 29(25):1605911, 2017. doi: <https://doi.org/10.1002/adma.201605911>.
- [2] S. I. Pekar. Local quantum states of electrons in an ideal ion crystal. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki*, 16(4):341–348, 1946.
- [3] L. D. Landau and S. I. Pekar. Effective Mass of a Polaron. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki*, 18(5):419–423, 1948, <http://archive.ujp.bitp.kiev.ua/files/journals/53/si/53SI15p.pdf>.
- [4] H. Fröhlich, H. Pelzer, and S. Zienau. Properties of slow electrons in polar materials. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(314):221–242, 1950. doi: [10.1080/14786445008521794](https://doi.org/10.1080/14786445008521794).
- [5] R. P. Feynman. Slow Electrons in a Polar Crystal. *Physical Review*, 97(3):660–665, 1955. doi: [10.1103/PhysRev.97.660](https://doi.org/10.1103/PhysRev.97.660).
- [6] J. T. Devreese and A. S. Alexandrov. Fröhlich polaron and bipolaron: recent developments. *Rep. Prog. Phys.*, 72(6):066501, 2009. doi: [10.1088/0034-4885/72/6/066501](https://doi.org/10.1088/0034-4885/72/6/066501).
- [7] F. Giustino. Electron-phonon interactions from first principles. *Reviews of Modern Physics*, 89(1):015003, 2017. doi: [10.1103/RevModPhys.89.015003](https://doi.org/10.1103/RevModPhys.89.015003).
- [8] P. B. Allen and V. Heine. Theory of the temperature dependence of electronic band structures. *J. Phys. C: Solid State Phys.*, 9(12):2305, 1976. doi: [10.1088/0022-3719/9/12/013](https://doi.org/10.1088/0022-3719/9/12/013).
- [9] P. B. Allen and M. Cardona. Theory of the temperature dependence of the direct gap of germanium. *Physical Review B*, 23(4):1495–1505, 1981. doi: [10.1103/PhysRevB.23.1495](https://doi.org/10.1103/PhysRevB.23.1495).
- [10] P. B. Allen and M. Cardona. Temperature dependence of the direct gap of Si and Ge. *Physical Review B*, 27(8):4760–4769, 1983. doi: [10.1103/PhysRevB.27.4760](https://doi.org/10.1103/PhysRevB.27.4760).
- [11] H. Y. Fan. Temperature Dependence of the Energy Gap in Semiconductors. *Phys. Rev.*, 82(6):900–905, 1951. doi: [10.1103/PhysRev.82.900](https://doi.org/10.1103/PhysRev.82.900).
- [12] E. Antončík. On the theory of temperature shift of the absorption curve in non-polar crystals. *Czechoslovakij fiziceskij zurnal*, 5(4):449–461, 1955. doi: [10.1007/BF01687209](https://doi.org/10.1007/BF01687209).
- [13] A. Miglio, V. Brousseau-Couture, E. Godbout, G. Antonius, Y. H. Chan, S. G. Louie, M. Côté, M. Giantomassi, and X. Gonze. Predominance of non-adiabatic effects in zero-point renormalization of the electronic band gap. *npj Computational Materials*, 6(1):167, 2020. doi: [10.1038/s41524-020-00434-z](https://doi.org/10.1038/s41524-020-00434-z).
- [14] B. Guster, P. Melo, B. A. A. Martin, V. Brousseau-Couture, J. C. de Abreu, A. Miglio, M. Giantomassi, M. Côté, J. M. Frost, M. J. Verstraete, and X. Gonze. Fröhlich polaron effective mass and localization length in cubic materials: Degenerate and

- anisotropic electronic bands. *Phys. Rev. B*, 104(23):235123, 2021. doi: [10.1103/PhysRevB.104.235123](https://doi.org/10.1103/PhysRevB.104.235123).
- [15] W. H. Sio, C. Verdi, S. Ponc e, and F. Giustino. Polarons from first principles, without supercells. *Phys. Rev. Lett.*, 122(24):246403, 2019. doi: [10.1103/PhysRevLett.122.246403](https://doi.org/10.1103/PhysRevLett.122.246403).
- [16] W. H. Sio, C. Verdi, S. Ponc e, and F. Giustino. *Ab initio* theory of polarons: Formalism and applications. *Physical Review B*, 99(23):235139, 2019. doi: [10.1103/PhysRevB.99.235139](https://doi.org/10.1103/PhysRevB.99.235139).
- [17] V. Vasilchenko, A. Zhugayevych, and X. Gonze. Variational polaron equations applied to the anisotropic frohlich model. *Phys. Rev. B*, 105(21):214301, 2022. doi: [10.1103/PhysRevB.105.214301](https://doi.org/10.1103/PhysRevB.105.214301).
- [18] J. Lafuente-Bartolome, C. Lian, W. H. Sio, I. G. Gurtubay, A. Eiguren, and F. Giustino. Unified approach to polarons and phonon-induced band structure renormalization. *Phys. Rev. Lett.*, 129(7):076402, 2022. doi: [10.1103/PhysRevLett.129.076402](https://doi.org/10.1103/PhysRevLett.129.076402).
- [19] J. Lafuente-Bartolome, C. Lian, W. H. Sio, I. G. Gurtubay, A. Eiguren, and F. Giustino. Ab initio self-consistent many-body theory of polarons at all couplings. *Phys. Rev. B*, 106(7):075119, 2022. doi: [10.1103/PhysRevB.106.075119](https://doi.org/10.1103/PhysRevB.106.075119).
- [20] M. Houtput and J. Tempere. Beyond the Fr ohlich Hamiltonian: Path-integral treatment of large polarons in anharmonic solids. *Phys. Rev. B*, 103(18):184306, 2021. doi: [10.1103/PhysRevB.103.184306](https://doi.org/10.1103/PhysRevB.103.184306).
- [21] N. Kandolf, C. Verdi, and F. Giustino. Many-body Green’s function approaches to the doped Fr ohlich solid: Exact solutions and anomalous mass enhancement. *Phys. Rev. B*, 105(8):085148, 2022. doi: [10.1103/PhysRevB.105.085148](https://doi.org/10.1103/PhysRevB.105.085148).
- [22] F. Macheda, P. Barone, and F. Mauri. Fr ohlich electron-phonon interaction and LO-TO splitting in doped semiconductors. *arXiv.2202.02835 [cond-mat.mtrl-sci]*, 2022. doi: [10.48550/arXiv.2202.02835](https://doi.org/10.48550/arXiv.2202.02835).
- [23] M. Schlipf, S. Ponc e, and F. Giustino. Carrier Lifetimes and Polaronic Mass Enhancement in the Hybrid Halide Perovskite $\text{CH}_3\text{NH}_3\text{PbI}_3$ from Multiphonon Fr ohlich Coupling. *Phys. Rev. Lett.*, 121(8):086402, 2018. doi: [10.1103/PhysRevLett.121.086402](https://doi.org/10.1103/PhysRevLett.121.086402).
- [24] C. Verdi and F. Giustino. Fr ohlich electron-phonon vertex from first principles. *Phys. Rev. Lett.*, 115(17):176401, 2015. doi: [10.1103/PhysRevLett.115.176401](https://doi.org/10.1103/PhysRevLett.115.176401).
- [25] P. M. M. C. de Melo, J. C. de Abreu, B. Guster, M. Giantomassi, Z. Zanolli, X. Gonze, and M. J. Verstraete. High-throughput analysis of Fr ohlich-type polaron models. *arXiv:2207.00364 [cond-mat]*, 2022. doi: [10.48550/arxiv.2207.00364](https://doi.org/10.48550/arxiv.2207.00364).
- [26] B. S. Kandemir and D. Akay. Photoinduced dynamical band gap in graphene: The effects of electron–phonon and spin–orbit interaction. *Physica Status Solidi (b)*, 255(10):1800163, 2018. doi: <https://doi.org/10.1002/pssb.201800163>.
- [27] Z. Li, Z. Ma, A. R. Wright, and C. Zhang. Spin-orbit interaction enhanced polaron effect in two-dimensional semiconductors. *Applied Physics Letters*, 90(11):112103, 2007.

doi: [10.1063/1.2713346](https://doi.org/10.1063/1.2713346).

- [28] L. Chen, Z. Ma, J. C. Cao, T. Y. Zhang, and C. Zhang. Phonon-limited mobility in two-dimensional semiconductors with spin-orbit coupling. *Applied Physics Letters*, 91(10):102115, 2007. doi: [10.1063/1.2783203](https://doi.org/10.1063/1.2783203).
- [29] E. Cappelluti, C. Grimaldi, and F. Marsiglio. Electron-phonon effects on spin-orbit split bands of two-dimensional systems. *Phys. Rev. B*, 76(8):085334, 2007. doi: [10.1103/PhysRevB.76.085334](https://doi.org/10.1103/PhysRevB.76.085334).
- [30] C. Grimaldi. Large polaron formation induced by Rashba spin-orbit coupling. *Phys. Rev. B*, 81(7):075306, 2010. doi: [10.1103/PhysRevB.81.075306](https://doi.org/10.1103/PhysRevB.81.075306).
- [31] C. Grimaldi. Energy levels of a two-dimensional hydrogen atom with spin-orbit Rashba interaction. *Phys. Rev. B*, 77(11):113308, 2008. doi: [10.1103/PhysRevB.77.113308](https://doi.org/10.1103/PhysRevB.77.113308).
- [32] K. Vardanyan, A. Vartanian, and A. Kirakosyan. Two-dimensional Fröhlich polaron with Rashba and Dresselhaus spin-orbit coupling. *Eur. Phys. J. B*, 85(11):367, 2012. doi: [10.1140/epjb/e2012-30534-y](https://doi.org/10.1140/epjb/e2012-30534-y).
- [33] H. R. Trebin and U. Rössler. Polarons in the Degenerate-Band Case. *Physica Status Solidi (b)*, 70(2):717–726, 1975. doi: [10.1002/pssb.2220700232](https://doi.org/10.1002/pssb.2220700232).
- [34] M. J. Verstraete, M. Torrent, F. Jollet, G. Zérah, and X. Gonze. Density functional perturbation theory with spin-orbit coupling: Phonon band structure of lead. *Physical Review B*, 78(4):045119, 2008. doi: [10.1103/PhysRevB.78.045119](https://doi.org/10.1103/PhysRevB.78.045119).
- [35] A. Dal Corso. Density functional perturbation theory for lattice dynamics with fully relativistic ultrasoft pseudopotentials: Application to fcc-Pt and fcc-Au. *Phys. Rev. B*, 76(5):054308, 2007. doi: [10.1103/PhysRevB.76.054308](https://doi.org/10.1103/PhysRevB.76.054308).
- [36] B. Monserrat. Electron–phonon coupling from finite differences. *J. Phys.: Condens. Matter*, 30(8):083001, 2018. doi: [10.1088/1361-648X/aaa737](https://doi.org/10.1088/1361-648X/aaa737).
- [37] M. Zacharias and F. Giustino. Theory of the special displacement method for electronic structure calculations at finite temperature. *Phys. Rev. Research*, 2(1):013357, 2020. doi: [10.1103/PhysRevResearch.2.013357](https://doi.org/10.1103/PhysRevResearch.2.013357).
- [38] F. Zheng, L. Z. Tan, S. Liu, and A. M. Rappe. Rashba Spin–Orbit Coupling Enhanced Carrier Lifetime in $\text{CH}_3\text{NH}_3\text{PbI}_3$. *Nano Lett.*, 15(12):7794–7800, 2015. doi: [10.1021/acs.nanolett.5b01854](https://doi.org/10.1021/acs.nanolett.5b01854).
- [39] A. Molina-Sánchez, M. Palumbo, A. Marini, and L. Wirtz. Temperature-dependent excitonic effects in the optical properties of single-layer MoS_2 . *Phys. Rev. B*, 93(15):155435, 2016. doi: [10.1103/PhysRevB.93.155435](https://doi.org/10.1103/PhysRevB.93.155435).
- [40] B. Monserrat and D. Vanderbilt. Temperature dependence of the bulk Rashba splitting in the bismuth tellurohalides. *Phys. Rev. Materials*, 1(5):054201, 2017. doi: [10.1103/PhysRevMaterials.1.054201](https://doi.org/10.1103/PhysRevMaterials.1.054201).

- [41] B. Monserrat and D. Vanderbilt. Temperature Effects in the Band Structure of Topological Insulators. *Phys. Rev. Lett.*, 117(22):226801, 2016. doi: [10.1103/PhysRevLett.117.226801](https://doi.org/10.1103/PhysRevLett.117.226801).
- [42] V. Brousseau-Couture, G. Antonius, and M. Côté. Temperature dependence of the topological phase transition of BiTeI from first principles. *Phys. Rev. Research*, 2(2):023185, 2020. doi: [10.1103/PhysRevResearch.2.023185](https://doi.org/10.1103/PhysRevResearch.2.023185).
- [43] J. Cao, J. Querales-Flores, S. Fahy, and I. Savić. Thermally induced band gap increase and high thermoelectric figure of merit of n-type PbTe. *Materials Today Physics*, 12:100172, 2020. doi: <https://doi.org/10.1016/j.mtphys.2019.100172>.
- [44] J. D. Querales-Flores, J. Cao, S. Fahy, and I. Savić. Temperature effects on the electronic band structure of PbTe from first principles. *Phys. Rev. Materials*, 3(5):055405, 2019. doi: [10.1103/PhysRevMaterials.3.055405](https://doi.org/10.1103/PhysRevMaterials.3.055405).
- [45] W. A. Saidi, S. Poncé, and B. Monserrat. Temperature dependence of the energy levels of methylammonium lead iodide perovskite from first-principles. *The Journal of Physical Chemistry Letters*, 7(24):5247–5252, 2016. doi: <https://doi.org/10.1021/acs.jpcclett.6b02560>.
- [46] I. Bravić and B. Monserrat. Finite temperature optoelectronic properties of BAs from first principles. *Phys. Rev. Materials*, 3(6):065402, 2019. doi: [10.1103/PhysRevMaterials.3.065402](https://doi.org/10.1103/PhysRevMaterials.3.065402).
- [47] R. Heid, K. P. Bohnen, I. Y. Sklyadneva, and E. V. Chulkov. Effect of spin-orbit coupling on the electron-phonon interaction of the superconductors Pb and Tl. *Phys. Rev. B*, 81(17):174527, 2010. doi: [10.1103/PhysRevB.81.174527](https://doi.org/10.1103/PhysRevB.81.174527).
- [48] S. Gołab and B. Wiendlocha. Electron-phonon superconductivity in CaBi₂ and the role of spin-orbit interaction. *Phys. Rev. B*, 99(10):104520, 2019. doi: [10.1103/PhysRevB.99.104520](https://doi.org/10.1103/PhysRevB.99.104520).
- [49] H. M. Tütüncü, H. Y. Uzunok, E. c. v. Karaca, E. Arslan, and G. P. Srivastava. Effects of spin-orbit coupling on the electron-phonon superconductivity in the cubic Laves-phase compounds CaIr₂ and CaRh₂. *Phys. Rev. B*, 96(13):134514, 2017. doi: [10.1103/PhysRevB.96.134514](https://doi.org/10.1103/PhysRevB.96.134514).
- [50] K. Hu, B. Gao, Q. Ji, Y. Ma, W. Li, X. Xu, H. Zhang, G. Mu, F. Huang, C. Cai, X. Xie, and M. Jiang. Effects of electron correlation, electron-phonon coupling, and spin-orbit coupling on the isovalent Pd-substituted superconductor SrPt₃P. *Phys. Rev. B*, 93(21):214510, 2016. doi: [10.1103/PhysRevB.93.214510](https://doi.org/10.1103/PhysRevB.93.214510).
- [51] G. Dresselhaus. Spin-Orbit Coupling Effects in Zinc Blende Structures. *Physical Review*, 100(2):580–586, 1955. doi: [10.1103/PhysRev.100.580](https://doi.org/10.1103/PhysRev.100.580).
- [52] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12):121102(R), 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).

- [53] S. Poncé, G. Antonius, Y. Gillet, P. Boulanger, J. Laflamme Janssen, A. Marini, M. Côté, and X. Gonze. Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation. *Phys. Rev. B*, 90(21):214304, 2014. doi: [10.1103/PhysRevB.90.214304](https://doi.org/10.1103/PhysRevB.90.214304).
- [54] X. Gonze, P. Boulanger, and M. Côté. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Annalen der Physik*, 523(1-2):168–178, 2011. doi: [10.1002/andp.201000100](https://doi.org/10.1002/andp.201000100).
- [55] L. Kleinman and D. Bylander. Efficacious Form for Model Pseudopotentials. *Physical Review Letters*, 48(20):1425, 1982. doi: <https://doi.org/10.1103/PhysRevLett.48.1425>.
- [56] X. Gonze, J. M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G. M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, P. Ghosez, J. Y. Raty, and D. Allan. First-principles computation of material properties: the ABINIT software project. *Computational Materials Science*, 25(3):478–492, 2002. doi: [https://doi.org/10.1016/S0927-0256\(02\)00325-7](https://doi.org/10.1016/S0927-0256(02)00325-7).
- [57] C. Hartwigsen, S. Goedecker, and J. Hutter. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Physical Review B*, 58(7):3641–3662, 1998. doi: [10.1103/PhysRevB.58.3641](https://doi.org/10.1103/PhysRevB.58.3641).
- [58] M. Born and K. Huang. *Dynamical theory of crystal lattices*. Oxford University Press, Oxford, 1954.
- [59] B. Guster, P. Melo, B. A. A. Martin, V. Brousseau-Couture, J. C. de Abreu, A. Miglio, M. Giantomassi, M. Côté, J. M. Frost, M. J. Verstraete, and X. Gonze. Erratum: Fröhlich polaron effective mass and localization length in cubic materials: Degenerate and anisotropic electronic bands [phys. rev. b 104, 235123 (2021)]. *Phys. Rev. B*, 105(11):119902, 2022. doi: [10.1103/PhysRevB.105.119902](https://doi.org/10.1103/PhysRevB.105.119902).
- [60] A. Manchon, H. C. Koo, J. Nitta, S. M. Frolov, and R. A. Duine. New perspectives for Rashba spin-orbit coupling. *Nature Materials*, 14(9):871–882, 2015. doi: [10.1038/nmat4360](https://doi.org/10.1038/nmat4360).
- [61] G. Mahan. *Many-Particle Physics*. Plenum, New York, 2nd edition, 1990.
- [62] S. Cho, W. Jung, J. Hong, B. Kim, G. Han, M. Leandersson, T. Balasubramanian, M. Arita, K. Shimada, J. H. Shim, C. Kim, and S. R. Park. Observation of Dresselhaus type spin splitting of zinc blende structure semiconductors by circular dichroic photoemission study. *Current Applied Physics*, 30:96–101, 2021. doi: <https://doi.org/10.1016/j.cap.2021.06.011>.
- [63] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J. M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J. B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig,

- N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. L. Marques, A. Martin, C. Martins, H. P. C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G. M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. Van Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger. The ABINIT project: Impact, environment and recent developments. *Computer Physics Communications*, 248:107042, 2020. doi: [10.1016/j.cpc.2019.107042](https://doi.org/10.1016/j.cpc.2019.107042).
- [64] X. Gonze. First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm. *Phys. Rev. B*, 55(16):10337–10354, 1997. doi: [10.1103/PhysRevB.55.10337](https://doi.org/10.1103/PhysRevB.55.10337).
- [65] X. Gonze and C. Lee. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B*, 55(16):10355–10368, 1997. doi: [10.1103/PhysRevB.55.10355](https://doi.org/10.1103/PhysRevB.55.10355).
- [66] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese. The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Computer Physics Communications*, 226:39–54, 2018. doi: [10.1016/j.cpc.2018.01.012](https://doi.org/10.1016/j.cpc.2018.01.012).
- [67] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, 1996. doi: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [68] G. Antonius. ElectronPhononCoupling: Post-processing of electron-phonon coupling calculations from Abinit, 2018, <https://github.com/GkAntonius/ElectronPhononCoupling>.
- [69] See Supplemental Material for complete results of our low SOC analysis for CdTe, as well as numerical tables containing relevant computational parameters for all materials and the numerical results discussed in the main text, which include Ref. [13, 14, 74, 83, 85–89].
- [70] G. Antonius, S. Poncé, E. Lantagne-Hurtubise, G. Auclair, X. Gonze, and M. Côté. Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure. *Phys. Rev. B*, 92(8):085137, 2015. doi: [10.1103/PhysRevB.92.085137](https://doi.org/10.1103/PhysRevB.92.085137).
- [71] S. Poncé, Y. Gillet, J. L. Janssen, A. Marini, M. Verstraete, and X. Gonze. Temperature dependence of the electronic structure of semiconductors and insulators. *The Journal of Chemical Physics*, 143:102813, 2015. doi: [10.1063/1.4927081](https://doi.org/10.1063/1.4927081).
- [72] J. Laflamme Janssen, Y. Gillet, S. Poncé, A. Martin, M. Torrent, and X. Gonze. Precise effective masses from density functional perturbation theory. *Physical Review B*, 93(20):205147, 2016. doi: [10.1103/PhysRevB.93.205147](https://doi.org/10.1103/PhysRevB.93.205147).
- [73] J. M. Luttinger and W. Kohn. Motion of Electrons and Holes in Perturbed Periodic Fields. *Phys. Rev.*, 97(4):869–883, 1955. doi: [10.1103/PhysRev.97.869](https://doi.org/10.1103/PhysRev.97.869).

- [74] C. Kittel. *Introduction to Solid State Physics*. Wiley, Hoboken, 8th edition, 2004.
- [75] J. Ma, A. S. Nissimagoudar, and W. Li. First-principles study of electron and hole mobilities of Si and GaAs. *Phys. Rev. B*, 97(4):045201, 2018. doi: [10.1103/PhysRevB.97.045201](https://doi.org/10.1103/PhysRevB.97.045201).
- [76] S. Poncé, E. R. Margine, and F. Giustino. Towards predictive many-body calculations of phonon-limited carrier mobilities in semiconductors. *Phys. Rev. B*, 97(12):121201, 2018. doi: [10.1103/PhysRevB.97.121201](https://doi.org/10.1103/PhysRevB.97.121201).
- [77] S. Poncé, F. Macheda, E. R. Margine, N. Marzari, N. Bonini, and F. Giustino. First-principles predictions of hall and drift mobilities in semiconductors. *Phys. Rev. Research*, 3(4):043022, 2021. doi: [10.1103/PhysRevResearch.3.043022](https://doi.org/10.1103/PhysRevResearch.3.043022).
- [78] E. O. Kane. Band structure of indium antimonide. *Journal of Physics and Chemistry of Solids*, 1(4):249–261, 1957. doi: [10.1016/0022-3697\(57\)90013-6](https://doi.org/10.1016/0022-3697(57)90013-6).
- [79] P. B. Allen. Theory of thermal expansion: Quasi-harmonic approximation and corrections from quasiparticle renormalization. *Mod. Phys. Lett. B*, 34(02):2050025, 2020. doi: [10.1142/S0217984920500256](https://doi.org/10.1142/S0217984920500256).
- [80] V. Brousseau-Couture, E. Godbout, M. Côté, and X. Gonze. Zero-point lattice expansion and band gap renormalization: Grüneisen approach versus free energy minimization. *Phys. Rev. B*, 106(8):085137, 2022. doi: [10.1103/PhysRevB.106.085137](https://doi.org/10.1103/PhysRevB.106.085137).
- [81] Y. A. Bychkov and E. I. Rashba. Properties of a 2D electron gas with lifted spectral degeneracy. *Journal of Experimental and Theoretical Physics*, 39(2):78–81, 1984, http://www.jetpletters.ac.ru/ps/1264/article_19121.pdf.
- [82] G. Bihlmayer, O. Rader, and R. Winkler. Focus on the Rashba effect. *New J. Phys.*, 17(5):050202, 2015. doi: [10.1088/1367-2630/17/5/050202](https://doi.org/10.1088/1367-2630/17/5/050202).
- [83] G. Dresselhaus, A. F. Kip, and C. Kittel. Cyclotron Resonance of Electrons and Holes in Silicon and Germanium Crystals. *Phys. Rev.*, 98(2):368–384, 1955. doi: [10.1103/PhysRev.98.368](https://doi.org/10.1103/PhysRev.98.368).
- [84] G. Grimvall. *Thermophysical properties of materials*. Number XVIII in Selected topics in solid state physics. North-Holland Physics Publishing, Amsterdam, 1986.
- [85] D. Karaiskaj, M. Thewalt, T. Ruf, M. Cardona, and M. Konuma. Photoluminescence studies of isotopically enriched silicon: isotopic effects on the indirect electronic band gap and phonon energies. *Solid State Communications*, 123(3):87–92, 2002. doi: [https://doi.org/10.1016/S0038-1098\(02\)00249-1](https://doi.org/10.1016/S0038-1098(02)00249-1).
- [86] F. Manjón, M. Mollar, B. Marí, N. Garro, A. Cantarero, R. Lauck, and M. Cardona. Effect of isotopic mass on the photoluminescence spectra of β zinc sulfide. *Solid State Communications*, 133(4):253–258, 2005. doi: <https://doi.org/10.1016/j.ssc.2004.10.031>.
- [87] J. M. Zhang, T. Ruf, R. Lauck, and M. Cardona. Isotope effects on exciton energies in CdS. *Phys. Rev. B*, 57(16):9716–9722, 1998. doi: [10.1103/PhysRevB.57.9716](https://doi.org/10.1103/PhysRevB.57.9716).

- [88] C. Parks, A. K. Ramdas, S. Rodriguez, K. M. Itoh, and E. E. Haller. Electronic band structure of isotopically pure germanium: Modulated transmission and reflectivity study. *Phys. Rev. B*, 49(20):14244–14250, 1994. doi: [10.1103/PhysRevB.49.14244](https://doi.org/10.1103/PhysRevB.49.14244).
- [89] R. Pässler. Dispersion-related description of temperature dependencies of band gaps in semiconductors. *Phys. Rev. B*, 66(8):085201, 2002. doi: [10.1103/PhysRevB.66.085201](https://doi.org/10.1103/PhysRevB.66.085201).

Intermède

Dans les trois chapitres précédents, nous avons analysé l'impact de différentes approximations sur la renormalisation du point zéro de l'énergie de la bande interdite des semi-conducteurs. Dans un premier temps, nous avons montré que les effets non adiabatiques dominent le ZPR des matériaux polaires et que le traitement explicite de ce type d'interaction améliore significativement la prédictibilité des résultats numériques obtenus. Nous avons ensuite mis en évidence que la contribution de la dilatation du point zéro du réseau cristallin au ZPR peut s'avérer substantielle, tant pour les matériaux plus lourds que pour des matériaux comportant des noyaux légers. Du même coup, nous avons observé que la ZPLE constitue le mécanisme de renormalisation dominant à température nulle pour le GaAs dans le contexte de la DFT semi-locale. Nous avons enfin montré que la diminution du ZPR induite par le couplage spin-orbite pour un grand nombre de matériaux cubiques est dominée par l'effet du SOC sur les énergies électroniques statiques et peut être estimée à l'aide d'un modèle de Fröhlich généralisé.

Au-delà des conclusions spécifiques, ces études mettent en lumière que ces trois mécanismes d'interaction peuvent avoir un impact substantiel sur la fiabilité des prédictions numériques du ZPR, particulièrement pour les matériaux plus lourds et ceux présentant un caractère ionique important. Ceci transparaît dans l'exemple du CdTe, un matériau à la fois lourd et ionique, pour lequel l'inclusion des effets non adiabatiques, de la dilatation du point zéro et du couplage spin-orbite sont nécessaires pour obtenir une prédiction numérique en accord avec la valeur expérimentale. En étudiant des ensembles de 20 à 30 matériaux, nos résultats font ressortir des tendances générales qui pourront être étendues à d'autres familles de matériaux et orienter les choix d'approximation dans un contexte où l'on s'intéresserait à prédire la dépendance en température de la structure électronique de matériaux pour lesquels aucune donnée expérimentale n'est disponible.

Par ailleurs, bien que nos études se soient concentrées sur l'effet des différentes approximations à température nulle, il ne faut pas perdre de vue que l'équation de la self-énergie à température non nulle (éq. (3.3.13) et (3.3.29)) ne diffère de celle à $T = 0$ K que par la valeur des fonctions d'occupation des phonons et des électrons, qui viennent modifier le poids de chacun des couplages contribuant à la self-énergie. Ainsi, une erreur en apparence faible à

$T = 0$ K pourrait être significativement amplifiée par la population thermique des différents modes phononiques lorsqu'on passe à température non nulle.

Ceci nous mène à introduire la seconde partie de cette thèse, qui se veut une application des techniques utilisées dans les trois chapitres précédents à l'étude des transitions de phase topologiques des matériaux. Au chapitre suivant, nous verrons que les phases topologiques de la matière sont caractérisées par des invariants topologiques non nuls et se manifestent concrètement par l'existence d'états électroniques conducteurs à la surface d'un matériau autrement isolant dans son volume. Nous constaterons aussi que la variation de la grandeur de l'énergie de la bande interdite, plus précisément le passage par une bande interdite d'énergie nulle, peut permettre d'identifier une transition de phase topologique. Ces phases exotiques sont particulièrement prometteuses pour le développement d'applications technologiques de pointe. Dans ce contexte, il est primordial de bien caractériser comment les propriétés singulières associées aux phases topologiques résistent aux fluctuations de température qui sont inévitables dans les dispositifs utilisés à l'échelle industrielle.

On cherchera donc à déterminer comment la température, qui modifie l'énergie de la bande interdite, peut affecter la phase topologique des matériaux. Dans cette optique, les chapitres précédents nous indiquent que l'interaction électron-phonon et la dilatation thermique devront être traitées sur un même pied d'égalité pour modéliser correctement la dépendance en température de l'énergie de la bande interdite. De même, l'inclusion des effets non adiabatiques sera importante pour que la valeur quantitative de la renormalisation calculée soit fiable.

Nous verrons également qu'une proportion importante des matériaux topologiques tridimensionnels correspondent à ce qu'on appellera des phases topologiques protégées par la symétrie de renversement du temps. Ce type de matériau est souvent formé d'éléments lourds tels que le bismuth, associés à un fort couplage spin-orbite. Le SOC jouera dès lors un rôle central dans la réalisation de la topologie non triviale de la structure de bande. Ceci implique que l'inclusion du SOC sera nécessaire dès le calcul des propriétés électroniques dans l'approximation du cristal statique. Il en découle évidemment que le calcul de la dépendance en température des énergies électroniques devra aussi le prendre en compte.

Chapitre 7

Topologie des structures de bande

Depuis leur découverte théorique au début des années 1980, l'étude des phases topologiques de la matière est devenue un champ de recherche florissant en physique de la matière condensée. Ces phases ont pour particularité qu'elles ne peuvent pas être décrites par un paramètre d'ordre dans le cadre de la théorie de Landau. Au contraire des transitions de phase thermodynamiques, telles que l'apparition d'un état magnétiquement ordonné, de supraconductivité ou simplement l'ébullition de l'eau, la transition d'une phase topologiquement triviale vers une phase non triviale ne brise aucune symétrie du système. Les premières réalisations expérimentales de ces phases exotiques de la matière, allant de l'effet Hall quantique dès 1980 [1] à celle des isolants topologiques \mathbb{Z}_2 au début des années 2000 [2], ont décuplé l'intérêt de la communauté scientifique envers la recherche, la caractérisation et la classification de nouveaux matériaux topologiques. L'importance fondamentale de cette découverte a été soulignée par l'octroi du prix Nobel de physique 2016 à David J. Thouless, Duncan M. Haldane et J. Michael Kosterlitz.

Ce chapitre a pour objectif d'introduire les différents concepts associés à la topologie des structures de bande qui seront abordés au chapitre 8. On commencera par définir intuitivement le concept de phase topologique du point de vue mathématique et examiner comment il se manifeste dans le modèle SSH du polyacétylène. Nous exposons ensuite les principales caractéristiques des isolants topologiques et des semi-métaux de Weyl en les reliant aux différentes observations expérimentales, pour enfin discuter comment l'interaction électron-phonon peut influencer la phase topologique d'un matériau.

7.1. Concept mathématique

D'un point de vue mathématique, la topologie étudie les propriétés des familles d'objets invariants sous déformation continue. Imaginons les objets solides comme faits de pâte à modeler ; deux objets topologiquement équivalents pourront être transformés l'un dans l'autre par une succession de petites déformations, dites adiabatiques, sans modifier le nombre de

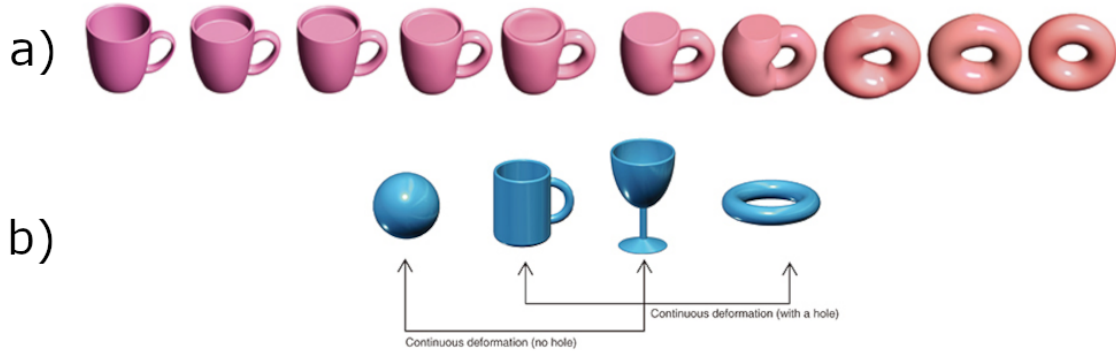


Figure 7.1 – Équivalence topologique. a) Une tasse peut être déformée de façon continue vers un tore. b) Le genre, ou nombre de trous présents dans l’objet, permet de classifier les objets topologiquement équivalents.¹

trous présents au départ dans l’objet. Le nombre de trous tient ici le rôle d’invariant topologique, c’est-à-dire d’une quantité entière qui soit suffisante pour caractériser la topologie d’un objet et prédire certaines de ses propriétés, qui ne dépendent pas de sa géométrie locale. Ainsi, comme on peut voir à la figure 7.1, une tasse et un beigne (tore) sont topologiquement équivalents, tout comme le sont une sphère et une coupe à vin (fig. 7.1b), bien que leurs géométries particulières soient différentes. L’invariant topologique distinct caractérisant la sphère et le tore implique qu’on ne peut transformer l’un dans l’autre sans percer ou combler un trou. Ce concept est capturé mathématiquement par le théorème de Gauss-Bonnet,

$$\frac{1}{2\pi} \oint dA K = 2 - 2g. \quad (7.1.1)$$

L’invariant topologique g , appelé le genre, relie le nombre de trous dans l’objet, soit $g = 0$ pour la sphère et $g = 1$ pour le tore, à l’intégrale de sa courbure gaussienne, K , sur une surface fermée. On voudra donc appliquer un concept similaire pour identifier les structures de bande électroniques topologiquement équivalentes.

7.2. Modèle SSH du polyacétylène

La structure de bande d’un matériau 3D est typiquement représentée par des bandes d’énergie le long de chemins reliant les points de haute symétrie de la première zone de Brillouin ou par un ensemble de surfaces isoénergétiques dans l’espace réciproque tridimensionnel. Dans un cas comme dans l’autre, ce type de représentation est beaucoup plus abstrait que le concept de courbure géométrique utilisé dans le théorème de Gauss-Bonnet. Par souci de simplicité et pour bâtir notre intuition sur la façon dont le concept de phase topologique se manifeste dans les hamiltoniens, considérons tout d’abord un système unidimensionnel.

1. Images tirés des pages Web a) <https://cems.riken.jp/en/laboratory/qmtrt> et b) <https://www.ntt-review.jp>.

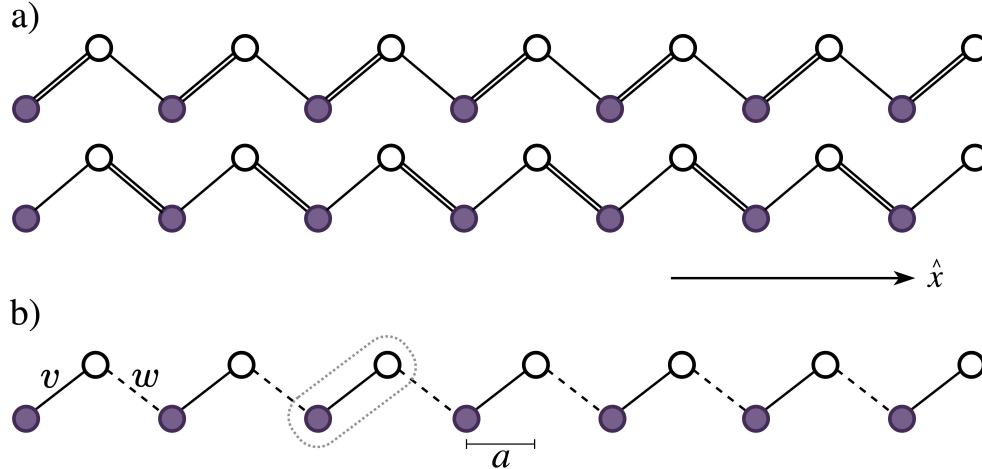


Figure 7.2 – Représentation schématique du modèle SSH du polyacétylène. La chaîne dimérisée par une déformation de Peierls montre une alternance de liens carbone-carbone simples et doubles. (Les atomes d’hydrogène liés à chaque atome de carbone sont absents de ce schéma). a) Il existe deux configurations symétriquement équivalentes, qui diffèrent par la nature du premier lien. b) Dans le modèle de liaisons fortes, v et w décrivent les paramètres de saut intra-dimère (traits solides) et inter-dimère (traits interrompus). Les pointillés gris délimitent un dimère. Image inspirée des références [3, 6].

Le modèle développé par Su, Shrieffer et Heeger (SSH) pour caractériser les excitations électroniques dans les chaînes de polymères [3–5] compte parmi les premiers modèles que l’on interprète aujourd’hui en terme des propriétés topologiques des matériaux, bien que leurs travaux d’origine ne discutaient pas de cet aspect en termes de phases topologiquement distinctes de la manière dont nous les aborderons ici². À l’aide de la méthode des liaisons fortes (*tight binding*), ce modèle étudie l’état fondamental du polyacétylène $((\text{CH})_x)$, la forme la plus simple des polymères organiques conjugués.

L’état fondamental de la structure *trans* du polyacétylène est une séquence unidimensionnelle de monomères (CH) reliés entre eux par une alternance de liens covalents simples et doubles, comme on peut voir à la figure 7.2a. Les différentes chaînes sont faiblement couplées entre elles. Trois des quatre électrons de valence du carbone font partie d’orbitales hybridées de type sp^2 qui assurent la cohésion structurelle de la chaîne le long de son axe et avec les atomes d’hydrogène. Les états électroniques les plus près du niveau de Fermi sont de symétrie $2p_z$ et forment des orbitales de type π . Dans l’état fondamental, toutes les orbitales occupées comptent un électron par site (*half filling*).

La dimérisation du polyacétylène est une manifestation concrète d’une instabilité de Peierls [7]. En effet, si on construit un modèle de liaisons fortes avec un paramètre de saut identique pour l’ensemble des liens carbone-carbone en supposant les atomes immobiles,

². Leurs travaux se concentraient plutôt sur la description de structures appelées « solitons topologiques », que l’on mentionnera un peu plus loin.

on obtient un état fondamental métallique non dégénéré³. Si l'on considère plutôt les positions des atomes comme des variables dynamiques, on constate qu'il est plus favorable énergétiquement que les atomes adjacents soient légèrement déplacés par rapport à leur position d'équilibre initiale dans une alternance de liens courts (doubles) et longs (simples). Ce type de déformation spontanée périodique est caractéristique des métaux unidimensionnels. Le polyacétylène en est l'exemple paradigmatique, puisque la déformation élastique double exactement la dimension de la cellule unitaire.

La dimérisation a deux conséquences importantes sur le spectre du polyacétylène. D'une part, elle ouvre une bande interdite, dont l'énergie dépend de la différence entre les paramètres de saut associés aux liaisons simples et doubles. Ceci se produit puisque le gain énergétique associé à l'ouverture de la bande interdite surpasse l'énergie élastique nécessaire au déplacement des ions de carbone. D'autre part, l'état fondamental acquiert une double dégénérescence, qui découle de la symétrie du système ; les deux configurations, illustrées à la figure 7.2a correspondent aux deux manières d'alterner les liaisons simples et doubles en partant d'une des extrémités de la chaîne et sont équivalentes à renverser l'orientation de l'axe \hat{x} de la chaîne. On verra plus tard que c'est l'existence de cette dégénérescence entre deux états topologiquement distincts qui permet la présence d'excitations topologiquement protégées dans ce matériau.

Considérons uniquement la portion « liaisons fortes » du modèle SSH. On veut décrire un système de fermions sans spins pouvant se déplacer sur un réseau dont la cellule unitaire est formée de deux sites non équivalents, A et B . En traitant seulement les sauts au premier voisin, l'hamiltonien d'une chaîne dimérisée comprenant N cellules unitaires peut s'écrire [6, 8]

$$H = v \sum_{n=1}^N (c_{n,B}^\dagger c_{n,A} + c_{n,A}^\dagger c_{n,B}) + w \sum_{n=1}^{N-1} (c_{n+1,A}^\dagger c_{n,B} + c_{n,B}^\dagger c_{n+1,A}), \quad (7.2.1)$$

où $c_{n,A}^\dagger$ et $c_{n,A}$ sont respectivement les opérateurs qui créent et annihilent un électron sur le site A du dimère n de la chaîne. Le premier terme décrit le déplacement d'un électron à l'intérieur d'un même dimère, avec paramètre de saut v , alors que le second capture le déplacement d'un électron entre deux dimères voisins, avec paramètre de saut w (fig. 7.2b). On suppose les deux paramètres de saut réels et positifs. Les paramètres de saut sont constants pour l'ensemble des dimères afin d'assurer l'invariance de la chaîne sous translation. Les deux configurations de la figure 7.2a correspondent respectivement au cas $v > w$ (haut) et $v < w$ (bas).

3. On suppose également des fermions sans spin.

7.2.1. Hamiltonien du cœur de la chaîne

En supposant la chaîne suffisamment longue pour pouvoir négliger les effets de bord, on peut obtenir l'hamiltonien du cœur de la chaîne⁴ en appliquant des conditions frontières périodiques, ce qui correspond à refermer la chaîne sur elle-même pour former un cercle. Ceci implique de remplacer $c_{n+1,A}^\dagger$ dans l'équation (7.2.1) par $c_{(n \bmod N)+1,A}^\dagger$ et de laisser l'indice de la seconde somme varier de 1 à N . On définit ensuite les opérateurs d'annihilation associés aux sites A et B dans l'espace réciproque,

$$\begin{aligned} a_k &= \frac{1}{\sqrt{N}} \sum_n e^{-ikn(2a)} c_{n,A}, \\ b_k &= \frac{1}{\sqrt{N}} \sum_n e^{-ikn(2a)} c_{n,B}, \end{aligned} \quad (7.2.2)$$

où $2a$ est la longueur de la cellule unitaire. Les opérateurs de création a_k^\dagger et b_k^\dagger sont obtenus par conjugaison hermitienne des expressions précédentes.

En inversant l'équation (7.2.2) et en substituant le résultat dans l'équation (7.2.1) périodique, on obtient

$$H = \sum_k v (a_k^\dagger b_k + b_k^\dagger a_k) + w (e^{-2ika} a_k^\dagger b_k + e^{2ika} b_k^\dagger a_k). \quad (7.2.3)$$

La somme sur le vecteur d'onde k couvre la première zone de Brillouin, $k \in [-\pi/2a, \pi/2a[$. Pour la suite, on simplifie la notation en posant $2ka \rightarrow k$, de sorte que la première zone de Brillouin est maintenant un cercle de circonférence 2π .

On réécrit enfin l'équation (7.2.3) de façon compacte,

$$H = \sum_k \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} H(k) \begin{pmatrix} a_k \\ b_k \end{pmatrix}, \quad (7.2.4)$$

de sorte que notre chaîne de N dimères se ramène à un problème à deux bandes, avec

$$H(k) = \begin{pmatrix} 0 & v + w e^{-ik} \\ v + w e^{ik} & 0 \end{pmatrix}. \quad (7.2.5)$$

On obtient la structure de bande en diagonalisant l'expression précédente,

$$\begin{aligned} E_\pm(k) &= \pm |v + w e^{-ik}| \\ &= \pm \sqrt{v^2 + w^2 + 2vw \cos(k)}. \end{aligned} \quad (7.2.6)$$

4. Puisque l'on traite un système unidimensionnel, la traduction de *bulk* comme « volume » ne nous semble pas appropriée. On préférera le terme *cœur* pour désigner la portion de la chaîne SSH qui se trouve suffisamment loin des extrémités pour qu'on puisse négliger les effets de bord.

L'énergie de la bande interdite correspond à la différence d'énergie minimale entre E_+ et E_- , qui se produit lorsque $\cos(k) = -1$,

$$\begin{aligned}\Delta &= 2 \min(E_+) \\ &= 2|v - w|.\end{aligned}\tag{7.2.7}$$

Notons au passage que l'on retrouve bien un état fondamental métallique (i.e. $\Delta = 0$) dans le cas $v = w$, qui équivaut à la chaîne non dimérisée.

Si l'on restreint notre analyse à leurs spectres d'énergie, les deux états fondamentaux⁵ possèdent a priori les mêmes caractéristiques : l'énergie de leur bande interdite est non nulle, leur dispersion électronique est identique et ils possèdent les mêmes symétries spatiales. Pour distinguer topologiquement ces deux configurations, on doit prendre en compte la structure interne de chaque état, qui est capturée par l'hamiltonien ou, de façon équivalente, par les fonctions d'onde [6]. Pour ce faire, réécrivons l'équation (7.2.5) sous une forme plus générique. En effet, tout hamiltonien à deux bandes peut s'écrire comme

$$H(k) = d_0(k)\sigma_0 + \mathbf{d}(k) \cdot \boldsymbol{\sigma},\tag{7.2.8}$$

où $\sigma_0 = \mathbb{1}_2$, la matrice identité de dimension 2, et $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ sont les matrices de Pauli. Les quantités $d_0(k)$ et $\mathbf{d}(k) = (d_x(k), d_y(k), d_z(k))$ sont des nombres réels.

Sous forme matricielle, l'équation précédente devient

$$H(k) = \begin{pmatrix} d_0 + d_z & d_x - id_y \\ d_x + id_y & d_0 - d_z \end{pmatrix}.\tag{7.2.9}$$

Par comparaison avec l'équation (7.2.5), on trouve, pour le modèle SSH,

$$\begin{aligned}d_0 &= 0, & d_x &= v + w \cos(k), \\ d_y &= w \sin(k), & d_z &= 0.\end{aligned}\tag{7.2.10}$$

La structure interne de l'hamiltonien à un vecteur d'onde k donné est donc entièrement encodée par la direction dans laquelle pointe le vecteur $\mathbf{d}(k)$ dans l'espace bidimensionnel (d_x, d_y) . Notons également que la norme de $\mathbf{d}(k)$ donne les énergies propres du système à un signe près.

Considérons la trajectoire formée par le vecteur $\mathbf{d}(k)$ alors que k parcourt la première zone de Brillouin, variant de $-\pi$ à π . Par périodicité, cette trajectoire est une courbe fermée dans le plan (d_x, d_y) . L'équation (7.2.10) nous indique qu'il s'agit d'un cercle de rayon w centré en $(d_x = v, d_y = 0)$. Comparons maintenant les courbes obtenues pour les deux configurations de la figure 7.2, dont la dispersion électronique (haut) et la trajectoire du vecteur $\mathbf{d}(k)$ (bas) sont illustrées à la figure 7.3 pour les paires de paramètres a) ($v = 1, w = 0.6$) et b) ($v = 0.6, w = 1$). Dans les deux cas, le cercle formé par $\mathbf{d}(k)$ ne croise pas l'origine du

⁵ Les deux états fondamentaux correspondent aux deux configurations de la chaîne illustrées à la figure 7.2a.

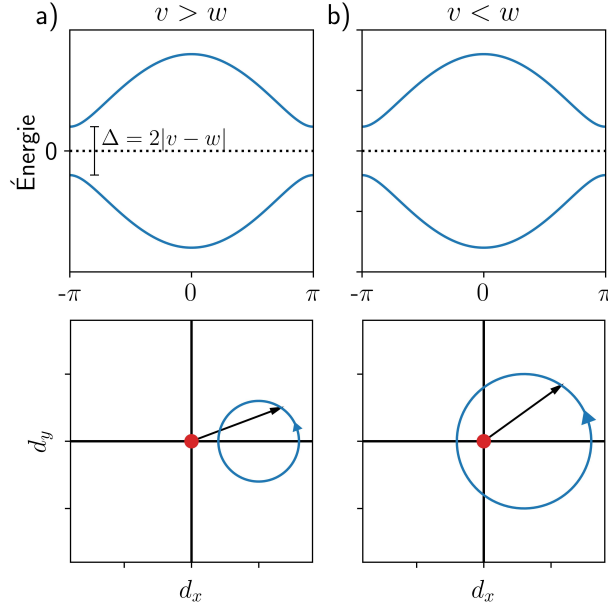


Figure 7.3 – Dispersion électronique (haut) et représentation bidimensionnelle de l'hamiltonien SSH (bas). Les cercles bleus décrivent la trajectoire du vecteur $\mathbf{d}(k)$ lors que k traverse la zone de Brillouin. La flèche noire représente $\mathbf{d}(k)$ pour un k spécifique. a) Dans la phase triviale ($v = 1, w = 0.6$), le cercle décrit par $\mathbf{d}(k)$ n'inclut jamais l'origine (marqueur rouge des figures du bas). b) Dans la phase topologique ($v = 0.6, w = 1.0$), l'origine se trouve à l'intérieur du cercle décrit par $\mathbf{d}(k)$. Cet enroulement de la phase de l'hamiltonien est à l'origine des propriétés topologiques. Image inspirée de la référence [6].

plan (d_x, d_y) (marqueur rouge), qui correspond à l'énergie propre $E(k) = 0$: les hamiltoniens décrivent donc tous deux des isolants électriques. De surcroît, leurs dispersions électroniques sont identiques, car leurs paramètres de sauts sont équivalents sous permutation $v \leftrightarrow w$. On peut toutefois remarquer une différence fondamentale quant à la position des cercles dans le plan. Dans le cas $v < w$ (droite), l'origine du plan (d_x, d_y) se trouve à l'intérieur du cercle paramétrisé par $\mathbf{d}(k)$, alors qu'elle est à l'extérieur du cercle pour $v > w$ (gauche). En se rappelant que les éléments de matrice non nuls sont de la forme $d_x \pm id_y$ (éq. (7.2.9) et (7.2.10)), on peut déduire que, pour $v < w$, la phase de $H(k)$ s'enroule autour de l'origine lorsque k parcourt une boucle fermée dans la zone de Brillouin. Cet enroulement de la phase ne se produit pas si $v > w$.

Cette distinction fondamentale entre les deux états dégénérés est capturée mathématiquement par l'invariant topologique appelé « nombre d'enroulement » (*winding number*), défini comme [6]

$$\nu = \frac{1}{2\pi} \int_{\text{BZ}} \left(\hat{\mathbf{d}}(k) \times \frac{d}{dk} \hat{\mathbf{d}}(k) \right)_z dk, \quad (7.2.11)$$

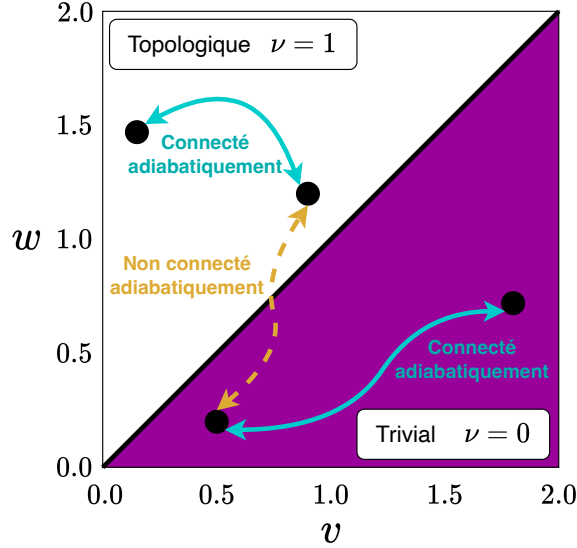


Figure 7.4 – Diagramme de phase topologique du modèle SSH. Deux paires de paramètres associés à la même valeur du nombre d’enroulement sont connectées adiabatiquement (flèches cyan). La ligne noire $v = w$ correspond à un état métallique et délimite la frontière entre les deux phases isolantes topologiquement distinctes. Ainsi, toute transformation qui relie deux paires de paramètres situées de part et d’autre de la ligne noire devra nécessairement fermer la bande interdite (flèche jaune). Image adaptée de la référence [6].

où $\hat{\mathbf{d}}(k) = \mathbf{d}(k)/|\mathbf{d}(k)|$ est la projection de $\mathbf{d}(k)$ sur un cercle de rayon 1 et \times représente le produit vectoriel. Cette expression relie donc l’intégrale d’une courbure⁶ à un invariant du système, le nombre d’enroulement de la phase de $H(k)$, en analogie directe avec le théorème de Gauss-Bonnet (éq. (7.1.1)). On verra à la section 7.2.3 qu’une valeur non nulle de ν est associée à la présence d’états de bord localisés aux extrémités de la chaîne.

7.2.2. Déformation adiabatique de l’hamiltonien et équivalence topologique

L’espace des paramètres du modèle SSH regroupe donc deux phases isolantes topologiquement distinctes : $\nu = 0$ pour $v > w$ et $\nu = 1$ pour $v < w$, tel qu’illustré à la figure 7.4. La frontière entre les deux phases correspond à l’état métallique $v = w$. Revenons maintenant au concept intuitif d’équivalence topologique énoncé en début de chapitre, voulant que deux objets soient topologiquement équivalents s’il existe une transformation adiabatique pouvant transformer l’un dans l’autre sans modifier la valeur d’un invariant topologique, et appliquons-le à l’hamiltonien SSH.

Notons au passage que la notion de déformation adiabatique sous-entend que l’hamiltonien est déformé suffisamment lentement pour empêcher les électrons d’effectuer une

6. Plus exactement, la composante \hat{z} d’un rotationnel (*curl*).

transition vers des états excités⁷. Cette interprétation diffère de la définition thermodynamique d'une transformation adiabatique, qui décrit les transformations qui se produisent sans échange d'énergie sous forme de chaleur avec l'extérieur du système⁸. Ainsi, un système qui se trouve à l'origine dans son état de plus basse énergie demeurera tout au long de la transformation dans l'état fondamental de l'hamiltonien instantané. En comparaison, une transformation non adiabatique impliquerait que l'état final soit une combinaison linéaire des états propres du nouvel hamiltonien.

Considérons deux hamiltoniens SSH caractérisés par $\nu = 0$, ayant respectivement $p_1 = (v, w)$ et $p_2 = (v + \delta v, w + \delta w)$ pour paramètres de saut. Les cercles correspondants dans le plan (d_x, d_y) , C_1 et C_2 , sont distincts et n'encerclent pas l'origine. On peut imaginer une séquence de transformations infinitésimales translatant la position du centre du cercle ou modifiant son rayon qui permettent de transformer C_1 en C_2 sans jamais que le cercle ne croise l'origine. Un exemple est illustré à la figure 7.5a. Cette transformation altérera la forme spécifique de la dispersion électronique, puisque les paramètres de saut changent, mais la bande interdite maintiendra une valeur finie en tout temps. Puisqu'on peut transporter C_1 vers C_2 sans altérer la valeur de l'invariant ν , les deux hamiltoniens sont connectés adiabatiquement et sont associés à la même phase topologique. Par un argument similaire, n'importe quelle paire de paramètres situés dans la portion $v < w$ de la figure 7.4 sont connectés adiabatiquement. En ce sens, on peut considérer un invariant topologique comme une constante adiabatique, puisque sa valeur est robuste sous toute transformation adiabatique de l'hamiltonien.

Imaginons maintenant une séquence de transformations qui transporterait un cercle C_1 caractérisé par $\nu = 0$ vers un cercle C_2 ayant $\nu = 1$. En restreignant la transformation au plan (d_x, d_y) , ce qui revient à se limiter à modifier la valeur des paramètres de saut intradimère et interdimer, on constate qu'il est impossible de relier C_1 à C_2 sans que le cercle ne croise l'origine (fig. 7.5b et c), indiquant que les deux hamiltoniens ne sont pas topologiquement équivalents. Bien que les structures de bande initiales et finales soient toutes les deux isolantes, la bande interdite devra nécessairement devenir nulle au cours de la transformation. Ceci se produira au moment où les paramètres de saut atteignent la condition critique $v = w$, juste avant que le nombre d'enroulement passe de $\nu = 0$ à $\nu = 1$. En ce sens, fermer la bande interdite est conceptuellement équivalent à « percer un trou » dans l'hamiltonien et constitue une signature d'une transition de phase topologique.

On pourrait par ailleurs, en apparente contradiction avec l'argument précédent, imaginer transporter C_1 vers C_2 en le soulevant légèrement le long de l'axe d_z , de sorte que la bande interdite demeure non-nulle tout au long de la transformation d'un état $\nu = 0$ vers un état

7. Le temps caractéristique de la déformation doit être beaucoup plus grand que $\hbar/\Delta E$, où ΔE est la différence d'énergie entre les états propres du système.

8. Du point de vue thermodynamique, le changement d'énergie interne qui se produit lors d'une transformation adiabatique est entièrement compensé par le travail fait sur le système.

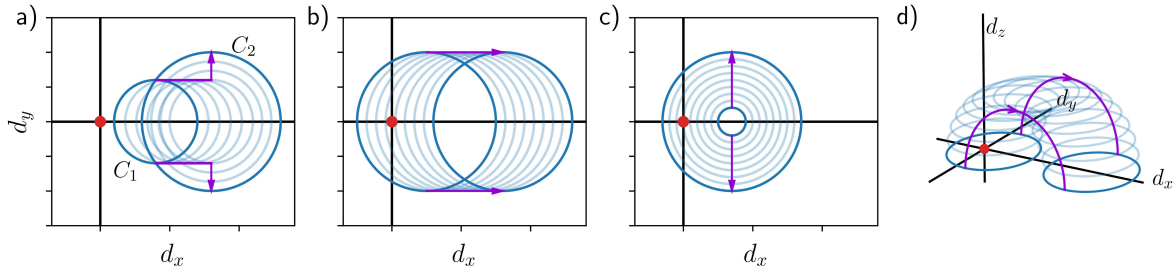


Figure 7.5 – Déformations de l’hamiltonien. a) Une déformation adiabatique relie les cercles C_1 et C_2 sans croiser l’origine. L’énergie de la bande interdite demeure non nulle tout au long de la transformation. b), c) Aucune transformation ne peut relier entre eux deux cercles appartenant à des phases topologiquement distinctes sans croiser l’origine. La fermeture de la bande interdite est une signature de la transition de phase topologique. d) L’introduction momentanée d’un potentiel local sur les sites de la chaîne permet de relier les deux cercles de topologie différente sans fermer la bande interdite, mais brise la symétrie chirale qui protège les états de bord topologiques. Image adaptée de la référence [6].

$\nu = 1$, tel qu’illustré à la figure 7.5d. Ceci revient à introduire dans l’équation (7.2.9) un terme d_z non nul, qui représente physiquement l’ajout d’un potentiel localisé sur les différents sites du dimère [6]. Bien que continue, une telle transformation n’est pas adiabatique puisqu’elle relie des états associés à un invariant topologique distinct. La bande interdite ne se ferme pas puisque la transformation brise une des symétries fondamentales de l’hamiltonien SSH, à savoir la symétrie chirale⁹, qui se manifeste par $d_z = 0$ dans l’équation (7.2.10). Dans ce cas, les états de bord topologiques caractérisant la phase $\nu = 1$, qui seront discutés à la prochaine section, ne survivent pas à l’ajout du potentiel. Ceci nous amène à introduire la notion de phase topologique protégée par symétrie [9]. En effet, dans une forte proportion des matériaux topologiques réels, la robustesse des états topologiques contre les perturbations est garantie par une symétrie fondamentale, typiquement la symétrie de renversement du temps. Une transition de phase topologique sera dès lors associée à une fermeture de la bande interdite dans la mesure où les symétries protégeant les états topologiques sont préservées.

7.2.3. États de bord et correspondance cœur-frontière

Par souci de simplicité, considérons le cas limite où la chaîne est entièrement dimérisée, c’est-à-dire que l’un des deux paramètres de saut est nul, comme illustré à la figure 7.6a. Dans cette situation, la dispersion électronique est constante, $E = \pm 1$. Il est également possible de définir des états propres parfaitement localisés sur chaque dimère de la chaîne en

9. La symétrie chirale combine la symétrie sous renversement du temps et la symétrie particule-trou.

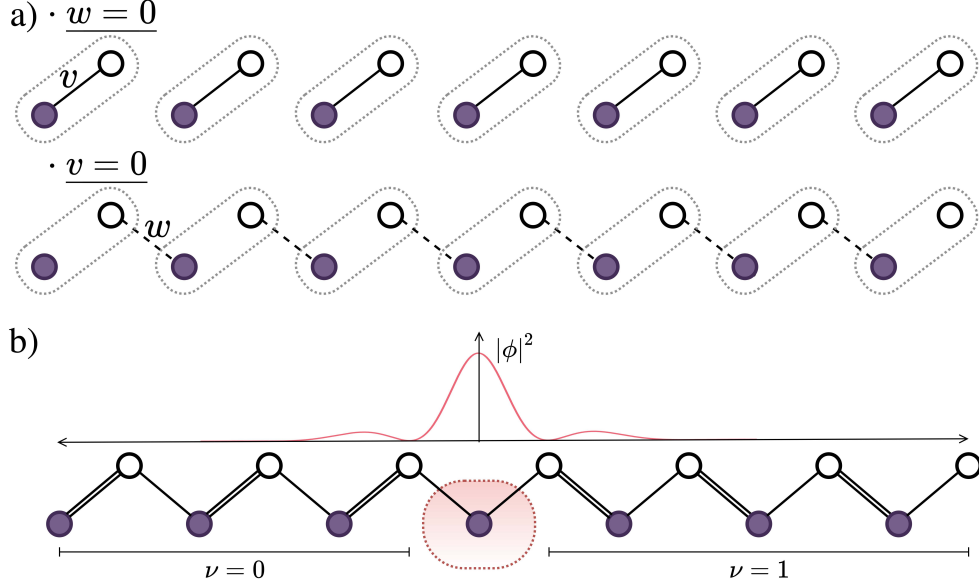


Figure 7.6 – Chaîne entièrement dimérisée et solitons topologiques. a) Les deux configurations possibles lorsque la chaîne est entièrement dimérisée se distinguent par la présence ($v = 0$) ou l'absence ($w = 0$) de site non appariés aux extrémités de la chaîne, qui hébergent des états d'énergie nulle. b) Un soliton topologique est localisé à l'interface entre une région $\nu = 0$ et $\nu = 1$ (région encerclée en rouge). La courbe rouge représente schématiquement la densité de probabilité associée au soliton le long de la chaîne. Image inspirée des références [4, 6].

construisant des combinaisons linéaires paires et impaires d'états à un électron situés sur le dimère en question.

Dans le cas trivial, ($v = 1, w = 0$), on aura, à normalisation près,

$$H(|n,A\rangle \pm |n,B\rangle) = \pm(|n,A\rangle \pm |n,B\rangle), \quad (7.2.12)$$

où l'état $|n,A\rangle$ représente un électron localisé sur le site A du dimère n . De tels états existent pour tout $n \in [1, N]$. Les états de bord correspondent ici aux combinaisons linéaires impliquant $n = 1$ à gauche et $n = N$ à droite. Ces états étant inclus dans l'équation précédente, le spectre du cœur de la chaîne décrit l'ensemble des états propres de H .

Lorsque ($v = 0, w = 1$), les différents dimères sont partagés entre deux cellules unitaires adjacentes, contrairement au cas précédent, où les divisions en dimère et cellule unitaire coïncident. Pour le cœur de la chaîne, les états localisés sur un dimère donné vérifient

$$H(|n,B\rangle \pm |n+1,A\rangle) = \pm(|n,B\rangle \pm |n+1,A\rangle), \quad (7.2.13)$$

avec, cette fois-ci, $n \in [1, N-1]$. Il y a donc deux sites non appariés aux deux extrémités de la chaîne, $|1,A\rangle$ à gauche et $|N,B\rangle$ à droite. En appliquant H (éq. (7.2.1)) sur chacun de ces états, on trouve que

$$H|1,A\rangle = H|N,B\rangle = 0. \quad (7.2.14)$$

Le spectre de H n'est donc pas entièrement décrit par celui des états de cœur. Contrairement à la phase triviale, la phase topologique du polyacétylène possède deux états de bord d'énergie nulle. Dans les cas généraux $v > w$ et $v < w$, l'inclusion progressive des interactions interdimère ou intra-dimère va permettre l'hybridisation des états propres définis pour la chaîne entièrement dimérisée, de sorte que la solution deviendra beaucoup plus complexe sur le plan mathématique. Il est néanmoins possible de montrer que les états de bord caractéristiques de la phase topologique demeurent exponentiellement localisés aux extrémités de la chaîne [6].

Le nombre d'états de bord à chaque extrémité de la chaîne est donc directement relié à la valeur du nombre d'enroulement. Pour s'en convaincre, imaginons l'existence d'un état localisé à l'extrémité gauche, dont l'énergie $+\varepsilon \neq 0$ se trouve dans la bande interdite. La symétrie du spectre SSH (éq. (7.2.6)) impose l'existence d'un état partenaire d'énergie $-\varepsilon$, localisé dans la même cellule unitaire. Dans l'esprit de la section 7.2.2, on peut imaginer une déformation adiabatique de H qui modifie les énergies propres du système et expulse progressivement cette paire d'états en dehors de la bande interdite. Puisque les configurations initiales et finales sont connectées adiabatiquement, elles doivent être associées au même invariant topologique, $\nu = 0$. La présence d'un unique état d'énergie nulle, au contraire, est protégée par la symétrie chirale, car, malgré la déformation, il n'existe aucun état partenaire permettant de modifier son énergie. La seule manière d'expulser cet état en dehors de la bande interdite est d'introduire suffisamment d'interactions intradimère pour qu'un partenaire symétrique apparaisse dans la cellule unitaire, ce qui revient à traverser la frontière critique $v = w$. On peut généraliser cet argument à tout nombre impair d'états de bord.

L'existence des états d'énergie nulle lorsque $\nu = 1$ est une manifestation du principe de correspondance cœur-frontière (*bulk-boundary correspondance*), qui dit que l'émergence d'états conducteurs liés aux frontières d'un matériau est une conséquence des propriétés topologiques du volume du système [2]. Ce concept est fondamental dans la théorie des matériaux topologiques.

Soulignons enfin l'existence d'excitations topologiquement non triviales à l'interface entre une chaîne de type $v > w$ et une de type $v < w$, tel qu'illustré en rouge à la figure 7.6b. La courbe du panneau supérieur représente schématiquement la densité de probabilité associée au soliton de long de la chaîne. Démonstrées et caractérisées par les travaux de Su, Schrieffer et Heeger, ces excitations exponentiellement localisées ont notamment permis d'expliquer l'observation expérimentale de défauts neutres mobiles dans le polyacétylène non-dopé [3–5]. Sans entrer dans les détails, on peut comprendre intuitivement l'existence des solitons en considérant la limite de la chaîne entièrement dimérisée. Dans ce cas, comme on peut observer à la figure 7.6b, on peut interpréter une interface entre deux domaines topologiquement distincts comme un nombre impair d'états non appariés qui, par le principe de correspondance cœur-frontière, sont associés à l'émergence d'états d'énergie nulle localisés à l'interface.

En d'autres mots, la chaîne interpole spatialement entre deux configurations isolantes $\nu = 0$ et $\nu = 1$, ce qui implique nécessairement de fermer la bande interdite à l'interface.

7.2.4. Phases topologiques et phases de Landau

Terminons notre analyse du modèle SSH en considérant les deux phases topologiquement distinctes du point de vue de la théorie des transitions de phase développée par Landau [10], qui est fondamentale à la compréhension moderne des transitions de phase. La beauté de cette théorie est d'expliquer les changements de phase observés dans la matière comme la manifestation d'une brisure de symétrie. Par exemple, l'alignement des spins d'un matériau ferromagnétique aimanté brise la symétrie de rotation, alors que la cristallisation de l'eau discrétise les symétries de translation et de rotation. Cette brisure de symétrie est également associée à une valeur non nulle d'un certain paramètre d'ordre. Un exemple simple de paramètre d'ordre est l'aimantation nette d'un ferroaimant. Pour la cristallisation de la glace, on pourrait utiliser la transformée de Fourier de la densité, qui capturera la périodicité du système cristallin. Ce paramètre d'ordre est utilisé pour paramétriser l'énergie libre du système, $F = U - TS$, qui met en évidence une compétition entre l'ordre, qui diminue l'énergie interne U , et le désordre, favorisé par l'entropie S à haute température¹⁰. Une transition de phase se manifeste alors par une discontinuité dans une des dérivées de l'énergie libre, ou fonctions de réponse, dans l'esprit de la classification d'Ehrenfest [11].

Nous avons montré que la phase $\nu = 1$ du modèle SSH se distingue de la phase $\nu = 0$ par l'existence d'un état métallique localisé aux extrémités de la chaîne, qui apparaît lorsque la force du paramètre de saut interdimère surpasse celle du paramètre intradimère. Aucune brisure de symétrie ou paramètre d'ordre ne peut expliquer l'émergence de cette nouvelle propriété, montrant bien que la transition de phase qui a eu lieu va au-delà du paradigme de Landau. Pour comprendre la différence fondamentale entre les transitions de phase décrites par la théorie de Landau et les transitions de phase topologiques, rappelons-nous que l'équation de Schrödinger est un problème de valeurs propres [8]. Un hamiltonien est donc non seulement caractérisé par son spectre d'énergie, mais aussi par ses états propres. La théorie de Landau étant construite sur l'énergie libre du système, elle utilise uniquement l'information contenue dans les valeurs propres de l'hamiltonien, qui capture les changements de symétrie (voir section 2.4). Or, l'hamiltonien et ses états propres sont définis dans l'espace d'Hilbert, qui a une dimensionnalité plus grande et une structure plus complexe que les énergies propres, qui sont des scalaires réels. Les propriétés topologiques d'un système, quant à elles, sont encodées dans les états propres du système, qui contiennent plus d'information que les valeurs propres, ce pour quoi elles sont invisibles du point de vue de la théorie de Landau.

¹⁰. Par simplicité, on se limite ici à l'énergie libre de Helmholtz. Dans le cas de l'énergie libre de Gibbs, $G = U + PV - TS = H - TS$, c'est l'enthalpie, $H = U + PV$, qui caractérise le niveau d'ordre du système.

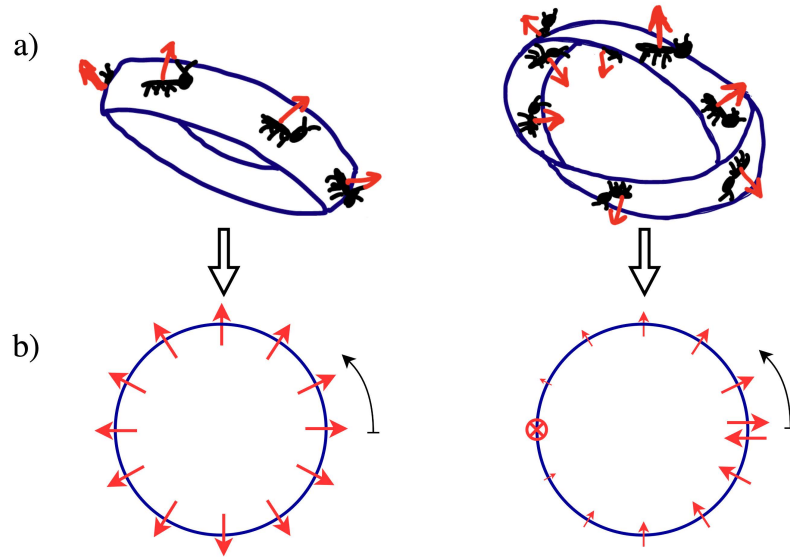


Figure 7.7 – Distinction entre les phases topologiques et les phases de Landau.

a) La phase triviale (gauche) et la phase topologique (droite) du modèle SSH sont représentées par un cylindre et un ruban de Möbius. Les fourmis qui se déplacent sur leur surface représentent l'information encodée dans les vecteurs propres de l'hamiltonien. b) Dans l'esprit de la théorie de Landau, les valeurs propres du système sont représentées par la projection du cylindre et du ruban de Möbius dans un plan 2D, qui forment deux cercles identiques. Pour distinguer les deux phases, il est nécessaire d'y projeter l'orientation des flèches attachées aux fourmis. Dans la phase triviale, les flèches pointent toujours vers l'extérieur du cylindre. Dans la phase topologique, une trajectoire de 2π sur le cercle mène la fourmi sur la face opposée du ruban par rapport à son point de départ, qui se manifeste par un enroulement de la flèche autour du cercle.

La figure 7.7 présente une analogie simple pour illustrer cet argument. La phase triviale y est représentée par un cylindre et la phase topologique, par un ruban de Möbius. L'approche de Landau équivaut conceptuellement à projeter ces structures dans un plan 2D. Ceci produit deux cercles identiques, tout comme le spectre du modèle SSH est invariant si on échange les valeurs de v et w . Or, en effectuant cette projection, nous avons masqué l'information contenue dans les états propres. Imaginons maintenant les états propres de H comme des fourmis auxquelles sont fixées une flèche, qui indique l'orientation normale de la fourmi par rapport à la structure sur laquelle elle se trouve. Si on laisse la fourmi parcourir le cylindre dans une direction donnée jusqu'à ce qu'elle réalise une rotation de 2π sur sa projection 2D^{11 12}, la flèche pointera toujours vers l'extérieur du cylindre. En revanche, la même trajectoire effectuée sur le ruban de Möbius verra la flèche s'enrouler autour du ruban, capturant

11. On supposera bien entendu que la fourmi vit dans un monde idéal où la gravité n'a pour seul effet que de maintenir l'orientation de la flèche perpendiculaire à la surface.

12. Sur le ruban de Möbius, la fourmi doit parcourir 4π pour revenir à son point de départ. Après 2π , elle se retrouve sur la face opposée du ruban.

ainsi sa topologie non triviale, tout comme l'invariant topologique ν capture l'enroulement de la phase de H autour de l'origine dans le plan (d_x, d_y) ¹³.

En conclusion, le modèle SSH du polyacétylène se veut l'un des modèles les plus simples pour illustrer le concept de phase topologique des hamiltoniens, en raison de sa basse dimensionnalité. Notre objectif consistait à introduire les concepts d'équivalence topologique des hamiltoniens, de déformation adiabatique et de correspondance cœur-frontière de manière à la fois simple et rigoureuse, afin de façonner une intuition physique qui facilitera la compréhension des systèmes tridimensionnels, dont le développement mathématique dépasse le cadre de cette thèse.

7.3. Isolants topologiques

La découverte de l'effet Hall quantique en 1980 [1] représente un tournant de la littérature. Dans cette expérience, l'application d'un fort champ magnétique à un gaz d'électrons confinés en deux dimensions à très basse température contraint les électrons du volume aux orbites circulaires induites par la force de Lorentz, mais crée simultanément un nombre discret d'états de conduction localisés au bord de l'échantillon [8, 12]. Cette découverte constitue une preuve irréfutable que la distinction fondamentale entre métaux et isolants dans l'esprit de la théorie des bandes n'est pas suffisante pour décrire l'ensemble des systèmes réalisés par la nature. Bien que l'expression « isolant topologique » due à Moore et Balents [13] ne soit apparue qu'au début des années 2000, le concept existe dans la littérature depuis le milieu des années 1980 pour identifier des systèmes qui se distinguent des isolants standards décrits par la théorie des bandes par l'existence d'états électroniques conducteurs localisés aux frontières [2]. La classification de ce type de matériau en termes de leur dimensionnalité et de leurs symétries tant spatiales que non spatiales s'est depuis profondément raffinée [14]. Dans la présente section, nous nous limiterons à mettre en avant les caractéristiques principales des isolants topologiques de type \mathbb{Z}_2 et définir l'invariant topologique qui sera utilisé au chapitre suivant, pour ensuite discuter brièvement des applications potentielles de ce type de matériau. Notons que l'appellation \mathbb{Z}_2 réfère aux valeurs que pourront prendre l'invariant topologique, soit 0 ou 1. On suppose de plus que les interactions entre les électrons sont suffisamment faibles pour que le système puisse être traité dans l'approximation du champ moyen, ce qui implique que les excitations électroniques peuvent être adéquatement décrites par des états à une particule.

13. Pour un isolant topologique, on peut interpréter la direction des flèches rouges comme une représentation du caractère orbital dominant des bandes électroniques lorsque l'on parcourt un chemin périodique dans la zone de Brillouin. Cette propriété sera discutée à la section 7.3.2.

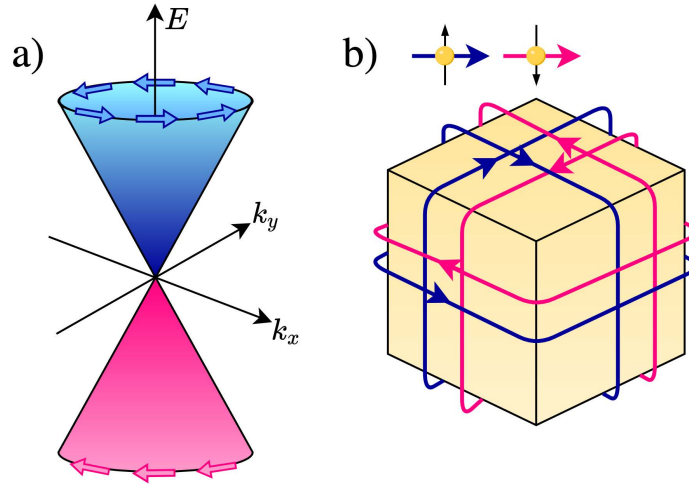


Figure 7.8 – Représentation schématique des états de surface d’un isolant topologique. a) La dispersion des états de surface forme un cône de Dirac. À énergie constante, on observe une texture de spin hélicoïdale, dont la chiralité dépend de l’énergie de l’état par rapport au niveau de Fermi. b) Les courants de surface associés à des spins différents circulent dans des directions opposées. Image inspirée de la référence [15].

7.3.1. États de surface conducteurs

La plus connue des propriétés des isolants topologiques est sans contredit la présence d’états métalliques localisés sur la surface alors que le volume possède une bande interdite d’énergie non nulle. Comme discuté dans le cadre du modèle SSH (section 7.2), l’existence de ces états est une manifestation du principe de correspondance cœur-frontière, qui nous dit que la surface d’un matériau reflète la topologie de son volume. La relation intrinsèque entre le volume et la surface se manifeste concrètement dans l’invariant topologique : dans le modèle SSH, le nombre d’enroulement capture de nombre d’états localisés à une extrémité de la chaîne, alors que le nombre de Chern caractérisant l’état Hall quantique correspond au nombre de niveaux de Landau contribuant à la conductance de Hall [8]. On peut également interpréter les états de surface dans l’esprit des solitons du modèle SSH : la surface du matériau topologique correspond à une interface avec le vide ou, en d’autres mots, à la juxtaposition de systèmes dont l’invariant topologique est différent, qui doit nécessairement être accompagnée par une fermeture de la bande interdite.

Dans les matériaux tridimensionnels, les états de surface topologiques sont caractérisés par une dispersion linéaire autour du niveau de Fermi, résultant en une structure appelée cône de Dirac qui s’apparente à la dispersion bien connue du graphène, tel qu’illustré schématiquement à la figure 7.8a. Leur chiralité se manifeste dans la corrélation entre l’orientation

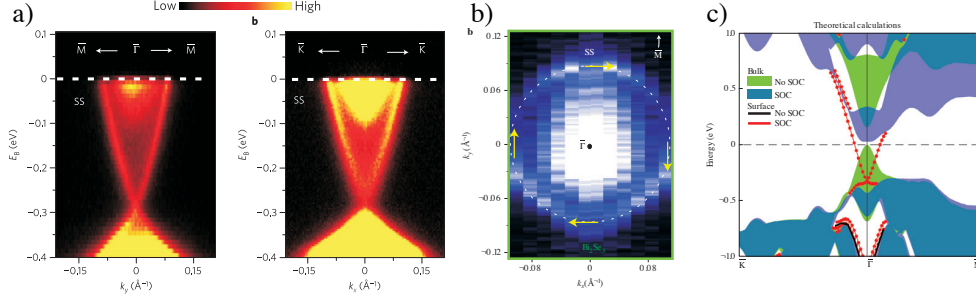


Figure 7.9 – États de surface de l’isolant topologique Bi_2Se_3 . a) Le spectre ARPES des états de surface montre une dispersion de type cône de Dirac. b) Le spectre ARPES de la surface de Fermi associée aux états de surface, indiquée en tirets blancs dans la figure de gauche, révèle une polarisation en spin hélicoïdale, de chiralité gauche. c) La dispersion des états de surface prédite par la DFT en présence du SOC représente bien les observations expérimentales. Les calculs sans SOC ne prédisent aucun état de surface dont l’énergie se trouve dans la bande interdite, confirmant ainsi le rôle fondamental du SOC dans la topologie non triviale. Images tirées de la référence [16].

du spin de l’électron et son impulsion¹⁴ (*spin-momentum locking*) et produit une polarisation en spin hélicoïdale dans les plans bidimensionnels qui intersectent le cône de Dirac. Le courant de surface qui en résulte circule donc dans une direction déterminée par le spin de l’électron (fig. 7.8b).

Les états de surface topologiques d’un matériau 3D ont été observés expérimentalement à de nombreuses reprises à l’aide de spectroscopie de photoémission résolue en angle (ARPES) [2]. Cette technique permet de déduire l’énergie et l’impulsion initiale des électrons éjectés du matériau par effet photoélectrique, et, par conséquent, d’imager directement la structure de bande. La figure 7.9 montre le cône de Dirac (a) et la polarisation en spin hélicoïdale (b) mesurés par ARPES dans le Bi_2Se_3 , de même que la dispersion en surface (c, points rouges) et dans le volume (c, bleu) calculée avec la DFT [16]. Ce matériau fait partie de la famille $(\text{Bi},\text{Sn})_2(\text{Se}, \text{Te})_3$, qui constitue aujourd’hui le système de référence pour les isolants topologiques \mathbb{Z}_2 , à l’instar de YBCO et BSSCO pour les cuprates supraconducteurs.

Du point de vue expérimental, la robustesse de l’invariant topologique aux déformations adiabatiques de l’hamiltonien, discutée dans le cadre du modèle SSH, implique que les états de surface résistent aux impuretés et aux défauts structurels, tant et aussi longtemps que la symétrie sous renversement du temps est préservée. On peut le démontrer très simplement à l’aide des propriétés de l’opérateur de renversement du temps, \mathcal{T} (éq. (2.4.17)). Supposons un potentiel de diffusion non magnétique V invariant sous renversement du temps. On cherche

14. On réfère ici à l’impulsion du cristal (*crystal momentum*) qui identifie les états propres dans l’espace réciproque, et non aux valeurs propres de l’opérateur d’impulsion.

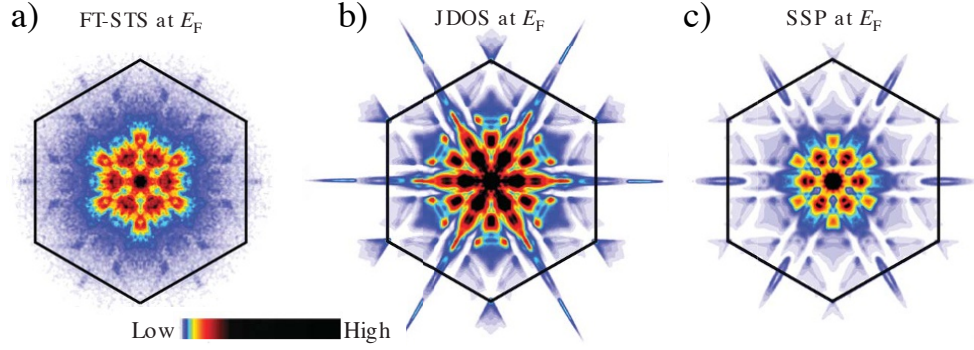


Figure 7.10 – Absence de rétrodiffusion dans les états de surface topologiques.

a) Transformée de Fourier du patron d'interférence mesuré par STS à la surface du $\text{Bi}_{0.92}\text{Se}_{0.08}$. b) Probabilité de diffusion modélisée en incluant la rétrodiffusion. L'amplitude de probabilité prédite est significativement plus élevée que celle observée. c) Probabilité de diffusion modélisée en supprimant la rétrodiffusion $\mathbf{k} \rightarrow -\mathbf{k}$, qui impliquerait une diffusion entre deux états de spins opposés. L'amplitude prédite reproduit adéquatement les observations, confirmant la protection des états de surface topologiques contre la rétrodiffusion. Images tirées de la référence [17].

à déterminer l'amplitude de diffusion entre les états partenaires de Kramers, $|\psi\rangle$ et $|\mathcal{T}\psi\rangle$,

$$\begin{aligned}
 \langle \mathcal{T}\psi | V | \psi \rangle &= \langle \mathcal{T}^2 \psi | V | \mathcal{T}\psi \rangle^* \\
 &= - \langle \psi | V | \mathcal{T}\psi \rangle^* \\
 &= - \langle \mathcal{T}\psi | V | \psi \rangle,
 \end{aligned} \tag{7.3.1}$$

ce qui implique que

$$\langle \mathcal{T}\psi | V | \psi \rangle = \langle -\mathbf{k} \downarrow | V | \mathbf{k} \uparrow \rangle = 0. \tag{7.3.2}$$

Ainsi, comme on peut déduire de la figure 7.8a, les états de surface ne subissent aucune rétrodiffusion puisqu'il n'existe aucun état électronique à $-\mathbf{k}$ qui soit compatible en spin.

Cette protection topologique a été investiguée dans des alliages de bismuth et d'antimoine [17] en comparant la transformée de Fourier du patron d'interférence créé par la diffusion sur la surface mesuré par spectroscopie à effet tunnel (STS, fig. 7.10a) avec la probabilité de diffusion calculée avec et sans suppression de la rétrodiffusion des états électroniques de \mathbf{k} vers $-\mathbf{k}$ (resp. fig. 7.10c et b). La grande similarité entre la figure 7.10c et les données expérimentales confirme l'absence de rétrodiffusion dans les états de surface, malgré un niveau de désordre important au niveau atomique. Le rôle fondamental de la symétrie de renversement du temps dans la protection des états de surface a quant à lui été démontré en étudiant l'effet de différents types d'impuretés sur les états de surface du $\text{Bi}_2(\text{Se}, \text{Te})_3$ [18]. En comparant des spectres ARPES avant et après adsorption d'atomes de potassium (fig. 7.11a), on constate que le cône de Dirac résiste à l'ajout de charges coulombiennes. En revanche,

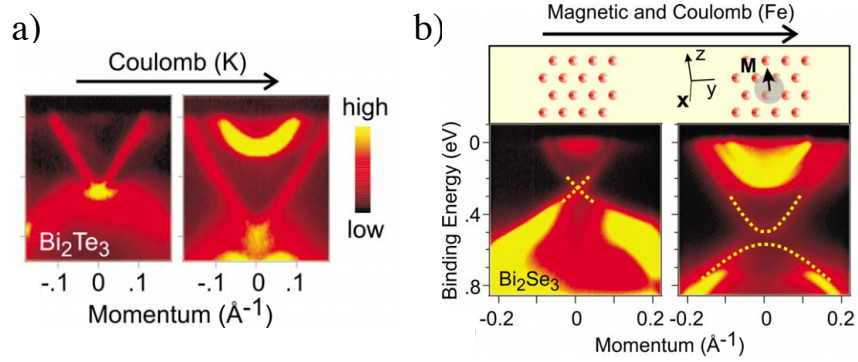


Figure 7.11 – Protection topologique des états de surface. a) La dispersion en cône de Dirac des états de surface du Bi_2Te_3 résiste à l’ajout d’impuretés chargées. b) Le magnétisme des atomes de fer brise la symétrie de renversement du temps et ouvre une bande interdite dans la dispersion des états de surface du Bi_2Se_3 . Images tirées de la référence [2].

le magnétisme intrinsèque des atomes de fer brise la symétrie de renversement du temps et ouvre une bande interdite dans le cône de Dirac (fig. 7.11b).

7.3.2. Inversion du caractère des bandes

La seconde caractéristique fondamentale des isolants topologiques est quant à elle une propriété du volume. Rappelons-nous que, dans le volume, la nature topologique d’un système est encodée dans les états propres de l’hamiltonien. Dans cette optique, on peut se demander s’il est possible de décomposer les différents états électroniques définis dans l’espace réciproque en termes de représentations simples, correspondant à des orbitales de type hydrogénique centrées sur les différents atomes de la cellule unitaire dans l’espace réel. Ce type d’interprétation des bandes électroniques en terme de leur caractère orbital dominant est courant dans la littérature; on n’a qu’à penser à la description des bandes de valence et de conduction des matériaux de structure zinc-blende en termes d’orbitales p de l’anion et d’orbitales s du cation, où s et p réfèrent aux valeurs propres du moment cinétique orbital [19]. Ainsi, pour un isolant standard dans l’esprit de la théorie des bandes, on peut voir les orbitales atomiques comme des blocs fondamentaux que l’on peut combiner pour construire les différents états électroniques, en respectant les restrictions imposées par les symétries du système [20, 21]. Une telle décomposition n’est pas possible dans le cas d’un isolant topologique. La nature topologique du volume se manifeste par une intrication locale entre les caractères orbitaux des bandes de valence et de conduction au voisinage de leurs extrema respectifs. En pratique, on observe une inversion entre leurs caractères orbitaux dominants, appelée communément « inversion des bandes ». Cette intrication est d’ordinaire interprétée comme une conséquence d’un fort couplage spin-orbite, ce qui explique pourquoi la majorité des isolants topologiques connus à ce jour sont constitués d’éléments lourds. Le concept

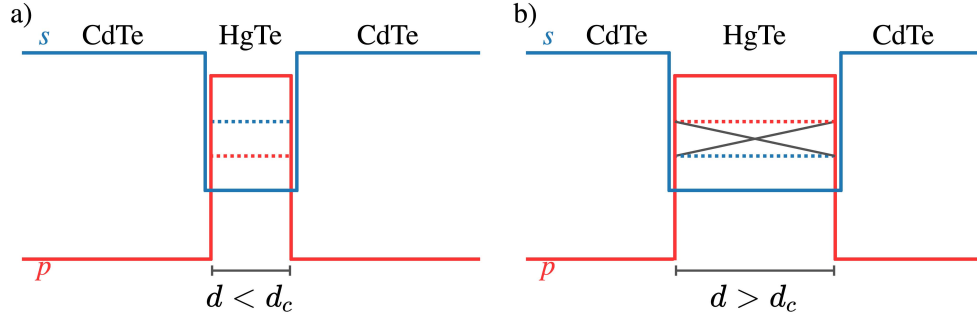


Figure 7.12 – Inversion du caractère des bandes dans le puits quantique CdTe/HgTe. Le VBM et CBM du CdTe dans le volume ont respectivement des caractères orbitaux de type p (rouge) et s (bleu). Le couplage spin-orbite plus prononcé du HgTe inverse cet ordonnement. a) Lorsque $d < d_c = 6.3$ nm, les états électroniques liés au puits quantique de CdTe/HgTe sont ordonnés en énergie de la même façon que dans le CdTe. b) Lorsque $d > d_c$, cet ordonnement est inversé par rapport à la phase triviale. Les bords du puits quantiques possèdent alors une paire d'états de surface hélicoïdaux polarisés en spin, montrés en gris. Image adaptée de la référence [22].

d'inversion des bandes a d'ailleurs mené au développement d'une méthode appelée chimie quantique topologique (*topological quantum chemistry*), qui prédit la topologie d'un matériau uniquement à partir des représentations en termes d'orbitales atomiques et des relations de compatibilité dans la zone de Brillouin issues de la théorie des groupes. Cette méthode a récemment été appliquée dans un contexte *high-throughput* pour révéler que plus du quart des cristaux inorganiques répertoriés possèderaient des propriétés topologiques intrinsèques [21].

La notion d'inversion des bandes est également au cœur du modèle développé par Bernevig, Hughes et Zhang, qui prédit l'existence d'une phase topologique dans un puits quantique 2D formé par une hétérostructure de CdTe et HgTe [22], illustrée à la figure 7.12. Le CdTe est un semi-conducteur de structure zinc-blende, dont les extrema des bandes de valence et de conduction sont alignés au centre de la zone de Brillouin. Les états électroniques du VBM du CBM ont respectivement les symétries p et s . Le fort couplage spin-orbite présent dans le HgTe inverse la position relative des états s et p par rapport au CdTe. Lorsque l'épaisseur de la couche de HgTe est inférieure à la valeur critique $d_c = 6.3$ nm (fig. 7.12a), les deux états électroniques 2D liés au puits quantique dont l'énergie se trouve dans la bande interdite sont ordonnés de la même manière que pour le CdTe, résultant en un système isolant standard. Pour $d > d_c$ (fig. 7.12b), la structure électronique du HgTe domine et l'ordonnement de ces états est inversé par rapport à la phase triviale. La nature topologique de la configuration $d > d_c$ a été confirmée par la détection expérimentale d'une paire d'états de bord conducteurs chiraux de spins opposés dans la conductance de Hall [23]. Cette caractéristique des phases topologiquement non triviales est particulièrement intéressante du point de vue du calcul

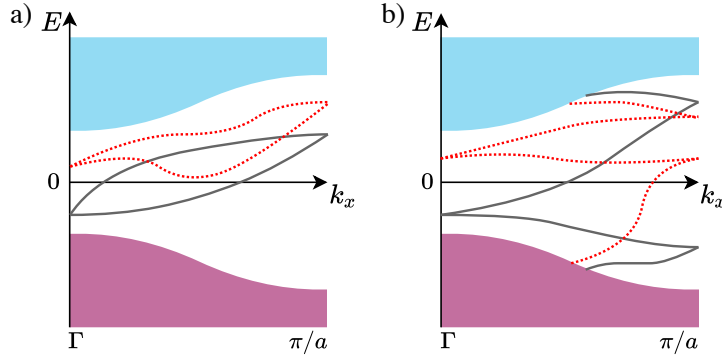


Figure 7.13 – Correspondance cœur-frontière. La topologie non triviale des isolants topologiques \mathbb{Z}_2 se manifeste dans le nombre d'états de surface (gris foncé) qui croisent le niveau de Fermi (arbitrairement posé à $E = 0$). Aux \mathbf{k} -points invariants sous renversement du temps, le théorème de Kramers impose que les états électroniques soient doublement dégénérés. La portion $[-\pi/a, \Gamma[$ de la zone de Brillouin regroupe les partenaires de Kramers des états illustrés. Le spectre électronique est donc symétrique par rapport à $k = \Gamma$. (a) Dans la phase triviale ($\mathbb{Z}_2 = 0$), les partenaires de Kramers à Γ se rejoignent à $\mathbf{k} = \pi/a$. Entre les deux, ils intersectent deux fois le niveau de Fermi. On peut dès lors imaginer une transformation adiabatique de l'hamiltonien qui les repousse au-dessus du niveau de Fermi (pointillés rouges). b) Dans la phase topologique ($\mathbb{Z}_2 = 1$), les partenaires de Kramers à Γ sont connectés à des paires différentes à $\mathbf{k} = \pi/a$. Ce faisant, les états de surface intersectent une seule fois le niveau de Fermi. Leur existence est protégée topologiquement contre les déformations adiabatiques. Image adaptée de la référence [2].

numérique, puisqu'elle dépend uniquement des propriétés électroniques du volume. Elle ne nécessite donc pas de modéliser les états de surface du matériau de façon explicite.

7.3.3. Invariant topologique \mathbb{Z}_2

Les isolants topologiques appartiennent à la famille des matériaux topologiques protégés par la symétrie de renversement du temps, qui sont caractérisés par un invariant topologique qui ne peut prendre que 0 ou 1 comme valeur, d'où son nom d'invariant \mathbb{Z}_2 [24]. L'interprétation physique la plus naturelle de cet invariant fait appel au principe de correspondance cœur-frontière. Rappelons-nous tout d'abord que la symétrie de renversement du temps implique que, pour chaque valeur propre électronique $\varepsilon_{\mathbf{k}n}$, il existe un état partenaire $\varepsilon_{-\mathbf{k}n}$ de même énergie, formant une paire de Kramers¹⁵. Le spectre d'énergie dans une direction $-\hat{\mathbf{k}}$ donnée est donc une copie miroir de celui dans la direction $\hat{\mathbf{k}}$. De même, les états propres dont le vecteur d'onde dans la zone de Brillouin est invariant sous renversement du temps sont tous doublement dégénérés.

¹⁵. La relation entre les états de spin de chacun des partenaires de Kramers dépend de la présence de la symétrie d'inversion, comme discuté aux sections 2.4.4 et 2.5.3.

Considérons maintenant un système isolant possédant des états de surface dont l'énergie croise le niveau de Fermi. La figure 7.13 montre les deux façons génériques de relier ces états entre eux sur une moitié de la zone de Brillouin dans la direction $\hat{\mathbf{k}}$. D'une part, si ces états sont reliés par paires (gauche), on peut imaginer une transformation adiabatique qui modifie le potentiel chimique et repousse ces états au-dessus ou au-dessous du niveau de Fermi (pointillés rouges), impliquant que la structure de bande a la même topologie qu'un isolant standard, pour lequel $\mathbb{Z}_2 = 0$. Dans le cas contraire (droite), les états de surface issus d'une paire de Kramers à $k = 0$ se connectent respectivement à des paires différentes à $k = \pi/a$, intersectant alors le niveau de Fermi un nombre impair de fois. Dans ce cas, une transformation adiabatique pourra modifier les détails de la dispersion, mais sera incapable de les expulser de la bande interdite. L'état de surface est donc topologiquement protégé et $\mathbb{Z}_2 = 1$. Ainsi, l'invariant \mathbb{Z}_2 capture la parité du nombre d'états de surface qui croisent le niveau de Fermi dans la moitié de la zone de Brillouin.

Si cet argument basé sur le principe de correspondance cœur-frontière est très intuitif, la définition mathématique générale de l'invariant \mathbb{Z}_2 en termes des états de Bloch du volume, quant à elle, est remarquablement complexe [2, 25]. Si l'on fait abstraction de nombreuses subtilités mathématiques, la méthode développée par Soluyanov et Vanderbilt [26, 27] permet néanmoins de visualiser comment la topologie non triviale se manifeste dans le volume des isolants topologiques. La présentation de cette méthode nécessite d'introduire la phase de Berry,

$$\gamma_L = i \sum_n \oint_L d\mathbf{k} \cdot \langle u_{\mathbf{k}n} | \nabla_{\mathbf{k}} | u_{\mathbf{k}n} \rangle, \quad (7.3.3)$$

qui est une phase géométrique que la fonction d'onde acquiert lors de son transport parallèle¹⁶ adiabatique le long d'un chemin fermé dans l'espace réciproque, L . $|u_{\mathbf{k}n}\rangle$ correspond à la partie périodique des états de Bloch. En plus d'être une notion centrale de la théorie des bandes topologique, la phase de Berry est au cœur de l'effet Aharonov-Bohm et de la théorie moderne de la polarisation électrique [28]. Un exemple de phase géométrique est illustré à la figure 7.14, où la géométrie de la sphère fait en sorte que le transport parallèle du vecteur bleu sur la courbe magenta le ramène à son point de départ avec une orientation différente par rapport au vecteur initial. De façon analogue, la phase de Berry capture la rotation de la phase de la fonction d'onde le long d'un chemin dans la zone de Brillouin, ce qui rappelle la définition du nombre d'enroulement qui caractérise le modèle SSH. Pour un chemin fermé,

16. En géométrie, le transport parallèle consiste à déplacer un vecteur le long d'une courbe définie dans un espace non euclidien, de manière à ce qu'il demeure localement parallèle à lui-même. Pour ce faire, on commence par définir un ensemble de plans localement tangents à chaque point de l'espace (plans beiges de la fig. 7.14). Pour pouvoir définir une notion de transport parallèle, il faut donc connaître comment les plans tangents voisins sont orientés les uns par rapport aux autres. Cette transformation est encodée dans une quantité appelée connexion. Ainsi, à l'instar du gradient d'une fonction, la connexion capture les variations de la géométrie locale aux différents points de l'espace. Dans l'équation (7.3.3), l'objet $i \langle u_{\mathbf{k}n} | \nabla_{\mathbf{k}} | u_{\mathbf{k}n} \rangle$ correspond à la connexion de Berry, qui nous indique quelle rotation on doit appliquer à la phase d'un vecteur d'état lorsqu'on le déplace sur la zone de Brillouin qu'on aurait repliée sur elle-même.

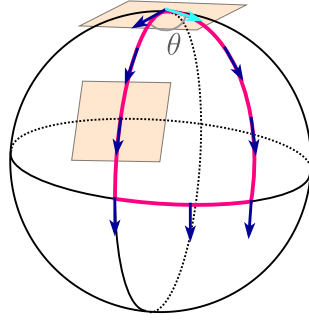


Figure 7.14 – Phase géométrique. Lors de son transport parallèle le long de la courbe magenta sur une sphère, le vecteur montré en bleu revient à son point de départ avec une orientation différente (cyan) de son orientation initiale. À chaque étape de son déplacement, le vecteur est déplacé de sorte qu’il demeure parallèle à lui-même sur les différents plans localement tangents à chaque point de la courbe (beige). L’angle θ correspond à la phase géométrique acquise lors du transport parallèle sur un chemin fermé.

Berry a démontré que l’invariance des propriétés physiques sous multiplication des états électroniques par une phase globale ($|\mathbf{k}n\rangle \rightarrow e^{i\phi} |\mathbf{k}n\rangle$) impose que γ_C soit un multiple entier de 2π . On peut donc l’utiliser pour définir un invariant topologique, connu sous le nom de nombre de Chern, $C = \gamma_L/2\pi$, qui caractérise la topologie des surfaces 2D [28].

La méthode de Soluyanov et Vanderbilt repose sur l’analyse des centres de charge de fonctions de Wannier hybrides. Avant de définir ces dernières, rappelons tout d’abord que les fonctions de Wannier sont une base orthonormale permettant d’exprimer les états électroniques en termes d’orbitales localisées dans l’espace réel. Elles sont obtenues par transformée de Fourier inverse des états de Bloch,

$$|\mathbf{R}n\rangle = \frac{V}{2\pi^3} \int_{\text{BZ}} e^{-i\mathbf{k}\cdot(\mathbf{R}-\mathbf{r})} |u_{\mathbf{k}n}\rangle, \quad (7.3.4)$$

où \mathbf{R} est un vecteur du réseau de Bravais. Ainsi, la fonction de Wannier $|\mathbf{R}n\rangle$ décrit un électron dont la densité de probabilité est exponentiellement restreinte à la cellule unitaire identifiée par le vecteur de réseau \mathbf{R} . Il s’agit donc, en quelque sorte, d’un analogue quantique de la position d’un électron dans le cristal [29].

Dans le même esprit, les fonctions de Wannier hybrides localisent les états électroniques dans une direction et demeurent délocalisées dans les deux autres [30],

$$|R_x k_y k_z n\rangle = \frac{a_x}{2\pi} \int_{-\pi/a_x}^{\pi/a_x} d^3\mathbf{k} e^{-ik_x(R_x-r_x)} |u_{\mathbf{k}n}\rangle. \quad (7.3.5)$$

a_x est le paramètre de maille dans la direction x . On définit le centre de charge de chaque fonction de Wannier hybride (HWCC) comme la position moyenne de cette orbitale dans les

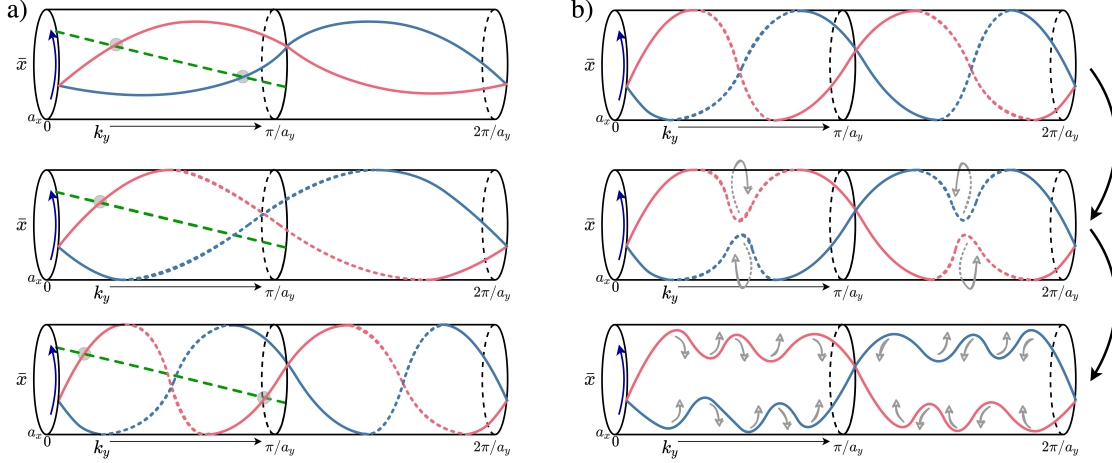


Figure 7.15 – Invariant topologique \mathbb{Z}_2 et trajectoires des centres de charge des fonctions de Wannier hybrides. Chaque cylindre montre la trajectoire des HWCC d’une paire de Kramers dans la direction \hat{x} lorsque k_y traverse la zone de Brillouin, pour un k_z fixe invariant sous renversement du temps. L’axe vertical, qui correspond à la position des centres de charges, a été replié sous forme de cylindre pour mettre en valeur la périodicité du cristal. a) L’invariant \mathbb{Z}_2 se manifeste dans la parité du nombre de croisements entre les HWCC et n’importe quelle ligne droite reliant $k_y = 0$ et $k_y = \pi/a_y$ sur la surface du cylindre (tirets verts). Les croisements sont mis en évidence par les cercles gris. De manière équivalente, on pourrait évaluer la parité du nombre de demi-périodes sur lesquelles les trajectoires s’enroulent entre $k_y = 0$ et $k_y = \pi/a_y$. Les trajectoires du haut et du bas croisent les tirets verts deux fois et possèdent une topologie triviale ($\mathbb{Z}_2 = 0$). La trajectoire du centre, qui n’intersecte les tirets verts qu’une seule fois, correspond à la phase topologique ($\mathbb{Z}_2 = 1$). b) Bien que les trajectoires du bas de la figure 7.15a s’enroulent deux fois autour du cylindre en traversant la zone de Brillouin, on peut les relier adiabatiquement à celles du haut de la même figure par une succession de transformations adiabatiques. Elles doivent donc être décrites par la même valeur de \mathbb{Z}_2 . La protection topologique des croisements des trajectoires est limitée aux k_y invariants sous renversement du temps, ce qui permet de relier adiabatiquement la configuration du haut à celle du milieu. Images adaptées de la référence [31].

deux autres directions,

$$\bar{x}_n(k_y, k_z) = \langle 0 k_y k_z n | \hat{X} | 0 k_y k_z n \rangle. \quad (7.3.6)$$

Le point central de la méthode est que l’on peut exprimer la phase de Berry, et donc un invariant topologique, en terme de la position des HWCC des différentes bandes occupées,

$$\gamma_L = \frac{2\pi}{a} \sum_n \bar{x}_n. \quad (7.3.7)$$

Le chemin L représente ici la direction pour laquelle des fonctions de Wannier hybrides sont exprimées dans l’espace réel.

Avec ces outils en main, nous sommes maintenant en mesure d’évaluer un invariant topologique pour n’importe quel plan 2D de la zone de Brillouin tridimensionnelle qui contienne

sa propre image sous renversement du temps. Pour ce faire, fixons k_z dans un plan donné et « wanniérisons » dans la direction x . On peut dès lors calculer la variation des HWCC lorsque k_y traverse la moitié de la zone de Brillouin, de l'origine ($k_y = 0$) à la frontière de zone ($k_y = \pi/a_y$)¹⁷. La figure 7.15a illustre les comportements possibles. Pour mettre en valeur la périodicité du cristal, l'axe vertical de la figure, correspondant à la position du centre de charge, \bar{x} , a été repliée pour former un cylindre. Par simplicité, nous considérons une seule paire de Kramers (lignes rouges et bleues) et montrons l'évolution des HWCC comme une fonction continue¹⁸. Notons que la symétrie de renversement du temps impose que les HWCC soient dégénérés à $k_y = 0$ et $k_y = \pi/a_y$. Dans le premier cas (haut), les HWCC oscillent sur le devant du cylindre, tandis que dans les deux autres (milieu, bas), ils s'enroulent respectivement une et deux fois autour du cylindre lorsque k_y parcourt l'ensemble de la zone de Brillouin. En analogie avec le nombre d'enroulement, on pourrait penser que ces deux derniers cas sont des manifestations de l'invariant non trivial. Or, dans l'esprit de notre discussion du principe de correspondance cœur-frontière, l'invariant \mathbb{Z}_2 a pour particularité de capturer la *parité* de l'enroulement. À la figure 7.15a, ceci revient à compter le nombre de fois que les trajectoires des HWCC s'enroulent autour du cylindre sur une période complète ou, de manière équivalente, le nombre de fois qu'une ligne droite arbitraire qui relie $k_y = 0$ et $k_y = \pi/a_y$ sur le cylindre intersecte la trajectoire du HWCC (tirets verts), L . Le cas $L = 1$ (milieu) correspond donc bien à une topologie non triviale. En revanche, le cas $L = 2$ (bas) sera associé à une topologie triviale. On peut interpréter ce cas en considérant que les croisements entre les HWCC ne sont topologiquement protégés qu'aux deux extrémités du cylindre, qui sont invariants sous renversement du temps. Ainsi, tel qu'illustré à la figure 7.15b, on peut connecter le cas $L = 2$ au cas $L = 0$ par une succession de transformations adiabatiques, ce qui implique qu'ils ont le même invariant topologique. De ces observations, on peut donc définir un invariant de surface,

$$\Delta = L \bmod 2, \quad (7.3.8)$$

qui compte la parité du nombre de croisements entre $k_y = 0$ et $k_y = \pi/a$ et prendra les valeurs de 0 ou 1.

À partir de cette définition, on peut caractériser la topologie d'un isolant 3D à l'aide de celle des différents plans 2D de la zone de Brillouin qui sont invariants sous renversement du temps en définissant quatre indices [31],

$$\nu; (\nu_x, \nu_y, \nu_z), \quad (7.3.9)$$

17. L'autre moitié est une copie miroir par symétrie de renversement du temps et n'apporte aucune information supplémentaire.

18. Lors d'un calcul numérique pour un matériau réel, la discrétisation des valeurs de k_y et le plus grand nombre de bandes occupées donnent lieu à plusieurs défis techniques, qui sont exposés dans la référence [31] dans et admirablement bien vulgarisés dans la référence [30].

tels que

$$\begin{aligned}\nu &= \Delta(k_i = 0) + \Delta(k_i = 0.5) \pmod{2}, \\ \nu_i &= \Delta(k_i = 0.5),\end{aligned}\tag{7.3.10}$$

où $i \in [x, y, z]$ et k_i est exprimé en coordonnées réduites. L'indice ν correspond directement à l'invariant \mathbb{Z}_2 .

Le cas $\nu = \mathbb{Z}_2 = 1$ correspond à un isolant topologique fort, auquel nous référons implicitement depuis le début de ce chapitre. Il existe par ailleurs des systèmes non triviaux pour lesquels $\nu = \mathbb{Z}_2 = 0$ mais au moins un ν_i est non nul, appelés isolants topologiques faibles [32]. Ce type de système peut être visualisé comme une superposition de structures 2D non triviales (analogues à celle réalisant l'effet Hall quantique) faiblement couplées entre elles. La direction de l'indice ν_i non nul indique l'axe de superposition. Les états de surface sont donc localisés au bord des surfaces 2D, ce qui implique que certaines surfaces n'en possèdent aucun. Les états de surface d'un isolant topologique faible manifestent par ailleurs une plus grande sensibilité au désordre, puisque leur protection topologique ne provient que de plans 2D et non du volume en entier, comme c'est le cas pour les isolants topologiques forts [33].

En revenant à la figure 7.15 et à la relation entre les HWCC et la phase de Berry, on peut enfin bâtir une interprétation physique de l'invariant \mathbb{Z}_2 en termes des états du volume. Lorsqu'on balaie un plan de la zone de Brillouin ayant k_z fixe en laissant varier l'impulsion dans une direction, k_y , la position moyenne des électrons se déplace dans la direction orthogonale. Puisque la zone de Brillouin est périodique, elle doit revenir à son point de départ dans la cellule unitaire une fois la traversée complétée (de 0 à 2π). Il n'y a par contre aucune restriction qui impose qu'elle demeure dans la même cellule unitaire. L'enroulement de la trajectoire du HWCC illustré à la figure 7.15a peut donc être interprété comme le « pompage » d'un électron d'une cellule unitaire à l'autre pour chaque cycle parcouru en k_y , qui induit un état de surface conducteur [30]. Ce type d'interprétation d'une topologie non triviale en termes d'un pompage de charge (*charge pumping*), que l'on doit initialement au modèle de Laughlin expliquant l'effet Hall quantique [34], est au cœur de nombreux modèles qui ont façonné notre compréhension des phases topologiques de la matière [6, 8, 35].

Terminons en soulignant que dans le cas d'un matériau centrosymétrique, il est possible de calculer l'invariant \mathbb{Z}_2 d'une façon étonnamment simple, proposée par Fu et Kane [32]. En effet, si l'hamiltonien est à la fois invariant sous inversion et sous renversement du temps, les états électroniques pour lesquels le vecteur d'onde est invariant sous renversement du temps sont aussi des états propres de parité, de valeur propre $\xi = \pm 1$. Fu et Kane ont montré que l'invariant \mathbb{Z}_2 pouvait alors s'exprimer en termes du produit des valeurs propres de parité des bandes occupées évaluées aux points \mathbf{k} invariants sous renversement du temps (TRIM,

pour *time-reversal invariant momenta*),

$$(-1)^\nu = \prod_{\mathbf{k}_j \in \text{TRIM}} \left(\prod_m \xi_m(\mathbf{k}_j) \right). \quad (7.3.11)$$

Cette relation facilite significativement l'analyse de la topologie dans les matériaux centrosymétriques.

7.3.4. Applications

Les propriétés intrinsèques des états conducteurs topologiques en font des matériaux prometteurs pour le développement d'applications technologiques de haute performance. Une première classe d'applications vise à exploiter la robustesse intrinsèque des courants de surface face à la rétrodiffusion. Au contraire des semi-conducteurs conventionnels, dont les propriétés électroniques et de transport peuvent être significativement affectées par la présence d'impuretés ou de défauts structuraux, les électrons topologiquement protégés « contournent » les défauts plutôt qu'effectuer des collisions. On peut imaginer construire des dispositifs électroniques à haute efficacité énergétique dans lesquels le courant circule essentiellement sans dissipation [2, 25], dans le même esprit que les supraconducteurs à haute température. Cette avenue fort prometteuse comporte néanmoins des défis importants sur le plan de l'ingénierie. Entres autres, plusieurs isolants topologiques connus, dont la famille Bi_2Se_3 , existent dans une forme naturellement « auto-dopée » (*self-doped*) par des défauts dans la structure atomique, alors que la protection topologique tire son origine de l'absence de conduction dans le volume. Des techniques de croissance des cristaux ont été développées pour compenser la présence des dopants et ramener le potentiel chimique à l'intérieur de la bande interdite [36]. Dans le même ordre d'idées, la chiralité et la polarisation en spin des états de surface en font des candidats idéaux pour la spintronique, qui vise à contrôler les propriétés de transport d'un matériau en manipulant le courant associé aux différents canaux de spin [37].

Le domaine de l'information quantique s'intéresse également aux matériaux topologiques. La course effrénée menée par les milieux de recherche tant académiques qu'industriels vers la concrétisation de l'ordinateur quantique n'a plus besoin de présentation [38]. À ce jour, l'un des grands défis à surmonter est le contrôle de la décohérence des qubits, qui survient lorsqu'ils interagissent avec leur environnement, entraînant du même coup la perte de l'information qui y est encodée. À ce sujet, il a été prédit que les excitations électroniques localisées à l'interface entre un isolant topologique et un supraconducteur conventionnel réaliseraient des fermions de Majorana [2]. Ces quasi-particules, qui ont pour étrange particularité d'être leur propre antiparticule, ont été proposées par Kitaev [39] pour réaliser des qubits résistants à la décohérence. L'utilisation de matériaux topologiques pour générer des fermions de Majorana par effet de proximité conférerait en plus une protection topologique aux qubits, en faisant ainsi un médium de prédilection pour le développement de l'ordinateur quantique.

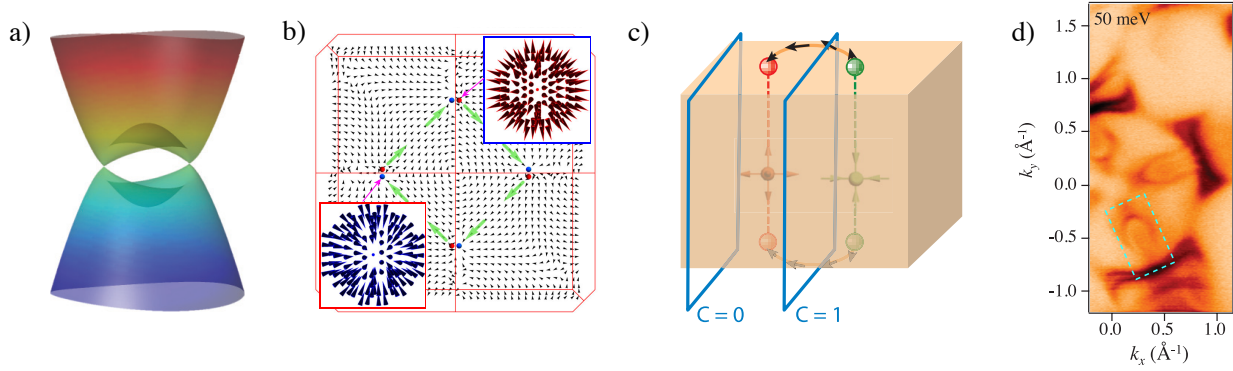


Figure 7.16 – Caractéristiques des semi-métaux de Weyl. a) La dispersion électronique d'un semi-métal de Weyl croise le niveau de Fermi à un nombre discret de paires de points \mathbf{k} . Elle est linéaire au voisinage des points de croisement. Image tirée de la référence [40]. b) Les paires de nœuds de Weyl de chiralité opposée (points rouges et bleus) sont séparés spatialement dans l'espace réciproque. Ils forment les sources et les drains de la courbure de Berry, dont le champ est représenté par les flèches noires. Au voisinage des nœuds de Weyl, elle présente une structure caractéristique en « porc-épic » orientée vers l'extérieur pour les sources (rouge) et vers l'intérieur pour les drains (bleu). Image tirée de la référence [41]. c) Les arcs de Fermi (orange) localisés à la surface des WSM sont la projection du flux de Berry non nul présent dans le volume entre une source et un drain. Le nombre d'arcs de Fermi reliant une même paire de nœuds de Weyl dépend du nombre de quanta de flux de Berry qui traverse un plan 2D dans l'espace réciproque (carrés bleus ouverts). Image tirée de la référence [40]. d) Le spectre ARPES des états de surface du WSM TaAs montre des arcs de Fermi doubles (encadré). Image tirée de la référence [42].

7.4. Semi-métaux de Weyl

Jusqu'ici, nous avons uniquement caractérisé la topologie de structures de bande qui possèdent une bande interdite non nulle. Des phases topologiquement non triviales ont également été observées dans les semi-métaux [42] et prédites théoriquement pour les métaux [43]. À l'instar du graphène, la surface de Fermi des semi-métaux de Weyl est composée de points discrets, caractérisés par une dispersion électronique linéaire formant un cône de Dirac, comme on peut voir à la figure 7.16a. Les excitations électroniques qui y sont associées sont des fermions chiraux sans masse, dans l'esprit de la solution à l'équation de Dirac proposée par Hermann Weyl en 1929 [44]. En plus d'un fort SOC, la réalisation d'un WSM nécessite de briser la symétrie d'inversion ou celle de renversement du temps, ce qui assure que les nœuds de Weyl sont séparés spatialement dans l'espace réciproque [40].

La topologie des WSM nécessite d'introduire le concept de courbure de Berry [6],

$$\Omega_{\mathbf{k}n} = \nabla \times i \langle u_{\mathbf{k}n} | \nabla_{\mathbf{k}} | u_{\mathbf{k}n} \rangle. \quad (7.4.1)$$

Étant définie dans l'espace réciproque, l'interprétation physique de la courbure de Berry est beaucoup moins intuitive que celle de la courbure gaussienne discutée en tout début de chapitre. On peut cependant la relier à la phase de Berry, que nous avons définie précédemment dans le contexte de l'invariant \mathbb{Z}_2 (éq. (7.3.3)). En observant les équations (7.4.1) et (7.3.3), on peut construire une analogie avec l'électromagnétisme : le second terme du rotationnel de l'équation (7.4.1) devient un potentiel vecteur et la phase de Berry, un flux magnétique. Dans cette optique, la courbure de Berry tient le rôle d'un champ de jauge associé à la phase de Berry [6]. Elle se comporte comme l'analogie d'un champ magnétique dans l'espace réciproque, dont les sources et les drains, l'équivalent de monopôles et antimonopôles magnétiques, correspondent aux nœuds de Weyl [35] de chiralité opposée¹⁹ (fig. 7.16b). Il est par ailleurs fascinant de constater que, bien qu'aucun monopôle magnétique n'ait été détecté expérimentalement, leur analogie existe réellement dans l'espace réciproque de plusieurs matériaux.

La topologie particulière des WSM se manifeste à la surface du matériau par la présence de structures ouvertes appelées « arcs de Fermi », qui diffèrent des boucles fermées que l'on retrouve typiquement à la surface des métaux ou des isolants topologiques. Par le principe de correspondance cœur-frontière, ces structures tirent leurs origines de la topologie du volume. En effet, on peut les interpréter comme une projection du flux de Berry non nul²⁰ dans certaines régions de la zone de Brillouin [40]. Pour y voir plus clair, regardons le schéma de la figure 7.16c. On y observe une paire de nœuds de Weyl, agissant respectivement comme une source et un drain pour la courbure de Berry. Si on évalue maintenant la quantité de flux qui traverse une surface 2D de la zone de Brillouin, on constate deux cas de figure possibles, illustrés par les carrés ouverts bleus. Si la surface se trouve entre la source et le drain, alors le flux net qui traverse la surface est non-nul et induit un état de surface topologique ($C = 1$). Dans le cas contraire, la résultante des flux entrants et sortants est nulle ($C = 0$). Puisque seuls les plans situés entre deux nœuds de Weyl de chiralité opposée possèdent une quantité non triviale de flux de Berry, l'arc de Fermi n'existe donc qu'entre leurs projections respectives sur la surface.

Les arcs de Fermi constituent une signature expérimentale distinctive des WSM, qui a été observée pour la première fois par ARPES en 2015 dans le TaAs [42] (fig. 7.16d). Les WSM sont de plus reconnus pour leurs propriétés de transport exotiques. Ils possèdent notamment une ultra-haute mobilité et une très grande magnétorésistance transverse [40, 45] ($\mathbf{E} \perp \mathbf{B}$), qui ont été reliées à la dispersion linéaire caractéristique des semi-métaux, sans égard à leur

19. D'un point de vue qualitatif, la chiralité d'une particule décrit si sa fonction d'onde se comporte comme la « règle de la main droite » ou la « règle de la main gauche » lorsqu'on lui applique une rotation de 2π .

20. La quantité de flux de Berry traversant une surface 2D, $W = \oint d^2\mathbf{k} \cdot \Omega_{\mathbf{k}}$, correspond à la définition modulo 2π de l'invariant topologique appelé nombre de Chern, $C = W/2\pi$ [35]. À l'instar du théorème de Gauss-Bonnet, il est obtenu par l'intégrale d'une courbure sur une surface fermée.

topologie. Lorsque combinées à une protection topologique, ces propriétés sont fort prometteuses pour le domaine du stockage d'information [40]. Une autre particularité des WSM qui attire considérablement l'intérêt est leur magnétorésistance longitudinale négative, qui pourrait constituer une réalisation physique de l'anomalie chirale en matière condensée [40]. En d'autres mots, l'interaction des fermions de Weyl avec des champs \mathbf{E} et \mathbf{B} parallèles abaisse le niveau de Fermi des particules de chiralité gauche par rapport à celui des particules de chiralité droite [46, 47]. Cette différence en énergie favorise une plus grande population de particules de chiralité droite à l'intérieur d'une même paire de nœuds de Weyl, ce qui induit un courant parallèle au champ électrique. La conductivité du matériau est donc plus élevée en présence du champ magnétique, contrairement aux métaux standards [48]. Bien que des signatures expérimentales cohérentes avec l'anomalie chirale aient été détectées dans des WSM [49, 50], d'autres mécanismes sont susceptibles d'expliquer les observations [51, 52], de sorte que la question demeure ouverte.

7.5. Transition de phase topologique

Dans le cadre de notre analyse du modèle SSH, nous avons observé qu'une variation des paramètres de sauts qui traverse la valeur critique $v = w$ modifiait l'invariant topologique et se manifestait par la fermeture de la bande interdite dans le volume. Dans le même esprit, on peut se demander s'il est possible de recourir à des paramètres contrôlables sur le plan expérimental pour induire une phase topologique dans un isolant standard. La perspective de manipuler avec précision l'apparition et la suppression des états de surface topologiques est d'ailleurs au cœur de plusieurs propositions d'applications, notamment le transistor topologique [53].

Sachant que la fermeture de la bande interdite est une des signatures d'une transition de phase topologique, tout paramètre dont la variation peut affecter l'énergie de la bande interdite pourrait donc, en principe, être utilisé pour induire une inversion des bandes et mener le système dans une phase topologique. Le comportement des bandes électroniques lors d'une transition de phase topologique contrôlée, que nous illustrons schématiquement à la figure 7.17 sera similaire à celui discuté pour le puits quantique de CdTe/HgTe. Dans la phase triviale, l'énergie de la bande interdite est non-nulle et l'ordonnement des bandes est tel que l'on peut les représenter en termes d'orbitales atomiques (gauche), représentées par les couleurs cyan et magenta. En augmentant progressivement le paramètre externe, l'énergie de la bande interdite diminue, jusqu'à devenir nulle à la transition de phase, lorsque le paramètre externe atteint sa valeur critique, P_C (centre). La dispersion électronique du volume est alors linéaire autour du niveau de Fermi et forme un cône de Dirac. Toute augmentation subséquente du paramètre externe repousse les états de conduction sous les états de valence, réalisant l'inversion des bandes caractéristique de la phase topologique au voisinage des

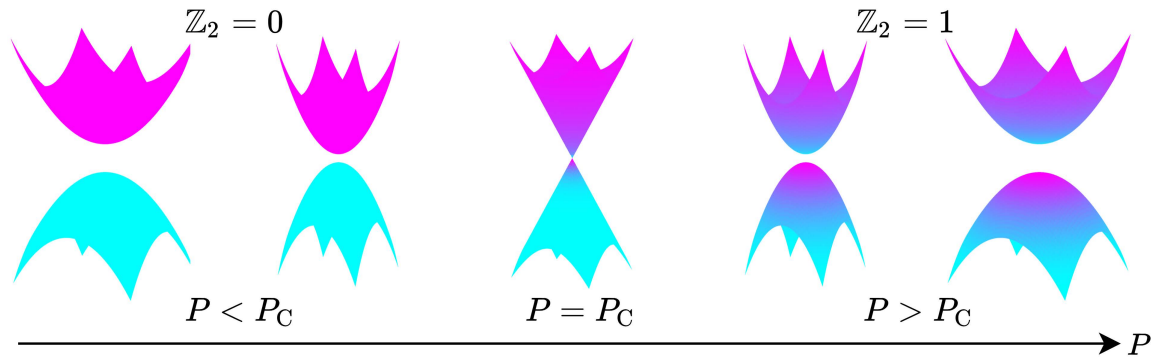


Figure 7.17 – Représentation schématique d’une transition de phase topologique. Dans la phase triviale ($\mathbb{Z}_2 = 0$), le caractère orbital des bandes au voisinage de leurs extrema respectifs peut être décomposé en terme d’orbitales atomiques, représenté ici par les couleurs cyan (VBM) et magenta (CBM). En augmentant la valeur du paramètre externe P , l’énergie de la bande interdite diminue et la courbure des bandes électroniques devient plus prononcée, jusqu’à devenir linéaire à la transition de phase ($P = P_C$). Cette dernière se produit lorsque l’énergie de la bande interdite est nulle. En augmentant à nouveau P , le système entre dans la phase topologique ($\mathbb{Z}_2 = 1$), caractérisée par une inversion du caractère orbital des bandes (cyan \leftrightarrow magenta) au voisinage des extrema de bande.

extrema de bande (droite). Cette configuration des bandes est souvent interprétée comme une bande interdite d’énergie négative (à l’instar des bandes du volume de la fig. 7.12b), à laquelle le couplage spin-orbite confère une valeur finie.

Ce type de transition de phase topologique a été observé à de nombreuses reprises dans la littérature tant théorique qu’expérimentale. Un des premiers paramètres explorés fut la modification des constituants atomiques du matériau : la première observation expérimentale d’états de surface topologiques dans un matériau 3D a été réalisée dans un alliage de bismuth et d’antimoine ($\text{Bi}_{1-x}\text{Sb}_x$) [54]. Le dopage stoechiométrique a également permis d’identifier des phases non-triviales dans le $\text{BiTl}(\text{S}_{1-x}\text{Se}_x)$ [55] et le $(\text{Pb}_{1-x}\text{Sn}_x)\text{Se}$ [56]. La signature d’une phase topologique a de plus été observée dans le phosphorène plongé dans un champ électrique [57] ainsi que dans le Sb_2Se_3 [58] et le SrAs_3 [59] soumis à une pression hydrostatique. La riche phénoménologie de l’interaction lumière-matière n’est pas en reste ; des expériences de type pompe-sonde dans le Bi_2Se_3 ont en effet démontré que la photoexcitation de la surface d’un isolant topologique par un laser dont la polarisation brise la symétrie de renversement du temps crée des répliques de l’état de surface séparées par l’énergie des photons et peut même ouvrir dynamiquement la bande interdite du cône de Dirac [60, 61]. De telles observations suggèrent qu’il pourrait un jour être possible de contrôler sélectivement les propriétés topologiques d’un matériau sur une échelle de temps de l’ordre de la picoseconde, ce qui est particulièrement intéressant du point de vue de l’ingénierie.

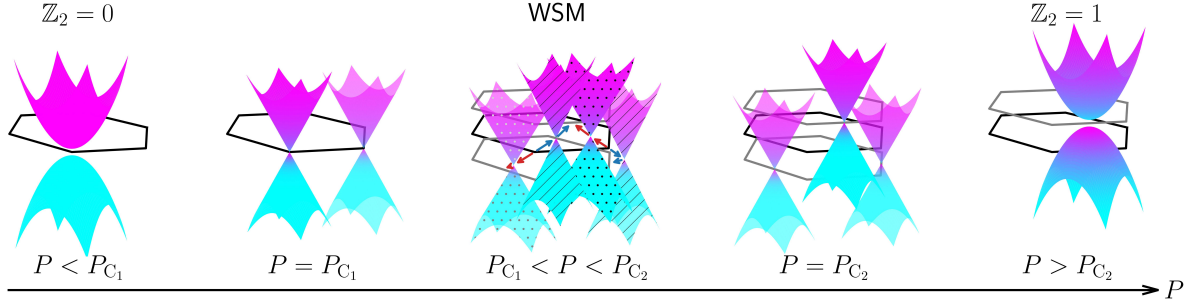


Figure 7.18 – Représentation schématique d’une transition de phase topologique dans un matériau sans symétrie d’inversion. Comme à la figure 7.17, la phase d’isolant topologique \mathbb{Z}_2 est associée à une inversion du caractère orbital dominant des bandes de valence et de conduction (couleurs cyan et magenta). L’absence de symétrie d’inversion impose toutefois l’existence d’une phase intermédiaire de type semi-métal de Weyl (WSM), bornée par deux valeurs critiques du paramètre externe P . À $P = P_{C_1}$, la dispersion électronique forme un nombre discret de cônes de Dirac, correspondant au nombre de partenaires de Kramers reliés par une symétrie spatiale du cristal (3 paires de Kramers pour la symétrie hexagonale illustrée ci-dessus). Le schéma se concentre sur la trajectoire des bandes électroniques opaques. Pour faciliter la compréhension, certains de ses partenaires symétriques sont montrés avec une opacité réduite. Pour $P_{C_1} < P < P_{C_2}$, le point de contact des bandes de chaque cône de Dirac se sépare en deux nœuds de Weyl de chiralité opposée, représentés ici par les structures hachurées et pointillées. En augmentant la valeur de P à l’intérieur de la phase WSM, les paires de nœuds de Weyl migrent dans la zone de Brillouin, le long du chemin indiqué par les flèches bleues et rouges. À $P = P_{C_2}$, les différents nœuds de Weyl sont annihilés lorsqu’ils rejoignent un nœud de chiralité opposée issu d’une autre paire. En augmentant à nouveau P , la bande interdite s’ouvre et le matériau entre dans la phase d’isolant topologique. À ce moment, la localisation du VBM et du CBM n’est plus la même que dans la phase isolante initiale.

7.5.1. Transition de phase intermédiaire dans les matériaux sans symétrie d’inversion

Pour les matériaux non centrosymétriques, la représentation schématique d’une transition de phase topologique présentée à la section précédente est incomplète. En effet, le modèle développé par Liu et Vanderbilt [62] a révélé que l’absence de symétrie d’inversion impose que le passage d’un état trivial $\mathbb{Z}_2 = 0$ à un état topologique $\mathbb{Z}_2 = 1$ soit accompagné d’une phase intermédiaire de type semi-métal de Weyl, dont l’énergie de la bande interdite est nulle. En fonction du système étudié, la largeur de cette phase dans l’espace des paramètres peut toutefois être arbitrairement petite, rendant ainsi difficile sa détection expérimentale. La conclusion de Liu et Vanderbilt est indépendante des autres symétries spatiales spécifiques au matériau.

La figure 7.18 illustre le comportement des bandes lors de ce type de transition de phase, en utilisant le même code de couleur qu’à la figure 7.17. Rappelons-nous qu’en présence du

SOC, les extrema de bande des matériaux non centrosymétriques ne sont pas exactement localisés à un point de haute symétrie, tel que mis en évidence par les modèles de Dresselhaus et Rashba (section 2.5.4)²¹. Ainsi, à une première valeur critique du paramètre, P_{C1} , les bandes électroniques du volume croisent le niveau de Fermi à un nombre discret de points, qui correspond au nombre de partenaires de Kramers reliés par la symétrie particulière du matériau. Par exemple, il y aurait 3 paires de Kramers, donc 6 points de contact, pour la symétrie hexagonale illustrée sur le schéma. La dispersion électronique y est linéaire et forme un cône de Dirac. La transition de phase représentée à la figure 7.18 se concentre sur la paire de bandes opaque (qui est la seule illustrée dans les phases isolantes, aux extrêmes gauche et droit). Pour faciliter la compréhension, nous montrons également certains partenaires symétriques avec une opacité réduite.

En augmentant à nouveau le paramètre externe, chaque point de contact devient une paire de nœuds de Weyl de chiralité opposée (identifiés par la texture pointée ou hachurée des bandes), qui se séparent en impulsion dans l'espace réciproque. L'énergie de la bande interdite du volume demeure donc nulle. La migration des nœuds de Weyl dans la zone de Brillouin se poursuit avec la variation subséquente du paramètre externe (flèches rouges et bleues), jusqu'à ce que chaque monopôle rencontre un antimonopôle issu d'un point de contact voisin. (On parle alors d'échange de partenaires.) L'annihilation mutuelle des nœuds de Weyl marque la seconde valeur critique du paramètre, P_{C2} , au-delà de laquelle l'inversion des bandes et l'apparition de la phase topologique isolante se produisent, à l'instar de la figure 7.17. L'absence de symétrie d'inversion a donc deux conséquences importantes sur la transition de phase topologique. D'une part, il existe deux valeurs critiques du paramètre externe, marquant respectivement le début et la fin de la phase WSM. D'autre part, la migration des nœuds de Weyl implique que l'ouverture et la fermeture de la bande interdite du volume ne se produit pas au même endroit dans la zone de Brillouin (plans hexagonaux en noir et en gris).

7.6. Conséquences de l'interaction électron-phonon sur les phases topologiques

La DFT a été abondamment utilisée pour prédire et caractériser de nouveaux matériaux topologiques [41, 63–70]. Un avantage indéniable de cette méthode est qu'elle ne dépend d'aucun paramètre externe, mis à part la structure cristalline et, dans une certaine mesure, le choix de la fonctionnelle d'échange-corrélation, qui déterminera l'énergie de la bande interdite. Or, nous avons vu aux chapitres 2 et 3 que la DFT est une théorie de l'état fondamental à $T = 0$ K reposant sur l'approximation des noyaux statiques, alors que les

21. Pour simplifier la figure schématique, on les a néanmoins positionnés aux frontières de plans 2D dans l'espace réciproque.

expériences capturent inévitablement les effets du mouvement des atomes, même à température nulle. Pour prédire correctement la phase topologique d'un matériau à température non nulle, il faudra donc considérer comment la structure électronique est modifiée par la température. Nous venons par ailleurs d'argumenter qu'une transition de phase topologique (TPT) implique nécessairement que l'énergie de la bande interdite devienne nulle. Dans ce contexte, on peut se demander si l'interaction électron-phonon, qui domine généralement la dépendance en température des énergies électroniques, est capable de fermer la bande interdite et, par conséquent, d'induire une transition de phase topologique dans un matériau autrement trivial.

Cette possibilité a initialement été proposée en 2013 par Garate *et al.* [71, 72] et explorée à l'aide de modèles hamiltoniens. Les premières études de ce type basées sur la DFT, réalisées par Antonius et Louie [73] et par Monserrat et Vanderbilt [74], sont quant à elles apparues en 2016. Dans le premier cas, les auteurs considèrent simultanément l'effet de la température et du dopage stœchiométrique dans la famille $\text{BiTl}(\text{S}_{1-x}\text{Se}_x)_2$. Leurs résultats indiquent que l'augmentation de la température détruit la phase topologique à dopage fort et suggèrent l'existence d'une phase non-triviale photo-induite à dopage faible. Dans le second, les auteurs n'observent aucune disparition de la phase non-triviale du Bi_2Se_3 , Bi_2Te_3 et Sb_2Te_3 jusqu'à 600 K. Ils investiguent de plus la TPT induite par pression hydrostatique dans le Sb_2Se_3 et constatent une augmentation de la pression critique avec la température. Ces deux études détectent la TPT en recherchant la valeur critique du paramètre externe pour laquelle l'énergie de la bande interdite devient nulle, comme illustré à la figure 7.17. Toutefois, puisque l'énergie de la bande interdite dépend désormais à la fois de la température et du paramètre externe, cette valeur critique acquiert une dépendance en température.

Depuis la publication de ces travaux, les méthodes numériques basées sur les principes premiers sont devenues une approche de prédilection pour étudier la dépendance en température des phases topologiques de la matière. Elles ont été appliquées à différentes classes de matériaux topologiques, notamment aux semi-métaux de Weyl [75], aux semi-métaux de Kondo [76], aux semi-métaux à ligne nodale [77, 78], en plus des isolants topologiques [79, 80]. Entre autres, Monserrat et Narayan ont relié le signe de la variation de l'énergie de la bande interdite en fonction de la température aux phases d'isolant topologique faible et fort du ZrTe_5 [81], alors que Lihm et Park [82] ont démontré l'importance de l'hybridisation des fonctions d'onde électroniques au proche voisinage des transitions de phase topologiques. L'observation expérimentale d'une transition de phase topologique induite par la température a également été rapportée dans des puits quantiques de CdTe/HgTe [83, 84] et dans le ZrTe_5 [85].

7.6.1. Une compétition entre les couplages interbandes et intra-bandés

Pour comprendre qualitativement comment l'interaction électron-phonon peut affecter la phase topologique d'un matériau, nous allons maintenant considérer un modèle très simple formé de deux bandes paraboliques au voisinage de leurs extrema respectifs. Cet argument heuristique, publié originellement dans ma contribution au compte rendu du XI^e Symposium international Théorie quantique et Symétries [86], s'inspire des modèles développés par Garate [71] et par Saha et Garate [72].

Rappelons-nous que les contributions de Fan et de Debye-Waller à la self-énergie de l'interaction électron-phonon (éq. (3.3.13) et (3.3.29)) sont de signe opposé et en général de grandeur similaire. Dans ces équations, la somme sur l'indice de bande, n' , peut être divisée en deux groupes, soit les contributions respectives du sous-ensemble des bandes occupées et de celui des bandes inoccupées. Dans un semi-conducteur typique, c'est habituellement le terme de Fan qui dicte le signe de la renormalisation nette associée à chacun de ces sous-espaces. Notre modèle se concentrera sur cette contribution et représentera les couplages à l'intérieur d'un même sous-ensemble de bandes par un couplage intrabande, tandis que les couplages entre les deux sous-ensembles seront associés au couplage interbande.

En analysant l'équation (3.3.13), que nous réécrivons ici par commodité,

$$\begin{aligned} \Sigma_{\mathbf{k}n}^{\text{Fan}}(\varepsilon_{\mathbf{k}n}^0, T) &= \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \sum_{n'}^{\text{BZ}} |\langle \mathbf{k} + \mathbf{q}n' | \nabla_{\mathbf{q}\nu} V^{\text{KS}} | \mathbf{k}n \rangle|^2 \\ &\times \left[\frac{n_{\mathbf{q}\nu}(T) + 1 - f_{\mathbf{k}+\mathbf{q}n'}(T)}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 - \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} + \frac{n_{\mathbf{q}\nu}(T) + f_{\mathbf{k}+\mathbf{q}n'}(T)}{\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 + \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} \right], \end{aligned} \quad (7.6.1)$$

on peut constater que le signe de chaque terme contribuant à la somme est entièrement déterminé par le signe des dénominateurs. En effet, les fréquences phononiques, les fonctions d'occupation Bose-Einstein et Fermi-Dirac, de même que la norme carrée des éléments de matrice du couplage électron-phonon, sont toutes des quantités définies positives. Le signe de chaque contribution dépendra donc de celui de la différence entre les énergies propres des états électroniques en jeu, $\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0$. On suppose ici que l'état électronique corrigé, $\varepsilon_{\mathbf{k}n}(T)$, est un extrema de bande et que les fréquences phononiques sont faibles par rapport aux énergies électroniques. Ainsi, pour un couplage intrabande (fig. 7.19a, gauche), l'énergie du VBM augmente puisqu'il est couplé à des états électroniques d'énergie plus basse, tandis que l'énergie du CBM diminue, car il est couplé à des états de plus haute énergie. En combinant ces deux corrections, on constate que le couplage intrabande réduit l'énergie de la bande interdite. Dans le cas d'un couplage interbande (fig. 7.19a, droite), le comportement de chaque extrema est inversé par rapport à l'argument précédent, de sorte que l'énergie

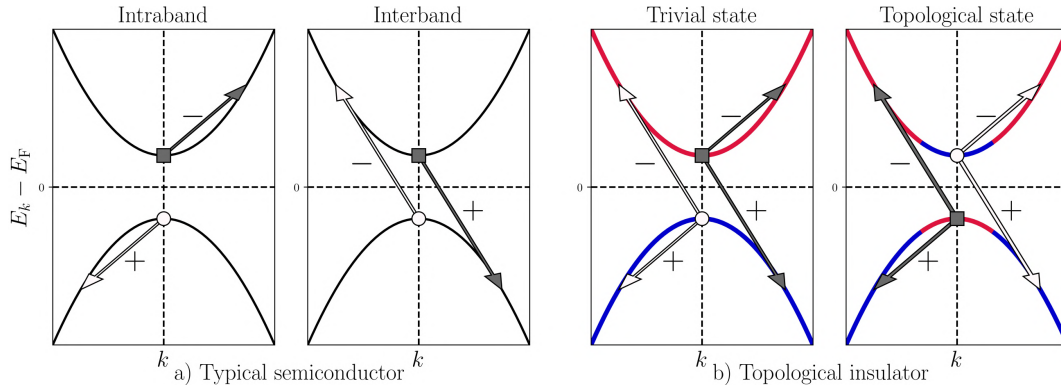


Figure 7.19 – Modèle heuristique de la renormalisation du point zéro dans un isolant topologique. a) Couplages intrabande and interbande dans un semi-conducteur typique (isolant trivial). b) Dans un isolant topologique, le caractère orbital dominant des extrema de bande électronique diffère de part et d’autre de la transition de phase topologique, en raison du phénomène d’inversion des bandes. Les symboles « + » et « - » indiquent le signe de la contribution de chaque type de couplage dans la self-énergie de Fan (éq. (7.6.1)). Image tirée de la référence [86].

de la bande interdite augmente. L’interaction dominante dictera donc le signe net de la renormalisation de l’énergie de la bande interdite.

Dans un semi-conducteur typique, l’énergie de la bande interdite est habituellement de l’ordre de $\sim 1\text{eV}$, de sorte que les couplages interbandes sont fortement défavorisés par la très grande différence entre les énergies électroniques présentes au dénominateur. Par conséquent, les couplages intrabandes dominent l’EPI et induisent une diminution de l’énergie de la bande interdite lorsque la température augmente. Ce comportement est connu dans la littérature sous le nom d’effet Varshni.

L’analyse est beaucoup plus subtile dans le cas d’un isolant topologique, puisque l’on doit également prendre en compte le caractère orbital dominant des bandes de valence et de conduction, ainsi que l’inversion du caractère des bandes associée à la phase topologique. Considérons le cas du VBM de caractère orbital « bleu » montré à la figure 7.19b (gauche, cercle blanc). Supposons que, dans la phase triviale, les couplages intrabandes dominent, de sorte que la somme des couplages entre les bandes de même caractère (« bleu-bleu ») et entre les bandes de caractères différents (« bleu-rouge ») augmente l’énergie du VBM. Ces comportements demeurent inchangés dans la phase topologique. Or, à cause de l’inversion des bandes, cet extrema est devenu le CBM (fig. 7.19b, droite). Par un argument similaire, l’énergie de l’extrema de bande de caractère orbital « rouge » (carré gris) diminue, tant dans la phase triviale où il constitue le CBM que dans la phase topologique, où il est désormais le VBM.

Cet argument heuristique nous permet donc de déduire deux scénarios possibles. Dans un premier cas, l’interaction électron-phonon réduit l’énergie de la bande interdite dans la phase

triviale, facilitant ainsi une inversion du caractère des bandes, et a l'effet opposé dans la phase topologique, ce qui stabilise l'inversion des bandes déjà établie. Dans le cas contraire, elle augmente l'énergie de la bande interdite dans la phase triviale, inhibant alors une éventuelle inversion des bandes, et la diminue dans la phase topologique, ce qui pourrait ramener les bandes inversées dans leur configuration triviale. En somme, les couplages intrabandes favorisent l'apparition de phases topologiques avec l'augmentation de la température, tandis que les couplages interbandes favorisent globalement la phase triviale.

Dans un matériau réel, la situation est bien entendu plus complexe que dans notre modèle simplifié. On doit notamment prendre en compte la contribution du terme de Debye-Waller à la self-énergie EPI (éq. (3.3.29) et (3.3.30)), ainsi que la contribution de la dilatation thermique à la dépendance en température de l'énergie de la bande interdite. En effet, une compétition similaire à celle discutée dans notre modèle pour les couplages intrabandes et interbandes existe entre l'interaction électron-phonon et la dilatation thermique dans un matériau réel. Ainsi, comme on verra au chapitre suivant, la dilatation thermique peut renforcer ou contrer la renormalisation induite par l'interaction électron-phonon. Une caractérisation complète du diagramme de phase topologique d'un matériau dépendra par conséquent tant du signe que de la force relative de ces deux contributions. Des travaux récents tant théoriques [79, 81] qu'expérimentaux [87] ont d'ailleurs souligné l'importance de considérer l'EPI et la dilatation thermique sur un même pied d'égalité pour prédire correctement la phase topologique d'un matériau à température non nulle²². Néanmoins, l'objectif de notre modèle heuristique est de favoriser une compréhension intuitive de la manière dont les mécanismes qui gouvernent l'interaction électron-phonon peuvent affecter la phase topologique d'un matériau, afin de faciliter la lecture du prochain chapitre.

Bibliographie

- [1] K. v. Klitzing, G. Dorda, and M. Pepper. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance. *Phys. Rev. Lett.*, 45(6) :494–497, 1980. doi: [10.1103/PhysRevLett.45.494](https://doi.org/10.1103/PhysRevLett.45.494).
- [2] M. Z. Hasan and C. L. Kane. Colloquium : Topological insulators. *Rev. Mod. Phys.*, 82(4) :3045–3067, 2010. doi: [10.1103/RevModPhys.82.3045](https://doi.org/10.1103/RevModPhys.82.3045).
- [3] W. P. Su, J. R. Schrieffer, and A. J. Heeger. Solitons in polyacetylene. *Phys. Rev. Lett.*, 42(25) :1698–1701, 1979. doi: [10.1103/PhysRevLett.42.1698](https://doi.org/10.1103/PhysRevLett.42.1698).

²². Dans le cas d'un isolant topologique \mathbb{Z}_2 tridimensionnel, dont la topologie non triviale est protégée par la symétrie de renversement du temps, la présence du couplage spin-orbite est nécessaire pour réaliser l'inversion des bandes [2, 25]. L'inclusion du SOC s'impose alors d'emblée dès le calcul de l'état fondamental statique (et, par conséquent, des propriétés vibrationnelles) puisqu'il est nécessaire pour modéliser correctement la topologie non triviale de ce type de matériau. Dans ce cas, il est donc implicite que le SOC est indispensable pour évaluer correctement la dépendance en température de l'énergie de la bande interdite d'un isolant topologique.

- [4] W. P. Su, J. R. Schrieffer, and A. J. Heeger. Soliton excitations in polyacetylene. *Phys. Rev. B*, 22(4) :2099–2111, 1980. doi: [10.1103/PhysRevB.22.2099](https://doi.org/10.1103/PhysRevB.22.2099).
- [5] A. J. Heeger, S. Kivelson, J. R. Schrieffer, and W. P. Su. Solitons in conducting polymers. *Rev. Mod. Phys.*, 60(3) :781–850, 1988. doi: [10.1103/RevModPhys.60.781](https://doi.org/10.1103/RevModPhys.60.781).
- [6] J. K. Asbóth, L. Oroszlány, and A. Pályi. *A Short Course on Topological Insulators*, volume 919 of *Lecture Notes in Physics*. Springer, Cham, 2016. doi: [10.1007/978-3-319-25607-8](https://doi.org/10.1007/978-3-319-25607-8).
- [7] M. P. Marder. *Condensed Matter Physics*. John Wiley & Sons, Hoboken, 2nd edition, 2010.
- [8] S. Q. Shen. *Topological Insulators*, volume 187 of *Springer Series in Solid-State Sciences*. Springer Singapore, Singapore, 2017. doi: [10.1007/978-981-10-4606-3](https://doi.org/10.1007/978-981-10-4606-3).
- [9] X. G. Wen. Colloquium : Zoo of quantum-topological phases of matter. *Rev. Mod. Phys.*, 89(4) :041004, 2017. doi: [10.1103/RevModPhys.89.041004](https://doi.org/10.1103/RevModPhys.89.041004).
- [10] L. Landau. On the Theory of Phase Transitions. *Zh. Eksp. Teor. Fiz.*, 7 :19–32, <http://archive.ujp.bitp.kiev.ua/files/journals/53/si/53SI08p.pdf>.
- [11] G. Jaeger. The Ehrenfest Classification of Phase Transitions : Introduction and Evolution. *Archive for History of Exact Sciences*, 53(1) :51–81, 1998. doi: [10.1007/s004070050021](https://doi.org/10.1007/s004070050021).
- [12] D. Tong. Lectures on the Quantum Hall Effect. *arXiv :1606.06687 [cond-mat, physics :hep-th]*, 2016. doi: [10.48550/arXiv.1606.06687](https://doi.org/10.48550/arXiv.1606.06687).
- [13] J. E. Moore and L. Balents. Topological invariants of time-reversal-invariant band structures. *Phys. Rev. B*, 75(12) :121306, 2007. doi: [10.1103/PhysRevB.75.121306](https://doi.org/10.1103/PhysRevB.75.121306).
- [14] C. K. Chiu, J. C. Y. Teo, A. P. Schnyder, and S. Ryu. Classification of topological quantum matter with symmetries. *Rev. Mod. Phys.*, 88(3) :035005, 2016. doi: [10.1103/RevModPhys.88.035005](https://doi.org/10.1103/RevModPhys.88.035005).
- [15] K. Mazumder and P. M. Shirage. A brief review of Bi₂Se₃ based topological insulator : From fundamentals to applications. *Journal of Alloys and Compounds*, 888 :161492, 2021. doi: <https://doi.org/10.1016/j.jallcom.2021.161492>.
- [16] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan. Observation of a large-gap topological-insulator class with a single Dirac cone on the surface. *Nat Phys*, 5(6) :398–402, 2009. doi: [10.1038/nphys1274](https://doi.org/10.1038/nphys1274).
- [17] P. Roushan, J. Seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, and A. Yazdani. Topological surface states protected from backscattering by chiral spin texture. *Nature*, 460(7259) :1106–1109, 2009. doi: [10.1038/nature08308](https://doi.org/10.1038/nature08308).
- [18] L. A. Wray, S. Y. Xu, Y. Xia, D. Hsieh, A. V. Fedorov, Y. S. Hor, R. J. Cava, A. Bansil, H. Lin, and M. Z. Hasan. A topological insulator surface under strong Coulomb, magnetic and disorder perturbations. *Nature Phys*, 7(1) :32–37, 2011. doi: [10.1038/nphys1838](https://doi.org/10.1038/nphys1838).

- [19] C. Kittel. *Introduction to Solid State Physics*. Wiley, 8th edition, 2004.
- [20] J. Cano, B. Bradlyn, Z. Wang, L. Elcoro, M. G. Vergniory, C. Felser, M. I. Aroyo, and B. A. Bernevig. Topology of disconnected elementary band representations. *Phys. Rev. Lett.*, 120(26) :266401, 2018. doi: [10.1103/PhysRevLett.120.266401](https://doi.org/10.1103/PhysRevLett.120.266401).
- [21] M. G. Vergniory, L. Elcoro, C. Felser, N. Regnault, B. A. Bernevig, and Z. Wang. A complete catalogue of high-quality topological materials. *Nature*, 566(7745) :480–485, 2019. doi: [10.1038/s41586-019-0954-4](https://doi.org/10.1038/s41586-019-0954-4).
- [22] B. A. Bernevig, T. L. Hughes, and S. C. Zhang. Quantum spin hall effect and topological phase transition in HgTe quantum wells. *Science*, 314(5806) :1757–1761, 2006. doi: [10.1126/science.1133734](https://doi.org/10.1126/science.1133734).
- [23] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X. L. Qi, and S. C. Zhang. Quantum spin hall insulator state in HgTe quantum wells. *Science*, 318(5851) :766–770, 2007. doi: [10.1126/science.1148047](https://doi.org/10.1126/science.1148047).
- [24] C. L. Kane and E. J. Mele. Z_2 topological order and the quantum spin hall effect. *Phys. Rev. Lett.*, 95(14) :146802, 2005. doi: [10.1103/PhysRevLett.95.146802](https://doi.org/10.1103/PhysRevLett.95.146802).
- [25] X. L. Qi and S. C. Zhang. Topological insulators and superconductors. *Rev. Mod. Phys.*, 83(4) :1057–1110, 2011. doi: [10.1103/RevModPhys.83.1057](https://doi.org/10.1103/RevModPhys.83.1057).
- [26] A. A. Soluyanov and D. Vanderbilt. Wannier representation of Z_2 topological insulators. *Phys. Rev. B*, 83(3) :035108, 2011. doi: [10.1103/PhysRevB.83.035108](https://doi.org/10.1103/PhysRevB.83.035108).
- [27] A. A. Soluyanov and D. Vanderbilt. Computing topological invariants without inversion symmetry. *Phys. Rev. B*, 83(23) :235401, 2011. doi: [10.1103/PhysRevB.83.235401](https://doi.org/10.1103/PhysRevB.83.235401).
- [28] D. Xiao, M. C. Chang, and Q. Niu. Berry phase effects on electronic properties. *Rev. Mod. Phys.*, 82(3) :1959–2007, 2010. doi: [10.1103/RevModPhys.82.1959](https://doi.org/10.1103/RevModPhys.82.1959).
- [29] N. Marzari, I. Souza, and D. Vanderbilt. An Introduction to Maximally-Localized Wannier Functions. *Highlight of the Month, Psi-K Newsletter*, 57, 2003, <https://cfm.ehu.es/ivo/publications/marzari-psik03.pdf>.
- [30] D. Gresch and A. Soluyanov. Calculating Topological Invariants with Z2Pack. In D. Bercioux, J. Cayssol, M. G. Vergniory, and M. Reyes Calvo, editors, *Topological Matter*, volume 190, pages 63–92. Springer International Publishing, Cham, 2018. doi: [10.1007/978-3-319-76388-0_3](https://doi.org/10.1007/978-3-319-76388-0_3).
- [31] D. Gresch, G. Autès, O. V. Yazyev, M. Troyer, D. Vanderbilt, B. A. Bernevig, and A. A. Soluyanov. Z2Pack : Numerical implementation of hybrid Wannier centers for identifying topological materials. *Phys. Rev. B*, 95(7) :075146, 2017. doi: [10.1103/PhysRevB.95.075146](https://doi.org/10.1103/PhysRevB.95.075146).
- [32] L. Fu and C. L. Kane. Topological insulators with inversion symmetry. *Phys. Rev. B*, 76(4) :045302, 2007. doi: [10.1103/PhysRevB.76.045302](https://doi.org/10.1103/PhysRevB.76.045302).
- [33] L. Fu, C. L. Kane, and E. J. Mele. Topological Insulators in Three Dimensions. *Phys. Rev. Lett.*, 98(10) :106803, 2007. doi: [10.1103/PhysRevLett.98.106803](https://doi.org/10.1103/PhysRevLett.98.106803).

- [34] R. B. Laughlin. Quantized Hall conductivity in two dimensions. *Phys. Rev. B*, 23(10) : 5632–5633, 1981. doi: [10.1103/PhysRevB.23.5632](https://doi.org/10.1103/PhysRevB.23.5632).
- [35] A. Akhmerov, J. Sau, B. van Heck, S. Rubbert, R. Skolasiński, B. Nijholt, I. Muhammad, and T. O. Rosdahl. Online course on topology in condensed matter. <https://topocondmat.org>.
- [36] O. Breunig and Y. Ando. Opportunities in topological insulator devices. *Nat Rev Phys*, 4 :184–193, 2021. doi: [10.1038/s42254-021-00402-6](https://doi.org/10.1038/s42254-021-00402-6).
- [37] I. Žutić, J. Fabian, and S. Das Sarma. Spintronics : Fundamentals and applications. *Rev. Mod. Phys.*, 76(2) :323–410, 2004. doi: [10.1103/RevModPhys.76.323](https://doi.org/10.1103/RevModPhys.76.323).
- [38] E. Gibney. Quantum computer race intensifies as alternative technology gains steam. *Nature News*, 2020. doi: [10.1038/d41586-020-03237-w](https://doi.org/10.1038/d41586-020-03237-w).
- [39] A. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1) : 2–30, 2003. doi: [https://doi.org/10.1016/S0003-4916\(02\)00018-0](https://doi.org/10.1016/S0003-4916(02)00018-0).
- [40] B. Yan and C. Felser. Topological Materials : Weyl Semimetals. *Annu. Rev. Condens. Matter Phys.*, 8(1) :337–354, 2017. doi: [10.1146/annurev-conmatphys-031016-025458](https://doi.org/10.1146/annurev-conmatphys-031016-025458).
- [41] H. Weng, C. Fang, Z. Fang, B. A. Bernevig, and X. Dai. Weyl Semimetal Phase in Noncentrosymmetric Transition-Metal Monophosphides. *Phys. Rev. X*, 5(1) :011029, 2015. doi: [10.1103/PhysRevX.5.011029](https://doi.org/10.1103/PhysRevX.5.011029).
- [42] S. Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C. C. Lee, S. M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan. Discovery of a Weyl fermion semimetal and topological Fermi arcs. *Science*, 349(6248) :613–617, 2015. doi: [10.1126/science.aaa9297](https://doi.org/10.1126/science.aaa9297).
- [43] X. Ying and A. Kamenev. Symmetry-Protected Topological Metals. *Phys. Rev. Lett.*, 121(8) :086810, 2018. doi: [10.1103/PhysRevLett.121.086810](https://doi.org/10.1103/PhysRevLett.121.086810).
- [44] P. B. Pal. Dirac, Majorana and Weyl fermions. *American Journal of Physics*, 79(5) : 485–498, 2011. doi: [10.1119/1.3549729](https://doi.org/10.1119/1.3549729).
- [45] C. Shekhar, A. K. Nayak, Y. Sun, M. Schmidt, M. Nicklas, I. Leermakers, U. Zeitler, Y. Skourski, J. Wosnitza, Z. Liu, Y. Chen, W. Schnelle, H. Borrmann, Y. Grin, C. Felser, and B. Yan. Extremely large magnetoresistance and ultrahigh mobility in the topological Weyl semimetal candidate NbP. *Nature Phys*, 11(8) :645–649, 2015. doi: [10.1038/nphys3372](https://doi.org/10.1038/nphys3372).
- [46] X. Yuan, C. Zhang, Y. Zhang, Z. Yan, T. Lyu, M. Zhang, Z. Li, C. Song, M. Zhao, P. Leng, M. Ozerov, X. Chen, N. Wang, Y. Shi, H. Yan, and F. Xiu. The discovery of dynamic chiral anomaly in a Weyl semimetal NbAs. *Nat Commun*, 11(1) :1259, 2020. doi: [10.1038/s41467-020-14749-4](https://doi.org/10.1038/s41467-020-14749-4).
- [47] T. Meng. Coupled-wire constructions : a Luttinger liquid approach to topology. *Eur. Phys. J. Spec. Top.*, 229(4) :527–543, 2020. doi: [10.1140/epjst/e2019-900095-5](https://doi.org/10.1140/epjst/e2019-900095-5).

- [48] S. Jia, S. Y. Xu, and M. Z. Hasan. Weyl semimetals, Fermi arcs and chiral anomalies. *Nature Mater*, 15(11) :1140–1144, 2016. doi: [10.1038/nmat4787](https://doi.org/10.1038/nmat4787).
- [49] X. Huang, L. Zhao, Y. Long, P. Wang, D. Chen, Z. Yang, H. Liang, M. Xue, H. Weng, Z. Fang, X. Dai, and G. Chen. Observation of the Chiral-Anomaly-Induced Negative Magnetoresistance in 3D Weyl Semimetal TaAs. *Phys. Rev. X*, 5(3) :031023, 2015. doi: [10.1103/PhysRevX.5.031023](https://doi.org/10.1103/PhysRevX.5.031023).
- [50] C. L. Zhang, S. Y. Xu, I. Belopolski, Z. Yuan, Z. Lin, B. Tong, G. Bian, N. Alidoust, C. C. Lee, S. M. Huang, T. R. Chang, G. Chang, C. H. Hsu, H. T. Jeng, M. Neupane, D. S. Sanchez, H. Zheng, J. Wang, H. Lin, C. Zhang, H. Z. Lu, S. Q. Shen, T. Neupert, M. Zahid Hasan, and S. Jia. Signatures of the Adler–Bell–Jackiw chiral anomaly in a Weyl fermion semimetal. *Nat Commun*, 7(1) :10735, 2016. doi: [10.1038/ncomms10735](https://doi.org/10.1038/ncomms10735).
- [51] R. D. d. Reis, M. O. Ajeesh, N. Kumar, F. Arnold, C. Shekhar, M. Naumann, M. Schmidt, M. Nicklas, and E. Hassinger. On the search for the chiral anomaly in Weyl semimetals : the negative longitudinal magnetoresistance. *New Journal of Physics*, 18(8) :085006, 2016. doi: [10.1088/1367-2630/18/8/085006](https://doi.org/10.1088/1367-2630/18/8/085006).
- [52] N. P. Ong and S. Liang. Experimental signatures of the chiral anomaly in Dirac–Weyl semimetals. *Nat Rev Phys*, 3(6) :394–404, 2021. doi: [10.1038/s42254-021-00310-9](https://doi.org/10.1038/s42254-021-00310-9).
- [53] L. Fleet. Topological transistor. *Nat Phys*, 11(1) :5, 2015. doi: [10.1038/nphys3217](https://doi.org/10.1038/nphys3217).
- [54] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan. A topological Dirac insulator in a quantum spin Hall phase. *Nature*, 452(7190) :970–974, 2008. doi: [10.1038/nature06843](https://doi.org/10.1038/nature06843).
- [55] S. Y. Xu, Y. Xia, L. A. Wray, S. Jia, F. Meier, J. H. Dil, J. Osterwalder, B. Slomski, A. Bansil, H. Lin, R. J. Cava, and M. Z. Hasan. Topological phase transition and texture inversion in a tunable topological insulator. *Science*, 332(6029) :560–564, 2011. doi: [10.1126/science.1201607](https://doi.org/10.1126/science.1201607).
- [56] B. M. Wojek, P. Dziawa, B. J. Kowalski, A. Szczerbakow, A. M. Black-Schaffer, M. H. Berntsen, T. Balasubramanian, T. Story, and O. Tjernberg. Band inversion and the topological phase transition in (Pb,Sn)Se. *Phys. Rev. B*, 90(16) :161202, 2014. doi: [10.1103/PhysRevB.90.161202](https://doi.org/10.1103/PhysRevB.90.161202).
- [57] Q. Liu, X. Zhang, L. B. Abdalla, A. Fazzio, and A. Zunger. Switching a Normal Insulator into a Topological Insulator via Electric Field with Application to Phosphorene. *Nano Lett.*, 15(2) :1222–1228, 2015. doi: [10.1021/nl5043769](https://doi.org/10.1021/nl5043769).
- [58] A. Bera, K. Pal, D. V. S. Muthu, S. Sen, P. Guptasarma, U. V. Waghmare, and A. K. Sood. Sharp Raman Anomalies and Broken Adiabaticity at a Pressure Induced Transition from Band to Topological Insulator in Sb₂Se₃. *Phys. Rev. Lett.*, 110(10), 2013. doi: [10.1103/PhysRevLett.110.107401](https://doi.org/10.1103/PhysRevLett.110.107401).

- [59] E. Cheng, W. Xia, X. Shi, Z. Yu, L. Wang, L. Yan, D. C. Peets, C. Zhu, H. Su, Y. Zhang, D. Dai, X. Wang, Z. Zou, N. Yu, X. Kou, W. Yang, W. Zhao, Y. Guo, and S. Li. Pressure-induced superconductivity and topological phase transitions in the topological nodal-line semimetal SrAs₃. *npj Quantum Materials*, 5 :38, 2020. doi: [10.1038/s41535-020-0240-6](https://doi.org/10.1038/s41535-020-0240-6).
- [60] Y. H. Wang, H. Steinberg, P. Jarillo-Herrero, and N. Gedik. Observation of Floquet-Bloch States on the Surface of a Topological Insulator. *Science*, 342(6157) :453–457, 2013. doi: [10.1126/science.1239834](https://doi.org/10.1126/science.1239834).
- [61] F. Mahmood, C. K. Chan, Z. Alpichshev, D. Gardner, Y. Lee, P. A. Lee, and N. Gedik. Selective scattering between Floquet-Bloch and Volkov states in a topological insulator. *Nat Phys*, 12(4) :306–310, 2016. doi: [10.1038/nphys3609](https://doi.org/10.1038/nphys3609).
- [62] J. Liu and D. Vanderbilt. Weyl semimetals from noncentrosymmetric topological insulators. *Phys. Rev. B*, 90(15) :155316, 2014. doi: [10.1103/PhysRevB.90.155316](https://doi.org/10.1103/PhysRevB.90.155316).
- [63] H. Zhang, C. X. Liu, X. L. Qi, X. Dai, Z. Fang, and S. C. Zhang. Topological insulators in Bi₂Se₃, Bi₂Te₃ and Sb₂Te₃ with a single Dirac cone on the surface. *Nature Physics*, 5(6) :438–442, 2009. doi: [10.1038/nphys1270](https://doi.org/10.1038/nphys1270).
- [64] M. Bahramy, B. Yang, R. Arita, and N. Nagaosa. Emergence of non-centrosymmetric topological insulating phase in BiTeI under pressure. *Nat. Commun.*, 3 :679, 2012. doi: [10.1038/ncomms1679](https://doi.org/10.1038/ncomms1679).
- [65] G. Chang, B. Singh, S. Y. Xu, G. Bian, S. M. Huang, C. H. Hsu, I. Belopolski, N. Alidoust, D. S. Sanchez, H. Zheng, H. Lu, X. Zhang, Y. Bian, T. R. Chang, H. T. Jeng, A. Bansil, H. Hsu, S. Jia, T. Neupert, H. Lin, and M. Z. Hasan. Magnetic and noncentrosymmetric weyl fermion semimetals in the RAlGe family of compounds ($R =$ rare earth). *Phys. Rev. B*, 97(4) :041104, 2018. doi: [10.1103/PhysRevB.97.041104](https://doi.org/10.1103/PhysRevB.97.041104).
- [66] B. Bian, Guang, T. R. Chang, R. Sankar, S. Y. Xu, H. Zheng, T. Neupert, C. K. Chiu, S. M. Huang, G. Chang, I. Belopolski, D. S. Sanchez, M. Neupane, N. Alidoust, C. Liu, B. Wang, and M. Z. Hasan. Topological nodal-line fermions in spin-orbit metal PbTaSe₂. *Nat Commun*, 7(10556), 2016. doi: [10.1038/ncomms10556](https://doi.org/10.1038/ncomms10556).
- [67] S. Liu, Y. Kim, L. Z. Tan, and A. M. Rappe. Strain-Induced Ferroelectric Topological Insulator. *Nano Letters*, 16(3) :1663–1668, 2016. doi: [10.1021/acs.nanolett.5b04545](https://doi.org/10.1021/acs.nanolett.5b04545).
- [68] Z. F. Wang, N. Su, and F. Liu. Prediction of a Two-Dimensional Organic Topological Insulator. *Nano Letters*, 13(6) :2842–2845, 2013. doi: [10.1021/nl401147u](https://doi.org/10.1021/nl401147u).
- [69] Y. Ma, L. Kou, Y. Dai, and T. Heine. Proposed two-dimensional topological insulator in SiTe. *Phys. Rev. B*, 94(20) :201104, 2016. doi: [10.1103/PhysRevB.94.201104](https://doi.org/10.1103/PhysRevB.94.201104).
- [70] S. Chowdhury, K. F. Garrity, and F. Tavazza. Prediction of Weyl semimetal and anti-ferromagnetic topological insulator phases in Bi₂MnSe₄. *npj Computational Materials*, 5(1) :33, 2019. doi: [10.1038/s41524-019-0168-1](https://doi.org/10.1038/s41524-019-0168-1).
- [71] I. Garate. Phonon-Induced Topological Transitions and Crossovers in Dirac Materials. *Phys. Rev. Lett.*, 110(4) :046402, 2013. doi: [10.1103/PhysRevLett.110.046402](https://doi.org/10.1103/PhysRevLett.110.046402).

- [72] K. Saha and I. Garate. Phonon-induced topological insulation. *Phys. Rev. B*, 89(20) : 205103, 2014. doi: [10.1103/PhysRevB.89.205103](https://doi.org/10.1103/PhysRevB.89.205103).
- [73] G. Antonius and S. G. Louie. Temperature-Induced Topological Phase Transitions : Promoted versus Suppressed Nontrivial Topology. *Phys. Rev. Lett.*, 117(24) :246401, 2016. doi: [10.1103/PhysRevLett.117.246401](https://doi.org/10.1103/PhysRevLett.117.246401).
- [74] B. Monserrat and D. Vanderbilt. Temperature Effects in the Band Structure of Topological Insulators. *Phys. Rev. Lett.*, 117(22) :226801, 2016. doi: [10.1103/PhysRevLett.117.226801](https://doi.org/10.1103/PhysRevLett.117.226801).
- [75] L. L. Wang, N. H. Jo, Y. Wu, Q. Wu, A. Kaminski, P. C. Canfield, and D. D. Johnson. Phonon-induced topological transition to a type-II Weyl semimetal. *Phys. Rev. B*, 95(16) :165114, 2017. doi: [10.1103/PhysRevB.95.165114](https://doi.org/10.1103/PhysRevB.95.165114).
- [76] T. S. Nam, J. Kim, C. J. Kang, K. Kim, and B. I. Min. Temperature-dependent electronic structure and topological property of the Kondo semimetal CeFe₂Al₁₀. *Phys. Rev. B*, 103(4) :045101, 2021. doi: [10.1103/PhysRevB.103.045101](https://doi.org/10.1103/PhysRevB.103.045101).
- [77] B. Peng, I. Bravić, J. L. MacManus-Driscoll, and B. Monserrat. Topological semimetallic phase in PbO₂ promoted by temperature. *Phys. Rev. B*, 100(16) :161101, 2019. doi: [10.1103/PhysRevB.100.161101](https://doi.org/10.1103/PhysRevB.100.161101).
- [78] P. Zhang, H. Yuan, and C. Cao. Electron-phonon coupling and nontrivial band topology in noncentrosymmetric superconductors LaNiSi, LaPtSi, and LaPtGe. *Phys. Rev. B*, 101(24) :245145, 2020. doi: [10.1103/PhysRevB.101.245145](https://doi.org/10.1103/PhysRevB.101.245145).
- [79] J. D. Querales-Flores, P. Aguado-Puente, D. Dangić, J. Cao, P. Chudzinski, T. N. Todorov, M. Grüning, S. Fahy, and I. Savić. Towards temperature-induced topological phase transition in SnTe : A first-principles study. *Phys. Rev. B*, 101(23) :235206, 2020. doi: [10.1103/PhysRevB.101.235206](https://doi.org/10.1103/PhysRevB.101.235206).
- [80] R. Heid, I. Sklyadneva, and E. Chulkov. Electron-phonon coupling in topological surface states : The role of polar optical modes. *Scientific Reports*, 7(1) :1095, 2017. doi: [10.1038/s41598-017-01128-1](https://doi.org/10.1038/s41598-017-01128-1).
- [81] B. Monserrat and A. Narayan. Unraveling the topology of ZrTe₅ by changing temperature. *Phys. Rev. Research*, 1(3) :033181, 2019. doi: [10.1103/PhysRevResearch.1.033181](https://doi.org/10.1103/PhysRevResearch.1.033181).
- [82] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12) :121102(R), 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).
- [83] S. Wiedmann, A. Jost, C. Thienel, C. Brüne, P. Leubner, H. Buhmann, L. W. Molenkamp, J. C. Maan, and U. Zeitler. Temperature-driven transition from a semiconductor to a topological insulator. *Phys. Rev. B*, 91(20) :205311, 2015. doi: [10.1103/PhysRevB.91.205311](https://doi.org/10.1103/PhysRevB.91.205311).
- [84] A. M. Kadykov, S. S. Krishtopenko, B. Jouault, W. Desrat, W. Knap, S. Ruffenach, C. Consejo, J. Torres, S. V. Morozov, N. N. Mikhailov, S. A. Dvoretiskii, and F. Teppe. Temperature-induced topological phase transition in HgTe quantum wells. *Phys. Rev.*

- Lett.*, 120(8) :086401, 2018. doi: [10.1103/PhysRevLett.120.086401](https://doi.org/10.1103/PhysRevLett.120.086401).
- [85] B. Xu, L. X. Zhao, P. Marsik, E. Sheveleva, F. Lyzwa, Y. M. Dai, G. F. Chen, X. G. Qiu, and C. Bernhard. Temperature-driven topological phase transition and intermediate dirac semimetal phase in ZrTe_5 . *Phys. Rev. Lett.*, 121(18) :187401, 2018. doi: [10.1103/PhysRevLett.121.187401](https://doi.org/10.1103/PhysRevLett.121.187401).
- [86] V. Brousseau-Couture and M. Côté. Influence of the Electron–Phonon Interaction on the Topological Phase Transition in BiTeI . In M. B. Paranjape, R. MacKenzie, Z. Thomova, P. Winternitz, and W. Witczak-Krempa, editors, *Quantum Theory and Symmetries*, CRM Series in Mathematical Physics, pages 305–313, Cham, 2021. Springer International Publishing. doi: [10.1007/978-3-030-55777-5_29](https://doi.org/10.1007/978-3-030-55777-5_29).
- [87] T. Imai, J. Chen, K. Kato, K. Kuroda, T. Matsuda, A. Kimura, K. Miyamoto, S. V. Ereameev, and T. Okuda. Experimental verification of a temperature-induced topological phase transition in TlBiS_2 and TlBiSe_2 . *Phys. Rev. B*, 102(12) :125151, 2020. doi: [10.1103/PhysRevB.102.125151](https://doi.org/10.1103/PhysRevB.102.125151).

Chapitre 8

Dépendance en température de la transition de phase topologique dans le BiTeI

8.0.1. Mise en contexte

Cet article se veut une application concrète des outils permettant d'évaluer la dépendance en température des énergies électroniques pour analyser l'effet de la température sur la topologie des structures de bande. Les méthodes discutées dans les chapitres précédents y sont utilisées pour construire le diagramme de phase topologique en pression et température du BiTeI. Semi-conducteur standard dans les conditions ambiantes, ce matériau devient un isolant topologique \mathbb{Z}_2 avec l'application d'une pression hydrostatique.

Ce projet a germé suite à la lecture des travaux pionniers de Garate *et al.* [1, 2] démontrant théoriquement la possibilité d'une transition de phase topologique induite par la température. Ces modèles, tout comme les premières études numériques explorant cette question [3, 4], considéraient des matériaux centrosymétriques ayant une bande d'énergie interdite directe localisée au centre de la zone de Brillouin. Dans ce cas, la présence de la symétrie d'inversion simplifie l'interprétation des résultats, puisque l'effet de l'EPI sur la phase topologique peut être expliqué en termes de la parité des modes de phonons et des états électroniques. En l'absence de cette symétrie, la parité n'est plus un bon nombre quantique ; la situation devient plus complexe à analyser.

Par ailleurs, l'absence de symétrie d'inversion couplée à une asymétrie du potentiel dans une direction axiale complexifie la dispersion électronique du BiTeI en induisant une importante séparation Rashba. Ce faisant, la localisation des extrema de bande qui déterminent l'énergie de la bande interdite indirecte change tant avec la pression qu'avec la température. La construction du diagramme de phase topologique nécessitera donc implicitement de capturer l'évolution de la séparation Rashba.

Enfin, nous avons vu au chapitre précédent que la transition d'un isolant trivial vers une phase d'isolant topologique \mathbb{Z}_2 dans un matériau non centrosymétrique est caractérisée

par l'existence d'une phase intermédiaire de type semi-métal de Weyl. Le BiTeI possède donc deux pressions critiques pouvant évoluer différemment en fonction de la température. Bien que prédite par la théorie, la phase WSM du BiTeI n'a toujours pas été observée expérimentalement ; d'une part, la plage de pression prédite pour cette phase est très faible (~ 0.2 GPa), et d'autre, la contrainte de pression hydrostatique limite le type de mesure qui peut être effectué¹. De plus, la valeur de la pression critique associée à l'apparition de la phase topologique n'est pas uniforme dans la littérature expérimentale. Puisque les différentes expériences rapportées ont été effectuées dans une plage de température allant de quelques kelvins à la température de la pièce, on peut se demander si la plage de pression critique observée expérimentalement peut être corrélée avec la dépendance en température de la structure électronique. L'objectif de ce projet était donc, d'une part, d'explorer comment la température affecte la pression critique à partir de laquelle les signatures de la phase topologique deviennent détectables, et d'autre, de déterminer si une augmentation de la température faciliterait la détection expérimentale de la phase WSM. À notre connaissance, il s'agit de la première étude basée sur les principes premiers qui a produit un diagramme de phase topologique incluant la dépendance en température pour un isolant topologique \mathbb{Z}_2 non centrosymétrique.

Rappelons finalement que, tel que mentionné dans l'**avant-propos**, cet article a été écrit avant les trois articles déjà discutés. À la lumière des chapitres précédents, l'analyse de ce matériau pourrait être améliorée en prenant en compte la correction du point zéro à la dilatation thermique ainsi qu'avec un traitement plus approfondi des effets non adiabatiques sur l'interaction électron-phonon. Cependant, compte tenu des résultats présentés précédemment, l'ajout de ces effets n'affectera pas les conclusions générales de l'article.

8.0.2. Contributions des auteurs

J'ai proposé la ligne directrice de ce projet au début de mon doctorat, en plus d'effectuer l'ensemble des calculs et analyses présentées dans cet article et d'en assurer la rédaction. Gabriel Antonius a fourni un soutien théorique et m'a guidée dans l'utilisation d'ABINIT et de l'outil de post-traitement `ElectronPhononCoupling` pour l'évaluation de la renormalisation des énergies électroniques induite par l'interaction électron-phonon. Michel Côté a supervisé le projet. Tous les auteurs ont contribué à la révision du manuscrit.

1. Les WSM sont généralement confirmés par l'observation d'arcs de Fermi dans le spectre ARPES des états de surface du matériau. Au meilleur de notre connaissance, ce type de mesure n'est pas compatible avec l'application une pression hydrostatique isotrope.

Quatrième article.

Temperature dependence of the topological phase transition of BiTeI from first principles

par

Véronique Brousseau-Couture¹, Gabriel Antonius² et Michel Côté³

- (¹) Département de physique and Regroupement Québécois sur les Matériaux de Pointe, Université de Montréal, Montréal, Québec, Canada
- (²) Département de chimie, biochimie et physique, Institut de recherche sur l'hydrogène, Université du Québec à Trois-Rivières, Trois-Rivières, Québec, Canada
- (³) Département de physique and Regroupement Québécois sur les Matériaux de Pointe, Université de Montréal, Montréal, Québec, Canada

Cet article a été publié dans Physical Review Research.

RÉSUMÉ. L'apparition d'une transition de phase topologique d'un isolant trivial à un isolant topologique de type \mathbb{Z}_2 nécessite que l'énergie de la bande interdite (*gap*) devienne nulle au point de transition. Pour les matériaux sans symétrie d'inversion, ces deux phases sont séparées par une phase sans bande interdite appelée semi-métal de Weyl. À température finie, l'énergie de la bande interdite est toutefois affectée par le mouvement des atomes, à travers l'interaction électron-phonon, ainsi que par la dilatation thermique du réseau cristallin. Conséquemment, l'espace des phases topologiquement non-triviales est modifié par la température. Dans cet article, la dépendance en pression et en température de l'énergie de la bande interdite indirecte du BiTeI est étudiée à l'aide de méthodes basées sur les principes premiers. Nous évaluons tant la contribution de l'interaction électron-phonon (EPI) que celle de la dilatation thermique, et démontrons que leurs effets combinés repoussent la transition de phase topologique à des pressions plus élevées lorsque la température augmente. Dans le cas de l'EPI, nous trouvons notamment que la sensibilité des extrema de bande à la pression et à la topologie varie significativement selon la composante dominante de leur décomposition en terme d'orbitales atomiques. Nos résultats indiquent que la plage de pression associée à la phase de semi-métal de Weyl augmente avec la température, ayant pratiquement doublé dès 100 K par rapport au résultat obtenu avec un réseau statique. Nos observations fournissent ainsi une ligne directrice pour la détection expérimentale des phases non-triviales du BiTeI et démontrent comment l'espace de phase associé à la phase semi-métal de Weyl dans les matériaux sans symétrie d'inversion peut être affecté de manière significative par la température.

Mots clés : semi-métaux de Weyl, isolants topologiques, transitions de phase topologiques, théorie de la fonctionnelle de la densité perturbative, structure électronique

ABSTRACT. A topological phase transition from a trivial insulator to a \mathbb{Z}_2 topological insulator requires the bulk band gap to vanish. In the case of noncentrosymmetric materials, these phases are separated by a gapless Weyl semimetal phase. However, at finite temperature, the gap is affected by atomic motion, through electron-phonon interaction, and by thermal expansion of the lattice. As a consequence, the phase space of topologically nontrivial phases is affected by temperature. In this paper, the pressure and temperature dependence of the indirect band gap of BiTeI is investigated from first principles. We evaluate the contribution from both electron-phonon interaction and thermal expansion, and show that their combined effect drives the topological phase transition towards higher pressures with increasing temperature. Notably, we find that the sensitivity of both band extrema to pressure and topology for electron-phonon interaction differs significantly according to their leading orbital character. Our results indicate that the Weyl semimetal phase width is increased by temperature, having almost doubled by 100 K when compared to the static lattice results. Our findings thus provide a guideline for experimental detection of the nontrivial phases of BiTeI and illustrate how the phase space of the Weyl semimetal phase in noncentrosymmetric materials can be significantly affected by temperature.

Keywords: Weyl semimetals, topological insulators, topological phase transitions, density-functional perturbation theory, electronic structure

8.1. Introduction

Topological phases of matter have become a thriving field of condensed matter physics, for both fundamental and applied research [5]. In three-dimensional (3D) materials, these phases are characterized by the existence of metallic surface states with peculiar properties, such as protection against backscattering, spin-momentum locking and dissipationless spin-polarized currents, as well as inverted bulk band gaps [5, 6]. The discovery of experimentally tunable topological phases, whether through stoichiometric doping [7–10], hydrostatic pressure [11, 12], strain [13, 14], external electric fields [15, 16] or interaction with light [17, 18], has led to a continually growing number of proposals for promising and innovative applications relying on the refined engineering of these robust states and their associated phase transitions [19–22].

Another widely studied class of materials is the bulk Rashba semiconductors, in which a strong spin-orbit interaction combined with the absence of inversion symmetry leads to a splitting of electronic bands of opposite spin polarization [23, 24]. The band extrema are shifted away from the time-reversal invariant points in the Brillouin zone, both in energy and momentum, in the plane perpendicular to the potential gradient [25]. This Rashba effect gives rise to many quantum phenomena, such as the Edelstein, spin Hall and spin galvanic effects, as well as noncentrosymmetric superconductivity [see, for example, Bahramy and Ogawa [25] and references therein]. Just like the topological surface states, the bulk Rashba split bands have their spin orientation locked perpendicular to their momentum, leading to a helical spin texture [26]. This feature makes these materials promising candidates for realizing devices involving the active manipulation of the spin degree of freedom, for both spintronics [27] and quantum computing [28].

Because it exhibits both of these phenomena, there is no wonder that BiTeI has attracted such a wide interest for the last decade, both from the experimental and first-principles communities. Besides displaying one of the largest Rashba splittings known so far [29–34], it was predicted to turn into a strong \mathbb{Z}_2 topological insulator under hydrostatic pressure [35]. Following this prediction, BiTeI has been widely investigated through optical [11, 36, 37], electrical transport [38, 39] and Shubnikov-de Haas oscillations experiments [26, 40–43]. Many experiments revealed signatures supporting the existence of the topological phase transition (TPT) [11, 38, 39, 42, 44].

While at first it was thought that the TPT of BiTeI occurred at a single critical pressure P_C , it was later demonstrated that the lack of inversion symmetry imposed the existence of an intermediate Weyl semimetal (WSM) phase [45] between the trivial band insulator and topological insulator phases, yielding *two* critical pressures. Tight-binding [45] and first-principles calculations [46, 47] predicted that this WSM phase could exist only within a narrow pressure range of 0.1–0.2 GPa, making its experimental detection technically challenging.

The precise value of the critical pressure of BiTeI is still elusive to this day. Experiments have located it between 2.0 and 3.5 GPa [11, 38–40, 42], while first-principles calculations have predicted it in a slightly wider range of pressures, from 1.6 to 4.5 GPa [35, 46–48]. From a theoretical point of view, predicting the critical pressure comes down to finding the gap closing pressure, which is inherently dependent on the accuracy of the calculated band gap at ambient pressure, which can be biased by the well-established underestimation of the band gap by density functional theory (DFT) [49].

On the other hand, one should not forget that most first-principles calculations are done under the assumption of a static lattice, while experiments are inherently done at finite temperature. This completely overlooks the fact that electrons can interact with thermally activated phonons, resulting in a temperature-dependent shift in the electronic eigenenergies. Moreover, at $T = 0$ K, it neglects the contribution of the zero-point renormalization (ZPR) to the eigenenergies, due to the zero-point motion of the ions. For narrow gap materials, this renormalization could, in principle, push the system towards a band inversion. In the case of BiTeI, this shift will directly affect *both* critical pressures at which the band gap will effectively close or reopen at a given temperature.

The ability of electron-phonon interaction (EPI) to induce a change in the bulk topology was first demonstrated with model Hamiltonians [1, 2, 50], and later analyzed by first-principles calculations for BiTl(S_{1-x}Se_x)₂ [3] and for the Bi₂Se₃ family [4]. These studies were able to capture the whole complexity of the EPI, with fewer approximations. They demonstrated that, depending on the strength and sign of the different types of couplings, nontrivial topology could either be promoted or suppressed by temperature in real materials. The suppression of the topological surface state signatures at sufficiently high temperatures was observed experimentally in Pb_{1-x}Sn_xSe [9, 51, 52]. The temperature-dependent band structure of BiTeI was previously investigated by focusing on the variation of the Rashba splitting rather than on tracking the TPT [53].

In this paper, we study the temperature dependence of the topological phase transition in BiTeI using first-principles methods, assessing for both the EPI and thermal expansion (TE) contributions. We explicitly assess the temperature dependence of the band-gap renormalization as a function of pressure, both in the trivial and the topological phases. We observe that the EPI contribution changes sign as the system undergoes the TPT, and link this behavior to the band inversion phenomenon, showing that the band extrema exhibit a distinct pressure dependence related to their leading orbital character. In turn, the TE contribution is unaffected by the leading orbital character of the band extrema, but manifests sensitivity to the topological invariant by changing sign as the system undergoes the TPT. We finally evaluate how the pressure width of the WSM phase is affected by temperature by extrapolating the total renormalization trends of each band towards the TPT, and find

that the stronger renormalization of the band gap in the TI phase widens the WSM with increasing temperature.

The remainder of this paper is organized as follows. After presenting the theoretical formalism used to obtain the EPI and TE contributions to the electronic structure in Sec. 8.2, we summarize the computational details in Sec. 8.3. Section 8.4.1 presents our analysis of the TPT and WSM phase for a static lattice at $T = 0$ K, while Sec. 8.4.2 and 8.4.3 focus on the effects of temperature on both the band gap and the TPT. We finally discuss the implications and limitations of our results in Sec. 8.4.4 and summarize our findings.

8.2. Methodology

The temperature dependence of the band-gap energy, E_g , for a given constant pressure, can be approximated at first order by the sum of two distinct contributions [54], namely the thermal expansion, ΔE_g^{TE} , and the renormalization of the electronic eigenenergies by electron-phonon interactions at constant volume, ΔE_g^{EPI} :

$$\Delta E_g \simeq \Delta E_g^{\text{EPI}} + \Delta E_g^{\text{TE}}. \quad (8.2.1)$$

8.2.1. Electron-phonon interaction

In many-body perturbation theory [55], the temperature dependence of the electronic eigenenergies, $\varepsilon_{\mathbf{k}n}$, induced by electron-phonon interaction can be captured, under certain assumptions [56], by the real part of the electron-phonon self-energy, Σ^{EPI} , evaluated at the quasiparticle energy:

$$\varepsilon_{\mathbf{k}n}(T) = \varepsilon_{\mathbf{k}n}^0 + \Re \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}(T)) \right], \quad (8.2.2)$$

where $\varepsilon_{\mathbf{k}n}^0$ is the unperturbed eigenvalue of a Kohn-Sham (KS) eigenstate with wavevector \mathbf{k} and band index n , computed at the fixed, relaxed geometry. The Hartree atomic unit system is used throughout this paper ($\hbar = m_e = e = 1$).

In the nonadiabatic Allen-Heine-Cardona (AHC) framework [54, 57, 58], the self-energy at the second-order of perturbation is the sum of the Fan and Debye-Waller (DW) contributions [Fig. 8.1]:

$$\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega, T) = \Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega, T) + \Sigma_{\mathbf{k}n}^{\text{DW}}(T), \quad (8.2.3)$$

where the former captures a frequency-dependent interaction with two first-order EPI vertices, and the latter captures a static one with a single second-order vertex. Evaluating the previous equation at absolute zero temperature yields the ZPR, which arises from the zero-point motion of the ions. For a detailed review of the AHC methodology, we refer our readers to the works of Ponc e *et al.* [59, 60] and to the extensive review by Giustino [61].

Assuming that the quasiparticle energy is close to the unperturbed electronic energy, one can apply the on-the-mass-shell approximation [62], in which the Fan self-energy is evaluated

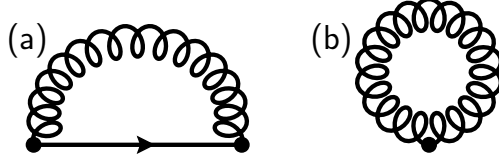


Figure 8.1 – The Fan (a) and Debye-Waller (b) diagrams describing the lowest-order EPI in the AHC framework.

at the poles of the Green's function, namely at the frequency ω corresponding to the bare eigenvalue $\varepsilon_{\mathbf{k}n}^0$, yielding

$$\varepsilon_{\mathbf{k}n}(T) \approx \varepsilon_{\mathbf{k}n}^0 + \Re \varepsilon \left[\Sigma_{\mathbf{k}n}^{\text{EPI}}(\omega = \varepsilon_{\mathbf{k}n}^0, T) \right]. \quad (8.2.4)$$

Furthermore approximating the fully interacting electronic wavefunction by the noninteracting KS-DFT wavefunction, one obtains the standard result for the retarded Fan self-energy [59, 61]:

$$\begin{aligned} \Sigma_{\mathbf{k}n}^{\text{Fan}}(\omega, T) &= \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \sum_{n'} \frac{1}{2\omega_{\mathbf{q}\nu}} |\langle \mathbf{k} + \mathbf{q}n' | V_{\mathbf{q}\nu}^{(1)} | \mathbf{k}n \rangle|^2 \\ &\times \left[\frac{n_{\mathbf{q}\nu}(T) + 1 - f_{\mathbf{k}+\mathbf{q}n'}(T)}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 - \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} + \frac{n_{\mathbf{q}\nu}(T) + f_{\mathbf{k}+\mathbf{q}n'}(T)}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 + \omega_{\mathbf{q}\nu} + i\eta_{\mathbf{k}}} \right], \end{aligned} \quad (8.2.5)$$

in which the contributions from all phonon modes with frequency $\omega_{\mathbf{q}\nu}$ are summed for all wavevectors \mathbf{q} and branches ν in the Brillouin zone (BZ). The whole temperature dependence of this expression is captured by the bosonic and fermionic occupation factors, respectively $n_{\mathbf{q}\nu}$ and $f_{\mathbf{k}+\mathbf{q}n'}$. The parameter $\eta_{\mathbf{k}} = \eta \text{sgn}(\varepsilon_{\mathbf{k}n}^0 - \mu)$, where η is real and positive, preserves causality by correctly shifting the poles of the Green's function in the complex plane.

Equation (8.2.5) implies the limit of an infinite number of phonon wavevectors ($N_{\mathbf{q}} \rightarrow \infty$), leading to a vanishingly small value of η . $V_{\mathbf{q}\nu}^{(1)}$ is the first-order self-consistent change of the local KS potential induced by the collective atomic displacement $\mathbf{R}_{\kappa\alpha}^{\nu}(\mathbf{q})$ along a given phonon mode [56, 59, 61]:

$$V_{\mathbf{q}\nu}^{(1)} = \sum_{\kappa\alpha} \frac{\partial V^{\text{KS}}}{\partial \mathbf{R}_{\kappa\alpha}^{\nu}(\mathbf{q})} = \sum_{\kappa,\alpha} V_{\kappa\alpha}^{(1)}(\mathbf{q}\nu), \quad (8.2.6)$$

with

$$\frac{\partial}{\partial \mathbf{R}_{\kappa\alpha}^{\nu}(\mathbf{q})} = U_{\nu,\kappa\alpha}(\mathbf{q}) \sum_l e^{i\mathbf{q}\cdot\mathbf{R}_l} \frac{\partial}{\partial \mathbf{R}_{l\kappa\alpha}}. \quad (8.2.7)$$

In these expressions, $\mathbf{R}_{l\kappa\alpha}$ denotes the displacement of atom κ , located in unit cell l with lattice vector \mathbf{R}_l , in cartesian direction α ; $\mathbf{R}_{\kappa\alpha}^{\nu}(\mathbf{q})$ therefore describes the collective atomic displacement of along the $\mathbf{q}\nu$ -phonon mode. $V_{\kappa\alpha}^{(1)}(\mathbf{q}\nu)$ refers to the contribution of a displacement of atom κ in direction α to the full first-order potential $V_{\mathbf{q}\nu}^{(1)}$. $U_{\nu,\kappa\alpha}(\mathbf{q})$ is the $\mathbf{q}\nu$ -phonon

displacement vector, which is related to the phonon eigenvector, $\xi_{\kappa\alpha}^\nu(\mathbf{q})$, through:

$$\xi_{\kappa\alpha}^\nu(\mathbf{q}) = \sqrt{M_\kappa} U_{\nu,\kappa\alpha}(\mathbf{q}), \quad (8.2.8)$$

with M_κ being the mass of atom κ .

The static Debye-Waller self-energy is defined as the second derivative of the potential with respect to two atomic displacements, evaluated at the first order in perturbation theory [56] :

$$\Sigma_{\mathbf{kn}}^{\text{DW}}(T) = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \frac{1}{2\omega_{\mathbf{q}\nu}} \langle \mathbf{kn} | V_{\mathbf{q}\nu}^{(2)} | \mathbf{kn} \rangle [2n_{\mathbf{q}\nu}(T) + 1], \quad (8.2.9)$$

where the second-order perturbation potential is

$$V_{\mathbf{q}\nu}^{(2)} = \frac{1}{2} \sum_{\substack{\kappa\alpha \\ \kappa'\alpha'}} \frac{\partial^2 V^{\text{KS}}}{\partial \mathbf{R}_{\kappa'\alpha'}^\nu(-\mathbf{q}) \partial \mathbf{R}_{\kappa\alpha}^\nu(\mathbf{q})}, \quad (8.2.10)$$

with the derivatives defined as in Eq. (8.2.7). Only the phonon occupation factor contributes to the temperature dependence of this term since it does not involve any intermediate electronic state [as can be seen from Fig. 8.1(b)]. The constant term inside the brackets accounts for the ZPR contribution.

The numerical evaluation of this second-order derivative is a computational bottleneck of density-functional perturbation theory (DFPT) [60]. It can, however, be circumvented by applying the rigid-ion approximation (RIA), which supposes that the potentials created by each nucleus are independent of each other. Within this approximation, thanks to the translational invariance of the lattice, $V_{\mathbf{q}\nu}^{(2)}$ can be expressed in terms of the first-order derivatives entering the Fan self-energy [59] [see Eq. (8.2.6)]. The consequences of this approximation on the ZPR of diamond have been discussed by Poncé *et al.* [59]. In this framework, the Debye-Waller self-energy becomes [3, 59]:

$$\begin{aligned} \Sigma_{\mathbf{kn}}^{\text{DW,RIA}} &= \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \sum_{n'}^{\text{BZ}} -\frac{1}{2\omega_{\mathbf{q}\nu}} \frac{|g_{\mathbf{knn}'}^{\text{DW}}(\mathbf{q}\nu)|^2}{\varepsilon_{\mathbf{kn}}^0 - \varepsilon_{\mathbf{kn}'}^0 + i\eta} \\ &\times \left[n_{\mathbf{q}\nu}(T) + \frac{1}{2} \right], \end{aligned} \quad (8.2.11)$$

with

$$\begin{aligned} |g_{\mathbf{knn}'}^{\text{DW}}(\mathbf{q}\nu)|^2 &= \\ &\sum_{\substack{\kappa\alpha \\ \kappa'\alpha'}} \left[U_{\nu,\kappa\alpha}(\mathbf{q}) U_{\nu,\kappa\alpha'}^*(\mathbf{q}) + U_{\nu,\kappa'\alpha}(\mathbf{q}) U_{\nu,\kappa\alpha'}^*(\mathbf{q}) \right] \\ &\times \langle \mathbf{kn} | V_{\kappa\alpha}^{\text{DW}*} | \mathbf{k} + \mathbf{qn}' \rangle \langle \mathbf{k} + \mathbf{qn}' | V_{\kappa'\alpha'}^{\text{DW}} | \mathbf{kn} \rangle, \end{aligned} \quad (8.2.12)$$

and

$$V_{\kappa\alpha}^{\text{DW}} = V_{\kappa\alpha}^{(1)}(\mathbf{q} = 0, \nu). \quad (8.2.13)$$

We finally evaluate the sum on band index n' in Eq. (8.2.5) and (8.2.11) using the semi-static approach described in Poncé *et al.* [59] and Antonius *et al.* [56]. We compute the full, nonadiabatic contribution for all bands up to a band index M , which we choose to be well-separated in energy (i.e., more than 10 eV) from the first conduction band. For the high-energy bands with band index $n' > M$, the phonon frequencies in the denominators of Eq. (8.2.5) and (8.2.11) are negligible with respect to the electronic eigenenergy difference. These bands can, therefore, be treated within the adiabatic approximation, neglecting the phonon frequencies. Furthermore, this sum over an arbitrarily large number of empty states can be replaced by the solution of a linear Sternheimer equation for the subspace orthonormal to the low-energy states ($n' \leq M$), thus significantly reducing the numerical cost of the calculation [63]. We estimate the relative error on the renormalization induced by this semi-static treatment of the high energy bands to be 1–2% at most.

8.2.2. Thermal expansion

The thermal expansion contribution was evaluated through the quasi-harmonic approximation (QHA) [64, 65]. Within this framework, the only departure from harmonicity occurs through the explicit volume dependence acquired by the phonon frequencies. It should be understood that the purpose of the QHA is to deliver the leading order expression of the thermal expansion coefficient α , rather than capturing the full anharmonic effects present in the crystal [66], as would do nonperturbative methods [56].

The thermal expansion coefficient, α , is a rank 2 tensor that relates a small temperature increment ΔT to the strain, ϵ , it induces on the lattice:

$$\epsilon_{ij} = \alpha_{ij} \Delta T, \quad (8.2.14)$$

where $i, j = 1, 2, 3$ are cartesian directions. Hence, the most general definition of α_{ij} is

$$\alpha_{ij} = \left(\frac{\partial \epsilon_{ij}}{\partial T} \right)_{\sigma}, \quad (8.2.15)$$

where the derivative is evaluated at constant stress σ , which is usually taken to be a constant pressure P . The diagonal components of the strain tensor describe the relative change in the length of the lattice parameters: $\epsilon_{ii} = \Delta a_i / a_i$. The volumic expansion coefficient β can therefore be obtained by taking the trace of α :

$$\beta = \frac{\partial}{\partial T} \left(\frac{\Delta V}{V} \right)_P = \sum_{i=1}^3 \alpha_{ii}. \quad (8.2.16)$$

In the most simple case of cubic symmetry, all α_{ii} are equal and can be expressed by [64]

$$\alpha(T) = \frac{\beta}{3} = \frac{\kappa_0}{3V_0 N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V c_{\mathbf{q}\nu}(T), \quad (8.2.17)$$

where κ_0 is the bulk compressibility at equilibrium volume, V_0 is the primitive cell equilibrium volume, $N_{\mathbf{q}}$ is the number of phonons in the homogenous grid used to sample the BZ and $c_{\mathbf{q}\nu}(T)$ is the specific heat of the $\mathbf{q}\nu$ -phonon mode at temperature T . We have also introduced the volumic mode Grüneisen parameters, defined as [64]

$$\gamma_{\mathbf{q}\nu}^V = -\frac{\partial \ln \omega_{\mathbf{q}\nu}(V)}{\partial \ln V}. \quad (8.2.18)$$

The bulk Grüneisen parameter, often referred to in the literature, is defined as

$$\gamma^V = \frac{\sum_{\mathbf{q}\nu} \gamma_{\mathbf{q}\nu}^V c_{\mathbf{q}\nu}}{\sum_{\mathbf{q}\nu} c_{\mathbf{q}\nu}}. \quad (8.2.19)$$

Thus, for cubic systems, the linear thermal expansion coefficient, $\alpha(T)$, Eq. (8.2.17), is simply proportional to two scalar parameters, γ^V and κ_0 , the former governing the sign of the thermal expansion.

For materials belonging to noncubic symmetry groups, one must consider the most general case, in which the mode Grüneisen parameters take a tensor form and capture the variation of the phonon frequencies with respect to a given strain:

$$\gamma_{\mathbf{q}\nu}^{ij} = -\frac{\partial \ln \omega_{\mathbf{q}\nu}}{\partial \epsilon_{ij}}. \quad (8.2.20)$$

One can also define bulk Grüneisen parameters γ^{ij} following the same procedure as in Eq. (8.2.19).

For axial crystals, which include the case of BiTeI's trigonal symmetry, the thermal expansion coefficient tensor has two distinct nonvanishing components: $\alpha_{11} = \alpha_{22}$ and α_{33} . Because of this anisotropy, the resulting linear thermal expansion coefficients along crystallographic directions \hat{a} and \hat{c} capture a subtle interplay between the vibrational and elastic properties of the material [64, 67, 68]:

$$\begin{aligned} \alpha_a = \alpha_{11} &= \frac{C_\sigma}{V_0} [(s_{11} + s_{12})\gamma^{a_1} + s_{13}\gamma^c], \\ \alpha_c = \alpha_{33} &= \frac{C_\sigma}{V_0} [2s_{13}\gamma^{a_1} + s_{33}\gamma^c], \end{aligned} \quad (8.2.21)$$

where C_σ is the heat capacity at constant stress and s_{ij} is the ij -coefficient of the elastic compliance tensor. In these expressions, the subscripts a and c refer to the length of the unit cell along (\hat{a}_1, \hat{a}_2) and \hat{c} . γ^{a_1} formally refers to a strain applied in only *one* direction perpendicular in the crystallographic axis.

From Eq. (8.2.20), the directional mode Grüneisen parameters $\gamma_{\mathbf{q}\nu}^{a1}$ and $\gamma_{\mathbf{q}\nu}^c$ take the form

$$\begin{aligned}\gamma_{\mathbf{q}\nu}^{a1} = \gamma_{\mathbf{q}\nu}^{11} &= -\frac{1}{2} \left(\frac{\partial \ln \omega_{\mathbf{q}\nu}}{\partial \ln a} \right)_c, \\ \gamma_{\mathbf{q}\nu}^c &= \gamma_{\mathbf{q}\nu}^{33} = - \left(\frac{\partial \ln \omega_{\mathbf{q}\nu}}{\partial \ln c} \right)_a.\end{aligned}\tag{8.2.22}$$

In the previous expressions, the phonon frequency derivatives are evaluated with respect to the variation of only one lattice parameter, the other one remaining fixed at its static equilibrium value. One should finally note that, in contrast to Eq. (8.2.20), the derivative entering $\gamma_{\mathbf{q}\nu}^{a1}$ is taken by changing the a cell dimension, which affects *both* lattice vectors in the basal plane, hence the factor 1/2. We finally note that the derivation of the Grüneisen formalism [64] neglects the zero-point energy of the phonons, such that there is no thermal expansion at $T = 0$ K.

8.3. Computational details

8.3.1. First-principles calculations

All first-principles calculations were performed using the ABINIT software package [69]. The bulk electronic structure was obtained within DFT, while response function and electron-phonon coupling calculations were performed within DFPT [70, 71]. We used a maximum plane-wave energy of 35 hartree and sampled the BZ using a $8 \times 8 \times 8$ Monkhorst-Pack \mathbf{k} -point grid. Spin-orbit coupling was taken into account as it is necessary to obtain the Rashba effect and since it has been shown to strongly affect both electronic [33] and vibrational [72] properties of BiTeI. We used Hartwigsen-Goedecker-Hutter (HGH) fully relativistic norm-conserving pseudopotentials [73], including explicitly the semi-core $5d$ electrons for Bi. The electron-phonon self-energy was computed with the ElectronPhononCoupling module.

Throughout this paper, we rely on the generalized gradient approximation of the Perdew-Burke-Ernzerhof (PBE-GGA) functional [74], although it has been shown to overestimate the lattice parameters at ambient pressure for this particular material [46, 53, 75]. Since the purpose of this paper is to investigate the temperature dependence of the TPT, we chose the functional that gives us the best overall agreement with experiment for the bare band gap and the projected orbital character of the band extrema at ambient pressure, as well as for the predicted critical pressure P_{C1} . For detailed results and a complete discussion about our choice of PBE-GGA functional, see Appendix A.

8.3.2. Structural properties

The unit-cell geometry has been optimized until the resulting forces on all atoms were lower than 10^{-5} hartree/bohr³. Due to the layered nature of BiTeI, the in-plane and normal

lattice parameters of BiTeI do not vary isotropically under hydrostatic pressure [11]: c decreases more rapidly than a within the first GPa applied. In order to allow for this nonhomogenous variation of the lattice parameters, the external pressure was modeled by applying an isotropic stress tensor. The lattice was fully optimized for nine different pressures between 0 and 5 GPa. The resulting cell volumes were then fitted with a Murnaghan equation of state (EOS) [76] in order to validate the theoretically applied pressure. For all pressures below 3.5 GPa, the discrepancy between the EOS pressure and the applied pressure was less than 0.1 GPa, and less than 0.06 GPa for pressures surrounding the TPT. The lattice structure at intermediate pressures was obtained by interpolating from the results of those nine calculations. The optimized lattice structure at all pressures considered in this paper can be found in Table 8.2 of Appendix A.

8.3.3. Lattice dynamics and electron-phonon coupling

Phonon frequencies and electron-phonon matrix elements were computed with a $12 \times 12 \times 12$ homogeneous, Γ -centered Monkhorst-Pack \mathbf{q} -point grid. Explicit calculations were done at seven different pressures, spanning both the trivial (0.0, 0.5, 1.0, and 1.5 GPa) and topological (3.0, 3.5 and 5.0 GPa) phases. The pressure dependence of the phonon frequencies for the Raman-active modes is in good agreement with experimental data [see Tables 8.4 and 8.5 of Appendix A].

In a typical nonadiabatic AHC calculation, one aims for a value of η smaller than the highest phonon energy (typically around $\eta = 0.01$ eV, see, for example, Nery *et al.* [77]). This however requires a very dense \mathbf{q} -point sampling. As discussed in the Appendix of Antonius *et al.* [56], when using a coarser \mathbf{q} -point grid, one risks artificially emphasizing the contribution of a discrete number of terms in Eq. (8.2.5) where, for a given $\mathbf{q}\nu$ -phonon, the value of $\varepsilon_{\mathbf{k}n}^0 - \varepsilon_{\mathbf{k}+\mathbf{q}n'}^0 \pm \omega_{\mathbf{q}\nu}$ gets vanishingly small, hence inducing rapid fluctuations of the self-energy [56]. These peaks are a numerical artefact of the \mathbf{q} -point sampling, and the presence of such fluctuations in the vicinity of the bare electronic eigenvalue $\varepsilon_{\mathbf{k}n}^0$ could lead to an overestimation of the renormalization. In contrast, increasing η without caution can lead to an unphysical flattening of the self-energy, suppressing physically relevant features. We therefore chose $\eta=0.1$ eV, which allows the self-energy to be a smooth function of frequency, without suppressing physically relevant features. We verified that this choice was suitable for the claims reported in this paper.

Lastly, our first-principles methodology relies on the assumption that the electron-phonon interaction can modify the electronic eigenenergies, but cannot by itself promote an electron to an excited state. Such a possibility would modify the resulting electronic density and would require one to evaluate the nondiagonal elements of the self-energy, $\Sigma_{\mathbf{k}n,n'}$ [see Eq. (4) of Antonius *et al.* [56]], allowing $n' \neq n$. This procedure was, however, out of reach

for the present work. Thus, we restricted our investigation of electron-phonon coupling to pressures where the predicted bare band gap was greater than the highest phonon energy ($E_g > 20$ meV).

8.3.4. Thermal expansion

The directional Grüneisen parameters γ^{a_1} and γ^c were evaluated using an $8 \times 8 \times 8$ Γ -centered Monkhorst-Pack \mathbf{q} -point grid. Note that the Grüneisen parameters require a smaller \mathbf{q} -point grid than the EPI interaction since they only depend on the phonon frequencies, which converge faster with the \mathbf{q} -point sampling.

The derivatives in Eq. (8.2.22) were computed by central finite difference using three different volumes per direction, per pressure. For each calculation, the internal atomic coordinates in the \hat{c} direction were optimized at fixed volume, using the same convergence criterion as for the equilibrium structure. The compliance constants s_{ij} were obtained by computing the strain-strain derivatives of the total energy within DFPT [78]. We used the so-called relaxed-ion compliance tensor, which takes into account the relaxation of the atomic coordinates within the unit cell. This typically lowers the components of the elastic stiffness tensor (usually referred to as the elastic constants), as the resulting stress on the atoms is reduced by the relaxation process. Since the compliance tensor is the inverse of the elastic stiffness tensor, it typically increases the elastic compliance when compared to the clamped-ion result [79].

8.4. Results and discussion

8.4.1. Topological phase transition in the static lattice

BiTeI is a layered material composed of alternating Bi, Te and I planes stacked along the high-symmetry crystallographic \hat{c} axis. This small band-gap semiconductor belongs to the noncentrosymmetric trigonal space group P3m1 (no. 156, mp-22965). The crystal structure and first BZ are shown in Fig. 8.2.

A common signature of a TPT is an inversion of the orbital character of the band structure in the vicinity of the band gap [5]. In the case of BiTeI, a schematic analysis of the band splitting showed that the inverted bands should be of p_z character [35]. Fig. 8.3 displays the leading orbital character of the last valence and first conduction bands, for different pressures throughout the studied range. We computed the p_z character by projecting the wavefunction on the $l = 1, m = 0$ spherical harmonic centered around the different atoms in the unit cell. To emphasize the band inversion, Fig. 8.3 shows the *relative* projected character on each band, namely the normalized difference between the Bi-6 p_z and Te/I-5 p_z projections.

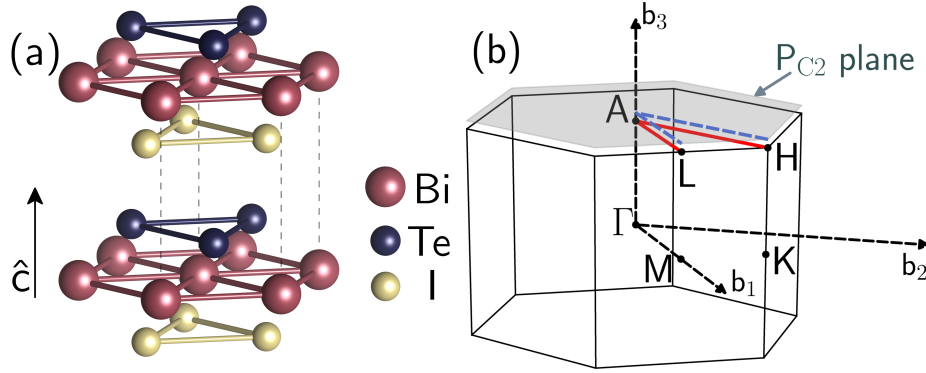


Figure 8.2 – Crystal structure and first Brillouin zone of BiTeI. (a) The different I-Bi-Te trilayers are weakly bound by van der Waals interaction along the crystallographic c axis. Dashed lines delimit a unit cell. (b) The path used for static band structure, EPI, and TE calculations in the trivial insulator phase is shown in solid red. The shaded gray area corresponds to the P_{C_2} plane (defined in Section 8.4.1), where the Weyl points annihilate and the system enters the TI phase. For $P > P_{C_2}$, the band structure was computed along the dashed blue path.

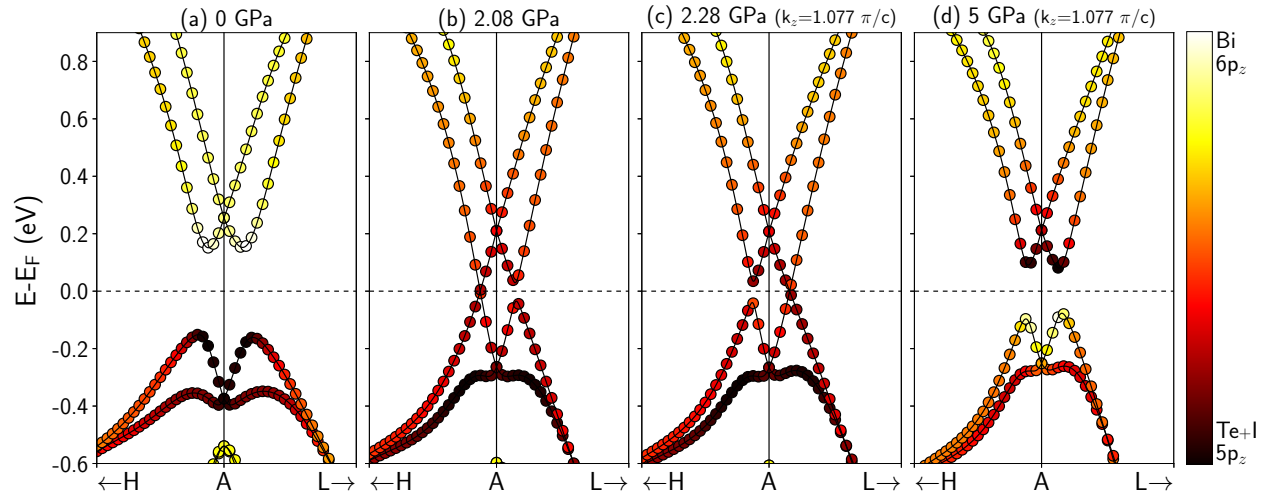


Figure 8.3 – Band inversion process and critical pressures. The solid black lines show the bulk band structure at a given pressure, while the colored markers describe the relative character of the projection of the wavefunction on the $l = 1, m = 0$ spherical harmonic centered around the different atoms in the unit cell. It displays the normalized difference between Bi- $6p_z$ and (Te+I)- $5p_z$ projections. A band with leading Bi character will show in light yellow, while dark brown refers to a leading Te+I character. The band inversion of the Bi and Te/I states observed at 5 GPa (d), when compared to ambient pressure (a), is a signature of the TPT. At the two critical pressures $P_{C_1} = 2.08$ GPa (b) and $P_{C_2} = 2.28$ GPa (c), the band structure displays the cross-section of a Dirac cone, along H-A for the former and in the A-L-M mirror plane for the latter. For all panels, the horizontal axis shows a path representing $\sim 15\%$ of the H-A and A-L distances. For $P \leq P_{C_1}$, the displayed band structure is located at $k_z = \pi/c$, while for $P \geq P_{C_2}$ it is shifted to the P_{C_2} plane [shaded area and dashed blue path of Fig. 8.2(b)].

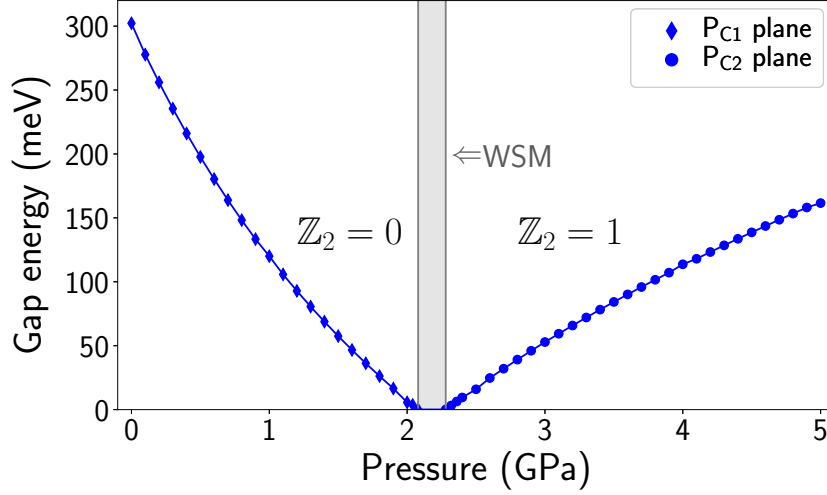


Figure 8.4 – Variation of the indirect band gap with pressure. In the trivial insulator phase ($\mathbb{Z}_2 = 0$), the minimal band gap is located in the H-A-L plane (diamonds), where the first band touching point will occur at $P_{C1}=2.08$ GPa. The system enters the topological phase ($\mathbb{Z}_2 = 1$) once the pairs of Weyl points have annihilated at the second band touching point, at $P_{C2}=2.28$ GPa. For higher pressures, the minimal band gap is now located in the A-L-M mirror plane, with $k_z = 1.07698 \pi/c$ (circles). The gray shaded area represents the WSM phase width in the static lattice approximation. See Section 8.4.1 for more details.

At ambient pressure [Fig. 8.3(a)], the indirect gap is located in the $k_z = \pi/c$ plane, close to the A high-symmetry point, in the direction of H [on the red path of Fig. 8.2(b)]. In the H-A and A-L directions, which both exhibit Rashba splitting, the valence band maximum (VBM) is dominated by p_z states of Te and I (dark brown), whereas the conduction band minimum (CBM) has Bi- p_z (light yellow) character. It is a trivial band insulator, with topological index $\mathbb{Z}_2=(0;000)$. The \mathbb{Z}_2 topological index was computed by tracking the evolution of the hybrid Wannier charge centers, using the Z2Pack software package [80, 81]. For more information about the definition of the \mathbb{Z}_2 topological index for 3D topological insulators, we refer to Sections II-C and IV-A of the review by Hasan and Kane [5].

As pressure is increased, the band gap progressively decreases until it closes at $P_{C1}=2.08$ GPa [Fig. 8.3(b)]. At this point, the BZ exhibits three pairs of Dirac cones, which are split into pairs of Weyl nodes upon further pressure increase; see Fig. 11 of Liu and Vanderbilt [45]. These Weyl nodes migrate within the BZ until they annihilate each other, at $P_{C2}=2.28$ GPa [Fig. 8.3(c)].

Fortunately, from previous studies [45, 46], it is known that the gap closing occurs along symmetry line H-A (with $k_z = \pi/c$), and that it will reopen along a symmetry line rotated by $\pi/6$, without symmetry constraints in the k_z direction (namely, in the A-L-M mirror plane). In the following, we define the P_{C2} plane as the plane parallel to the H-A-L plane, with $k_z = 1.07698 \pi/c$, which is the k_z coordinate of the second band touching point displayed

in Fig. 8.3(c), where the Weyl nodes' annihilation occurred. This plane corresponds to the shaded plane in Fig. 8.2(b). In a similar fashion, we identify the H-A-L plane as the P_{C1} plane.

In the vicinity of the TPT, the valence and conduction band extrema show a mixed character, with an almost equal contribution of Bi- p_z and Te/I- p_z states [Fig.8.3(b-c)], foreshadowing the band inversion [Fig. 8.3(d)] which occurs for $P > 2.28$ GPa. The system is then a strong topological insulator, with $\mathbb{Z}_2=(1;001)$. In the TI phase, the minimal band gap is shifted from the H-A-L plane ($k_z = \pi/c$) to the P_{C2} plane ($k_z = 1.07698 \pi/c$), where the Weyl nodes' annihilation occurred. The full pressure dependence of the band gap for the static lattice is shown in Fig. 8.4.

For the static lattice at $T = 0$ K, we thus obtain a WSM phase width of 0.2 GPa, in good agreement with previous calculations [45–47]. Our value for the first gap closing $P_{C1}=2.08$ GPa agrees with many experimental estimations [11, 38, 42] but is slightly lower than those of other experimental [39, 40] and computational works [46, 47, 82]. For comparison with experiments, one should note that, as the TPT of BiTeI occurs under hydrostatic pressure, direct observation of the topological surface states with angle-resolved photoemission spectroscopy (ARPES) would be challenging from an experimental point of view. As a consequence, experiments rather focused on indirect signatures of these states, for example, a broad minimum of resistivity [38, 39], which do not allow a precise estimation of P_C . The difference with other computational works can be explained by the use of experimental pressure-dependent lattice parameters instead of the DFT-relaxed ones [47], as well as by the choice of projector-augmented wave (PAW) pseudopotentials without semi-core states for Bi [46, 82] and the use of the Heyd-Scuseria-Ernzerhof (HSE) hybrid functional [82] to compute the static gap at ambient pressure. Indeed, the calculated critical pressure is quite sensitive to the choice of functional, as it may significantly modify the starting gap at ambient pressure. One must also recall that the DFT gap has fundamentally no physical meaning; a proper, theoretically adequate calculation of the ambient pressure gap would require using a many-body methodology like GW. Nevertheless, for the scope of this paper, we are more interested in the *relative* variation of the critical pressures rather than in their absolute values. These slight discrepancies with other works do not, therefore, undermine our conclusions.

Our value of P_{C1} corresponds to a relative volume (V/V_0) compression of $\sim 12\%$, in good agreement with other calculations relying on the PBE functional [35]. We also verified that the relative volume compression within the WSM phase was consistent with the original prediction of Liu and Vanderbilt [45].

8.4.2. Temperature dependence of the topological phase transition

Within this paper, we track the temperature dependence of the critical pressures P_{C1} and P_{C2} by evaluating how the electron-phonon interaction and the thermal expansion will affect the indirect band-gap value. Our analysis is based on several assumptions. First, we suppose that the \mathbb{Z}_2 topological index computed at $T = 0$ K for the static lattice remains unchanged unless the renormalization closes the gap. Second, since the system lacks inversion symmetry regardless of temperature, we assume that Liu and Vanderbilt’s argument [45] is still valid, such that a gapless WSM phase still separates the two insulating phases. We finally suppose that the k_z value of the P_{C2} plane remains unchanged by temperature, as it is unlikely that the Weyl nodes’ annihilation point in the BZ changes significantly from the static result. We verified that the position of both band extrema along the A-L line in the TI phase does not change significantly over the range of temperatures and pressures studied, between the P_{C1} and P_{C2} planes. A more thorough analysis of the temperature dependence of the band-gap location within the P_{C1} and P_{C2} planes, as well as the pressure-temperature dependence of the Rashba splitting, can be found in Appendix B.

A more formal approach would also require one to redefine the topological invariant for finite temperatures, relying on the density matrix rather than on the static ground state wavefunction, since temperature modifies the occupation of the electronic eigenstates [83–85]. The question of the persistence of quantum topological order at finite temperatures, compared to a nonquantized “classical” topological order, is still under investigation [86].

When considering the effect of temperature on the TPT within our framework, there are thus two possible outcomes, which depend on the sign of the band-gap renormalization. In the first case, a negative correction implies that the band gap closes with increasing temperature, which can both drive a topologically trivial system towards a band inversion and stabilize an already inverted band structure, thus globally promoting the topological phase. On the other hand, a positive correction will favor the trivial phase: the band gap opens with increasing temperature, preventing the band inversion in a trivial system and destabilizing a TI band structure until the band inversion is reversed. A more detailed version of this argument can be found in our proceedings paper [87].

In the case of a pressure-induced TPT with an intermediate WSM phase, both critical pressures, P_{C1} and P_{C2} , will be modified by temperature: a negative renormalization favoring the topological phase will diminish the amount of pressure required to close the band gap, while a positive renormalization promoting the trivial phase will increase it. The WSM phase width will also be affected by the relative strength of the renormalization between the trivial and topological phases. The combined effect of the strength and sign of the renormalization will determine if the WSM phase is widened or narrowed by increasing temperature.

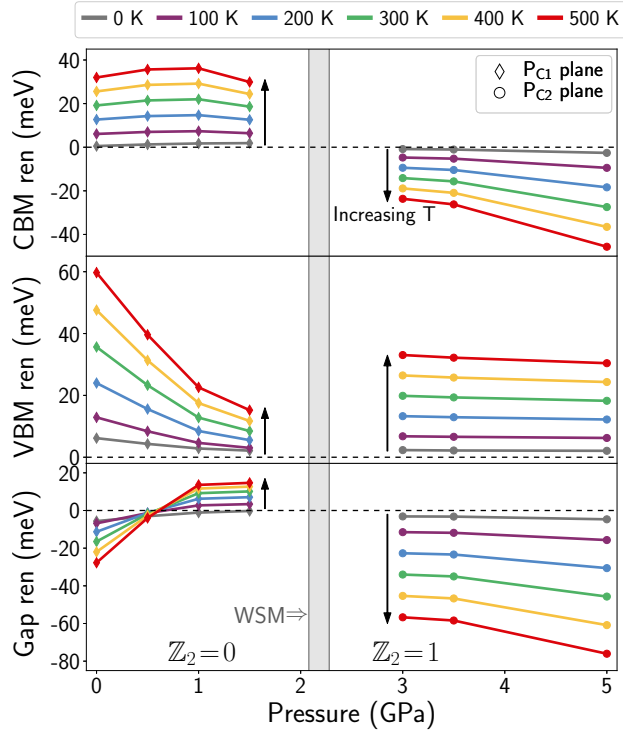


Figure 8.5 – EPI contribution to the temperature-dependent renormalization for the conduction-band minimum (CBM, top panel), valence-band maximum (VBM, center panel) and total band gap (lower panel) of BiTeI, as a function of pressure. A zero renormalization (black dashed lines) refers to the static lattice results. The gray curves (closest to the dashed lines) show the zero-point renormalization (ZPR). Black arrows indicate increasing temperature. For visual reference, the shaded gray area locates the WSM phase boundaries from the static calculation.

8.4.2.1. Electron-phonon interaction. In a typical semiconductor, with a nearly parabolic dispersion near well-defined band extrema and a sufficiently wide band gap, the Fan contribution usually dominates the self-energy. One then expects from Eq. (8.2.5) to observe a decrease of the band-gap energy with increasing temperature: the VBM is repelled by neighboring occupied states with similar but lower energy, leading to a positive renormalization. Similarly, the CBM is repelled by neighboring unoccupied states with higher energy, yielding a negative renormalization. Couplings between states with different occupation factors are disadvantaged by the large energy difference in the denominator, due to the presence of the gap. This behavior is sometimes referred to in the literature as the Varshni effect [88].

In BiTeI, the temperature dependence of the gap cannot be inferred by such a heuristic argument, since, on the one hand, the small band gap emphasizes the weight of couplings between the subsets of occupied and unoccupied bands compared to a typical semiconductor, and, on the other, the Rashba splitting creates regions in the BZ where a phonon with finite, nonvanishing wavevector \mathbf{q} can couple electronic states with very similar eigenenergies.

The EPI contribution to the VBM (center panel), CBM (top panel) and total band gap (bottom panel) is shown in Fig. 8.5. In order to track any temperature-induced change to the gap location and to accurately capture the renormalization for the minimal gap, the EPI was computed in the H-A-L plane for the trivial phase, and in the P_{C2} plane for the TI phase, for electronic states along the solid red and dashed blue paths shown in Fig. 8.2(b).

In the trivial insulator phase, both the VBM and the CBM are shifted towards higher energies. While the CBM renormalization is almost pressure independent for a given temperature, the VBM is strongly affected by pressure, its renormalization dropping by roughly a factor of 3 between 0 and 1 GPa. This behavior is most likely linked to the higher compressibility along the c axis at lower pressures [11]. The strength of this variation could, however, be amplified by the overestimated compressibility predicted by the PBE functional for BiTeI [75]. In our case, this results in a total gap renormalization that changes sign within the trivial phase, going from negative to positive shortly after 0.5 GPa. Nevertheless, the EPI contribution to the gap remains quite small compared to the bare gap energy at ambient pressure (at most 20 meV at all temperatures). As argued before, such an opening of the gap with temperature when approaching the TPT will delay the first gap closing, thus moving P_{C1} towards a slightly higher value.

In the TI phase, the renormalization displays a different behavior: the VBM and CBM renormalizations have opposite signs, both contributing to a decrease of the gap energy with increasing temperature. Again, this behavior is not favorable for the nontrivial case as it works to revert the already inverted gap, driving the system back to the trivial phase. One should also note that, in this phase, it is the VBM that exhibits a pressure-independent behavior, while, for the CBM, the negative renormalization increases steadily with increasing pressure. This seemingly distinct behavior is simply a signature of the band inversion: it can be understood by recalling that the orbital decomposition of the state being corrected at the VBM in the TI phase has the same character as the CBM in the trivial phase, as was emphasized in Fig. 8.3. Similarly, the CBM in the trivial phase has the same leading orbital character as the VBM in the TI phase. From this point of view, the band extremum dominated by Bi- p_z character exhibits a quasi-pressure-independent renormalization, while the extremum with Te/I- p_z character is gradually shifted towards lower energies throughout the studied pressure range. It is also rather peculiar that only the Te/I states are affected by the change of topology, which can be seen in the temperature-dependent renormalization at a given pressure going from positive to negative as the system goes through the phase transition. A more thorough explanation of this sign change would require one to probe the EPI in the gap-closing region, which was not possible in the scope of this paper. Nevertheless, if considering the gap behavior in both phases, we can infer that the EPI globally promotes the trivial phase, pushing the TPT towards higher pressures as temperature increases.

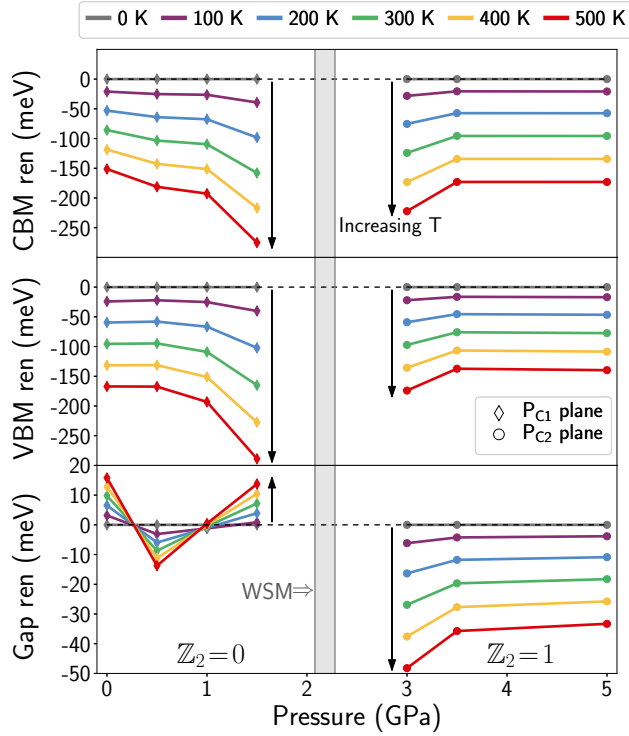


Figure 8.6 – TE contribution to the temperature-dependent renormalization for the conduction-band minimum (CBM, top panel), valence-band maximum (VBM, center panel) and total band gap (lower panel) of BiTeI, as a function of pressure. A zero renormalization (black dashed lines) refers to the static lattice results. Black arrows indicate increasing temperature. For visual reference, the shaded gray area locates the WSM phase boundaries from the static calculation. Note that the Grüneisen formalism neglects the zero-point phonon vibrational energy, hence the identically zero correction at 0 K (gray).

8.4.2.2. Thermal expansion. The TE contribution to the total renormalization is displayed in Fig. 8.6, for the CBM (top panel), VBM (center panel), and indirect band gap (bottom panel). When solely looking at the band renormalization, one finds a seemingly monotonic behavior, both bands being shifted towards lower absolute energies. The pressure-dependent renormalization at a given temperature does not seem affected by the leading orbital character of the band extrema, as was the case for EPI. In the TI phase, the renormalization rate at a given pressure is greater for the CBM than for the VBM, hence a global decrease of the gap energy with increasing temperature. In the trivial phase, the indirect band-gap renormalization shows a far more erratic behavior: while for most pressures the renormalization rate is greater for the VBM than for the CBM, hence a gap opening with increasing temperature, at 0.5 GPa we obtain the opposite trend. This is reminiscent of the sign change observed in this pressure range for the EPI contribution [see Fig. 8.5, lower panel], which we again attribute to an overestimated compressibility in the low-pressure regime. For a layered structure like BiTeI, a higher compressibility will mainly enhance

the s_{33} compliance constant entering the definition of α_c (Eq. (8.2.21)), which will, in turn, modify the thermal expansion along this axis and the resulting gap renormalization. As discussed more thoroughly in Appendix A.1, the lattice structure is reasonably well described by PBE for $P \geq 1$ GPa. Thus, in the following section, we only consider the normalization trends for pressures of 1 GPa and higher to evaluate the temperature-dependent critical pressures. One should finally note that the TE contribution has the same order of magnitude as the EPI contribution, such that it cannot be neglected when investigating the temperature dependence of the gap, as it is often done for materials with light ions.

We recall here that, as mentioned in Section 8.2.2, the Grüneisen formalism neglects the zero-point phonon vibrational energy, such that the $T = 0$ K lattice parameters are identical to the static ones. We estimated the missing zero-point thermal expansion contribution to the band-gap renormalization using the high-temperature extrapolation method displayed in Fig. 2 of Cardona and Thewalt [88]. In this method, the renormalized lattice parameters at $T = 0$ K are evaluated by extrapolating the high-temperature slope of $\Delta a(T)/a$ and $\Delta c(T)/c$ to 0 K. This led to a nearly constant contribution of ~ 2 meV in the trivial phase and ~ -5 meV in the topological phase. This missing contribution does not affect the qualitative trends previously described.

The renormalization trends observed when approaching the TPT in both insulating phases, a gap opening in the trivial phase and a gap closing in the topological phase, are precisely the trends observed for the EPI contribution. Consequently, just like EPI, TE delays the TPT, moving both critical pressures towards higher values with increasing temperature. Excluding $P < 1$ GPa from the analysis for previously stated reasons, these trends agree qualitatively with Monserrat and Vanderbilt [53], despite being smaller in absolute value. This could partly be attributed to the fact that Monserrat and Vanderbilt minimized the Gibbs free energy independently for each lattice parameter, while our methodology [see Eq. (8.2.22)], correlates the expansion along a given axis with the “thermal pressure” [67] induced on the lattice by varying both lattice parameters, through γ^{a1} and γ^c .

This discrepancy could also be explained by anharmonic contributions to the phonon free energy, which our perturbative methodology does not capture. A more refined and numerically accurate calculation of the anisotropic thermal expansion coefficients would rely on the full temperature-dependent phonon-perturbed potential, taking into account all anharmonic interactions, as implemented in the Temperature Dependent Effective Potential method (TDEP) method [89].

8.4.3. Weyl semimetal phase evolution

The temperature dependence of the WSM phase width was assessed by combining the results displayed in the two previous sections to evaluate the full gap energy renormalization

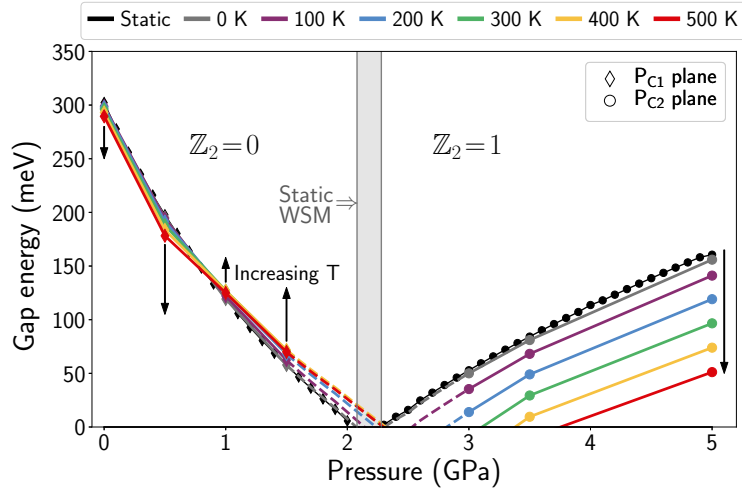


Figure 8.7 – Temperature-dependent band-gap energy. At a given temperature, the total renormalization of the band gap (EPI+TE, solid-colored lines) is extrapolated towards the TPT region and applied to the bare gap energies (black markers with denser pressure sampling). The resulting temperature-dependent gap energies within the pressure region excluded from our calculations are shown with colored dashed lines. Black arrows indicate increasing temperature. For visual reference, the static lattice WSM phase is shown in shaded gray.

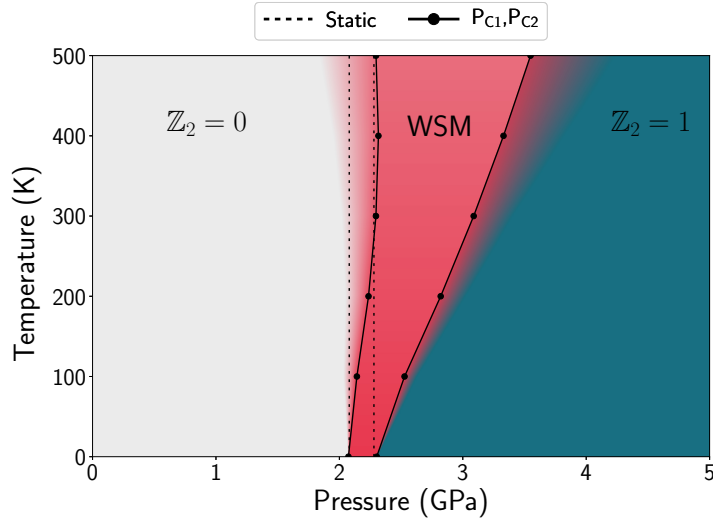


Figure 8.8 – Phase diagram of BiTeI. The temperature-dependent WSM phase boundaries are the intersections of the extrapolated gap closing energies from Fig. 8.7. Both critical pressures are pushed towards higher values. The renormalization is stronger for P_{C2} than for P_{C1} , resulting in a widening of the WSM phase with increasing temperature. At each temperature, we defined a crossover region between topologically distinct phases, since one cannot physically distinguish an insulating phase from a gapless phase if $E_g < k_B T$, due to thermal excitations. See Section 8.4.3 for more details.

at each pressure and temperature. For each \mathbf{k} point on the paths considered [see Fig. 8.2(b)], the EPI and TE corrections were added to the static band structure before computing the resulting temperature-dependent band gap.

Since, for technical reasons discussed in Section 8.4.1, we could not directly probe the EPI in the gap closing region, we tracked the temperature dependence of P_{C1} and P_{C2} by extrapolating the total renormalization behavior for each band extrema from neighboring pressures towards the TPT in the trivial and topological phases. This correction was then added to the bare gap energy computed with the static lattice (black markers with denser pressure sampling), as shown by the colored dashed lines in Fig 8.7.

The temperature-dependent critical pressures extracted from Fig. 8.7 are explicitly shown in the pressure-temperature phase diagram of Fig. 8.8 (black markers). For visual reference, the critical pressures obtained with the static lattice are shown with dotted lines. The numerical values for $P_{C1}(T)$ and $P_{C2}(T)$ are detailed in Table 8.6 of Appendix C, as well as the WSM phase boundaries obtained by considering only the contribution from EPI or TE.

As discussed in the previous sections, both EPI and TE are unfavorable to the topological phase. We therefore observe without surprise that both P_{C1} and P_{C2} are shifted towards higher pressures. However, as the renormalization from both contributions is stronger in magnitude in the topological phase, the temperature-dependent variation is much weaker for P_{C1} than for P_{C2} . Thus, the WSM phase width increases with temperature. This effect is already sizable at $T = 0$ K, where the zero-point motion of the ions induces a 15% increase in the WSM phase width, compared to the static lattice. It has almost doubled by $T = 100$ K, despite the absolute WSM phase width remaining small at 0.39 GPa.

The only deviation to this behavior is the slight decrease of P_{C1} above room temperature. This can be explained by the fact that, for $P = 1$ and 1.5 GPa, EPI and TE do not shift the band extrema in the same way in the BZ. Therefore, in this pressure range, the total renormalization does not sum up to the individual corrections.

One should also take into consideration the fact that, at finite temperature, thermal excitations do not allow one to physically distinguish between an insulating and a gapless phase, if the band-gap energy is smaller than $k_B T$. Thus, for each temperature, we defined a crossover region between topologically distinct phases by evaluating the pressure difference between a given static lattice critical pressure and the pressure at which $E_g = k_B T$ in the appropriate insulating phase, using the data from Fig. 8.4. In Fig. 8.8, these regions identify the phase space where, within our methodology, the topological index cannot be properly assessed. Lastly, we also added to Fig. 8.8 a color intensity gradient in the high-temperature regime of the WSM phase, to emphasize the fact that our calculations did not explicitly treat the temperature dependence within this pressure region. Therefore, there remains some uncertainty about the topological nature of BiTeI in this region of the pressure-temperature phase space.

From these results, we can conclude that temperature effects do not restrain the experimental detection of the WSM phase. In the low-temperature regime, where quantum effects can more easily be observed thanks to the reduction of thermal noise, the WSM phase widening remains however small, despite being sizable when compared to the static lattice phase width. The experimental signatures of the TI phase will be found at higher pressures with increasing temperature. These qualitative temperature dependencies can assist the experimental effort when designing experiments the purpose of which is to detect signatures of a topologically nontrivial phase. Moreover, they shed some light on the subtle interplay between EPI and TE and their effect on the phase space of topological phases in noncentrosymmetric materials. Since the presence of gapless modes characterizes both the WSM and \mathbb{Z}_2 TI phases, additional experiments focusing on observables that can physically distinguish between these two nontrivial phases, and that are feasible under isotropic hydrostatic pressure, could provide more insight about this phenomenon.

8.4.4. Discussion

The purpose of this paper was to determine the qualitative trend of the temperature-dependent renormalization of P_{C1} and P_{C2} : the global behavior of the TPT, more than the precise numerical values of the critical pressures, was our main target. From this point of view, there are several limitations to our analysis, the first of them being the \mathbf{q} -point sampling for the EPI calculation. BiTeI is a polar material, i.e., has nonvanishing Born effective charges, with infrared-active phonon modes. At long wavelength, $|q| \rightarrow 0$, atomic displacements along longitudinal optical (LO) modes generate a macroscopic electric field throughout the material, which can couple to the electrons through dielectric interaction. This particular type of EPI, known as the Fröhlich interaction [90], has been shown to dominate the ZPR for polar materials [91, 92] and to cause an unphysical divergence of the ZPR within the adiabatic AHC framework [59], in which Eq. (8.2.5) is approximated by neglecting the phonon frequencies in the denominators.

Thus, for polar materials, evaluating the precise contribution of the long-range Fröhlich interaction to the ZPR within the nonadiabatic AHC framework requires a very dense \mathbf{q} -point grid, which for technical reasons was not available to us at the time of this study. The methodology developed by Nery and Allen [91] to estimate the polar contribution missing from a coarse \mathbf{q} -point grid calculation could also not be straightforwardly applied to BiTeI for two reasons. First, the band extrema are not located at the zone center and spin-orbit coupling is mandatory to describe Rashba splitting correctly. Second, the validity of the parabolic effective mass approximation near the band extrema is also questionable for this

material, especially at higher pressures. Our results should consequently be understood as a lower bound to the full EPI contribution.

The extrapolation technique used to evaluate the temperature dependence of the critical pressures is also by itself limited by the assumption that the renormalization behavior we observe in both insulating phases varies monotonically towards the TPT. This assumption could only be validated by an explicit calculation of the EPI in the close vicinity of the TPT, which would require one to lift the diagonal self-energy approximation, as excitations across the gap will no longer be negligible. The addition of these nondiagonal coupling terms to the self-energy, which we conjecture will have a significant contribution, could alter the general trends observed in Fig. 8.7. These extra terms will also be crucial to fully characterize the effect of EPI on the Te/I- p_z band extrema near the critical pressures (VBM for $Z_2 = 0$, CBM for $Z_2 = 1$; see Fig. 8.5 and Section 8.4.2.1).

Lastly, our analysis supposed a perfect semiconductor crystal, while experimentally BiTeI has been shown to be naturally n -doped due to small deviations in stoichiometry [37]. The presence of such doping charge could slightly affect the phonon-perturbed crystal potential, which could, in return, modify the matrix elements entering the EPI self-energy. We do not however expect this self-doping to have a strong impact on the qualitative trends observed throughout this paper, as long as the amount of defects remains small compared to substitutional doping.

8.5. Conclusion

In the present paper, we have characterized the temperature dependence of the topological phase transition in BiTeI using first-principles methodologies based on DFPT. The electron-phonon interaction was obtained with the nonadiabatic AHC methodology, while the thermal expansion was computed using the Grüneisen parameters formalism.

The indirect band-gap renormalization induced by EPI changes sign as the system undergoes the phase transition, opening the gap in the trivial phase and closing it in the TI phase. The band extremum with leading Bi- $6p_z$ character displays a quasi pressure-independent behavior, while the Te/I- $5p_z$ extremum manifests a strong sensitivity to both pressure and topology.

The thermal expansion contribution to the band-gap renormalization has the same order of magnitude as the EPI contribution. At a given temperature, both band extrema are affected in a similar manner throughout the pressure range, regardless of their leading orbital character. The resulting band-gap renormalization nevertheless captures the change of topology: it mainly increases the indirect gap in the trivial phase and reduces it throughout the topological phase.

The combined effect of EPI and TE globally moves the TPT of BiTeI towards higher pressures. The resulting temperature-dependent renormalization is stronger in magnitude in the TI phase compared to the trivial phase, such that the intermediate WSM phase is widened by temperature. Clear signatures of the \mathbb{Z}_2 topological insulator phase should, therefore, appear at higher pressures as temperature is increased.

Overall, our results indicate that temperature effects are not negligible for materials with heavy ions and must be accounted for when engineering devices relying on topological properties. Our findings can also aid the search for experimental evidence of the topologically nontrivial phases of BiTeI, and more generally reveal how temperature can have a substantial influence on the phase space of topological phases of noncentrosymmetric materials.

Note added. Recently, we learned of an article by Lihm and Park [93] which is relevant to this research field.

Acknowledgments

This research was financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), under the Discovery Grants program grant No. RGPIN-2016-06666. Computations were made on the supercomputers Beluga and Briaree managed by Calcul Québec and Compute Canada. The operation of these supercomputers is funded by the Canada Foundation for Innovation (CFI), the Ministère de l'Économie, de la Science et de l'Innovation du Québec (MESI) and the Fonds de recherche du Québec - Nature et technologies (FRQ-NT). V.B.-C. acknowledges support by the NSERC Alexander Graham Bell Canada Graduate Scholarship doctoral program.

Appendix

A. Effect of Van der Waals dispersion correction

Since the first prediction of the TPT in BiTeI [35], numerous studies of ground-state properties of bismuth tellurohalides (BiTeX, X=Cl, Br, I) have appeared in the first-principles literature [46, 72, 94–97]. Many of them have discussed the fact that PBE-GGA does not capture the Van der Waals bonding properly along the normal axis for this class of materials, resulting in a general overestimation of the lattice parameters [46, 75], up to 7% for the \hat{c} axis [46]. PBE-GGA was also criticized as yielding an unrealistic compressibility for BiTeI at low pressures [75]. It has also been shown that dispersion-corrected DFT, which accounts for Van der Waals forces through a semi-empirical dispersion potential, reproduces more accurately the pressure dependence of the lattice structure for bismuth tellurohalides [48].

These studies did not, however, investigate the impact of these corrections on the phonon frequencies, nor on the electron-phonon interaction. In the following, we include the Van der Waals interaction within the DFT framework, using dispersion-corrected DFT-D3BJ, where a semi-empirical dispersion potential is added [98] in conjunction with Becke-Jonhson damping [99, 100] to avoid short-range diverging behavior. DFT-D3BJ was shown to be one of the most reliable dispersion-corrected methods to describe vibrational properties [101]. Both the trivial (0, 1.5 GPa) and the topological (3.5, 5 GPa) phases are considered.

A.1. Structural properties

The lattice parameters, bare band gap energies and equilibrium distance between the Te/I and Bi planes for 0 and 5 GPa are shown in Tables 8.1 and 8.3, with comparison to available experimental data. In good agreement with previous works [46, 53, 75], PBE-GGA

Table 8.1 – Static lattice properties at 0 and 5 GPa.

	0 GPa			5 GPa		
	Exp.	PBE	DFT-D3BJ	Exp.	PBE	DFT-D3BJ
a (Å)	4.34 [102]	4.44	4.36	4.15 [11]	4.23	4.12
c (Å)	6.85 [102]	7.31	6.78	6.52 [11]	6.55	6.49
c/a ratio	1.58 [102]	1.65	1.56	1.57 [11]	1.58	1.58
Gap (eV)	0.38 [29], 0.36 [103], 0.33 [37]	0.30	0.03		0.16	0.30
Bi-Te \hat{c} (Å)	2.10 [102]	1.70	1.73		1.75	1.78
Bi-I \hat{c} (Å)	1.72 [102]	2.11	2.15		2.16	2.17

Table 8.2 – Optimized static lattice structure

	0 GPa	0.5 GPa	1 GPa	1.5 GPa	3 GPa	3.5 GPa	5 GPa
a (Å)	4.44	4.41	4.39	4.37	4.30	4.28	4.23
c (Å)	7.31	7.05	6.90	6.81	6.66	6.62	6.55
Bi-Te \hat{c} (Å)	1.70	1.71	1.71	1.72	1.74	1.75	1.75
Bi-I \hat{c} (Å)	2.11	2.12	2.13	2.14	2.15	2.15	2.16

Table 8.3 – Static lattice 0 GPa Rashba parameters along A-L.

	Exp.	PBE	DFT-D3BJ
E_R (meV)	100 [29], 108 [32]	103	181
k_R (Å ⁻¹)	0.052 [29], 0.046 [41], 0.050 [32]	0.050	0.053
α_R (eV·Å)	3.85 [29], 4.3 [32]	4.13	6.86

overestimates the lattice parameters at ambient pressure more than what is typically expected from DFT calculations, at 2.3% for a and 6.7% for c , while the dispersion-corrected method reproduces more accurately the experimental structure. At 5 GPa, both methods agree within 2% of experimental value, for both a and c . The interplanar equilibrium distances at both pressures are in good agreement with experiment. Note that the positions of Te and I planes are inverted when compared to the experimental results [102]; this effect has been attributed to the fact that x-ray diffraction measurements cannot establish a clear distinction between the Te and I planes, due to their very similar atomic radii and charge [33]. This inversion is systematic throughout the first-principles literature [33, 53, 72, 75].

Analyzing the full pressure dependence of the lattice parameters for PBE-GGA and DFT-D3BJ between 0 and 5 GPa (not included here) lead to similar results as those of Güler-Kiliç and Kiliç (see Figs. 3(g) and 3(h) of [75]), where the overestimation of a remains around 2% throughout the whole pressure range, while for c it drops rapidly under 2% after 1 GPa. Thus, for $P \gtrsim 1$ GPa, the lattice structure is similarly well described by both methods. The resulting lattice parameters and interatomic distances along \hat{c} obtained by the optimization procedure are presented in Table 8.2 for all pressures included in this work.

The bulk band gap of BiTeI has been measured at ambient pressure via ARPES [29], optical spectroscopy [103] and optical spectral ellipsometry [37], yielding respectively 0.38, 0.36 and 0.33 eV. Another soft x-ray ARPES experiment obtained a smaller value of 0.26 eV [104]; they however argued that their measurement probes a region between the bulk and the subsurface, where some band bending effects could lower the measured band gap compared to the bulk value. Our results with PBE, 0.30 eV for the indirect gap and 0.33 eV for the direct gap, are in pleasantly good agreement with experiment, thanks to the overestimation of the lattice parameters counterbalancing PBE's well-know gap underestimation [46]. For DFT-D3BJ, the band gap is strongly underestimated: the system is almost metallic, with a band gap of about 30 meV at ambient pressure. We are not aware of any reference for experimental gap measurements in the topological phase.

Beyond the crystal structure, the parameters characterizing the Rashba splitting can also be used to assess the level of agreement of the calculated electronic structure with experiment. The Rashba energy, E_R , is defined as the energy difference between electronic states at the band extrema and at the time-reversal protected degeneracy point of spin-split bands, namely A for BiTeI [see Fig. 8.3]. The momentum offset between those two \mathbf{k} -points is captured by the Rashba wavevector, \mathbf{k}_R . The Rashba parameter,

$$\alpha_R = \frac{2E_R}{k_R}, \quad (\text{A1})$$

is the strength of the spin-orbit interaction in the Rashba Hamiltonian. These parameters are clearly displayed in the left panel of Fig. 2a and Eq. (1) of Monserrat and Vanderbilt [53].

Table 8.4 – Zone-center optical phonon frequencies ω (cm^{-1}) in the trivial phase. When available, LO-TO splitting is indicated with parentheses. LO-TO splitting occurs when the applied electric field is in plane for E modes, and normal for A_1 modes.

0 GPa				
Mode	Exp. [72], [105], [44], [36]	Calc. [95], [72]	PBE	DFT-D3BJ
E^1	52.5, 55, 56, n/a	59, 54 (73)	53 (74)	57 (77)
A_1^1	90.5, 93, 92, 92	92, 88 (94)	90 (94)	90 (100)
E^2	99, 102, 101, 101	101, 96 (115)	97 (118)	101 (117)
A_1^2	146.5, 150, 148, 146	152, 141 (142)	141 (142)	146 (148)
1.5 GPa				
Mode	Exp. [44],[36]	PBE	DFT-D3BJ	
E^1	60, n/a	58 (77)	62 (80)	
A_1^1	96, 96	92 (100)	95 (106)	
E^2	106, 107	102 (119)	106 (118)	
A_1^2	152, 153	147 (148)	151 (152)	

Table 8.5 – Zone-center optical phonon frequencies ω (cm^{-1}) in the topological phase. When available, LO-TO splitting is indicated with parentheses. LO-TO splitting occurs when the applied electric field is in plane for E modes, and normal for A_1 modes.

3.5 GPa				5 GPa		
Mode	Exp. [44], [36]	PBE	DFT-D3BJ	Exp. [44], [36]	PBE	DFT-D3BJ
E^1	66, n/a	63 (81)	67 (84)	70 ^a , n/a	68 (84)	71 (87)
A_1^1	105, 104	99 (108)	102 (111)	107, 108	102 (114)	109 (121)
E^2	112, 112	108 (120)	115 (120)	115, 115	111 (122)	115 (123)
A_1^2	162, 162	154 (155)	159 (151)	165, 167	159 (160)	164 (166)

^aExtrapolated to 5 GPa from the referenced experimental data.

Our calculations agree reasonably well with available experimental data [see Table 8.3], with a clear advantage for PBE-GGA.

A.2. Vibrational properties

At the zone center, Γ , BiTeI has C_{3v} point group symmetry, yielding four Raman-active modes, whose irreducible representations are two doubly-degenerate E modes and two A_1 modes [36, 105]. Since BiTeI has no inversion center, all modes are both Raman and infrared active [36]. Phonon frequencies for these modes are compared to available experimental data and previous first-principles calculations in Tables 8.4 and 8.5. At ambient pressure, we obtain a relative mean deviation with experiment of 4.1% for PBE and 1.6% for D3BJ. At

5 GPa, these deviations drop to 3.8% for PBE and 0.8% for D3BJ. We verified that the general pressure dependence of the Raman-active frequencies agreed well with experiment in both cases. Our calculated values for LO-TO splitting are in excellent agreement with previous calculations [72]. A small feature measured at $\omega=118.5 \text{ cm}^{-1}$ at ambient pressure [72] can be associated with the LO-TO split E^2 mode, thus validating our results.

A.3. Discussion

The purpose of this work is to analyze the temperature dependence of the TPT in BiTeI. The numerical value of the critical pressures is directly related to the bare band gap value, which will be corrected by electron-phonon interaction. From Eq. (8.2.5) and (8.2.11), we can see that the phonon frequencies and the bare electronic eigenvalues are key physical quantities in this correction. This dependence goes beyond the simple electronic dispersion: as discussed in the main text, the band gap also plays a crucial role. Indeed, the denominator captures the “unlikelyhood” of couplings between two eigenstates with a large energy difference, rooted in perturbation theory. Furthermore, considering the renormalization of both band extrema forming the fundamental band gap, the gap energy will act as a “barrier”, giving a lower bound on the smallest value of the denominator for couplings with eigenstates within the subset of bands located across the gap. Therefore, a significant underestimation of the bare band gap will artificially emphasize interband couplings and bias the resulting zero-point renormalization. While one could argue that this effect should be small at $T = 0 \text{ K}$, especially for a material with heavy atoms like BiTeI, it should not be forgotten that it will be amplified by the phonon occupation factor at higher temperatures, and thus could lead to significantly different conclusions. Lastly, since we are dealing with a TPT, the orbital character of the valence and conduction bands close to their extrema is subject to an inversion process. This phenomenon is captured by the wavefunction, such that a wrong band character could alter the different matrix elements entering the self-energy.

While DFT-D3BJ provides the most accurate descriptions of the cell geometry and Raman-active phonon frequencies, our results indicate that in our case, it is not a suitable choice to track the temperature dependence of the TPT. On the one hand, despite its accurate description of Van der Waals bonding between the atomic planes, DFT-D3BJ delivers a severe underestimation of the band gap at ambient pressure. This can be understood by considering the van der Waals correction as acting like an “effective pressure” along the \hat{c} axis. As previously argued, such a small value of the band gap would not only flaw our bare P_C predictions (by the same arguments as Rusinov *et al.* [46]) but could also artificially strengthen some interband couplings in the EPI self-energy.

One could easily argue that this band gap underestimation could be overturned simply by applying a scissor shift operator to the eigenvalues entering the self-energy, or be

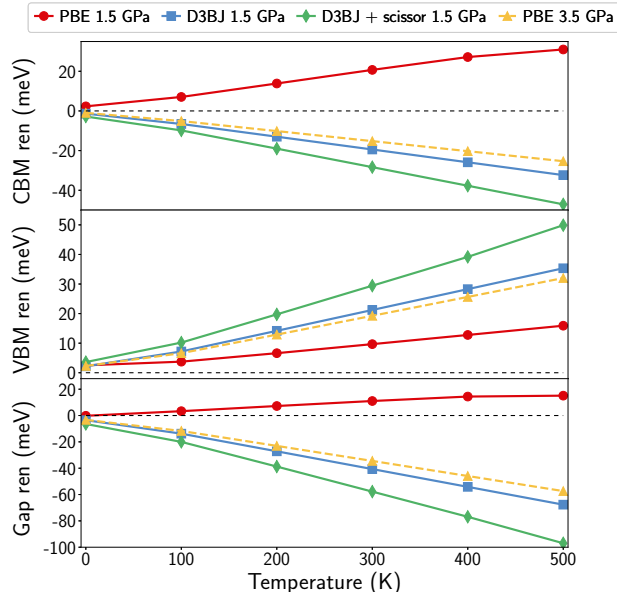


Figure 8.9 – Temperature-dependent renormalization of the band-gap edges at 1.5 GPa, using DFT-D3BJ (blue squares) and PBE-GGA (red circles). The green curve (diamond markers) shows the effect of adding a scissor shift operator to DFT-D3BJ to bring the ambient pressure gap at the experimental value. Both methods disagree on the sign of the renormalization for the conduction band: it is positive for PBE-GGA and negative for DFT-D3BJ. As shown in Fig. 8.5, this sign change occurs when the system enters the topological phase. Note that the DFT-D3BJ result at 1.5 GPa is almost identical to the PBE-GGA result at 3.5 GPa (dashed yellow line, triangle markers), that is, shifted by $P_0 = -2$ GPa.

compensated by evaluating the pressure shift that would bring the gap to the experimental value. We considered both these approaches. For the latter, we estimate that one must shift towards a negative pressure $P_0 \simeq -2$ GPa in order to bring the DFT-D3BJ gap close to the experimental value, thus bringing the lattice parameters very close to the PBE relaxed values at ambient pressure.

About the former, we have investigated the temperature-dependent renormalization at 0, 1.5, 3.5 and 5 GPa, with and without the scissor shift operator. The results for $P = 1.5$ GPa shown in Fig. 8.9 are quite striking: PBE-GGA (red circles) and DFT-D3BJ with (green diamonds) and without (blue squares) the scissor shift disagree about the sign of the renormalization for the CBM. This sign disagreement is also reflected in the total gap renormalization.

This discrepancy brings to light a fundamental aspect of the calculation, namely the fact that the EPI renormalization is affected by the system’s topology. From \mathbb{Z}_2 topological invariant calculations, we found that at 1.5 GPa, the DFT-D3BJ wavefunction already is in the *topological* state (yielding a *negative* correction), while for PBE-GGA it is still in the *trivial* state (where the correction is *positive*). These are *precisely* the trends observed in

Fig. 8.5. Moreover, if we compare the DFT-D3BJ without scissor shift (blue squares) to the PBE-GGA result at 3.5 GPa (dashed yellow line, triangle markers), which is roughly $P = 1.5 \text{ GPa} - P_0$, we can see that those two calculations, with both wavefunctions in the topological state and with roughly the same lattice parameters and bare band gap, deliver almost identical results.

Thus, even if DFT-D3BJ delivers the most accurate lattice parameters and Raman-active phonon frequencies at 1.5 GPa when compared to experiment, it predicts a \mathbb{Z}_2 topological index that contradicts all experimental evidences of the phase transition. Therefore, computing the renormalization with DFT-D3BJ and a scissor shift operator for a given pressure would roughly result in applying the electron-phonon contribution from a PBE-GGA calculation at another pressure, shifted by P_0 . This is especially crucial when investigating pressures neighboring the TPT, as it can lead to wrongly assign a renormalization computed with a TI wavefunction to a shifted band structure intended to describe the trivial phase. For pressures sufficiently far from the transition in the TI phase, we expect the trend of a DFT-D3BJ with scissor shift calculation to be correct, despite a discrepancy in the numerical value of the renormalization. For the trivial phase, a \mathbb{Z}_2 calculation showed that $0 < P_{C1} < 0.5 \text{ GPa}$ for DFT-D3BJ. Therefore, we cannot rely on such results to describe the ambient pressure renormalization since the resulting wavefunction does not predict the correct band character.

For these reasons, we chose to rely on PBE-GGA functional: despite a clear overestimation of the lattice parameters, it does deliver a band gap, Rashba splitting and Raman-active phonon frequencies that are in reasonable overall agreement with experiment. Moreover and most crucially, it exhibits the right band characters at ambient pressure and predicts a critical pressure P_{C1} in agreement with experimental evidence.

B. Pressure and temperature-dependent Rashba Splitting

As discussed in the main text, one of the assumptions underlying our work is that the Weyl nodes trajectory throughout the WSM phase is not qualitatively affected by temperature. In other words, we suppose that the k_z location of the P_{C1} and P_{C2} planes remains constant; hence we only investigated EPI and TE within these planes. Without having access to the pressure and temperature dependence of the band structure in the full BZ, we can nevertheless gain significant insight by tracking the pressure and temperature dependence of the different parameters characterizing Rashba splitting within these planes. The Rashba energy, E_R , captures the steepness of the band's slope in the vicinity of the band gap. The Rashba momentum, k_R , tracks the distance between the band extrema and the degeneracy point at $(0,0,k_z)$, that is, the temperature-dependent band gap location within a given k_z

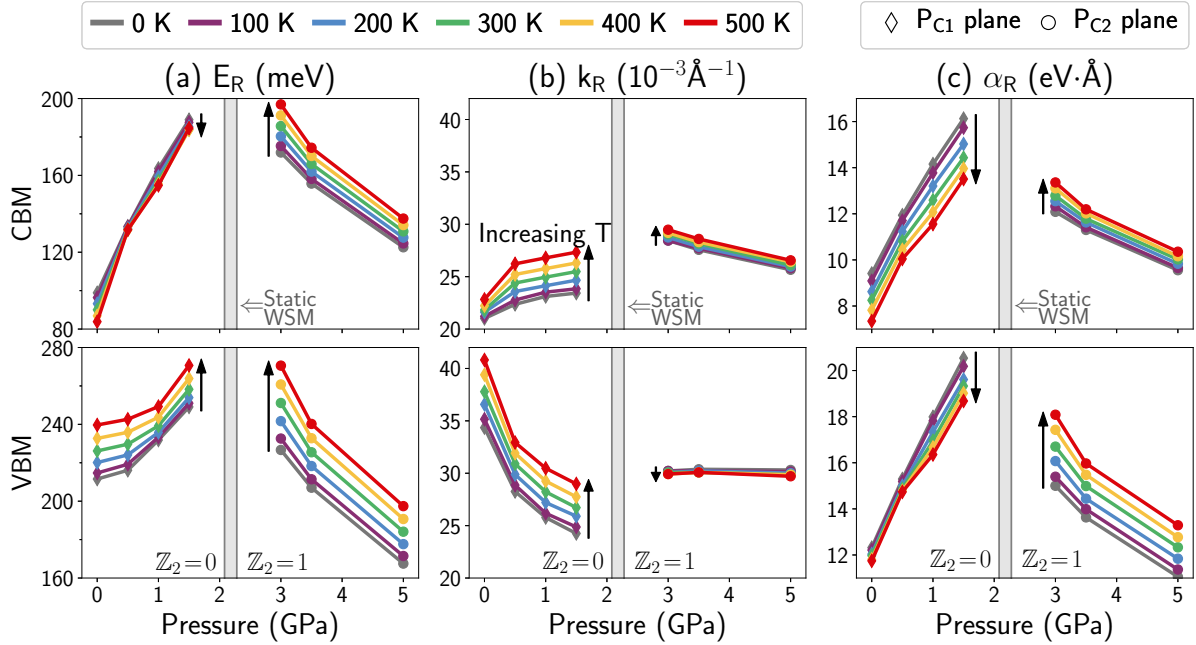


Figure 8.10 – Pressure and temperature-dependent Rashba parameters, for conduction (top panels) and valence bands (bottom panels), near the minimum band gap of the P_{C1} (diamond markers) and P_{C2} (round markers) planes, as defined in Fig. 8.2(b). Note that for $P < P_{C1}$ these quantities were computed in the H-A direction in the P_{C1} plane, hence the difference with the data in Table 8.3 for the conduction band at ambient pressure, where it was computed in the A-L direction for comparison with experimental data. For $P > P_{C2}$ it was computed in the A-L direction in the P_{C2} plane.

plane. Finally, the Rashba parameter, α_R , captures the strength of the Rashba interaction (Eq. (A1)), which is related to the potential gradient along the \hat{c} axis [106].

Fig. 8.10 presents the results for the Rashba energy (a), Rashba momentum (b) and Rashba parameter (c), for both the valence (bottom panels) and conduction bands (upper panels). We emphasize that these parameters were computed for the minimal band gap, namely on the (H-A, P_{C1}) line for the trivial phase and on the (A-L, P_{C2}) for the TI phase. Hence, the numerical values for ambient pressure in this figure should not be compared to experimental data, as available ARPES measurements were done on the (A-L, P_{C1}) line. The trends reported here nevertheless agree with those of Monserrat and Vanderbilt [53] for the (A-L, $k_z = \frac{\pi}{c}$) direction.

As can be seen in Fig. 8.10, pressure has a stronger effect than temperature on the Rashba parameters. This behavior can be qualitatively understood as increasing temperature enhances atomic vibrations around their slightly shifted equilibrium positions, while pressure modifies more significantly these equilibrium positions by reducing the distance between the atomic planes. Hence, for a given temperature, increasing pressure enhances the potential

gradient along the \hat{c} axis, thus increasing α_R . On the contrary, for a given pressure, temperature increases the complex interplay between the electrons and the lattice, yielding a reduction of α_R . These are the behaviors observed in the trivial phase [Fig. 8.10(c), diamond markers].

The TI phase displays the opposite behavior [Fig. 8.10(c), round markers]: α_R is reduced by pressure and increased by temperature. To understand this behavior, we must analyze more carefully both the Rashba energy and momentum, as α_R captures the interplay between those two quantities. On the one hand, for all pressures, the Rashba energy [Fig. 8.10(a)] almost steadily increases with temperature, except for the CB in the trivial phase, which shows a very weak decrease with temperature. On the other hand, the Rashba momentum is almost unaffected by pressure and temperature in the TI phase [Fig. 8.10(b), round markers], while it increases almost steadily with temperature for pressures below the TPT.

The opposite trends of α_R between the trivial and TI phase can, therefore, be understood by two observations. On the one hand, in both phases, the pressure dependence of α_R is governed by the Rashba energy, which increases in the trivial phase and decreases in the TI phase. On the other, their opposite temperature dependence is caused by the Rashba momentum. In the trivial phase, both E_R and k_R increase with temperature, the latter being more significant than the former, hence the decrease of α_R . In the TI phase, the increase of α_R captures the change of the Rashba energy, as k_R is scarcely affected by temperature. We cannot, however, determine if this behavior is linked to the TPT or merely a consequence of an increased bonding along \hat{c} induced by increased pressure. This question could be addressed by investigating the temperature dependence of the high-pressure Rashba splitting in other bismuth tellurohalides, BiTeBr and BiTeCl, which are not known to exhibit a pressure-induced TPT [46].

Lastly, we verified that the trends presented in Fig. 8.10 for the TI phase do not vary if the temperature-dependent band structure is evaluated in the P_{C1} plane rather than the P_{C2} plane, remaining in the A-L direction. Hence, it seems reasonable that these trends, including the quasi temperature-independence of the Rashba momentum, could be applied to any k_z between those planes. We also verified that the temperature-dependent values of P_{C2} are not significantly altered if considering the EPI and TE corrections computed in the P_{C1} plane.

C. Temperature-dependent Weyl semimetal phase width

In Table 8.6, we report the numerical values for the temperature-dependent WSM phase boundaries, namely $P_{C1}(T)$ and $P_{C2}(T)$. These values were extracted from Fig. 8.7 using the extrapolation procedure described in Sec. 8.4.3 and are explicitly plotted in Fig. 8.8.

Table 8.6 – Temperature-dependent WSM phase boundaries, obtained with the extrapolation procedure described in Sec. 8.4.3. P_{C1} marks the phase transition from a trivial insulator to a WSM, while P_{C2} marks the transition from WSM to Z_2 topological insulator. The EPI+TE values are plotted in Fig. 8.8 (black markers). The middle and left sections display the critical pressures obtained when considering only the EPI or TE contribution to the renormalization.

Temperature (K)	EPI only			TE only			EPI+TE		
	P_{C1} (GPa)	P_{C2} (GPa)	width (GPa)	P_{C1} (GPa)	P_{C2} (GPa)	width (GPa)	P_{C1} (GPa)	P_{C2} (GPa)	width (GPa)
Static	2.08	2.28	0.20	2.08	2.28	0.20	2.08	2.28	0.20
0	2.08	2.31	0.23	2.08	2.28	0.20	2.08	2.31	0.23
100	2.10	2.42	0.32	2.09	2.39	0.30	2.14	2.53	0.39
200	2.14	2.57	0.44	2.16	2.55	0.39	2.24	2.82	0.58
300	2.17	2.73	0.56	2.24	2.70	0.46	2.30	3.09	0.79
400	2.20	2.89	0.69	2.33	2.83	0.50	2.32	3.33	1.01
500	2.22	3.05	0.83	2.46	2.95	0.51	2.30	3.55	1.25

For comparison, we also include the temperature-dependent critical pressures obtained by considering the EPI and TE corrections independently.

Bibliography

- [1] I. Garate. Phonon-Induced Topological Transitions and Crossovers in Dirac Materials. *Phys. Rev. Lett.*, 110(4):046402, 2013. doi: [10.1103/PhysRevLett.110.046402](https://doi.org/10.1103/PhysRevLett.110.046402).
- [2] K. Saha and I. Garate. Phonon-induced topological insulation. *Phys. Rev. B*, 89(20):205103, 2014. doi: [10.1103/PhysRevB.89.205103](https://doi.org/10.1103/PhysRevB.89.205103).
- [3] G. Antonius and S. G. Louie. Temperature-Induced Topological Phase Transitions: Promoted versus Suppressed Nontrivial Topology. *Phys. Rev. Lett.*, 117(24):246401, 2016. doi: [10.1103/PhysRevLett.117.246401](https://doi.org/10.1103/PhysRevLett.117.246401).
- [4] B. Monserrat and D. Vanderbilt. Temperature Effects in the Band Structure of Topological Insulators. *Phys. Rev. Lett.*, 117(22):226801, 2016. doi: [10.1103/PhysRevLett.117.226801](https://doi.org/10.1103/PhysRevLett.117.226801).
- [5] M. Z. Hasan and C. L. Kane. Colloquium : Topological insulators. *Rev. Mod. Phys.*, 82(4):3045–3067, 2010. doi: [10.1103/RevModPhys.82.3045](https://doi.org/10.1103/RevModPhys.82.3045).
- [6] M. Z. Hasan, S. Y. Xu, and G. Bian. Topological insulators, topological superconductors and Weyl fermion semimetals: discoveries, perspectives and outlooks. *Phys. Scr.*, T164:014001, 2015. doi: [10.1088/0031-8949/2015/T164/014001](https://doi.org/10.1088/0031-8949/2015/T164/014001).
- [7] S. Y. Xu, Y. Xia, L. A. Wray, S. Jia, F. Meier, J. H. Dil, J. Osterwalder, B. Slomski, A. Bansil, H. Lin, R. J. Cava, and M. Z. Hasan. Topological Phase Transition and

- Texture Inversion in a Tunable Topological Insulator. *Science*, 332(6029):560–564, 2011. doi: [10.1126/science.1201607](https://doi.org/10.1126/science.1201607).
- [8] B. M. Wojek, P. Dziawa, B. J. Kowalski, A. Szcerbakow, A. M. Black-Schaffer, M. H. Berntsen, T. Balasubramanian, T. Story, and O. Tjernberg. Band inversion and the topological phase transition in (Pb,Sn)Se. *Phys. Rev. B*, 90(16):161202, 2014. doi: [10.1103/PhysRevB.90.161202](https://doi.org/10.1103/PhysRevB.90.161202).
- [9] P. Dziawa, B. J. Kowalski, K. Dybko, R. Buczko, A. Szcerbakow, M. Szot, E. Łusakowska, T. Balasubramanian, B. M. Wojek, M. H. Berntsen, O. Tjernberg, and T. Story. Topological crystalline insulator states in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$. *Nat. Mater.*, 11(12):1023–1027, 2012. doi: [10.1038/nmat3449](https://doi.org/10.1038/nmat3449).
- [10] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan. A topological Dirac insulator in a quantum spin Hall phase. *Nature*, 452(7190):970–974, 2008. doi: [10.1038/nature06843](https://doi.org/10.1038/nature06843).
- [11] X. Xi, C. Ma, Z. Liu, Z. Chen, W. Ku, H. Berger, C. Martin, D. B. Tanner, and G. L. Carr. Signatures of a Pressure-Induced Topological Quantum Phase Transition in BiTeI. *Phys. Rev. Lett.*, 111(15):155701, 2013. doi: [10.1103/PhysRevLett.111.155701](https://doi.org/10.1103/PhysRevLett.111.155701).
- [12] A. Bera, K. Pal, D. V. S. Muthu, S. Sen, P. Guptasarma, U. V. Waghmare, and A. K. Sood. Sharp Raman Anomalies and Broken Adiabaticity at a Pressure Induced Transition from Band to Topological Insulator in Sb_2Se_3 . *Phys. Rev. Lett.*, 110(10):107401, 2013. doi: [10.1103/PhysRevLett.110.107401](https://doi.org/10.1103/PhysRevLett.110.107401).
- [13] S. s. Li, W. x. Ji, C. w. Zhang, P. Li, and P. j. Wang. Robust room-temperature inversion-asymmetry topological transitions in functionalized HgSe monolayer. *J. Mater. Chem. C*, 4(11):2243–2251, 2016. doi: [10.1039/C6TC00020G](https://doi.org/10.1039/C6TC00020G).
- [14] Y. Liu, Y. Y. Li, S. Rajput, D. Gilks, L. Lari, P. L. Galindo, M. Weinert, V. K. Lazarov, and L. Li. Tuning Dirac states by strain in the topological insulator Bi_2Se_3 . *Nat. Phys.*, 10(4):294–299, 2014. doi: [10.1038/nphys2898](https://doi.org/10.1038/nphys2898).
- [15] Q. Liu, X. Zhang, L. B. Abdalla, A. Fazio, and A. Zunger. Switching a Normal Insulator into a Topological Insulator via Electric Field with Application to Phosphorene. *Nano Lett.*, 15(2):1222–1228, 2015. doi: [10.1021/nl5043769](https://doi.org/10.1021/nl5043769).
- [16] T. Zhang, J. Ha, N. Levy, Y. Kuk, and J. Stroscio. Electric-Field Tuning of the Surface Band Structure of Topological Insulator Sb_2Te_3 Thin Films. *Phys. Rev. Lett.*, 111(5):056803, 2013. doi: [10.1103/PhysRevLett.111.056803](https://doi.org/10.1103/PhysRevLett.111.056803).
- [17] Y. H. Wang, H. Steinberg, P. Jarillo-Herrero, and N. Gedik. Observation of Floquet-Bloch States on the Surface of a Topological Insulator. *Science*, 342(6157):453–457, 2013. doi: [10.1126/science.1239834](https://doi.org/10.1126/science.1239834).
- [18] H. Hübener, M. A. Sentef, U. D. Giovannini, A. F. Kemper, and A. Rubio. Creating stable Floquet–Weyl semimetals by laser-driving of 3D Dirac materials. *Nat. Commun.*, 8:13940, 2017. doi: [10.1038/ncomms13940](https://doi.org/10.1038/ncomms13940).

- [19] L. Fleet. Topological transistor. *Nat Phys*, 11(1):5, 2015. doi: [10.1038/nphys3217](https://doi.org/10.1038/nphys3217).
- [20] H. Zhu, C. A. Richter, E. Zhao, J. E. Bonevich, W. A. Kimes, H. J. Jang, H. Yuan, H. Li, A. Arab, O. Kirillov, J. E. Maslar, D. E. Ioannou, and Q. Li. Topological Insulator Bi₂Se₃ Nanowire High Performance Field-Effect Transistors. *Sci. Rep.*, 3: 1757, 2013. doi: [10.1038/srep01757](https://doi.org/10.1038/srep01757).
- [21] J. J. Zhou, W. Feng, Y. Zhang, S. A. Yang, and Y. Yao. Engineering Topological Surface States and Giant Rashba Spin Splitting in BiTeI/Bi₂Te₃ Heterostructures. *Sci. Rep.*, 4:3841, 2014. doi: [10.1038/srep03841](https://doi.org/10.1038/srep03841).
- [22] J. Alicea. Majorana fermions in a tunable semiconductor device. *Phys. Rev. B*, 81(12): 125318, 2010. doi: [10.1103/PhysRevB.81.125318](https://doi.org/10.1103/PhysRevB.81.125318).
- [23] Y. A. Bychkov and E. I. Rashba. Properties of a 2D electron gas with lifted spectral degeneracy. *J. Exp. Theor. Phys.*, 39(2):78–81, 1984, http://jetpletters.ru/ps/1264/article_19121.pdf.
- [24] A. Manchon, H. C. Koo, J. Nitta, S. M. Frolov, and R. A. Duine. New perspectives for Rashba spin–orbit coupling. *Nat. Mater.*, 14(9):871–882, 2015. doi: [10.1038/nmat4360](https://doi.org/10.1038/nmat4360).
- [25] M. S. Bahramy and N. Ogawa. Bulk Rashba Semiconductors and Related Quantum Phenomena. *Adv. Mater.*, 29(25):1605911, 2017. doi: <https://doi.org/10.1002/adma.201605911>.
- [26] H. Murakawa, M. S. Bahramy, M. Tokunaga, Y. Kohama, C. Bell, Y. Kaneko, N. Nagaosa, H. Y. Hwang, and Y. Tokura. Detection of Berry’s Phase in a Bulk Rashba Semiconductor. *Science*, 342(6165):1490–1493, 2013. doi: [10.1126/science.1242247](https://doi.org/10.1126/science.1242247).
- [27] I. Žutić, J. Fabian, and S. Das Sarma. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76(2):323–410, 2004. doi: [10.1103/RevModPhys.76.323](https://doi.org/10.1103/RevModPhys.76.323).
- [28] S. Das Sarma, J. Fabian, X. Hu, and I. Žutić. Spin electronics and spin computation. *Solid State Commun.*, 119(4):207–215, 2001. doi: [10.1016/S0038-1098\(01\)00111-9](https://doi.org/10.1016/S0038-1098(01)00111-9).
- [29] K. Ishikawa, M. Bahramy, S. Murakawa, M. Sakano, S. Shimojima, and T. Sonobe. Giant Rashba-type spin splitting in bulk BiTeI. *Nat. Mater.*, 10(3051):3051, 2011. doi: [10.1038/NMAT3051](https://doi.org/10.1038/NMAT3051).
- [30] A. Crepaldi, L. Moreschini, G. Autès, C. Tournier-Colletta, S. Moser, N. Virk, H. Berger, P. Bugnon, Y. J. Chang, K. Kern, A. Bostwick, E. Rotenberg, O. V. Yazyev, and M. Grioni. Giant Ambipolar Rashba Effect in the Semiconductor BiTeI. *Phys. Rev. Lett.*, 109(9):096803, 2012. doi: [10.1103/PhysRevLett.109.096803](https://doi.org/10.1103/PhysRevLett.109.096803).
- [31] G. Landolt, S. V. Eremeev, Y. M. Koroteev, B. Slomski, S. Muff, T. Neupert, M. Kobayashi, V. N. Strocov, T. Schmitt, Z. S. Aliev, M. B. Babanly, I. R. Amiraslanov, E. V. Chulkov, J. Osterwalder, and J. H. Dil. Disentanglement of Surface and Bulk Rashba Spin Splittings in Noncentrosymmetric BiTeI. *Phys. Rev. Lett.*, 109(11):116403, 2012. doi: [10.1103/PhysRevLett.109.116403](https://doi.org/10.1103/PhysRevLett.109.116403).

- [32] M. Sakano, M. S. Bahramy, A. Katayama, T. Shimojima, H. Murakawa, Y. Kaneko, W. Malaeb, S. Shin, K. Ono, H. Kumigashira, R. Arita, N. Nagaosa, H. Y. Hwang, Y. Tokura, and K. Ishizaka. Strongly Spin-Orbit Coupled Two-Dimensional Electron Gas Emerging near the Surface of Polar Semiconductors. *Phys. Rev. Lett.*, 110(10):107204, 2013. doi: [10.1103/PhysRevLett.110.107204](https://doi.org/10.1103/PhysRevLett.110.107204).
- [33] M. S. Bahramy, R. Arita, and N. Nagaosa. Origin of giant bulk Rashba splitting: Application to BiTeI. *Phys. Rev. B*, 84(4):041202, 2011. doi: [10.1103/PhysRevB.84.041202](https://doi.org/10.1103/PhysRevB.84.041202).
- [34] S. V. Eremeev, I. A. Nechaev, Y. M. Koroteev, P. M. Echenique, and E. V. Chulkov. Ideal Two-Dimensional Electron Systems with a Giant Rashba-Type Spin Splitting in Real Materials: Surfaces of Bismuth Tellurohalides. *Phys. Rev. Lett.*, 108(24):246802, 2012. doi: [10.1103/PhysRevLett.108.246802](https://doi.org/10.1103/PhysRevLett.108.246802).
- [35] M. Bahramy, B. Yang, R. Arita, and N. Nagaosa. Emergence of non-centrosymmetric topological insulating phase in BiTeI under pressure. *Nat. Commun.*, 3:679, 2012. doi: [10.1038/ncomms1679](https://doi.org/10.1038/ncomms1679).
- [36] M. K. Tran, J. Levallois, P. Lerch, J. Teyssier, A. B. Kuzmenko, G. Autès, O. V. Yazyev, A. Ubaldini, E. Giannini, D. v. d. Marel, and A. Akrap. Infrared- and Raman-Spectroscopy Measurements of a Transition in the Crystal Structure and a Closing of the Energy Gap of BiTeI under Pressure. *Phys. Rev. Lett.*, 112(4):047402, 2014. doi: [10.1103/PhysRevLett.112.047402](https://doi.org/10.1103/PhysRevLett.112.047402).
- [37] A. A. Makhnev, L. V. Nomerovannaya, T. V. Kuznetsova, O. E. Tereshchenko, and K. A. Kokh. Optical properties of BiTeI semiconductor with a strong Rashba spin-orbit interaction. *Opt. Spectrosc.*, 117(5):764–768, 2014. doi: [10.1134/S0030400X14110125](https://doi.org/10.1134/S0030400X14110125).
- [38] M. L. Jin, F. Sun, L. Y. Xing, S. J. Zhang, S. M. Feng, P. P. Kong, W. M. Li, X. C. Wang, J. L. Zhu, Y. W. Long, H. Y. Bai, C. Z. Gu, R. C. Yu, W. G. Yang, G. Y. Shen, Y. S. Zhao, H. K. Mao, and C. Q. Jin. Superconductivity Bordering Rashba Type Topological Transition. *Sci. Rep.*, 7:39699, 2017. doi: [10.1038/srep39699](https://doi.org/10.1038/srep39699).
- [39] Y. Qi, W. Shi, P. G. Naumov, N. Kumar, R. Sankar, W. Schnelle, C. Shekhar, F. C. Chou, C. Felser, B. Yan, and S. A. Medvedev. Topological Quantum Phase Transition and Superconductivity Induced by Pressure in the Bismuth Tellurohalide BiTeI. *Adv. Mater.*, 29(18):1605965, 2017. doi: [10.1002/adma.201605965](https://doi.org/10.1002/adma.201605965).
- [40] T. Ideue, J. G. Checkelsky, M. S. Bahramy, H. Murakawa, Y. Kaneko, N. Nagaosa, and Y. Tokura. Pressure variation of Rashba spin splitting toward topological transition in the polar semiconductor BiTeI. *Phys. Rev. B*, 90(16):161107, 2014. doi: [10.1103/PhysRevB.90.161107](https://doi.org/10.1103/PhysRevB.90.161107).
- [41] C. Martin, E. D. Mun, H. Berger, V. S. Zapf, and D. B. Tanner. Quantum oscillations and optical conductivity in Rashba spin-splitting BiTeI. *Phys. Rev. B*, 87(4):041104, 2013. doi: [10.1103/PhysRevB.87.041104](https://doi.org/10.1103/PhysRevB.87.041104).

- [42] J. Park, K. H. Jin, Y. J. Jo, E. S. Choi, W. Kang, E. Kampert, J. S. Rhyee, S. H. Jhi, and J. S. Kim. Quantum Oscillation Signatures of Pressure-induced Topological Phase Transition in BiTeI. *Sci. Rep.*, 5:15973, 2015. doi: [10.1038/srep15973](https://doi.org/10.1038/srep15973).
- [43] D. VanGennep, S. Maiti, D. Graf, S. W. Tozer, C. Martin, H. Berger, D. L. Maslov, and J. J. Hamlin. Pressure tuning the Fermi level through the Dirac point of giant Rashba semiconductor BiTeI. *J. Phys.: Condens. Matter*, 26(34):342202, 2014. doi: [10.1088/0953-8984/26/34/342202](https://doi.org/10.1088/0953-8984/26/34/342202).
- [44] Y. S. Ponosov, T. V. Kuznetsova, O. E. Tereshchenko, K. A. Kokh, and E. V. Chulkov. Dynamics of the BiTeI lattice at high pressures. *JETP Lett.*, 98(9):557–561, 2014. doi: [10.1134/S0021364013220074](https://doi.org/10.1134/S0021364013220074).
- [45] J. Liu and D. Vanderbilt. Weyl semimetals from noncentrosymmetric topological insulators. *Phys. Rev. B*, 90(15):155316, 2014. doi: [10.1103/PhysRevB.90.155316](https://doi.org/10.1103/PhysRevB.90.155316).
- [46] I. P. Rusinov, T. V. Menshchikova, I. Y. Sklyadneva, R. Heid, K. P. Bohnen, and E. V. Chulkov. Pressure effects on crystal and electronic structure of bismuth tellurohalides. *New J. Phys.*, 18(11):113003, 2016. doi: [10.1088/1367-2630/18/11/113003](https://doi.org/10.1088/1367-2630/18/11/113003).
- [47] J. I. Facio, D. Efremov, K. Koepernik, J. S. You, I. Sodemann, and J. van den Brink. Strongly Enhanced Berry Dipole at Topological Phase Transitions in BiTeI. *Phys. Rev. Lett.*, 121(24):246403, 2018. doi: [10.1103/PhysRevLett.121.246403](https://doi.org/10.1103/PhysRevLett.121.246403).
- [48] S. Güler-Kılıç and C. Kılıç. Pressure dependence of the band-gap energy in BiTeI. *Phys. Rev. B*, 94:165203, 2016. doi: <https://doi.org/10.1103/PhysRevB.94.165203>.
- [49] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, Cambridge, 2004. doi: [10.1017/CBO9780511805769](https://doi.org/10.1017/CBO9780511805769).
- [50] Z. Li and J. P. Carbotte. Conductivity of Dirac fermions with phonon-induced topological crossover. *Phys. Rev. B*, 88(19):195133, 2013. doi: [10.1103/PhysRevB.88.195133](https://doi.org/10.1103/PhysRevB.88.195133).
- [51] A. A. Reijnders, Y. Tian, L. J. Sandilands, G. Pohl, I. D. Kivlichan, S. Y. F. Zhao, S. Jia, M. E. Charles, R. J. Cava, N. Alidoust, S. Xu, M. Neupane, M. Z. Hasan, X. Wang, S. W. Cheong, and K. S. Burch. Optical evidence of surface state suppression in Bi-based topological insulators. *Phys. Rev. B*, 89(7):075138, 2014. doi: [10.1103/PhysRevB.89.075138](https://doi.org/10.1103/PhysRevB.89.075138).
- [52] B. M. Wojek, M. H. Berntsen, V. Jonsson, A. Szczerbakow, P. Dziawa, B. J. Kowalski, T. Story, and O. Tjernberg. Direct observation and temperature control of the surface Dirac gap in a topological crystalline insulator. *Nat. Commun.*, 6:8463, 2015. doi: [10.1038/ncomms9463](https://doi.org/10.1038/ncomms9463).
- [53] B. Monserrat and D. Vanderbilt. Temperature dependence of the bulk Rashba splitting in the bismuth tellurohalides. *Phys. Rev. Mater.*, 1(5):054201, 2017. doi: [10.1103/PhysRevMaterials.1.054201](https://doi.org/10.1103/PhysRevMaterials.1.054201).
- [54] P. B. Allen and M. Cardona. Temperature dependence of the direct gap of Si and Ge. *Phys. Rev. B*, 27(8):4760–4769, 1983. doi: [10.1103/PhysRevB.27.4760](https://doi.org/10.1103/PhysRevB.27.4760).

- [55] G. Mahan. *Many-Particle Physics*. Plenum, New York, 2nd edition, 1990.
- [56] G. Antonius, S. Poncé, E. Lantagne-Hurtubise, G. Auclair, X. Gonze, and M. Côté. Dynamical and anharmonic effects on the electron-phonon coupling and the zero-point renormalization of the electronic structure. *Phys. Rev. B*, 92(8):085137, 2015. doi: [10.1103/PhysRevB.92.085137](https://doi.org/10.1103/PhysRevB.92.085137).
- [57] P. B. Allen and V. Heine. Theory of the temperature dependence of electronic band structures. *J. Phys. C: Solid State Phys.*, 9(12):2305, 1976. doi: [10.1088/0022-3719/9/12/013](https://doi.org/10.1088/0022-3719/9/12/013).
- [58] P. B. Allen and M. Cardona. Theory of the temperature dependence of the direct gap of germanium. *Phys. Rev. B*, 23(4):1495–1505, 1981. doi: [10.1103/PhysRevB.23.1495](https://doi.org/10.1103/PhysRevB.23.1495).
- [59] S. Poncé, G. Antonius, Y. Gillet, P. Boulanger, J. Laflamme Janssen, A. Marini, M. Côté, and X. Gonze. Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation. *Phys. Rev. B*, 90(21):214304, 2014. doi: [10.1103/PhysRevB.90.214304](https://doi.org/10.1103/PhysRevB.90.214304).
- [60] S. Poncé, Y. Gillet, J. L. Janssen, A. Marini, M. Verstraete, and X. Gonze. Temperature dependence of the electronic structure of semiconductors and insulators. *J. Chem. Phys.*, 143(10):102813, 2015. doi: [10.1063/1.4927081](https://doi.org/10.1063/1.4927081).
- [61] F. Giustino. Electron-phonon interactions from first principles. *Rev. Mod. Phys.*, 89(1):015003, 2017. doi: [10.1103/RevModPhys.89.015003](https://doi.org/10.1103/RevModPhys.89.015003).
- [62] A. Marini, S. Poncé, and X. Gonze. Many-body perturbation theory approach to the electron-phonon interaction with density-functional theory as a starting point. *Phys. Rev. B*, 91(22):224310, 2015. doi: [10.1103/PhysRevB.91.224310](https://doi.org/10.1103/PhysRevB.91.224310).
- [63] X. Gonze, P. Boulanger, and M. Côté. Theoretical approaches to the temperature and zero-point motion effects on the electronic band structure. *Ann. Phys.*, 523(1-2): 168–178, 2011. doi: [10.1002/andp.201000100](https://doi.org/10.1002/andp.201000100).
- [64] G. Grimvall. *Thermophysical properties of materials*. Number XVIII in Selected topics in solid state physics. North-Holland Physics Publishing, Amsterdam, 1986.
- [65] G. Srivastava. *The Physics of Phonons*. A. Hilger, New York, NY, USA, 1990.
- [66] P. B. Allen. Theory of thermal expansion: Quasi-harmonic approximation and corrections from quasiparticle renormalization. *Modern Physics Letters B*, 34(02):2050025, 2020. doi: [10.1142/S0217984920500256](https://doi.org/10.1142/S0217984920500256).
- [67] R. W. Munn. Role of the elastic constants in negative thermal expansion of axial solids. *J. Phys. C: Solid State Phys.*, 5(5):535–542, 1972. doi: [10.1088/0022-3719/5/5/005](https://doi.org/10.1088/0022-3719/5/5/005).
- [68] E. T. Ritz and N. A. Benedek. Interplay between Phonons and Anisotropic Elasticity Drives Negative Thermal Expansion in PbTiO₃. *Phys. Rev. Lett.*, 121(25):255901, 2018. doi: [10.1103/PhysRevLett.121.255901](https://doi.org/10.1103/PhysRevLett.121.255901).
- [69] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J. M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac,

- J. B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. L. Marques, A. Martin, C. Martins, H. P. C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G. M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. Van Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger. The ABINIT project: Impact, environment and recent developments. *Comput. Phys. Commun.*, page 107042, 2020. doi: [10.1016/j.cpc.2019.107042](https://doi.org/10.1016/j.cpc.2019.107042).
- [70] X. Gonze. First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm. *Phys. Rev. B*, 55(16):10337–10354, 1997. doi: [10.1103/PhysRevB.55.10337](https://doi.org/10.1103/PhysRevB.55.10337).
- [71] X. Gonze and C. Lee. Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. *Phys. Rev. B*, 55(16):10355–10368, 1997. doi: [10.1103/PhysRevB.55.10355](https://doi.org/10.1103/PhysRevB.55.10355).
- [72] I. Y. Sklyadneva, R. Heid, K. P. Bohnen, V. Chis, V. A. Volodin, K. A. Kokh, O. E. Tereshchenko, P. M. Echenique, and E. V. Chulkov. Lattice dynamics of bismuth tellurohalides. *Phys. Rev. B*, 86(9):094302, 2012. doi: [10.1103/PhysRevB.86.094302](https://doi.org/10.1103/PhysRevB.86.094302).
- [73] C. Hartwigsen, S. Goedecker, and J. Hutter. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B*, 58(7):3641–3662, 1998. doi: [10.1103/PhysRevB.58.3641](https://doi.org/10.1103/PhysRevB.58.3641).
- [74] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77(18):3865–3868, 1996. doi: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [75] S. Güler-Kılıç and C. Kılıç. Crystal and electronic structure of BiTeI, AuTeI, and PdTeI compounds: A dispersion-corrected density-functional study. *Phys. Rev. B*, 91(24):245204, 2015. doi: [10.1103/PhysRevB.91.245204](https://doi.org/10.1103/PhysRevB.91.245204).
- [76] F. D. Murnaghan. The compressibility of media under extreme pressures. *PNAS*, 30(9):244–247, 1944. doi: [10.1073/pnas.30.9.244](https://doi.org/10.1073/pnas.30.9.244).
- [77] J. P. Nery, P. B. Allen, G. Antonius, L. Reining, A. Miglio, and X. Gonze. Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron. *Phys. Rev. B*, 97(11):115145, 2018. doi: [10.1103/PhysRevB.97.115145](https://doi.org/10.1103/PhysRevB.97.115145).
- [78] D. R. Hamann, X. Wu, K. M. Rabe, and D. Vanderbilt. Metric tensor formulation of strain in density-functional perturbation theory. *Phys. Rev. B*, 71(3):035117, 2005. doi: [10.1103/PhysRevB.71.035117](https://doi.org/10.1103/PhysRevB.71.035117).
- [79] X. Wu, D. Vanderbilt, and D. R. Hamann. Systematic treatment of displacements, strains, and electric fields in density-functional perturbation theory. *Phys. Rev. B*, 72(3):035105, 2005. doi: [10.1103/PhysRevB.72.035105](https://doi.org/10.1103/PhysRevB.72.035105).

- [80] D. Gresch, G. Autès, O. V. Yazyev, M. Troyer, D. Vanderbilt, B. A. Bernevig, and A. A. Soluyanov. Z2pack: Numerical implementation of hybrid Wannier centers for identifying topological materials. *Phys. Rev. B*, 95(7):075146, 2017. doi: [10.1103/PhysRevB.95.075146](https://doi.org/10.1103/PhysRevB.95.075146).
- [81] A. A. Soluyanov and D. Vanderbilt. Computing topological invariants without inversion symmetry. *Phys. Rev. B*, 83(23):235401, 2011. doi: [10.1103/PhysRevB.83.235401](https://doi.org/10.1103/PhysRevB.83.235401).
- [82] Y. Chen, X. Xi, W. L. Yim, F. Peng, Y. Wang, H. Wang, Y. Ma, G. Liu, C. Sun, C. Ma, Z. Chen, and H. Berger. High-Pressure Phase Transitions and Structures of Topological Insulator BiTeI. *J. Phys. Chem. C*, 117(48):25677–25683, 2013. doi: [10.1021/jp409824g](https://doi.org/10.1021/jp409824g).
- [83] J. C. Budich and S. Diehl. Topology of density matrices. *Phys. Rev. B*, 91(16):165140, 2015. doi: [10.1103/PhysRevB.91.165140](https://doi.org/10.1103/PhysRevB.91.165140).
- [84] Z. Huang and D. P. Arovas. Topological Indices for Open and Thermal Systems Via Uhlmann’s Phase. *Phys. Rev. Lett.*, 113(7):076407, 2014. doi: [10.1103/PhysRevLett.113.076407](https://doi.org/10.1103/PhysRevLett.113.076407).
- [85] O. Viyuela, A. Rivas, and M. A. Martin-Delgado. Two-Dimensional Density-Matrix Topological Fermionic Phases: Topological Uhlmann Numbers. *Phys. Rev. Lett.*, 113(7):076408, 2014. doi: [10.1103/PhysRevLett.113.076408](https://doi.org/10.1103/PhysRevLett.113.076408).
- [86] T. C. Lu and T. Grover. Singularity in entanglement negativity across finite-temperature phase transitions. *Phys. Rev. B*, 99(7):075157, 2019. doi: [10.1103/PhysRevB.99.075157](https://doi.org/10.1103/PhysRevB.99.075157).
- [87] V. Brousseau-Couture and M. Côté. Influence of the Electron–Phonon Interaction on the Topological Phase Transition in BiTeI. In M. B. Paranjape, R. MacKenzie, Z. Thomova, P. Winternitz, and W. Witczak-Krempa, editors, *Quantum Theory and Symmetries*, CRM Series in Mathematical Physics, pages 305–313, Cham, 2021. Springer International Publishing. doi: [10.1007/978-3-030-55777-5_29](https://doi.org/10.1007/978-3-030-55777-5_29).
- [88] M. Cardona and M. L. W. Thewalt. Isotope effects on the optical spectra of semiconductors. *Rev. Mod. Phys.*, 77(4):1173–1224, 2005. doi: [10.1103/RevModPhys.77.1173](https://doi.org/10.1103/RevModPhys.77.1173).
- [89] N. A. Pike and O. M. Løvvik. Calculation of the anisotropic coefficients of thermal expansion: A first-principles approach. *Comput. Mater. Sci.*, 167:257–263, 2019. doi: [10.1016/j.commatsci.2019.05.045](https://doi.org/10.1016/j.commatsci.2019.05.045).
- [90] H. Fröhlich. Electrons in Lattice Fields. *Adv. Phys.*, 3(11):325–361, 1954. doi: [10.1080/00018735400101213](https://doi.org/10.1080/00018735400101213).
- [91] J. P. Nery and P. B. Allen. Influence of Fröhlich polaron coupling on renormalized electron bands in polar semiconductors: Results for zinc-blende GaN. *Phys. Rev. B*, 94(11):115135, 2016. doi: [10.1103/PhysRevB.94.115135](https://doi.org/10.1103/PhysRevB.94.115135).
- [92] A. Miglio, V. Brousseau-Couture, E. Godbout, G. Antonius, Y. H. Chan, S. G. Louie, M. Côté, M. Giantomassi, and X. Gonze. Predominance of non-adiabatic effects in

- zero-point renormalization of the electronic band gap. *npj Computational Materials*, 6 (1):167, 2020. doi: [10.1038/s41524-020-00434-z](https://doi.org/10.1038/s41524-020-00434-z).
- [93] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12):121102(R), 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).
- [94] I. P. Rusinov, I. A. Nechaev, S. V. Eremeev, C. Friedrich, S. Blügel, and E. V. Chulkov. Many-body effects on the Rashba-type spin splitting in bulk bismuth tellurohalides. *Phys. Rev. B*, 87(20):205103, 2013. doi: [10.1103/PhysRevB.87.205103](https://doi.org/10.1103/PhysRevB.87.205103).
- [95] I. P. Rusinov, O. E. Tereshchenko, K. A. Kokh, A. R. Shakhmametova, I. A. Azarov, and E. V. Chulkov. Role of anisotropy and spin-orbit interaction in the optical and dielectric properties of BiTeI and BiTeCl compounds. *JETP Lett.*, 101(8):507–512, 2015. doi: [10.1134/S0021364015080147](https://doi.org/10.1134/S0021364015080147).
- [96] S. V. Eremeev, I. A. Nechaev, and E. V. Chulkov. Two- and three-dimensional topological phases in BiTeX compounds. *Phys. Rev. B*, 96(15):155309, 2017. doi: [10.1103/PhysRevB.96.155309](https://doi.org/10.1103/PhysRevB.96.155309).
- [97] S. Schwalbe, R. Wirnata, R. Starke, G. A. H. Schober, and J. Kortus. *Ab initio* electronic structure and optical conductivity of bismuth tellurohalides. *Phys. Rev. B*, 94(20):205130, 2016. doi: [10.1103/PhysRevB.94.205130](https://doi.org/10.1103/PhysRevB.94.205130).
- [98] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132(15):154104, 2010. doi: [10.1063/1.3382344](https://doi.org/10.1063/1.3382344).
- [99] E. R. Johnson and A. D. Becke. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *J. Chem. Phys.*, 124(17):174104, 2006. doi: [10.1063/1.2190220](https://doi.org/10.1063/1.2190220).
- [100] S. Grimme, S. Ehrlich, and L. Goerigk. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, 32(7):1456–1465, 2011. doi: [10.1002/jcc.21759](https://doi.org/10.1002/jcc.21759).
- [101] B. Van Troeye, M. Torrent, and X. Gonze. Interatomic force constants including the DFT-D dispersion contribution. *Phys. Rev. B*, 93(14):144304, 2016. doi: [10.1103/PhysRevB.93.144304](https://doi.org/10.1103/PhysRevB.93.144304).
- [102] A. Shevelkov, E. Dikarev, R. Shpanchenko, and B. Popovkin. Crystal Structures of Bismuth Tellurohalides from X-ray powder diffraction data. *J. Solid State Chem.*, 114: 379–384, 1995. doi: <https://doi.org/10.1006/jssc.1995.1058>.
- [103] J. S. Lee, G. A. H. Schober, M. S. Bahramy, H. Murakawa, Y. Onose, R. Arita, N. Nagaosa, and Y. Tokura. Optical Response of Relativistic Electrons in the Polar BiTeI Semiconductor. *Phys. Rev. Lett.*, 107(11):117401, 2011. doi: [10.1103/PhysRevLett.107.117401](https://doi.org/10.1103/PhysRevLett.107.117401).
- [104] M. Sakano, J. Miyawaki, A. Chainani, Y. Takata, T. Sonobe, T. Shimojima, M. Oura, S. Shin, M. S. Bahramy, R. Arita, N. Nagaosa, H. Murakawa, Y. Kaneko, Y. Tokura,

- and K. Ishizaka. Three-dimensional bulk band dispersion in polar BiTeI with giant Rashba-type spin splitting. *Phys. Rev. B*, 86(8):085204, 2012. doi: [10.1103/PhysRevB.86.085204](https://doi.org/10.1103/PhysRevB.86.085204).
- [105] V. Gnezdilov, P. Lemmens, D. Wulferding, A. Möller, P. Recher, H. Berger, R. Sankar, and F. C. Chou. Enhanced quasiparticle dynamics of quantum well states: The giant Rashba system BiTeI and topological insulators. *Phys. Rev. B*, 89(19):195117, 2014. doi: [10.1103/PhysRevB.89.195117](https://doi.org/10.1103/PhysRevB.89.195117).
- [106] M. I. Dyakonov, editor. *Spin physics in semiconductors*. Springer series in solid-state sciences; 157. Springer, Berlin, 2nd edition, 2017. doi: [10.1007/978-3-319-65436-2](https://doi.org/10.1007/978-3-319-65436-2).

Conclusion

Au cours de cette thèse, nous avons exploré le calcul de la dépendance en température de la structure électronique sous différents angles d'approche, en mettant l'accent sur la renormalisation du point zéro de l'énergie de la bande interdite des semi-conducteurs. Nous avons notamment identifié le mécanisme physique à l'origine de l'inexactitude de l'approximation adiabatique pour les matériaux polaires. Nous avons de plus clarifié l'importance de la dilatation du point zéro du réseau et du couplage spin-orbite sur le ZPR de la bande interdite, ainsi que déterminé à travers quelle quantité physique ce dernier affecte principalement la self-énergie de l'interaction électron-phonon. En étudiant systématiquement des ensembles de 20 à 30 matériaux, nous avons pu extraire des tendances générales qui contribueront à accroître notre compréhension de l'effet de ces différents mécanismes. L'exemple du CdTe, qui est un matériau à la fois lourd et ionique, a quant à lui montré que la prise en compte de ces trois interactions est nécessaire pour obtenir une prédiction du ZPR en accord avec la valeur expérimentale. Nous avons enfin illustré comment, à travers l'effet combiné de l'interaction électron-phonon et de la dilatation thermique, la température peut influencer le diagramme de phase des matériaux topologiques.

Plus précisément, nous avons tout d'abord mis en évidence la prédominance des effets non adiabatiques dans le ZPR induit par l'interaction électron-phonon pour les matériaux présentant un degré d'ionicté important et démontré que la prise en compte de ces effets s'avère essentielle pour obtenir un accord avec les données expérimentales. Ce faisant, nous avons remis en question la prédictibilité des méthodes basées sur les supercellules adiabatiques pour ce type de matériau, à moins que celles-ci ne soient accompagnées de techniques permettant d'estimer la contribution manquante provenant de l'interaction Fröhlich.

Nous avons en second lieu établi numériquement l'équivalence entre l'approche de Grüneisen et la minimisation de l'énergie libre pour l'évaluation de la dilatation du point zéro du réseau cristallin, pour des matériaux tant isotropes qu'anisotropes. La contribution du ZPLE peut représenter une fraction substantielle de la renormalisation du point zéro de l'énergie de la bande interdite, y compris dans des matériaux comprenant des noyaux légers, et peut même s'avérer aussi importante que celle de l'interaction électron-phonon. Les deux

contributions doivent donc être traitées sur un même pied d'égalité lorsque le ZPR théorique est comparé aux données expérimentales.

Nous avons ensuite démontré que l'inclusion du couplage spin-orbite réduit le ZPR d'un ensemble substantiel de matériaux cubiques de structure zinc-blende, diamant et rock-salt. L'origine physique de cette diminution est dominée par la variation des énergies propres électroniques induite par le SOC dans l'approximation des noyaux statiques et peut être estimée à l'aide d'un modèle de Fröhlich généralisé à l'intérieur duquel le SOC se manifeste par la diminution des masses effectives.

Nous avons finalement appliqué les méthodologies permettant d'évaluer la dépendance en température de la structure électronique pour construire le diagramme de phase topologique en pression et température du BiTeI, révélant par le fait même que, face à une augmentation de température, l'effet combiné de l'interaction électron-phonon et de la dilatation thermique repousse l'apparition de la phase d'isolant topologique \mathbb{Z}_2 des pressions plus élevées et élargit l'espace de phase correspondant à un semi-métal de Weyl.

Mes travaux sur le BiTeI appellent à considérer de manière plus systématique l'effet de la température lors de la caractérisation numérique des matériaux topologiques. Ce faisant, en s'approchant d'une transition de phase, on se retrouvera inévitablement dans des conditions où l'énergie de la bande interdite devient similaire à l'énergie typique des phonons, ce qui invalide du même coup tant l'approximation diagonale appliquée à la self-énergie de l'interaction électron-phonon que l'omission de la contribution entropique des électrons à l'énergie libre dont on extrait l'expansion thermique. L'étude de cette classe de matériaux demandera donc que l'on généralise l'utilisation de méthodes tenant compte de l'hybridation des états électroniques non perturbés induite par l'interaction électron-phonon, dans l'esprit de ce qu'ont proposé Lihm et Park [1]. De même, il sera important d'étudier plus en profondeur les conséquences de l'approximation qui décorrèle le calcul des contributions EPI et TE au ZPR de la bande interdite. On pourrait notamment se demander s'il serait possible de s'inspirer de l'approche de Grüneisen pour effectuer un développement en série de la self-énergie de l'interaction électron-phonon qui prendrait en compte la dilatation cristalline sans nécessiter une réévaluation complète de la self-énergie à chaque température d'intérêt. Il serait de plus particulièrement intéressant d'étudier si certaines propriétés des phonons (par exemple, leur largeur de bande) seraient en mesure de capturer la signature d'une topologie non triviale [2]. Ceci ouvrirait la porte à de nouvelles méthodes de détection de ces phases exotiques, plus facilement applicables à des matériaux soumis à des contraintes externes.

Par ailleurs, bien que mes travaux se soient essentiellement concentrés sur l'effet des vibrations atomiques sur les niveaux d'énergie électroniques, ces dernières modifient également l'énergie totale du cristal comparativement à l'approximation des noyaux statiques. Les techniques utilisées dans le cadre de cette thèse pourraient donc contribuer, entre autres, à raffiner la prédiction des barrières d'énergie régissant les mécanismes de diffusion d'ions,

qui ont des applications importantes pour la caractérisation des matériaux candidats pour les batteries de nouvelle génération, ainsi que pour l'étude de la propagation des défauts structurels. En effet, les barrières d'énergie sont habituellement évaluées à l'aide de la dynamique moléculaire, qui repose sur des équations classiques, négligeant par le fait même la nature quantique des noyaux. On peut raisonnablement supposer que l'effet risque d'être d'autant plus important pour des matériaux contenant des éléments légers comme de l'oxygène, l'azote ou le soufre, qui sont habituellement associés à des vibrations atomiques de plus grande amplitude et à des corrections dues au mouvement du point zéro plus significatives.

En tant que théoriciennes-numériciennes et théoriciens-numériciens, il est fondamental d'être en mesure de caractériser l'impact des différentes approximations et des niveaux de théorie que nous utilisons dans le cadre de nos calculs afin de positionner nos résultats de manière critique, à plus forte raison lorsque l'on souhaite effectuer des prédictions de propriétés jamais mesurées ou valider nos modèles aux résultats expérimentaux. En effet, le rôle scientifique des modèles théoriques et numériques est intimement lié à leur capacité à faire des prédictions fiables, qui pourront par la suite être vérifiées expérimentalement ¹.

Dans le cadre de mes travaux, j'ai levé le voile sur l'impact de trois approximations fréquemment utilisées dans le cadre méthodologique bien établi qu'est le formalisme AHC, tout en quantifiant leur impact sur l'accord avec les données expérimentales. Si l'étude de l'effet du SOC sur le ZPR confirme le comportement attendu, soit que l'effet gagne en importance lorsqu'on considère des matériaux de plus en plus lourds, l'analyse des conséquences du ZPLE et des effets non adiabatiques sur le ZPR de la bande interdite a plutôt soulevé des nuances importantes qui invitent à revisiter certaines idées établies ². Dans cette optique, je crois qu'il est primordial qu'un pan de la recherche se penche sur la caractérisation des erreurs induites par certaines approximations courantes et demeure à l'affût d'éventuelles limitations non répertoriées dans les méthodologies largement utilisées, en parallèle avec le développement de méthodologies nouvelles. Dans le contexte du développement très rapide des approches automatisées de type *high throughput*, qui accèdent à des propriétés de plus en plus haut niveau, il devient d'autant plus crucial d'être en mesure de bien appréhender les impacts des différents choix de paramétrisations et de niveaux de théorie, afin d'interpréter les résultats de façon critique et d'éviter que des erreurs de filtrage n'éliminent des matériaux autrement prometteurs pour les applications désirées. En ce sens, j'espère que cette modeste contribution saura servir de guide à de futures études.

Dans un tout autre ordre d'idées, de nombreux efforts ont été déployés au cours des deux dernières décennies pour développer des méthodes numériques qui incorporent de plus en plus

1. On pourrait philosopher longtemps sur la fonction scientifique d'un modèle de la théorie du tout, tout aussi élégant soit-il, qui ne fournirait aucune prédiction vérifiable, de sorte qu'il soit rigoureusement impossible de confirmer ou d'infirmer sa validité. À mon humble avis, un tel modèle constituerait plutôt une œuvre d'art s'exprimant à travers le médium qu'est le langage scientifique. Mais je digresse un peu...

2. Après tout, personne n'est à l'abri de la force de l'habitude.

précisément les effets des corrélations électroniques, dans le but d'améliorer la description de la structure électronique dans l'approximation des noyaux statiques et de pallier les limitations de la DFT. On peut penser par exemple à la méthode *GW*, qui évalue explicitement la self-énergie de l'interaction électron-électrons dans l'approximation des noyaux statiques, ainsi qu'à la théorie du champ moyen dynamique (DMFT), qui s'attaque au surécrantage des fortes corrélations locales dans les orbitales de caractère *d* ou *f* situées près du niveau de Fermi, que l'on retrouve entre autres dans les métaux de transition. Néanmoins, comme cette thèse met en évidence, le raffinement de ces méthodes statiques peut dévier de son objectif premier s'il n'est pas accompagné des corrections induites par le mouvement des noyaux, y compris à température nulle, particulièrement pour les matériaux comportant des noyaux légers. Le développement de méthodes numériques permettant de prédire avec fiabilité la structure électronique à n'importe quelle température s'avère donc primordial pour soutenir la recherche expérimentale, tant fondamentale qu'appliquée, vers la conception d'applications pour lesquelles l'énergie de la bande interdite joue un rôle prépondérant, notamment en photovoltaïque et en ce qui concerne le contrôle des phases topologiques des matériaux.

À cet égard, il est plus que souhaitable d'étendre les avancées réalisées dans le traitement des corrélations électroniques statiques aux méthodes utilisées dans le cadre de cette thèse pour évaluer l'interaction électron-phonon et la dilatation thermique. En effet, l'approche DFT+DFPT évalue l'écrantage du vertex d'interaction électron-phonon à partir d'un calcul statique, alors que cette quantité devrait plutôt être évaluée de manière dynamique, en considérant l'effet mutuel des phonons sur les électrons et des électrons sur les phonons. À ce sujet, le réflexe le plus naturel est de recourir aux méthodes à *N*-corps ; ainsi, une approche développée au cours des dernières années permet désormais d'évaluer la self-énergie de l'interaction électron-phonon en se basant sur l'état fondamental statique évalué en *GW* [3]. De manière générale, cette méthode révèle un renforcement de l'interaction électron-phonon, qui augmente la renormalisation des énergies électroniques.

On ne peut néanmoins passer sous silence que cette précision accrue ne s'obtient qu'au prix d'une augmentation considérable du coût numérique déjà lourd des calculs *ab initio*. Les ressources informatiques n'étant pas illimitées, il est d'autant plus pertinent de chercher à développer des méthodes qui offriraient un niveau de précision similaire à l'approche à *N*-corps tout en ayant un coût numérique plus proche de celui de la DFT, qu'il serait réaliste d'appliquer à des systèmes comprenant plus que quelques atomes. Dans cette optique, des travaux récents ont proposé une approche perturbative basée sur les matrices de densité pour évaluer la self-énergie non adiabatique à l'aide de fonctionnelles hybrides [4]. La méthode proposée par Paleari et Marini [5] pour s'affranchir du problème de surécrantage du vertex d'interaction électron-phonon lors de l'évaluation de la self-énergie phononique tout en demeurant dans le formalisme de la DFT+DFPT pourrait également servir de source d'inspiration.

Cette observation met en évidence un défi important lié à la recherche reposant sur des simulations numériques, à savoir l'utilisation responsable, efficace et équitable des ressources matérielles et énergétiques nécessaires à la réalisation de nos calculs. En tant que chercheuses et chercheurs, nous avons une responsabilité morale de nous assurer que les méthodes et algorithmes que nous développons et utilisons maintiennent un certain équilibre entre la recherche de précision et l'accessibilité du coût numérique.

En ce sens, il semble inévitable que le domaine de la simulation numérique des matériaux tende de plus en plus à incorporer des techniques développées dans le contexte de l'apprentissage automatique. En effet, les méthodes *ab initio*, et encore plus les méthodes à N -corps, ne sont applicables de manière réaliste que sur des systèmes de petite taille, ce qui restreint leur utilisation pour des systèmes de taille mésoscopique comme ceux pertinents pour des applications réalistes en ingénierie. On pourrait alors souhaiter que des modèles entraînés puissent permettre de prédire les propriétés de systèmes comprenant plusieurs centaines, voire milliers d'atomes dans une configuration hors-équilibre avec une précision de l'ordre de celle des calculs *ab initio*, incluant les propriétés thermodynamiques. Les capacités de généralisation des modèles issus de l'apprentissage automatique étant intimement liées à la représentativité de l'échantillonnage de configurations atomiques sur lequel ils ont été entraînés, l'atteinte d'un tel objectif ne se fera évidemment pas sans nécessiter le recours à des approches de modélisation multi-échelles, qui combinerait la précision des méthodes *ab initio* et l'efficacité numérique des méthodes basées sur des potentiels effectifs. On peut penser, par exemple, aux méthodes de type *seconds principes* [6] et à l'optimisation à la volée de potentiels [7]. Dans un contexte où les changements climatiques sont de plus en plus pré-occupants, de telles approches pourraient jouer un rôle significatif dans l'accélération de la recherche de nouveaux matériaux plus performants et plus respectueux de l'environnement, qui nous permettraient de stocker plus efficacement les énergies vertes et de diminuer notre dépendance envers les hydrocarbures et les métaux rares.

Bibliographie

- [1] J. M. Lihm and C. H. Park. Phonon-induced renormalization of electron wave functions. *Phys. Rev. B*, 101(12) :121102(R), 2020. doi: [10.1103/PhysRevB.101.121102](https://doi.org/10.1103/PhysRevB.101.121102).
- [2] K. Saha, K. Légaré, and I. Garate. Detecting Band Inversions by Measuring the Environment : Fingerprints of Electronic Band Topology in Bulk Phonon Linewidths. *Phys. Rev. Lett.*, 115(17) :176405, 2015. doi: [10.1103/PhysRevLett.115.176405](https://doi.org/10.1103/PhysRevLett.115.176405).
- [3] Z. Li, G. Antonius, M. Wu, F. H. da Jornada, and S. G. Louie. Electron-phonon coupling from *ab initio* linear-response theory within the *GW* method : Correlation-enhanced interactions and superconductivity in $\text{Ba}_{1-x}\text{K}_x\text{BiO}_3$. *Phys. Rev. Lett.*, 122(18) :186402, 2019. doi: [10.1103/PhysRevLett.122.186402](https://doi.org/10.1103/PhysRevLett.122.186402).

- [4] H. Yang, M. Govoni, A. Kundu, and G. Galli. Computational protocol to evaluate electron-phonon interactions within density matrix perturbation theory. *arXiv :2206.00891 [cond-mat]*, 2022. doi: [10.48550/arXiv.2206.00891](https://doi.org/10.48550/arXiv.2206.00891).
- [5] F. Paleari and A. Marini. Electron-phonon interaction without overscreening : a strategy for first-principles modelling. *arXiv :2105.09828 [cond-mat]*, 2021. doi: [10.48550/arXiv.2105.09828](https://doi.org/10.48550/arXiv.2105.09828).
- [6] P. García-Fernández, J. C. Wojdeł, J. Íñiguez, and J. Junquera. Second-principles method for materials simulations including electron and lattice degrees of freedom. *Phys. Rev. B*, 93(19) :195137, 2016. doi: [10.1103/PhysRevB.93.195137](https://doi.org/10.1103/PhysRevB.93.195137).
- [7] C. Verdi, F. Karsai, P. Liu, R. Jinnouchi, and G. Kresse. Thermal transport and phase transitions of zirconia by on-the-fly machine-learned interatomic potentials. *npj Comput Mater*, 7(1) :156, 2021. doi: [10.1038/s41524-021-00630-5](https://doi.org/10.1038/s41524-021-00630-5).