

Université de Montréal

Étude numérique des mécanismes d'autodiffusion dans les semiconducteurs

par
Fedwa El-Mellouhi

Département de physique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en Physique

Mai, 2006

© Fedwa El-Mellouhi, 2006.



QC

3

U54

2006

V.025

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

Étude numérique des mécanismes d'autodiffusion dans les semiconducteurs

présentée par:

Fedwa El-Mellouhi

a été évaluée par un jury composé des personnes suivantes:

Michel Côté
président-rapporteur

Normand Mousseau
directeur de recherche

Sjoerd Roorda
membre du jury

Yves Limoge
examineur externe

Anne Bourlioux
représentante du doyen de la FES

Thèse acceptée le

RÉSUMÉ

Le processus physique le plus prometteur et le plus naturel pour la formation des éléments essentiels en nanotechnologie comme les points quantiques et les nanocristaux reste l'autoassemblage. L'autoassemblage est régi par la diffusion atomique qui est le processus fondamental de transport de masse dû aux défauts (lacunes, interstitiels, impuretés, etc). Ce travail utilise les développements récents dans les méthodes *ab-initio* basées sur la théorie de la fonctionnelle de la densité afin de donner une description de la structure électronique et des mécanismes de diffusion des lacunes dans les semiconducteurs. Cette thèse fait état de quatre contributions dans ce domaine. Premièrement, des événements rares et d'autres plus fréquents associés à la lacune neutre de Si dans le silicium sont présentés. La barrière de migration calculée est en bon accord avec les résultats expérimentaux de Watkins *et al.* [1]. Par la suite, une étude détaillée qui s'intéresse à la stabilité et l'état de charge des lacunes d'As et de Ga dans le GaAs en fonction des conditions de préparation expérimentales (dopage, stœchiométrie) est présentée. Troisièmement, les événements activés associés aux lacunes de Ga dans le GaAs ont été étudiés en mettant l'accent sur l'effet de l'état de charge sur le chemin et la barrière de diffusion. Les lacunes de Ga peuvent diffuser en faisant des sauts aux seconds ainsi qu'aux premiers et quatrièmes voisins avec une barrière de diffusion qui augmente avec l'état de charge. En plus, un nouveau mécanisme de diffusion au second voisin a été découvert pour les lacunes négatives. Enfin, la stabilité de la lacune de Ga chargée à haute température est examinée. L'énergie libre de formation de Gibbs est calculée en incluant les effets d'entropie, elle révèle qu'il peut y avoir un renversement de la charge de la lacune vers des états moins chargés. Elle permet d'ajuster, en même temps, les profils de concentration obtenus par annihilation des positrons et les profils de diffusion qui découlent de l'interdiffusion dans les hétérostructures.

Mots clés : Nanostructure, Défauts, Calcul de principes premiers, Lacune, ab-initio, Simulation, GaAs, Si.

ABSTRACT

The most promising and most natural physical process for the formation of elementary structures in nanotechnology such as quantum dots and nanocrystals remains self-assembly. Self-assembly is governed by the atomic diffusion which is the fundamental process of mass transport due to defects (vacancies, interstitials, impurities, etc). This work exploits recent developments in *ab-initio* methods based on density functional theory in order to give a reliable description of the electronic structure and migration mechanisms of vacancy type defects in semiconductors. Four main achievements in this direction are covered by this thesis. First, frequent and rare events connected to the neutral Si vacancy in silicon are identified. The calculated migration barrier to the first neighbor agrees with the experimental results obtained by Watkins *et al.* [1]. Second, a detailed study focuses on the stability and the charge state of As and Ga vacancies in the binary semiconducting compound GaAs with respect to experimental preparation conditions (doping, stoichiometry). Next, activated events associated to Ga vacancies are studied together with the effect of charge states on the path and the migration barrier. Gallium vacancy can migrate by jumps to the second, first and fourth neighbors with an increasing barrier for more negatively charged vacancies. In addition, a new mechanism for migration to the second neighbor have been found for negatively charged vacancies. Finally, the debate about the stability of charged Ga vacancies in GaAs at elevated temperature is analyzed. Gibbs free energy of formation is calculated by adding entropic effects. The thermal dependence of the Fermi level and of the ionization levels lead to a reversal of the preferred charge state as the temperature increases. This single set of calculated energies is suitable to fit at the same time experimental concentration profiles from positron annihilation experiments and diffusion profiles obtained from heterostructure interdiffusion experiments.

Keywords: Nanostructure, Self-diffusion, Semiconductors, Defects, First principles calculation, Vacancy, *ab-initio*, Simulation, GaAs, Si.

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES APPENDICES	xv
LISTE DES SIGLES	xvi
NOTATION	xvii
DÉDICACE	xviii
REMERCIEMENTS	xix
AVANT-PROPOS	xx
INTRODUCTION	1
CHAPITRE 1 : ÉTUDE EXPÉRIMENTALE DES DÉFAUTS DANS LES SEMI- CONDUCTEURS	8
1.1 Diffusion assistée par les défauts dans les semiconducteurs	8
1.1.1 Échange direct ou concerté	8
1.1.2 Diffusion par lacunes	9
1.1.3 Diffusion par interstitiels	9
1.1.4 Diffusion par interstitiel indirect	10
1.2 Classification électronique des défauts	11

1.2.1	Les impuretés isoélectroniques	11
1.2.2	Les défauts hydrogéniques	11
1.2.3	Les défauts à niveaux profonds	12
1.3	Étude expérimentale des défauts	12
CHAPITRE 2 : ÉTUDE NUMÉRIQUE DES DÉFAUTS		14
2.1	Limitations, exigences et problèmes des calculs ab-initio	14
2.2	Les lacunes dans les semiconducteurs simples : silicium	16
2.3	Les lacunes dans les composés binaires	18
2.3.1	Arséniure de gallium	20
2.3.2	Carbure de silicium	23
2.3.3	Nitride de gallium	25
CHAPITRE 3 : COMMENT IDENTIFIER LES CHEMINS DE DIFFUSION DES DÉFAUTS ?		27
3.1	Méthodes dynamiques	27
3.2	Théorie de l'état de transition	28
3.3	Méthodes de recherche des trajectoires activées	32
3.3.1	Méthode d'entraînement	33
3.3.2	Méthode de l'arête	33
3.3.3	Méthode de la bande élastique supportée	35
3.3.4	Technique d'activation et de relaxation	38
3.3.5	Méthode du dimère	41
CHAPITRE 4 : RÉOLUTION DU PROBLÈME DE LA STRUCTURE ÉLEC- TRONIQUE		43
4.1	Approximation de Born-Oppenheimer	44
4.2	Théorie de la fonctionnelle de la densité (DFT)	45
4.3	Les pseudopotentiels	49
4.4	Construction de la fonction d'onde	50
4.4.1	Les ondes planes	51

4.4.2	Les orbitales atomiques	51
4.5	Fonctionnement du code SIEST-A-RT	52
CHAPITRE 5 : L'AUTODIFFUSION DES LACUNES DANS LES SEMICON-		
DUCTEURS ÉLÉMENTAIRES : SILICIUM		56
5.1	Introduction	57
5.2	Method: SIEST-A-RT	59
5.2.1	ART nouveau	59
5.2.2	Siesta	60
5.2.3	Optimization of the parameters	60
5.3	Details of the simulation	66
5.4	Results	67
5.4.1	Neutral vacancy relaxation	67
5.4.2	Simple diffusion paths	70
5.4.3	Complex diffusion paths	75
5.4.4	Size effects	77
5.5	Discussion and Conclusions	78
5.6	Acknowledgments	79
BIBLIOGRAPHY		80
CHAPITRE 6 : STABILITÉ DES LACUNES DANS LE GAAS EN FONC-		
TION DE LEUR ÉTAT DE CHARGE		84
6.1	Introduction	85
6.2	Simulation details and methodology	87
6.2.1	Total energy calculations	87
6.2.2	Defects formation energies in supercell calculations	89
6.2.3	Computing the ionization energy of charged defects	91
6.3	Converging defect formation and ionization energies	92
6.3.1	Local basis set effect	92
6.3.2	K-points effects	98

6.3.3	Supercell size effects	100
6.3.4	Summary	101
6.4	Discussion	102
6.4.1	Gallium vacancies	104
6.4.2	Arsenic vacancies	107
6.5	Conclusions	113
6.6	Acknowledgments	113
BIBLIOGRAPHY		115
CHAPITRE 7 : L'AUTODIFFUSION DÉPENDANTE DE LA CHARGE POUR LES LACUNES DE GA DANS LE GAAS		118
7.1	Introduction	120
7.2	Details of the simulation	122
7.3	Results	123
7.3.1	Gallium vacancies at the initial state	123
7.3.2	Diffusion path to first neighbor	125
7.3.3	Diffusion path to fourth neighbor	127
7.3.4	Diffusion path to second neighbor	127
7.4	Discussion and comparisons	136
7.4.1	Self-diffusion to first-neighboring sites	136
7.4.2	Diffusion to second neighbor	138
7.4.3	Diffusion in experimental systems	139
7.5	Conclusions	140
BIBLIOGRAPHY		142
CHAPITRE 8 : EFFET DE LA TEMPÉRATURE SUR LE RENVERSEMENT DE L'ÉTAT DE CHARGE DES LACUNES DE GALLIUM DANS LE GAAS		146
8.1	Introduction	147

8.2	Computational details	149
8.3	Analysis of the experimental results	150
8.3.1	Concentration model	150
8.3.2	Diffusion model	153
8.3.3	Fits of positron-annihilation measurements	154
8.3.4	Fits of interdiffusion measurements	155
8.4	Results and discussion	157
8.4.1	Uncertainties on the calculated Gibbs free energy	161
8.4.2	Fitting experimental measurements	163
8.5	Conclusion	165
	BIBLIOGRAPHY	168
	SYNTHÈSE ET CONCLUSION	171
	CONTRIBUTIONS	176
	BIBLIOGRAPHIE	177

LISTE DES TABLEAUX

5.1	Converged formation energies (E_f) for different supercells and k-points meshes.	64
5.2	Relaxation geometries for different supercells and k-points meshes. . .	65
5.3	Interatomic distances along the minimum energy path for a simple diffusion.	71
6.1	Comparison between converged basic parameters for bulk GaAs.	88
6.2	Convergence of the formation energy E'_f in eV with respect to the basis set for the Ga vacancy.	94
6.3	Convergence of the formation energy E'_f in eV with respect to the basis set for the As vacancy.	96
6.4	Convergence of the formation energies E'_f (in eV) for Ga and As vacancies with respect to the Brillouin zone sampling.	99
6.5	Comparison between ionization energies (measured from the valence band edge) of the Ga vacancy in GaAs.	105
6.6	Comparison between ionization energies (measured from the valence band edge) for the As vacancy in GaAs.	107
7.1	Nearest neighbor distances in 111 direction relevant for first nearest neighbor diffusion from the initial \rightarrow to the final state.	124
7.2	Evolution of the distance (in Å) between the initial position of the vacancy and the diffusing atom ($V_{Ga}^q - Ga^{2nd}$) in the initial state and the transition state for both diffusion mechanisms to the second neighbour. .	130
7.3	Calculated diffusion barriers (in eV) for $V_{Ga}^{0,-1,-2,-3}$ in GaAs for all possible migration paths identified.	134
8.1	Calculated formation entropies S_f and formation enthalpies H_f for $V_{Ga}^{0,-1,-2,-3}$ in GaAs are compared to earlier calculation and experimental fits (see the text)	161

LISTE DES FIGURES

1	Spectre de photoluminescence des points quantiques de InAs sur du $\text{Ga}_x\text{In}_{1-x}\text{P}$ pour x variant de 0.1 à 0.3 [2].	2
2	Images par microscope à force atomique (AFM) montrant les nanostructures de GaAs sur la surface de AlGaAs [3].	3
3	Les étapes de formation des nanostructures avec l'épitaxie par gouttelettes [3].	5
1.1	Quelques mécanismes de diffusion dans le GaAs.	10
2.1	Comparaison entre les résultats des calculs réalisés durant les 20 dernières années concernant les niveaux d'ionisation de V_{Ga} et V_{As} dans le GaAs. Tiré de l'article de Deepak <i>et al.</i> [4].	20
2.2	Illustration du mécanisme de diffusion en anneau proposé par Van Vechten [5].	21
2.3	Chemin de migration de V_{Ga} vers le second voisin tiré de l'article Bockstedt et Scheffler [6].	23
3.1	Schéma illustrant une transition de l'état A vers l'état B dans un espace à 2 dimensions. Le taux de transition, dans la TÉT, est le flux (représenté par une flèche) qui traverse la surface séparant les états A et B	29
3.2	Recherche du point de selle par la méthode d'entraînement (Tiré de la référence [7]).	34
3.3	Recherche du point de selle par la méthode de l'arête (Tiré de la référence [7]).	36
3.4	Le chemin de migration tel que reconstitué par la méthode de la bande élastique (Tiré de la référence [7]).	38
3.5	Exploration de la surface d'énergie par la technique d'activation et de relaxation	40

3.6	Une représentation du processus itératif dans la méthode du dimère (Tiré de la référence [7]).	41
4.1	Organigramme qui montre l'interaction entre TAR et SIESTA durant la recherche des trajectoires activées.	55
5.1	Convergence tests for the energy per atom as a function of Brillouin zone sampling for 64, 216 and 512 cubic supercells. The lines are guide to the eye. (see the text).	61
5.2	Top: (a) The total minimum-energy path for the simple diffusion. The saddle points identified in various ART events are indicated by the same labels (B, C and D) as in Table II. The path is generated by previous knowledge of the initial, the saddle point and final states, then the overall configuration is relaxed using the nudged-elastic-band method [28].	69
5.3	The vacant site is visible in the snapshots A and C. A is the initial minimum and C is the ideal split vacancy site, the arrow shows the direction of the diffusion.	70
5.4	The reorientation path passing over the saddle point configuration. (I→S→F), the tetragonal axis is fixed throughout this process. At the saddle point the local atomic coordinates of the distorted lattice shows a trigonal symmetry. The pairing mode changes to one of the equivalent modes.	72
5.5	A <i>spectator</i> WWW process involving atoms far from the vacant site.	76
6.1	Schematic representation of the convergence of ionization energies as function of the basis set.	98
6.2	Histogram of the change in the volume of the relaxed tetrahedron formed by atoms surrounding the vacancy in % of the ideal one for two densities of k-points in the Brillouin zone.	100
6.3	Formation energies as function of Fermi level in various charge states of Ga vacancies at 0 K calculated for stoichiometric GaAs ($\Delta\mu = 0$).	103

6.4	Formation energies as a function of the Fermi energy of various charge states of As vacancies at 0 K.	103
6.5	Formation energies of Ga (solid line) and As (dashed line) vacancies in GaAs as a function of the growth conditions ($\Delta\mu$).	111
7.1	Constant surface at 0.0004 electron/ \AA^3 of the difference in electronic charge density between the neutral and the -3 charged gallium vacancy. The gallium vacancy is located initially at the center of the box.	123
7.2	Migration path of V_{Ga}^q to the fourth neighbor along (100) direction (see the text).	126
7.3	Diffusion to second neighbor by plane-passing mechanism (refer to the text).	128
7.4	Migration trajectory to the second neighbor by the plane-passing mechanism for -3 charge state.	129
7.5	Mulliken population analysis of the difference in electronic charge between the neutral and the -1 (top panel), and -3 (low panel) charges at the saddle point of plane-passing mechanism as function of the distance from the centre of the vacancy.	131
7.6	Diffusion to second neighbor by cluster-assisted mechanism (refer to the text).	132
7.7	Migration trajectory to the second neighbor by the cluster-assisted mechanism for -3 charge state.	135
8.1	Ionization levels (solid lines) used by Gebauer <i>et al.</i> [2].	155
8.2	Ionization levels (solid lines) as obtained by Bracht <i>et al.</i> [1].	156
8.3	Entropy and enthalpy (inset) differences obtained after progressive ionization of V_{Ga}^q	157
8.4	Stability domains of V_{Ga}^q in GaAs as function of temperature, indicated by the charge state.	159

- 8.5 Experimental measurement (filled circle), and fit (dashed line) for temperature dependence of the vacancy concentration, as reported by Gebauer *et al.* [2]. 164
- 8.6 Gallium self diffusion constant versus inverse temperature for (a) N-type (b) intrinsic (c) P-type. Theoretical results for D_{Ga}^q (dashed lines) show the contribution of each charge state. The transition temperature lie at the crossing of the curves. The total D_{Ga} (solid) is in good agreement with the measured interdiffusion data (square) and fit by Bracht *et al.* [1](circle). 167
- I.1 Les conditions aux limites périodiques génèrent un système constitué de la cellule de base et un nombre infini d'images. ccxvii

LISTE DES APPENDICES

Annexe I :	Conditions aux limites périodiques	ccxv
Annexe II :	Algorithme de Lanczòsccxxi
Annexe III :	Construction de la matrice dynamique	ccxxiii

LISTE DES SIGLES

ART	Technique d'activation et de relaxation
DFT	Théorie de la fonctionnelle de la densité
GaAs	Arséniure de gallium
GaN	Nitride de gallium
GGA	Approximation du gradient généralisé
IR	Spectroscopie à infrarouge
LDA	Approximation de la densité locale
LSDA	Approximation de la densité locale dépendante du spin
MBC	Minimum de la bande de conduction
MBV	Maximum de la bande de valence
NEBM	Méthode de la bande élastique supportée
PA	Spectroscopie d'annihilation des positrons
REDS	La spectroscopie de diffusion modulé par radiation
SiC	Carbure de Silicium
TB-MD	méthode des liaisons fortes + dynamique moléculaire

NOTATION

Le symbole D_a^q est utilisé pour décrire les défauts ponctuels. D fait référence au type de défaut : il devient V dans le cas des lacunes et I pour les interstitiels. q désigne l'état de charge et a indique le sous-réseau original occupé par le défaut. Par exemple V_{Ga}^{-1} est une lacune occupant un site dans le sous-réseau du Ga de charge -1 . As_{Ga} désigne les défauts de type antisite. Un atome d'As est substitué dans un site du sous réseau du Ga.

Les citations dans le texte principal (Introduction, Chapitres 1 à 4 et les annexes) renvoient à la bibliographie de la page 177. Tandis que les citations dans chacun des articles renvoient à la bibliographie à la fin de chaque chapitre.

À mes parents et à mon cher mari

REMERCIEMENTS

Je tiens à remercier chaleureusement tous ceux qui ont contribué de près ou de loin à l'aboutissement de ce travail. Je cite en particulier :

- Mon directeur de recherche Prof. Normand Mousseau pour sa supervision, son soutien et son orientation tout au long de ce projet de recherche ;
- Les membres du groupe **Grofnun** pour l'ambiance qu'ils font régner au sein du laboratoire ainsi que mes collègues au département de physique et au RQMP ;
- Les responsables et analystes du RQCHP pour l'assistance technique et la disponibilité ainsi que pour le précieux temps de calcul sur les machines du RQCHP ;
- Tous les membres de ma famille, particulièrement mon frère Fouad et sa femme Assia pour leur support ;
- Mes amies en Algérie, en particulier Fouzia Boukabrine ainsi que mes collègues du ICTP.

AVANT-PROPOS

Les quatre premiers chapitres de cette thèse présentent les concepts de base en relation avec l'autodiffusion des défauts dans les semiconducteurs. Les résultats de mes travaux de recherche sur ce sujet sont présentés sous forme d'articles dans les quatre derniers chapitres. Les deux premiers articles ont été publiés dans des revues spécialisées. Il sont suivis par un article prêt à la soumission et un autre qui est déjà soumis.

INTRODUCTION

L'utilisation des semiconducteurs dans la fabrication des dispositifs électroniques et optiques nécessite l'introduction et le contrôle d'imperfections dans la structure cristalline [8]. Les défauts cristallins sont indispensables pour le dopage des matériaux semiconducteurs avec des donneurs ou des accepteurs. Ils influencent grandement le type de porteur majoritaire, sa concentration ainsi que sa mobilité. Les toutes premières expériences avec les semiconducteurs étaient très difficiles à reproduire car les semiconducteurs sont des matériaux dont les propriétés électroniques changent considérablement pour des petites différences dans les concentrations d'impuretés.

Une nouvelle stratégie, qui consiste en la préparation des cristaux semiconducteurs de haute pureté par purification et cristallisation, est donc apparue lors de la découverte du transistor en 1947. Ces cristaux purs sont ensuite dopés pour obtenir les propriétés électroniques désirées. Les défauts cristallins sont introduits durant les procédures de dopage et peuvent soit améliorer soit dégrader la performance des dispositifs électroniques.

Particulièrement, les défauts ponctuels affectent considérablement les propriétés électroniques et optiques des semiconducteurs en introduisant des états électroniques localisés dans la bande interdite [8]. La modification apportée aux propriétés du semiconducteur, par les défauts ponctuels, dépend fortement de leur concentration et de leur dynamique. La présence des défauts natifs dans un semiconducteur favorise le transport de masse causé par la diffusion atomique assistée par les défauts.

Il a été prouvé récemment [2] que la diffusion du Ga dans les points quantiques de InAs décale leur spectre d'émission vers le bleu. Le spectre de photoluminescence de 2.4 monocouches de points quantiques de InAs/Ga_xIn_{1-x}P dépend fortement de la composition en gallium (x). Le spectre montré dans la figure 1 se déplace vers le bleu au fur et à mesure que la concentration du Ga augmente. Ce phénomène est dû à la diffusion d'une quantité considérable de Ga à partir du substrat vers les points quantiques de InAs durant la croissance.

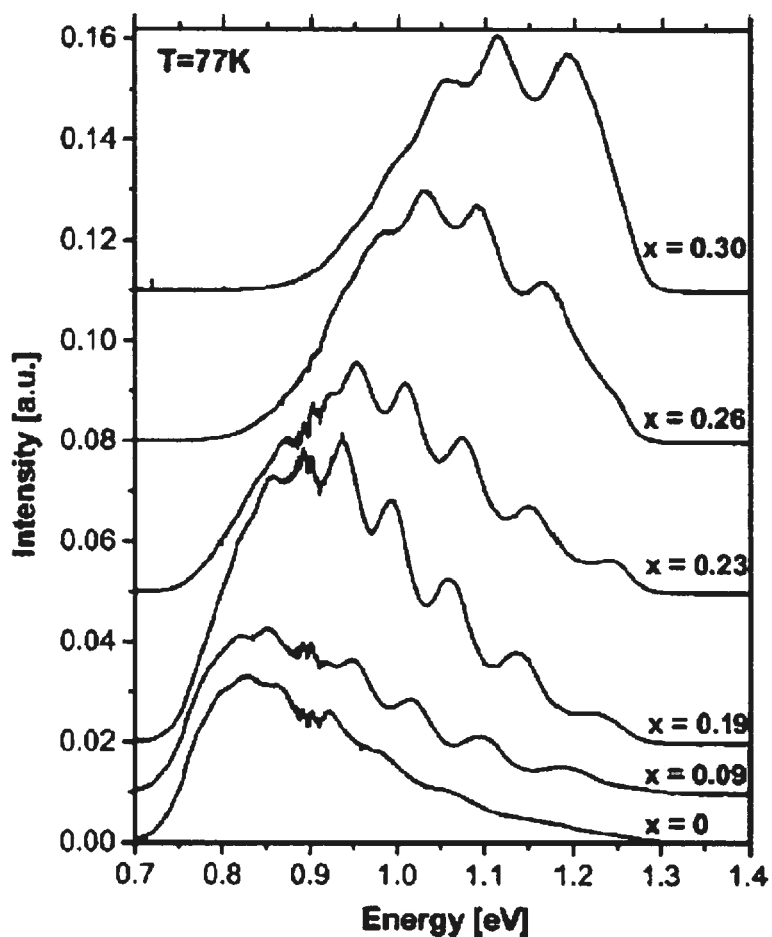


FIG. 1 – Spectre de photoluminescence des points quantiques de InAs sur du $\text{Ga}_x\text{In}_{1-x}\text{P}$ pour x variant de 0.1 à 0.3 [2].

Les défauts jouent aussi un rôle majeur dans la fabrication des matériaux nanostructurés destinés à des applications novatrices en nanotechnologie. Les propriétés optoélectroniques de ces nanostructures de semiconducteurs peuvent être modifiées en contrôlant leur structure et leur forme. Il reste qu'à cette échelle, la fabrication et le contrôle des nanostructures est un processus complexe qui est limité par l'avancement des techniques de fabrication. Or, l'autoassemblage est le processus physique principal utilisé de nos jours dans la fabrication de ces structures. Il se base principalement sur le transport de masse causé par l'autodiffusion assistée par les défauts ponctuels. Dans ce qui suit, nous présentons trois exemples qui montrent l'importance de l'autodiffusion en nanotechno-

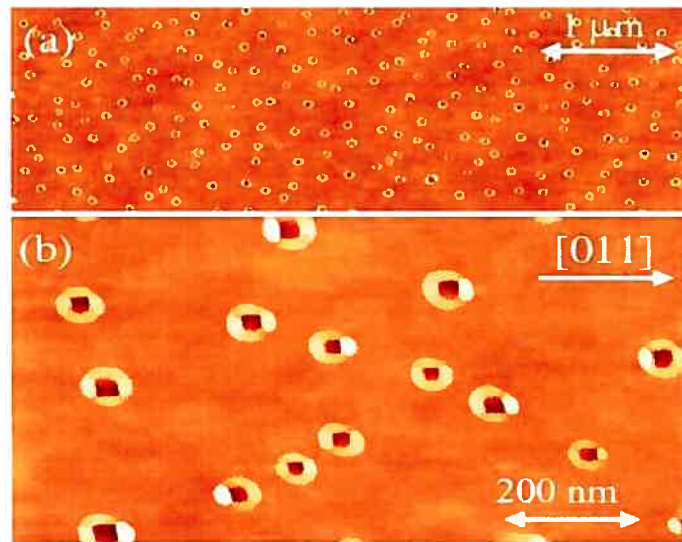


FIG. 2 – Images par microscope à force atomique (AFM) montrant les nanostructures de GaAs sur la surface de AlGaAs [3].

logie.

La méthode la plus répandue pour fabriquer des nanostructures par autoassemblage consiste en la croissance par la méthode de Stranski-Krastanow (SK) [9]. Un film mince de semiconducteur est déposé par croissance épitaxiale sur la surface d'un substrat semiconducteur ayant un désaccord de maille assez important avec le film déposé. Le désaccord de maille conduit spontanément à l'agglomération ou l'autoassemblage du film mince et à la formation des points quantiques ou des fils quantiques. Même si elle est largement utilisée, cette méthode n'est toutefois pas appropriée pour la croissance des nanostructures de GaAs sur le AlGaAs à cause de l'accord de maille entre les deux composés [3]. Ce système est très intéressant car les nanostructures formées ne sont pas contraintes et les interfaces sont nettes. Les nanostructures GaAs/AlGaAs peuvent, notamment, être utiles pour émettre dans le spectre comparable à celui des composés optiques les plus avancés [10].

Des méthodes plus récentes et plus complexes que la méthode de SK sont utili-

sées [11], dont le principe de base reste l'autodiffusion. Citons, à titre d'exemple, l'épitaxie par gouttelettes [3]. Cette méthode permet de fabriquer des nanostructures ayant des formes géométriques intéressantes (voir figure 2). Dans cet exemple particulier, illustré sur la figure 3 : (a) le gallium est déposé sur le substrat afin d'obtenir des gouttelettes liquides de Ga sur la surface du AlGaAs. (b) Il est ensuite exposé à un flux élevé d'arsenic à 380° C. L'arsenic réagit avec les gouttelettes pour former des nanostructures de GaAs. L'As ne pénètre pas directement dans les gouttelettes, le GaAs est d'abord formé à la surface concave de la gouttelette. (c) Le composé préfère s'agglomérer autour de la gouttelette sur la surface du AlGaAs. Le Ga des gouttelettes est progressivement consommé pour former le GaAs. (d) Une fois la source de Ga épuisée, les trous laissés par les gouttelettes sont progressivement remplis par la diffusion du GaAs. La forme finale des nanostructures de GaAs/AlGaAs dépend fortement de la direction de migration préférentielle du Ga sur la surface, ce qui rend cette méthode principalement basée sur l'autodiffusion du Ga et de l'As.

L'autodiffusion est aussi responsable de la fabrication de réseaux périodiques de nanobulles autoorganisées [12]. Ces nanobulles ou nanotrous se forment après irradiation continue de l'échantillon par des neutrons à haute température [13]. Les lacunes sont créées dans le matériau puis elles s'agglomèrent pour former des bulles. Les bulles apparaissent dans le matériau durant les premières étapes de l'irradiation, puis elles grossissent et commencent à s'organiser durant le processus d'irradiation (pendant 2 à 10 heures). Un des modèles utilisés pour expliquer la formation de ces nanobulles se base sur l'autodiffusion des interstitiels et des lacunes générées lors de l'irradiation [14].

Les détails de formation des nanostructures restent peu connus tant que nous n'avons pas une description complète de la thermodynamique et des mécanismes d'autodiffusion et d'autoassemblage associés aux défauts et leur vitesse. Pour les systèmes complexes, les techniques traditionnelles de simulation numérique, comme la dynamique moléculaire, ne permettent pas d'atteindre les temps de simulation nécessaires pour voir une lacune diffuser dans ces matériaux. Pour les cristaux élémentaires tel que le silicium, on peut faire des études sur la diffusion par la théorie de l'état de transition en considérant les propriétés de symétrie. Cette approche n'est plus valable dès qu'on passe à la famille

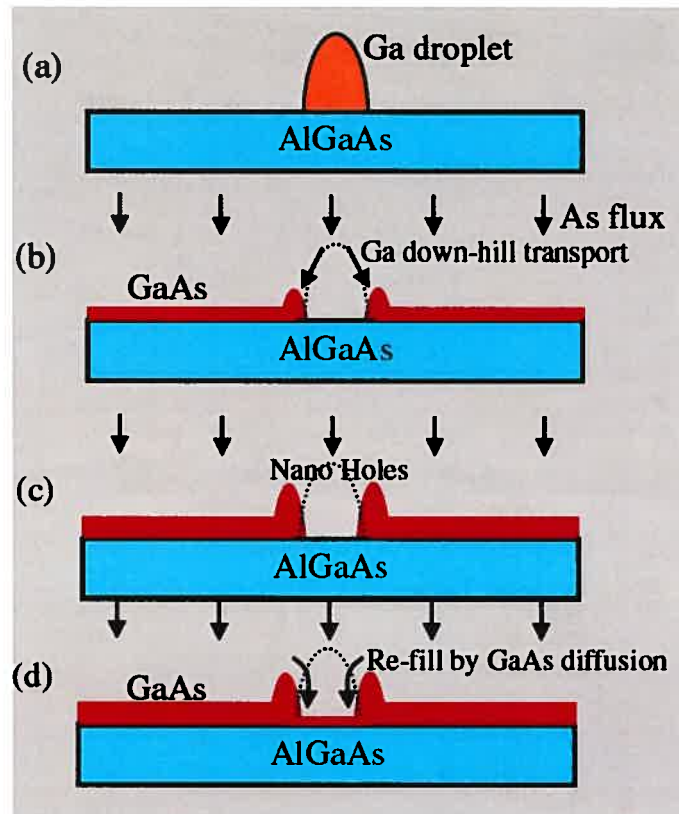


FIG. 3 – Les étapes de formation des nanostructures avec l'épitaxie par gouttelettes [3].

des composés binaires ou plus complexes, comme l'arséniure de gallium (GaAs), qui suscitent un intérêt technologique croissant.

D'autre part, l'état de charge des défauts natifs responsables des processus cités plus haut peut être connu en étudiant expérimentalement la dépendance de l'autodiffusion en fonction du dopage. La contribution relative de chacun des états de charge à l'autodiffusion est cependant inconnue. Afin de comprendre et pouvoir prédire les propriétés macroscopiques des semiconducteurs, il est crucial de déterminer les propriétés thermodynamiques et électroniques qui régissent la concentration des défauts et leur diffusion par la simulation numérique. À l'équilibre thermique, la concentration des défauts est déterminée par l'énergie libre de formation de Gibbs (G_f) du défaut qui est égale à la différence entre l'enthalpie de formation (H_f) et l'entropie de formation (S_f) multipliée par la température ($G_f = H_f - TS_f$). Sur le plan expérimental, il est très difficile de dé-

terminer l'énergie libre de formation de Gibbs pour les défauts dans les semiconducteurs. La lacune de Ga dans le GaAs est un excellent exemple. Les propriétés électroniques du défaut, en particulier son état de charge, contribuent fortement à l'énergie de formation libre de Gibbs. Or, la plupart des travaux sur ce défaut utilisent les expériences de diffusion pour extraire l'enthalpie et l'entropie de formation. Mais, ces expériences mesurent la diffusion assistée par les lacunes de Ga sans détecter les lacunes elles-mêmes. Les analyses faites sur ces profils de diffusion donnent des résultats qui dépendent fortement du modèle de diffusion utilisé.

En plus, les travaux de simulation numérique ont rarement calculé l'énergie libre de Gibbs et la plupart du temps, la contribution de l'entropie est négligée. Seuls Bockstedt et Scheffler [6] ont calculé l'enthalpie et l'entropie de formation de cette lacune dans l'état neutre seulement. Les résultats, concernant l'état de charge dominant de V_{Ga} , obtenus par les expériences ne s'accordent pas avec ceux qui sont obtenus par la simulation numérique. La plupart des calculs théoriques sont faits à 0° K et ne prennent pas en compte les effets entropiques. Ils prédisent que l'état -3 est le plus stable. À haute température, les expériences de diffusion menées par Bracht *et al.* [15] proposent les états -2 et -1 alors que les expériences par annihilation des positrons de Gebauer *et al.* [16] proposent l'état de charge -3 .

L'étude des défauts par la simulation numérique est un domaine qui est en progression continue. Il y a beaucoup de questions auxquelles nous pouvons répondre grâce au développement des méthodes de calcul plus précises et plus avancées et grâce à l'augmentation de la puissance des ordinateurs. Nous pouvons traiter des problèmes plus complexes ou même réexaminer des problèmes déjà étudiés. L'autodiffusion des défauts dans les semiconducteurs constitue le sujet principal de cette thèse. Les calculs de principes premiers, qui sont utilisés, sont basés sur la théorie de la fonctionnelle de densité de Kohn et Sham [17] et permettent d'éclaircir des points fondamentaux sur les propriétés électroniques, dynamiques et thermodynamiques des défauts natifs dans les semiconducteurs simples et composés, comme :

- Comprendre les paramètres qui influencent le calcul des énergies de formation des lacunes par les méthodes *ab-initio* ;

- Réexaminer les propriétés de la lacune de Si en utilisant les meilleurs paramètres de simulation ;
- Étudier les modes de relaxation et de diffusion des lacunes de gallium et d'arsenic dans le GaAs ;
- Étudier l'influence des conditions de préparation du GaAs (l'état de charge) sur le chemin de migration des lacunes ;
- Résoudre le débat qui concerne l'état de charge de la lacune de Ga à haute température dans le GaAs ;
- Comparer les résultats du GaAs avec d'autres composés semiconducteurs binaires.

Cette thèse est organisée comme suit : les quatre premiers chapitres exposent les concepts reliés à ce domaine de recherche et mettent l'emphase sur les différents volets en relation avec ce sujet. Le chapitre 1 traite la classification des défauts par les études expérimentales. Il expose aussi les différents types de défauts ponctuels et quelques modèles de diffusion simples associés à ces défauts. Le chapitre 2 est consacré aux études récentes réalisées par la simulation numérique axées sur les défauts ponctuels dans les semiconducteurs. Dans le chapitre 3, on y compare les différentes méthodes de recherche de trajectoires activées tout en discutant leur efficacité. Le chapitre 4 présente brièvement les concepts de base utilisés dans les calculs *ab-initio*.

Les quatre derniers chapitres rassemblent les résultats des travaux sur les défauts dans les semiconducteurs sous forme d'articles. Le chapitre 5 traite la formation et la diffusion de la lacune de Si dans le silicium. Il montre qu'un défaut aussi simple est très actif et peut être associé à plusieurs mécanismes. Le chapitre 6 présente une étude plus poussée qui s'intéresse aux propriétés statiques des lacunes d'As et de Ga dans le composé GaAs. Dans le chapitre 7, les différents chemins de migration des lacunes de Ga dans le GaAs sont étudiés en mettant l'emphase sur l'effet de l'état de charge sur les chemins de migration. L'étude sur l'état de charge de la lacune de Ga dans le GaAs à haute température fait l'objet du chapitre 8. Le résumé de l'aboutissement de ces travaux de simulation et l'exposition de notre vision pour des travaux encore plus poussés dans ce domaine concluent cette thèse.

CHAPITRE 1

ÉTUDE EXPÉRIMENTALE DES DÉFAUTS DANS LES SEMICONDUCTEURS

Les premières études expérimentales des défauts dans les semiconducteurs se sont intéressées à identifier les défauts cristallins et à étudier leurs propriétés optiques et électroniques fondamentales [18]. Dans la plupart des cas, il en a résulté une compréhension détaillée de la configuration atomique des défauts dans différents états de charge incluant les états excités métastables ou bistables qui montrent dans certains cas des propriétés très intéressantes. Des études plus détaillées ont cherché à mieux comprendre les mécanismes qui contrôlent les défauts, comme l'introduction des défauts durant la croissance cristalline, la diffusion de ces défauts dans le cristal, leur stabilité thermique, leur passivation ou la compensation etc.

1.1 Diffusion assistée par les défauts dans les semiconducteurs

Dans cette section, nous présentons les différents modèles de diffusion possibles qui sont dus à la présence de défauts ponctuels dans les semiconducteurs. Ces mécanismes contribuent aussi bien à l'autodiffusion qu'à la diffusion des impuretés et des dopants. Il s'agit de mécanismes intuitifs qui peuvent avoir lieu dans un semiconducteur simple tel que le silicium. L'importance de l'un ou l'autre des mécanismes dépend fortement du matériau étudié, des conditions de croissance et des conditions de l'expérience. Pour les matériaux composés, d'autres types de défauts peuvent surgir lors de la diffusion comme les complexes à base d'antisite (un cation qui occupe un site d'anion). La figure 1.1 illustre quelques uns des mécanismes expliqués ci-dessous :

1.1.1 Échange direct ou concerté

Deux atomes qui occupent des positions régulières dans la structure cristalline du semiconducteur peuvent échanger leur position en échangeant graduellement leurs voisins. Ce mécanisme nécessite un échange de voisin du type WWW [19]. Cet échange

engendre une diffusion atomique sur un anneau ou sur une ligne. Mais, la plupart du temps, ce type de diffusion est coûteux en énergie, car sa barrière est généralement très élevée, et ne se produit que rarement. L'échange concerté peut devenir plus difficile dans le cas des composés binaires comme le GaAs car il induit la formation d'une paire d'antisites $As_{Ga}-Ga_{As}$.

1.1.2 Diffusion par lacunes

Les lacunes sont aussi présentes dans la structure cristalline. Elles peuvent, par exemple se former lors de l'irradiation du matériau. Un atome quitte son site régulier pour aller se loger dans la région interstitielle ou s'évapore en laissant derrière lui un trou. Un atome peut diffuser en échangeant sa position avec la lacune, ce qui conduit à l'autodiffusion. D'un autre côté, une impureté qui occupe un site substitutionnel peut aussi utiliser ce mécanisme pour diffuser. On parle de diffusion assistée par les lacunes. Toutefois, le mécanisme de diffusion par lacune est subtil et très sensible à l'environnement structural et électronique de la lacune. Ce mécanisme peut être plus avantageux lors des échanges entre la lacune et le premier voisin ou des voisins éloignés.

1.1.3 Diffusion par interstitiels

Certains atomes ou ions se déplacent à travers la région interstitielle sans pour autant interagir avec les sites réguliers de la structure cristalline. Dans quelques cas particuliers, les interstitiels ne forment pas de liens covalents avec les atomes avoisinants. Au contraire, ils se trouvent expulsés d'un site interstitiel vers un autre sous l'effet de la répulsion de Pauli entre les électrons confinés dans les liens covalents saturés dans la structure et les électrons de valence de l'interstitiel. Il peut s'agir d'autointerstitiels¹ ou des impuretés qui occupent les différentes positions interstitielles possibles dans le cristal. En utilisant ce mécanisme, la diffusion peut devenir rapide à condition que la barrière d'énergie nécessaire soit assez faible.

¹Des atomes du même type que le semiconducteur qui s'installent dans les ouvertures de la structure cristalline.

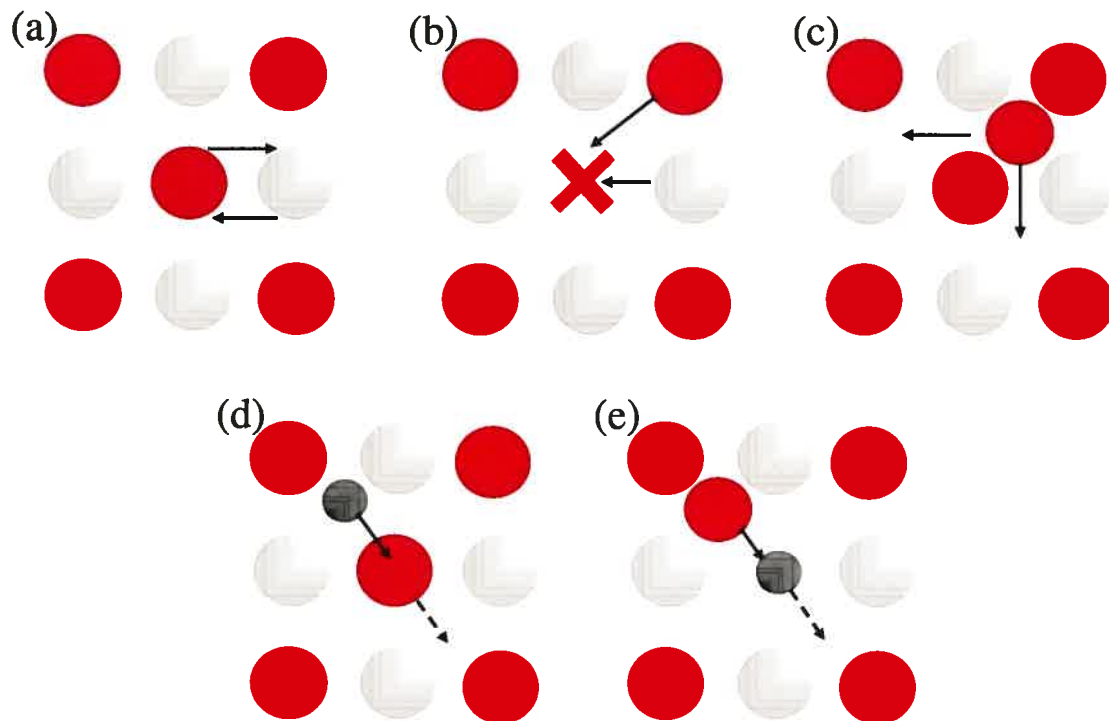


FIG. 1.1 – Quelques mécanismes de diffusion dans le GaAs. Les atomes de Ga sont en rouge, les atomes d'As sont en blanc et les impuretés sont les petits atomes en noir. (a) échange concerté qui implique une paire As–Ga (b) diffusion par lacune de Ga (X) soit par diffusion au premier ou au second voisin (c) diffusion dans la région interstitielle (d) diffusion par interstitiel-substitutionnel "kick-in" (e) "kick-out"

1.1.4 Diffusion par interstitiel indirect

Ce type de mécanisme implique un atome interstitiel et une impureté substitutionnelle. Dans le mécanisme "kick-in", une impureté très mobile qui se trouve dans une région interstitielle peut déloger un atome de Ga. Au lieu de diffuser vers le prochain site interstitiel, cette impureté pousse un des atomes de Ga substitutionnels les plus proches vers la région interstitielle et occupe sa place. La configuration finale est constituée d'un autointerstitiel et d'une impureté substitutionnelle. Pour le mécanisme "kick-out", c'est l'autointerstitiel qui déloge une impureté substitutionnelle, puis la pousse vers la région interstitielle. Il en résulte une impureté interstitielle.

1.2 Classification électronique des défauts

Dans les semiconducteurs purs et parfaits, la bande interdite est la fenêtre d'énergie qui ne compte pas d'états électroniques et qui est bornée par les bandes de conduction et de valence. Les dopants ou les défauts peuvent, par contre modifier considérablement la conductivité électronique en introduisant de nouveaux états dans la bande interdite [20]. Les défauts dans les semiconducteurs peuvent être classés en quatre catégories dépendamment de leurs propriétés électroniques :

1.2.1 Les impuretés isoélectroniques

Ces impuretés ne font paraître aucun niveau électronique dans la bande interdite du matériau semiconducteur. Elles possèdent une charge neutre indépendamment de la position du niveau de Fermi dans la bande interdite. Parmi les exemples les plus connus, nous citons les interstitiels d'oxygène dans le silicium et l'azote dans le GaAs et le GaP. Ces impuretés causent "généralement" un rétrécissement de la bande d'énergie interdite. Ceci est aussi vrai pour les alliages où l'aluminium ou l'indium sont introduits à des stœchiométries variables dans le GaAs, ou dans le cas de l'introduction du Germanium dans le Si, alors que, l'introduction de l'azote dans le GaAs cause un élargissement de la bande d'énergie interdite. Ce type de défauts peut donc être utilisé pour améliorer les propriétés mécaniques ou électroniques des semiconducteurs destinés à des applications spécifiques.

1.2.2 Les défauts hydrogéniques

Ces défauts possèdent des niveaux *délocalisés et peu profonds* dans la bande interdite. Il s'agit de dopants substitutionnels et d'impuretés introduites délibérément dans le semiconducteur lors du dopage. L'atome substitutionnel appartient généralement à une colonne adjacente à celle de l'atome qu'il remplace. Ces défauts restent généralement chargés pour la plupart des positions du niveau de Fermi dans la bande interdite. Leurs niveaux localisés se situent à quelques dizaines de millielectronvolts des limites de la bande interdite (MBV et MBC).

1.2.3 Les défauts à niveaux profonds

Il s'agit des impuretés ou des défauts cristallins (lacunes, interstitiels, antisites et leurs agrégats) qui peuvent générer des niveaux d'énergie *localisés* et *profonds* dans la bande interdite. Les niveaux d'ionisation sont de l'ordre du dixième de la largeur de la bande interdite. Les défauts à niveaux profonds possèdent une section efficace de capture des porteurs qui est nettement supérieure à celle des défauts hydrogéniques. Par conséquent, la concentration de ces défauts influence fortement la concentration des porteurs majoritaires dans le matériau semiconducteur.

1.3 Étude expérimentale des défauts

Afin d'étudier la diffusion des défauts dans les semiconducteurs, il est primordial d'avoir une distribution hétérogène de défauts dans le matériau pour détecter le changement dans les distributions mesurées et induites par la diffusion de ces défauts. Par exemple, les autointerstitiels peuvent être créés dans le matériau grâce à l'implantation ionique ou l'injection de dopant par déposition de source suivie de recuit. Les lacunes sont mieux introduites aux interfaces (dans les superréseaux) dans la couche epitaxiée lors d'une croissance epitaxiale à basse température ou dans les matériaux massifs par bombardement.

L'étude des défauts commence par l'observation et la caractérisation des défauts. Ceci consiste en l'identification du type de défaut qui est initialement totalement inconnu. L'information à propos d'un défaut ponctuel peut être obtenue par les méthodes spectroscopiques comme la résonance électronique paramagnétique (EPR), la spectroscopie à infrarouge (IR), la spectroscopie transitoire de niveaux profonds (*Deep Level Transient Spectroscopy*, DLTS) ou la photoluminescence. Pour les études à haute température, d'autres méthodes sont sollicitées telle que la spectroscopie d'annihilation des positrons (PA) mais cette méthode se limite à la détection des lacunes neutres ou chargées négativement. Les défauts étendus sont identifiés grâce à l'analyse structurale du matériau en utilisant la diffraction des rayons X ou le microscope à transmission électronique (TEM). Par contre, il est plus difficile d'identifier la nature réelle de certains

défauts ponctuels en utilisant ces méthodes seulement. La manière la plus efficace reste l'identification basée sur leur signature optique ou électrique.

L'identification des défauts est plus efficace en combinant plusieurs méthodes incluant la spectroscopie à résonance magnétique et l'étude des modes de vibration localisés. Les calculs de principes premiers constituent un moyen très efficace pour l'identification des défauts ². Les calculs quantiques donnent des résultats comparables aux résultats de la spectroscopie concernant la structure et l'énergie de formation des défauts.

Le contrôle des défauts commence d'abord par l'étude de la stabilité thermique d'un défaut complexe pendant les expériences de recuit. Une série de recuits isothermiques est effectuée pour engendrer soit la dissociation du défaut soit sa migration. Ceci permet de mieux comprendre le défaut pour pouvoir le contrôler en le créant ou le dissociant afin d'améliorer la performance des dispositifs électroniques.

²L'efficacité des méthodes *ab-initio* est discutée dans le chapitre 2.

CHAPITRE 2

ÉTUDE NUMÉRIQUE DES DÉFAUTS

L'objectif de ce chapitre est de discuter les derniers progrès réalisés dans le domaine de la simulation numérique *ab-initio* pour comprendre les propriétés des défauts natifs, et plus particulièrement les lacunes dans les semiconducteurs. Les résultats des calculs obtenus durant les dernières années se comparent de mieux en mieux aux expériences. Il est maintenant possible de comprendre la diffusion des défauts dans les matériaux réels en les recréant dans nos boîtes de simulation et en les faisant diffuser de manière efficace. Certains des travaux discutés dans ce chapitre ont été publiés durant le déroulement de mon projet de thèse ; ce qui montre l'importance et la pleine effervescence de ce domaine de recherche.

La première partie de ce chapitre résume les limitations, les problèmes et les exigences qui s'imposent à l'utilisation des méthodes *ab-initio* pour l'étude des défauts dans les semiconducteurs. La partie suivante discute les travaux récents réalisés dans ce domaine. La diffusion des lacunes dans les semiconducteurs simples est discutée en premier lieu, en prenant l'exemple du silicium. Des cas plus complexes, qui traitent la diffusion dans les semiconducteurs binaires, comme le GaAs, le SiC et le GaN sont examinés par la suite.

2.1 Limitations, exigences et problèmes des calculs *ab-initio*

Les méthodes numériques les plus fiables pour le calcul des propriétés des défauts sont les méthodes à base de calculs de principes premiers généralement basés sur la théorie de la fonctionnelle de la densité (DFT) expliquée en détail dans la section 4.2.

Les défauts dans les semiconducteurs posent plusieurs défis pour les calculs *ab-initio*. Ces calculs doivent expliquer l'origine des bandes de défauts observées lors des expériences, analyser la structure atomique des défauts pour les différents états de charge et

donner une bonne estimation de l'énergie de formation des défauts¹. La précision dans l'estimation de l'énergie de formation est affectée par plusieurs facteurs, à savoir :

- La taille et la géométrie de la boîte de simulation ² ;
- La densité des points spéciaux dans la zone de Brillouin ;
- La qualité de la base utilisée : le nombre d'ondes planes ou des orbitales atomiques localisées ;
- La relaxation des atomes autour du défaut : comment l'énergie totale est réduite ?
- La correction sur les énergies totales pour des défauts chargés ;
- L'approximation faite sur les potentiels chimiques ;
- Les erreurs introduites par l'usage des approximations à la DFT.

Les chapitres 5 et 6 traitent la plupart de ces facteurs en détail. Ils mettent en évidence les normes qui doivent être respectées pour avoir les résultats les mieux convergés. Des facteurs additionnels spécifiques aux défauts chargés sont discutés dans un article de Shim *et al.* [21] paru en 2005 (voir la discussion dans l'annexe I). Cet article étudie l'effet de la taille de la supercellule, la validité de la correction électrostatique appliquée sur les systèmes chargés ³ ainsi que les différentes méthodes d'alignement du maximum de la bande de valence.

D'autre part, les valeurs propres de de Kohn-Sham calculées à partir des méthodes DFT [17] ne correspondent généralement pas aux niveaux électroniques mesurés expérimentalement. Pour palier à cette limite, les niveaux thermodynamiques d'ionisation sont calculés à partir des différences d'énergies totales entre deux systèmes relaxés ayant des états de charge différents. C'est cette approche que nous utilisons dans les travaux de cette thèse.

Il y a aussi des limites de précision plus spécifiques à l'approximation de la densité locale (LDA). La LDA sous estime la largeur de la bande interdite, ce qui donne une faible précision pour le calcul des valeurs limites du potentiel chimique. Pour améliorer

¹La stabilité relative et la concentration des défauts dépendent de l'énergie de formation.

²L'utilisation des supercellules crée une interaction artificielle entre les défauts périodiques, tandis que l'utilisation d'agrégat fait apparaître des interactions artificielles avec les surfaces.

³Cette approche est largement utilisée pour compenser l'interaction entre le défaut chargé et ses images périodiques.

nos résultats, on procède à des corrections en utilisant la valeur expérimentale de la largeur de la bande interdite.

Mis à part les problèmes cités plus haut, les calculs *ab-initio* utilisés pour le calcul des énergies de formation des défauts ponctuels dans les semiconducteurs à 0°K donnent des résultats qui se comparent très bien avec les résultats de spectroscopie à basse température. Mais le problème se pose pour la comparaison avec les expériences d'interdiffusion qui sont généralement faites à haute température.

Un exemple concret est celui de l'autodiffusion de la lacune de Ga dans le GaAs discuté dans le chapitre 8 qui montre que les effets de température doivent être pris en considération pour bien refléter les résultats expérimentaux. Un autre exemple est celui de la diffusion de la lacune de silicium (V_{Si}) dans le silicium, discuté avec plus de détail dans la section 2.2. Les effets de température doivent donc être considérés dans les calculs *ab-initio* pour avoir des informations supplémentaires à propos de l'influence des défauts ponctuels sur le fonctionnement des dispositifs électroniques à haute température.

2.2 Les lacunes dans les semiconducteurs simples : silicium

Le silicium est le matériau de base pour la fabrication des dispositifs électroniques qui ont tendance à devenir de plus en plus miniaturisés. Pour les dispositifs métal-oxyde-semiconducteur (MOS) à base de Si, la concentration de charge dans la source et le drain doit augmenter pour minimiser la résistance du dispositif. En augmentant le niveau de dopage de type N ou de type P, la position du niveau de Fermi s'approche des limites de la bande interdite, soit le minimum de la bande de conduction et le maximum de la bande de valence. Par conséquent, la formation de défauts ponctuels qui présentent un caractère donneur ou accepteur devient plus favorable.

La connaissance de l'état de charge des défauts dans la bande interdite du Si et les niveaux d'énergie correspondants est d'une importance majeure pour le contrôle de la diffusion des dopants durant la fabrication des dispositifs électroniques à base de Si. Les lacunes de Si (V_{Si}) dominent l'autodiffusion au dessus d'une température critique

$T_C = 890^\circ \text{ C}$ [22] alors qu'au dessous de cette température, la diffusion se fait principalement par l'intermédiaire des interstitiels (I_{Si}). Un très grand nombre d'études a été axé sur les propriétés de V_{Si} dans le silicium. Il reste que c'est un matériau qui suscite un intérêt continuel et plusieurs questions nécessitent le développement des méthodes expérimentales et numériques pour y répondre.

La structure de V_{Si} est maintenant bien connue grâce aux calculs *ab-initio* réalisés pendant les dernières années [23–25]. V_{Si} peut exister dans les états de charge $q = +2, 0, -1, -2$ et présente une transition de type U-négatif [26, 27] de l'état $+2$ à 0 . La déformation structurelle due à la relaxation des atomes qui entourent la lacune, dépend fortement de l'état de charge. Pour $q = 0, -1$ la symétrie est brisée sous l'effet de la relaxation de type *Jahn-Teller*. La brisure de symétrie est principalement due à la structure électronique de la lacune. Les niveaux électroniques de la lacune neutre proviennent de l'état triplement dégénéré qui se trouve initialement au maximum de la bande de valence (dans le Si parfait). Après la création de la lacune, cet état est poussé vers la bande interdite sous l'effet d'un potentiel de Madelung répulsif. La distorsion atomique causée par le pairage des atomes entourant la lacune entraîne une levée de dégénérescence de ces niveaux. Le calcul de l'énergie de formation de V_{Si} neutre utilisant la LDA, tourne autour de la valeur $3.7 \pm 0.5 \text{ eV}$ [28–31]. Une étude récente utilisant une cellule de 256 atomes et la GGA donne 3.17 eV [32] alors qu'une étude utilisant la sx-LDA [33] donne une énergie de formation de 3.8 eV . Malheureusement, la taille de la supercellule utilisée dans le calcul sx-LDA était limitée à 32 atomes ; ce qui a pour effet d'introduire d'autres erreurs dues aux effets de taille. Les différences dans l'estimation des énergies de formation, entre les différentes simulations, sont principalement dues aux différences dans les paramètres numériques utilisés comme le pseudopotentiel, la fonctionnelle d'échange et de corrélation de la DFT, la taille de la supercellule etc.

Dans l'ensemble, ces calculs de l'énergie de formation de V_{Si} sont proches des estimations tirées des mesures expérimentales les plus récentes. Les mesures d'annihilation des positrons [34] trouve $E_F = 2.8 \pm 0.3 \text{ eV}$ pour la lacune neutre. Une autre étude de diffusion accélérée par l'irradiation (Radiation Enhanced Self-Diffusion, RESD) donne $E_F = 2.1 \pm 0.15$ [35], reste que l'état de charge de la lacune est indéterminé dans cette

expérience.

Il y a eu aussi plusieurs études axées sur l'autodiffusion dans le silicium. Certains de ces calculs utilisent la dynamique moléculaire quantique [30, 36, 37] et d'autres utilisent la méthode des liaisons fortes (TB-MD) [25, 38]. Ces méthodes sont capables de simuler la diffusion de V_{Si} au premier voisin. L'atome de Si premier voisin de la lacune commence à se déplacer dans la direction 111 le long de l'axe qui le relie avec la lacune. Au point de selle de la migration, l'atome de Si aura parcouru la moitié du chemin qui le mène vers la lacune. Il se forme ce qu'on appelle une *lacune divisée*. L'atome de Si continue sa migration en suivant l'axe de diffusion jusqu'à ce qu'il arrive à la position finale.

La barrière de migration au premier voisin, calculée dans cette thèse (voir chapitre 5), est de 0.4 eV pour la lacune neutre. Elle peut être directement comparée avec l'enthalpie de migration mesurée à très basse température (autour de 4.2° K) par Watkins [1] qui varie de 0.18 à 0.45 eV dépendamment de l'état de charge de la lacune. La barrière calculée est en bon accord avec les expériences et avec d'autres calculs *ab-initio* [31] qui affirment que l'enthalpie de migration de V_{Si} est inférieure à 1 eV.

Par contre, les études de diffusion à haute température par REDS donnent une enthalpie de migration de la lacune de 1.8 ± 0.5 eV [35]. Ceci suggère que les lacunes se comportent différemment à haute température. D'où la nécessité d'étudier la diffusion V_{Si} à haute température.

2.3 Les lacunes dans les composés binaires

Les mécanismes de diffusion des lacunes deviennent plus compliqués à étudier aussitôt qu'on passe à la famille des semiconducteurs composés comme le GaAs :

- La stœchiométrie ou en d'autre terme le potentiel chimique d'un des éléments devient un paramètre important qui influence les concentrations des défauts ponctuels à l'équilibre ⁴ ;
- Les chemins de migration sont plus complexes que dans le Si et donnent lieu à des

⁴On peut avoir des situations expérimentales où le matériau est riche en cations ou en anions.

barrières plus élevées. Certains types d'atomes préfèrent diffuser vers leur propre sous-réseau et sont même incapables de diffuser au premier voisin. Même si la diffusion au premier voisin est possible, elle engendre la formation d'une cascade de défauts comme les complexes *Lacune-Antisite-Interstitial*. Ces complexes nécessitent plusieurs cycles de diffusion avant de s'annihiler ;

- La distance de migration des lacunes au second voisin est au moins deux fois plus grande que pour les lacunes de Si.

Les groupes de recherche qui font la dynamique moléculaire quantique ont vite compris qu'elle était inefficace pour ce genre de problème. Même les études par la dynamique moléculaire classique [39] utilisant un potentiel de Stillinger-Weber paramétré pour le GaAs, ont échoué à simuler la diffusion des lacunes. Les méthodes de recherche de trajectoires activées⁵ sont nettement plus efficaces pour étudier la diffusion des défauts dans les semiconducteurs composés. La plupart des travaux cités dans les prochaines sections utilisent des méthodes de recherche de point de selle qui se basent sur la connaissance préalable de l'état final comme la méthode d'entraînement ou la méthode de l'arête. Dans ce genre de méthodes, on doit faire usage d'imagination pour reconstituer le chemin de migration. Ces méthodes se limitent à trouver le point de selle de la diffusion et la barrière d'énergie qui s'y associe.

Des méthodes de recherche des trajectoires activées plus sophistiquées, comme la technique d'activation et de relaxation ou la méthode du dimère combinées avec des codes *ab-initio* sont de plus en plus populaires. Elles procurent un pouvoir de prédiction des trajectoires activées de diffusion des défauts dans les semiconducteurs. Une étude complète menée sur le GaAs, utilisant ART et le code *ab-initio* Siesta, est présentée dans le chapitre 7. D'autres groupes de recherche utilisent de plus en plus la méthode du dimère combinée avec VASP ; comme pour étudier l'autodiffusion dans le GaSb [40] et l'influence de l'hydrogène sur la migration des lacunes d'azote dans le GaN [41].

⁵Ces méthodes sont discutées dans la section 3.3.

	Electronic transition	Ionization levels (eV)					
		B & S ^a	J & S ^a	S & L ^a	C & C ^a	P <i>et al</i> ^a	N & Z ^a
V _{Ga}	+ 1/- 1			0.035			
	0/- 1	0.20	0.10			0.11	0.19
	- 1/- 2	0.50	0.35		0.49	0.22	0.20
	- 2/- 3	0.72	0.50		0.69	0.33	0.32
	- 1/- 3			0.078			
V _{As}	+ 2/+ 1	0.10					
	+ 1/0		1.30	1.41			
	0/- 1		1.40	1.54			
	+ 1/- 1				0.785	0.86	

^aB & S: Baraff and Schluter (1985), J & S: Jansen and Sankey (1989), S & L: Seong and Lewis (1995), C & C: Cheong and Chang (1994), P *et al*: Poykko *et al* (1996), N & Z: Northrup and Zhang (1993).

FIG. 2.1 – Comparaison entre les résultats des calculs réalisés durant les 20 dernières années concernant les niveaux d'ionisation de V_{Ga} et V_{As} dans le GaAs. Tiré de l'article de Deepak *et al.* [4].

2.3.1 Arséniure de gallium

L'arséniure de gallium est un composé semiconducteur de la famille des III-V qui cristallise dans la structure zinc-blende. Le caractère covalent qui domine sur les liaisons Ga-As donne aux structures à base de GaAs une dureté moyenne. Le GaAs possède un gap direct, ce qui fait de lui un des composés les plus importants pour la fabrication des composés optoélectroniques de haute performance. Il possède également une grande mobilité électronique, ce qui le sollicite pour la fabrication des transistors de très haute vitesse. Il est généralement utilisé comme base pour fabriquer les hétérojonctions, les superréseaux ou des structures à puits quantiques.

Ainsi, pour contrôler la fabrication de ces structures, il est indispensable de comprendre les mécanismes de diffusion qui peuvent y avoir lieu, et qui causent soit la dégradation du dispositif soit l'amélioration de son rendement. Quelques défauts ponctuels présents dans les dispositifs électroniques à base de GaAs régissent la diffusion

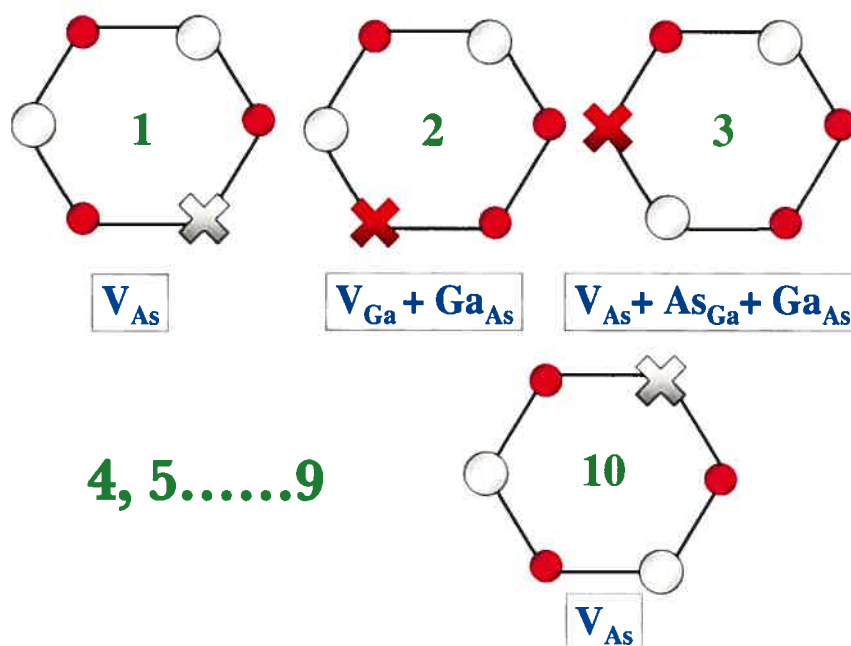


FIG. 2.2 – Illustration du mécanisme de diffusion en anneau proposé par Van Vechten [5]. Les atomes d'As sont en blanc, et les atomes de Ga sont en rouge. La lacune est représentée par un X. Lors des premiers sauts, une série d'antisites (As_{Ga} et Ga_{As}) est générée. Un second tour est nécessaire pour les annihiler. La lacune doit faire 10 sauts au premier voisin pour diffuser au second voisin.

à basse température tandis que d'autres types de défauts ponctuels deviennent actifs à haute température. Récemment, Bracht *et al.* [42] ont trouvé que les lacunes de Ga ont une contribution importante à l'autodiffusion qui est supérieure à celle des interstitiels à haute température.

Plusieurs études ont été menées afin d'étudier les propriétés statiques des lacunes d'arsenic (V_{As}) et de gallium (V_{Ga}). Le sous réseau de Ga est beaucoup plus étudié que le sous réseau de l'arsenic parce que les dopants utilisés en technologie (les accepteurs de Si, Zn, Be, Mg) préfèrent occuper des sites dans le sous réseau du Ga. L'usage du carbone comme accepteur occupant le sous réseau du As augmente l'intérêt à comprendre les mécanismes de diffusion des défauts à base d'As. On connaît peu de choses sur les lacunes d'arsenic. En plus, il y a très peu d'accord entre les travaux théoriques qui s'y sont intéressés.

Le tableau de la figure 2.1 est tiré d'un travail de synthèse de Deepack *et al.* [4]. Il montre le grand désaccord qui règne autour de la détermination des niveaux d'ionisation de V_{Ga} et V_{As} . Les résultats reportés dans ce tableau montrent que les calculs des 20 dernières années utilisaient des méthodes et des critères de calculs peu précis et insatisfaisants. Il montre aussi que la source de désaccord entre les différents travaux vient des différences dans les paramètres de simulation utilisés.

La plupart des calculs *ab-initio* [43–48] trouvent que les lacunes de Ga peuvent exister dans les états de charges $q = 0, -1, -2, -3$. Pour le GaAs de type N ou intrinsèque, la théorie affirme que c'est l'état de charge -3 qui domine, tandis que les expériences à haute température sont en désaccord sur ce point. Certains groupes proposent les états de charge -1 et -2 [49], alors que d'autres proposent la charge -3 [16].

Tous les mécanismes de diffusion directs possibles pour le Si sont *théoriquement* possibles pour le GaAs à l'exception de l'échange concerté mais ceci se passe de manière plus complexe. Van Vechten [5] a été le premier à imaginer un chemin de migration des lacunes dans les composés III-V. C'est un mécanisme dit *en anneau*, où il suppose que la diffusion au premier voisin est **toujours** possible et que les paires lacune-antisite qui en résultent sont **toujours** stables. Les sauts successifs au premier voisin (voir figure 2.2) entraînent l'accumulation d'antisites. Ces antisites sont par la suite annihilés lors d'un second tour effectué par la lacune. Ce chemin requiert 10 étapes au total pour diffuser au second voisin.

Bockstedte et Scheffler [6] ont essayé de simuler la diffusion de V_{Ga} dans une supercellule de 64 atomes en utilisant la dynamique moléculaire quantique Car-Parinello [50]. Le temps de simulation nécessaire pour voir la diffusion complète au second voisin s'est avéré interminable. Ils ont finalement utilisé la méthode d'entraînement pour calculer les barrières de migration au premier et au second voisin pour différents états de charge de la lacune de Ga. Ces simulations indiquent que la diffusion par sauts au premier voisin est impossible et que la diffusion a lieu par diffusion au second voisin seulement.

Le mécanisme de diffusion au second voisin est illustré dans la figure 2.3 : un atome de Ga second voisin de la lacune se déplace suivant la direction 110 vers la région interstitielle qui le sépare de la lacune. Pour que la migration ait lieu, il est *impératif* que

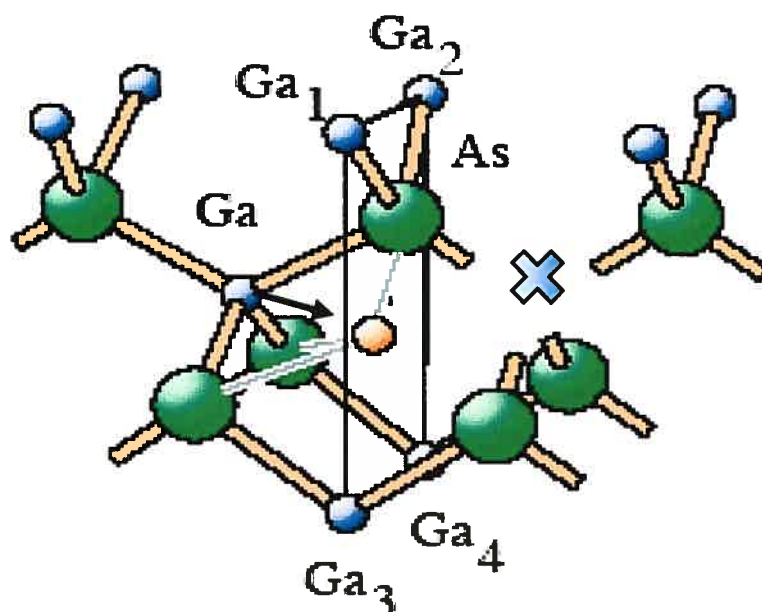


FIG. 2.3 – Chemin de migration de V_{Ga} vers le second voisin tiré de l'article Bockstedte et Scheffler [6]. Les quatre atomes $\text{Ga}_1, \dots, \text{Ga}_4$ forment le plan par lequel l'atome Ga doit passer. La flèche indique la direction du saut.

l'atome qui diffuse passe par le plan formé par les atomes $\text{Ga}_1, \dots, \text{Ga}_4$ dans la figure 2.3. La barrière de migration au second voisin est de 1.7 eV pour V_{Ga}^0 et 1.9 eV pour V_{Ga}^{-3} [6].

2.3.2 Carbure de silicium

Le carbure de silicium (SiC) est maintenant vu comme le candidat potentiel pour remplacer le silicium à cause de ses propriétés mécaniques et électroniques intéressantes. Il compte parmi les semiconducteurs ayant un large gap, ce qui le rend souhaitable pour la fabrication de dispositifs de haute puissance et de haute fréquence qui opèrent dans des conditions extrêmes (haute température, haute exposition aux radiations).

Le SiC est un composé IV-IV qui cristallise dans la structure zinc-blende. Il se distingue par sa dureté et sa haute résistance. Ceci est dû à la dureté des liens Si–C ayant un caractère purement covalent. Comme pour tous les semiconducteurs, les défauts ponctuels sont introduits dans le SiC durant l'implantation ionique ou le recuit. Donc, il est

intéressant de comprendre comment ces défauts diffusent.

Différents groupes de recherche ont mené plusieurs études par la simulation numérique sur les mécanismes de diffusion dans le SiC. Les travaux les plus récents sont cités ici et toute comparaison avec des résultats antérieurs se trouvent dans les articles en question.

Les propriétés statiques des défauts dans le SiC ont été étudiées par Torpo *et al.* [51] en utilisant les calculs *ab-initio* (LDA et LSDA). Ce groupe de recherche a mis l'emphasis sur la structure électronique et atomique des lacunes et des antisites en fonction de l'état de charge dans le polytype 4H-SiC. Ils ont donné une description détaillée des énergies de formation, des niveaux d'ionisation, et de la géométrie des structures à défauts relaxés. Ces calculs trouvent un effet U-négatif pour les lacunes de carbone en bon accord avec certaines expériences.

Bockstedte *et al.* [52] ont mené une étude détaillée sur les propriétés des lacunes et les interstitiels dans le SiC, entre autres, la relaxation, la stabilité, l'état de charge et les chemins de migration. Leurs résultats de relaxation sont en bon accord avec ceux de Torpo *et al.* [51]. Pour l'étude de la diffusion, ils utilisent la méthode d'entraînement et la méthode de l'arête. Un résultat fort intéressant révèle que la lacune de carbone V_C ne peut pas migrer au premier voisin. Le seul moyen qu'elle utilise pour diffuser est d'échanger sa position avec un atome de carbone qui est second voisin. Les barrières ainsi calculées sont de 5.2, 4.1 et 3.5 eV pour les états de charge +2, +1, 0 respectivement.

La lacune de Si (V_{Si}) serait stable pour les charges allant du positif au doublement négatif. Or, V_{Si} peut suivre le chemin vers le second voisin avec une barrière allant de 3.6 à 2.4 eV et qui diminue avec l'état de charge. Elle peut aussi sauter vers le premier voisin en franchissant une barrière plus faible pour former le complexe $V_C + C_{Si}$ qui est plus stable que V_{Si} . Ceci cause un sérieux problème car V_{Si} risque de se trouver piégée dans cet état plus profond $V_C + C_{Si}$. Par conséquent, la diffusion en anneaux par des sauts successifs [5] au plus proche voisin s'avère difficile.

En parallèle avec les travaux de Bockstedte *et al.* [52], il a y eu une série de travaux sur le SiC menés par Rurali *et al.* [53] en utilisant le code Siesta. Les résultats de ces

travaux sont en accord qualitatif avec les résultats de Bockstedte *et al.* [52] sur la plupart des points cités ci-dessus. En plus, Rurali *et al.* [53] proposent un nouveau mécanisme pour la diffusion de la lacune en trois étapes, mais les barrières à franchir sont trop élevées pour avoir un impact physique.

Enfin, étant donné que la barrière de diffusion des lacunes est aussi élevée que celle des interstitiels, l'autodiffusion dans le SiC est régie par les interstitiels qui sont plus mobiles que les lacunes.

2.3.3 Nitrure de gallium

Le nitrure de gallium (GaN) est un semiconducteur ayant une large bande interdite de 3.4 eV. C'est un composé binaire de la famille des III-V qui possède une semiconductivité intrinsèque. Le GaN cristallise dans la structure zinc-blende ou Wurtzite avec des liaisons atomiques Ga–N ayant un caractère fortement ionique. Il est utilisé en optoélectronique et dans les dispositifs de grande puissance ou de haute fréquence. Il est aussi peu sensible aux rayonnements ionisants, ce qui fait de lui un matériau approprié pour la fabrication des panneaux solaires des satellites. Le GaN est très dur, mécaniquement stable et a une capacité thermique élevée. Sous forme pure, il résiste à la fissuration et peut être déposé en film sur du saphir. Le GaN peut être dopé par du silicium pour devenir un semiconducteur de type N. Cependant, les atomes de silicium changent le mode de cristallisation de GaN en introduisant un grand nombre de défauts dans le matériau.

Une étude détaillée sur la formation et la diffusion des autodéfauts dans le GaN a été menée en 2004 par Limpiumnong et Van De Walle [54]. On y trouve une excellente comparaison avec les travaux théoriques et expérimentaux réalisés jusqu'à date sur ce matériau. Les barrières de diffusion des lacunes sont calculées en considérant les mécanismes de diffusion au premier et au second voisin seulement. Ici, la méthode d'entraînement combinée avec les calculs de principes premiers basés sur la LDA sont utilisées pour trouver les points de selle pour la diffusion.

Les principaux résultats révèlent que les interstitiels de Ga (I_{Ga}) dans le GaN se stabilisent dans le site octaédrique dans les états de charge +1, +3 mais sont métastables dans l'état de charge +2. Les interstitiels I_{Ga} diffusent avec une faible barrière d'environ 0.9

eV à travers la région interstitielle, ce qui confirme certains résultats expérimentaux [55]. Pour l'interstitiel d'azote I_N , il serait plus stable dans un site interstitiel dissocié et peut prendre la charge -1 , -2 et -3 . La migration a lieu par le même mécanisme que le précédent et requiert une barrière de 2.4 eV ou moins dépendamment de l'état de charge.

La lacune d'azote (V_N) dans le GaN existe dans deux états de charge $+1$ et $+3$. La barrière calculée pour la migration au second voisin est élevée (4.3 eV) pour l'état de charge $+1$ mais elle diminue considérablement avec la charge pour atteindre 2.6 eV pour $+3$. La lacune de Ga (V_{Ga}) dans le GaN existe dans les états de charge $q = 0, -1, -2, -3$. La lacune triplement négative est la plus stable à basse température, elle diffuse au second voisin en passant une barrière de 1.9 eV.

Finalement, une analyse de la compétition entre les différents défauts natifs dans le GaN par rapport aux conditions expérimentales de croissance montre que :

- Pour le GaN de type P, la diffusion dans le sous réseau du Ga est dominée par les interstitiels de Ga avec des barrières inférieures à 0.9 eV. La diffusion dans le sous réseau d'azote est régie par les lacunes d'azote V_N mais leur contribution baisse aussitôt qu'on commence à doper le matériau avec des donneurs.
- Pour le GaN de type N, les lacunes de Ga sont les plus favorisées et jouent le rôle principal pour la diffusion. Elles migrent préférentiellement au second voisin.

CHAPITRE 3

COMMENT IDENTIFIER LES CHEMINS DE DIFFUSION DES DÉFAUTS ?

Ce chapitre introduit les différentes méthodes couramment utilisées qui permettent de bien comprendre la dynamique d'un système en trouvant les trajectoires activées. La dynamique du système est déterminée par l'énergie nécessaire pour le passage d'une configuration donnée d'un minimum local vers le point de selle. En 1990, Feibelman [56] a découvert que la diffusion d'un atome d'aluminium sur la surface d'Al (100) est beaucoup plus complexe qu'on ne le pensait. L'intuition nous mène à penser que l'atome d'Al diffuserait en faisant des sauts d'un site à une autre, mais le vrai mécanisme est plus complexe. Les atomes d'Al diffusent à la surface grâce à un échange concerté avec un atome d'Al voisin situé à la surface. Dans plusieurs systèmes, les chemins de migration peuvent être plus complexes qu'on ne le pense et peu intuitifs. D'où la nécessité d'utiliser des méthodes exploratrices qui voyagent sur la surface d'énergie potentielle. Plusieurs méthodes numériques comme les méthodes dynamiques, les méthodes statistiques ou les méthodes de recherche des états de transition peuvent être utilisées.

3.1 Méthodes dynamiques

Nous discutons en premier lieu la dynamique moléculaire (DM) classique qui utilise des potentiels empiriques (Tersoff [57], Stillinger-Weber [58], etc.). Dans ce genre de calcul [59, 60], l'influence de la structure électronique du matériau est généralement mal approximée. Aussitôt que la coordination des atomes qui diffusent diffère trop de l'état cristallin, les potentiels empiriques sont incapables de bien représenter l'interaction entre les atomes. La DM classique ne peut prédire ni les configurations atomiques, ni les énergies et les barrières de migration. Elle peut, par contre, être utile pour explorer avec plus de détails des mécanismes prédits par d'autres méthodes car elle permet l'utilisation d'un grand nombre d'atomes.

Les insuffisances des potentiels empiriques peuvent être évitées en incluant les pro-

priétés électroniques soit par des méthodes peu coûteuses comme la méthode des liaisons fortes [61] soit par des méthodes plus précises comme la DFT [62]. Or l'augmentation de la précision limite le nombre d'atomes à utiliser et augmente le coût et le temps de calcul.

Ces méthodes dynamiques tant empiriques qu'*ab-initio* sont limitées par la période des vibrations atomiques qui est de l'ordre de 10^{-13} sec. Il devient donc nécessaire de faire des simulations à très haute température et pendant de longues périodes pour pouvoir observer des événements activés. À titre d'exemple, il est impossible d'observer le repliement des protéines avec la DM puisque la plus petite protéine se replie dans un temps caractéristique de l'ordre de la microseconde.

La tendance actuelle est d'accélérer les méthodes dynamiques pour trouver rapidement les événements activés et pour améliorer l'échantillonnage de l'espace des phases. Citons par exemple ¹ la méthode dynamique moléculaire accélérée par la température de Voter [64, 65], la méthode hyper dynamique [66], la dynamique des répliques parallèles [67], la dynamique autoguidée [68] et enfin la dynamique biaisée thermodynamiquement [69]. Jusqu'à date, l'application de ces méthodes dynamiques accélérées se limite aux systèmes simples et bien connus qui possèdent un nombre limité d'états de transitions. Ces méthodes dynamiques ne seront pas discutées avec plus de détails. Nous allons plutôt mettre l'emphase sur les méthodes d'exploration extensive de la surface d'énergie.

3.2 Théorie de l'état de transition

La théorie de l'état de transition (TÉT) est une méthode purement statistique particulièrement utile pour estimer les taux de transition pour les événements rares qui sont difficiles à obtenir avec la dynamique moléculaire. Le taux de transition d'un état A vers un état adjacent B (voir la figure 3.1) est représenté par le flux à travers la surface qui sépare les états A et B. Ce taux de transition est calculé dans la TÉT par la différence d'énergie libre entre le minimum A et le point de transition. Si les événements dyna-

¹Voter fait une bonne explication de quelques unes de ces méthodes dans un article publié en 2002 [63].

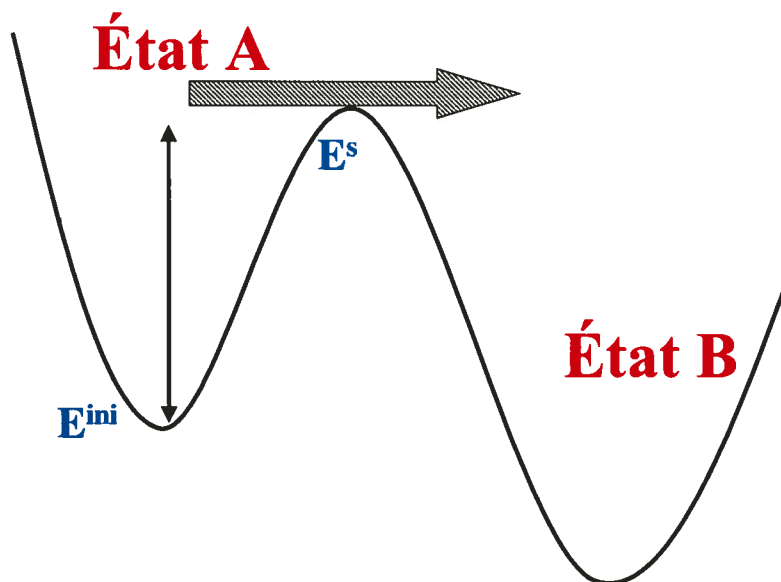


FIG. 3.1 – Schéma illustrant une transition de l'état A vers l'état B dans un espace à 2 dimensions. Le taux de transition, dans la TÉT, est le flux (représenté par une flèche) qui traverse la surface séparant les états A et B .

miques n'étaient pas corrélés, le taux calculé par la TÉT serait exact.

Supposons qu'on génère par la DM un événement simple de transition de l'état A vers l'état B . En examinant la trajectoire de transition, on peut déterminer la fraction de temps (χ_A) que la configuration passe dans l'état A ainsi que le nombre de fois, par unité de temps, que la surface séparant les états A et B a été traversée. Donc, le taux de sortie de l'état A ($\Gamma_{A \rightarrow}^{TÉT}$) n'est que la moitié du taux de transition $\Gamma^{TÉT}$ divisé par χ_A .

Dans le cas où l'événement de transition de A à B serait rare et très difficile à obtenir avec la DM, la TÉT de transition nous permet de calculer le taux de sortie ($\Gamma_{A \rightarrow}^{TÉT}$) de l'état A sans faire la simulation [63]. Cependant, la condition suivante doit être respectée [7] : "La surface qui relie les états A et B est de dimension $D - 1$ où D est le nombre de degrés de liberté dans le système. Cette surface est traversée une seule fois durant la transition de l'état A vers l'état B et elle représente le point de selle de la transition".

La moyenne d'ensemble est :

$$\Gamma_{A \rightarrow}^{TÉT} = \langle |dX/dt| \delta(X - S) \rangle_A \quad (3.1)$$

La fonction delta de Dirac $\delta(X - S)$ permet de localiser la configuration dont la coordon-

née réactionnelle se trouve à X par rapport au point de selle S . Elle permet également de trouver la probabilité que le système soit au point de selle.

Le terme entre crochets représente le rapport entre les intégrales pondérées par le facteur de Boltzmann dans l'espace des phases à $6N$ dimensions. L'indice A indique que l'intégrale sur l'espace des configurations \vec{r} est limitée par l'espace qu'occupe la configuration à l'état A . Pour une quantité donnée O :

$$\langle O \rangle = \frac{\iint O(\vec{r}, \vec{p}) \exp[-E_B(\vec{r}, \vec{p})/k_B T] d\vec{r} d\vec{p}}{\iint \exp[-E_B(\vec{r}, \vec{p})/k_B T] d\vec{r} d\vec{p}} \quad (3.2)$$

Où, k_B est le facteur de Boltzmann. Si la masse effective m de la coordonnée réactionnelle est constante (\vec{p} est aussi constant), l'équation 3.1 devient une moyenne d'ensemble sur l'espace des configurations seulement :

$$\Gamma_{A \rightarrow}^{TÉT} = \sqrt{2k_B T / \pi m} \langle \delta(X - S) \rangle_A \quad (3.3)$$

Aucune référence n'est faite à l'état final B ou à sa nature, mais il est impératif de connaître le point de selle S . À l'aide d'un Monte Carlo de type Metropolis [70], le taux de transition à une température donnée est directement calculé à partir de l'équation 3.3. Si la température est basse, le taux d'acceptation des sauts par l'algorithme de Métropolis n'est pas efficace. Il est donc nécessaire d'utiliser des méthodes plus avancées.

Maintenant, si les événements du système sont corrélés, les états de transition ne mènent pas toujours vers des événements activés. Donc, le taux de réaction, calculé par la TÉT, est surestimé par rapport au taux réel. Pour les événements de diffusion dans les matériaux à température moyenne, les événements dynamiques corrélés ne causent pas un grand changement au taux de transition calculé par la TÉT. Ce taux est même une excellente approximation contrairement aux systèmes chimiques (comme les réactions moléculaires à l'état gazeux ou en solution) où les corrections de la TÉT deviennent nécessaires. Citons, à titre d'exemple, le formalisme de correction dynamique de Chandler [71, 72] qui a été utilisé pour plusieurs systèmes [73].

La théorie de l'état de transition harmonique (TÉTh) est une approximation efficace

et très avantageuse qui peut être appliquée une fois que le point de selle est identifié. Dans la TÉTh, on suppose que les déplacements des atomes sont assez petits pour que l'énergie potentielle dans le bassin entourant le point A corresponde à une expansion du second ordre de l'énergie. En d'autres termes, les modes de vibration sont harmoniques au point A et au point de selle S .

La TÉTh est efficace pour l'étude des réactions dans les cristaux car les atomes sont proches les uns des autres et les réactions ont lieu à des températures nettement inférieures à la température de fusion. La recherche des états de transition se simplifie considérablement comparativement à la TÉT et se limite à la recherche de quelques points de selle qui entourent le bassin harmonique de la structure initiale A dans la surface d'énergie potentielle. Le taux de transition $\Gamma^{TÉTh}$ vers chacun des points de selle est calculé à partir des énergies et des fréquences des modes normaux de phonon au point de selle et au minimum initial [74] :

$$\Gamma^{TÉTh} = \frac{\prod_{i=1}^{3N} \nu_i^{(i)}}{\prod_{i=1}^{3N-1} \nu_i^{(s)}} \exp - \left(\frac{E^s - E^{init}}{k_B T} \right) \quad (3.4)$$

Où, E^s est l'énergie du point de selle et E^{init} est l'énergie au minimum initial. $\nu_i^{(s)}$ et $\nu_i^{(i)}$ correspondent aux modes normaux au point de selle et au minimum respectivement. L'utilisation de la TÉTh est conditionnée par l'identification des points de selle pertinents pour la diffusion ou la réaction. La coordonnée réactionnelle au point de selle correspond à la direction du mode de phonon instable ayant une valeur propre négative. Une fois que le point de selle est bien identifié, il suffit de suivre le gradient de l'énergie et minimiser l'énergie pour trouver l'état final. Il existe des cas particuliers comme la dynamique de repliement des protéines [75] où le chemin d'énergie minimum est très complexe. Les états final et initial sont reliés par plusieurs configurations métastables et plusieurs points de selle. Dans une telle situation, le point de selle le plus élevé en énergie est utilisé pour les calculs du taux de transition $\Gamma^{TÉTh}$.

3.3 Méthodes de recherche des trajectoires activées

Les méthodes de recherche des trajectoires activées jumelées avec les méthodes *ab-initio* sont beaucoup plus efficaces et plus appropriées au problème de diffusion. Elles offrent un compromis entre la précision et l'échantillonnage efficace de la surface d'énergie potentielle.

Les méthodes les plus simples ne peuvent pas prédire l'état final. Elles utilisent l'état initial et final pour trouver la trajectoire de transition. Parmi ces méthodes, nous présentons la méthode d'entraînement, la méthode de l'arête et la méthode de la bande élastique.

Les toutes premières méthodes développées pour **prédire** l'état final peuvent trouver les événements rares et voyager sur la surface d'énergie. Mais, ces méthodes sont guidées par la seconde dérivée de la force. Le principal désavantage de ces méthodes est qu'elles nécessitent la construction et la diagonalisation de la matrice dynamique du système. Pour un système de dimension D , le calcul des constantes de force nécessite D^3 évaluations de la force ce qui fait en sorte que ces méthodes sont limitées pour les systèmes de taille réduite comme les molécules.

Des méthodes plus sophistiquées [76] et plus rapides sont guidées dans leurs recherches par la valeur propre la plus basse seulement. La technique d'activation et de relaxation et la méthode du dimère (discutées ci-bas) font partie de cette catégorie. Ces méthodes sont très efficaces pour les systèmes continus comme les cristaux et il est possible de les utiliser tant avec des potentiels empiriques qu'avec les méthodes de principes premiers. La méthode du dimère et la technique d'activation et de relaxation sont formellement équivalentes car elles trouvent les points de selle du premier ordre bien que différents algorithmes soient utilisés pour cette fin. D'après une étude comparative de Olsen *et al.* [76] sur l'exemple de la diffusion d'un îlot de Pt sur la surface de Pt(111), les deux méthodes peuvent trouver les mêmes points de selle de migration en faisant un nombre équivalent d'appels de force.

3.3.1 Méthode d'entraînement

Communément connue sous le nom de *Drag method* [6], cette méthode nécessite la connaissance basée sur l'intuition de la direction de la transition ainsi que les états initial et final. La direction de la transition représente par exemple la direction de diffusion d'un défaut en volume ou en surface. Pour un système ayant D degrés de liberté, une seule coordonnée est choisie et utilisée comme coordonnée réactionnelle ou d'entraînement. Le point de selle est trouvé en suivant les étapes suivantes :

1. Faire évoluer la coordonnée d'entraînement avec un petit pas de l'état initial vers l'état final. Tout le système est donc entraîné durant cette recherche du point de selle.
2. Minimiser l'énergie du système à chaque pas d'entraînement par rapport aux $D - 1$ degrés de liberté.

Si aucune information sur la direction initiale de transition n'est disponible, la coordonnée réactionnelle est approximée par la ligne droite obtenue après interpolation entre les états initial et final. La figure 3.2 montre l'évolution du chemin de transition en utilisant la méthode d'entraînement. La coordonnée réactionnelle intuitive ou en ligne droite peut constituer un très mauvais choix, ce qui rend la méthode d'entraînement totalement inefficace pour certains systèmes. À titre d'exemple, Henkelman *et al.* [7] ont montré que cette méthode échoue à trouver le bon chemin de réaction pour la diffusion d'îlot de 7 atomes de Pt sur la surface de Pt (111) si la coordonnée réactionnelle en ligne droite est considérée.

3.3.2 Méthode de l'arête

Communément connue sous le nom de *Ridge method*, cette méthode a été implémentée par Ionova et Carter [77] en 1993. Dans cette méthode, il est indispensable de connaître l'état final de la réaction pour aller à la recherche du point de selle. Elle est plus développée que la méthode d'entraînement, car elle explore la surface d'énergie potentielle afin de trouver le point de selle. Le point de selle est approximé par l'arête qui relie les états initial et final (Voir figure 3.3) et sa recherche suit les étapes suivantes :

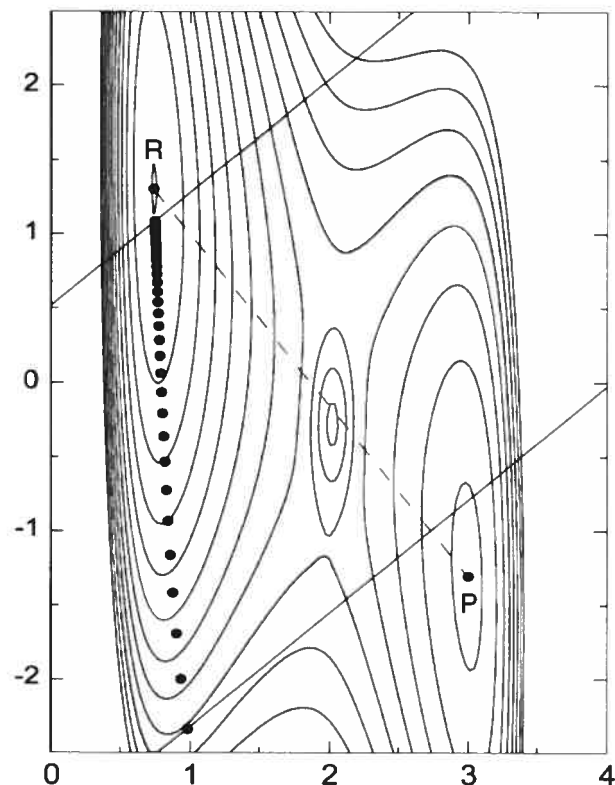


FIG. 3.2 – Recherche du point de selle par la méthode d’entraînement (Tiré de la référence [7]). Un exemple où la coordonnée d’entraînement est choisie en interpolant par la ligne en points tillés entre les états initial (R) et final (P). À chaque itération, la coordonnée d’entraînement est incrémentée d’un petit pas puis elle est gardée fixe tout en relaxant les autres degrés de liberté du système. La trajectoire est représentée par les points. Dans ce cas particulier de la diffusion d’un îlot de Pt sur la surface de Pt, la méthode d’entraînement est incapable de localiser le point de selle.

1. Faire une interpolation linéaire entre les deux minimums ;
2. Identifier l’arête comme étant le point ayant la plus haute énergie ;
3. Créer deux images de l’état initial et de l’état final sur la ligne droite, en appliquant une légère distorsion sur ces configurations ;
4. Faire évoluer les deux images dans un processus itératif en les rapprochant l’une de l’autre dans l’espace des phases ;
5. Minimiser la barrière d’énergie le long de la ligne qui relie les deux images durant le processus itératif ;

6. Arrêter quand les deux images convergent l'une vers l'autre.

Dans cette méthode, plusieurs évaluations de la force sont nécessaires surtout quand les deux images se rapprochent du point de selle. Le rendement de la méthode, à cette étape, peut être amélioré en suivant les suggestions proposées par Ionova et Carter. Cette méthode s'avère totalement inefficace dans quelques cas comme l'étude de la trajectoire de diffusion des lacunes dans le SiC [52]. Les lacunes de Si ou de C introduisent des niveaux électroniques dans la bande interdite ayant des symétries différentes et qui se croisent le long du chemin de migration. Le croisement de ces niveaux de la lacune conduit à une distorsion structurelle des images qui déstabilise le processus itératif de la méthode de l'arête.

3.3.3 Méthode de la bande élastique supportée

Communément connue sous le nom de *Nudged elastic band method* [78, 79], cette méthode s'inspire de l'idée des images comme dans la méthode de l'arête mais en utilisant en même temps, plusieurs images du système. Le chemin de transition est représenté par une série d'images connectées les unes aux autres par des ressorts. La série d'images initiales est créée en faisant une interpolation linéaire entre les états initial et final. Dans une seconde étape, cette ligne est relaxée itérativement en appliquant un algorithme d'optimisation. Les $N + 1$ images qui forment la bande élastique sont situées à $R_I, R_1, R_2, \dots, R_F$ dans l'espace des phases. Pour l'image $i = 1 \dots N$, la position initiale est définie comme $R_i = R_I + i/N(R_F - R_I)$. Les configurations aux états initial et final R_I et R_F sont gardées fixes durant l'optimisation alors que les configurations des $N - 1$ images intermédiaires peuvent varier.

La fonction S , définie ci-bas, est minimisée en fonction des coordonnées des images intermédiaires R_1, \dots, R_N .

$$S(R_1, \dots, R_N) = \sum_{i=1}^{N-1} E(R_i) + \sum_{i=1}^N \frac{k}{2} (R_i - R_{i-1})^2 \quad (3.5)$$

Où, $E(R_i)$ est l'énergie totale calculée pour chaque image et k est la constante de ressort.

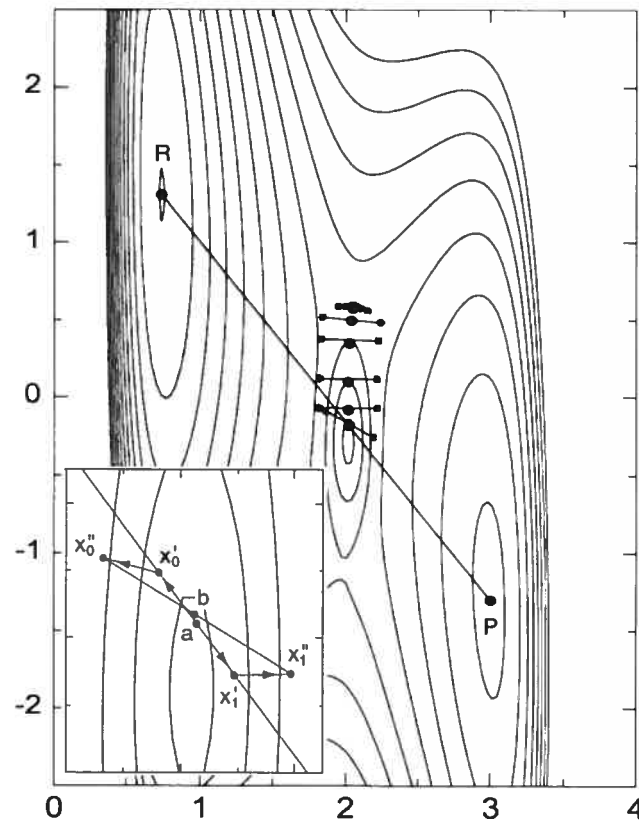


FIG. 3.3 – Recherche du point de selle par la méthode de l'arête (Tiré de la référence [7]). Dans la figure agrandie en bas à gauche, l'arête a est initialement identifié, les deux images sont situées initialement sur la ligne droite qui relie l'état initial (R) et l'état final (P) à x_0^i et x_1^i . Après un pas d'optimisation, la position des deux images devient x_0^{ii} et x_1^{ii} et le nouvel arête b est identifié. Durant l'optimisation, les deux images se rapprochent l'une de l'autre et le nouvel arête est identifié à chaque fois jusqu'à ce que les images convergent vers le point de selle.

Cette approche ressemble beaucoup au modèle de la bande élastique formée par des balles et des ressorts. Pour un système aussi simple, deux problèmes importants peuvent surgir durant l'optimisation :

1. Si la constante de ressort k est trop petite, il est très difficile de trouver le point de selle car les images ont tendance à se concentrer sur la partie inférieure de la trajectoire d'énergie. Ceci conduit à une sous-estimation de la barrière de transition parce que les parties supérieures de la bande proche du point de selle sont mal représentées.

2. Si k est trop grande, la force de ressort adoucit la courbure de la bande élastique et l'empêche de relaxer correctement. Ceci donne une surestimation de la barrière de transition.

Il n'est pas toujours facile de trouver un choix judicieux de la constante de ressort. Ces problèmes sont résolus en ajoutant un certain support aux images et en éliminant l'influence néfaste des composantes de la force de ressorts et de la vraie force. La force, sur chaque image, est donc modifiée et elle comporte seulement la composante parallèle de la force de ressorts et la composante perpendiculaire de la vraie force de l'image

$$F_i = -\Delta E(R_i) |_{\perp} + F_{\parallel}^{\text{ressort}} \quad (3.6)$$

$\Delta E(R_i)$ est le gradient de l'énergie en fonction des coordonnées des atomes dans un système composé de i images. Dans cette version de la méthode de la bande élastique "supportée", les images convergent mieux vers le chemin de migration minimum et la distance entre ces images est constante (comme illustré sur la figure 3.4). Le point de selle est obtenu en faisant une interpolation linéaire entre les coordonnées des deux images ayant les plus hautes énergies. Cette estimation est parfois peu efficace. Henkelman *et al.* [80] corrigent cette situation en introduisant un nouveau concept ; celui de "l'image grimpante" (*Climbing image*). Dans cette méthode, on libère une seule image des constantes de ressort pour lui permettre de converger vers le vrai point de selle. Après quelques itérations avec la méthode de la bande élastique, l'image qui semble avoir l'énergie la plus élevée—qui mènera probablement au point de selle— est poussée encore plus haut en énergie le long de la bande élastique. On annule complètement la constante de ressort sur cette image ; seule la composante parallèle de la vraie force contribue à la force totale de la bande élastique. La méthode de la bande élastique a subi plusieurs améliorations, elle est, de nos jours, largement utilisée aussi bien avec des potentiels empiriques qu'avec des codes *ab-initio*. Elle est généralement utilisée pour trouver les états de transition ou pour raffiner les états de transition trouvées par d'autres méthodes. Cette méthode se distingue par sa capacité de reconstituer entièrement le chemin d'énergie minimum. Elle est particulièrement utile pour détecter si d'autres points

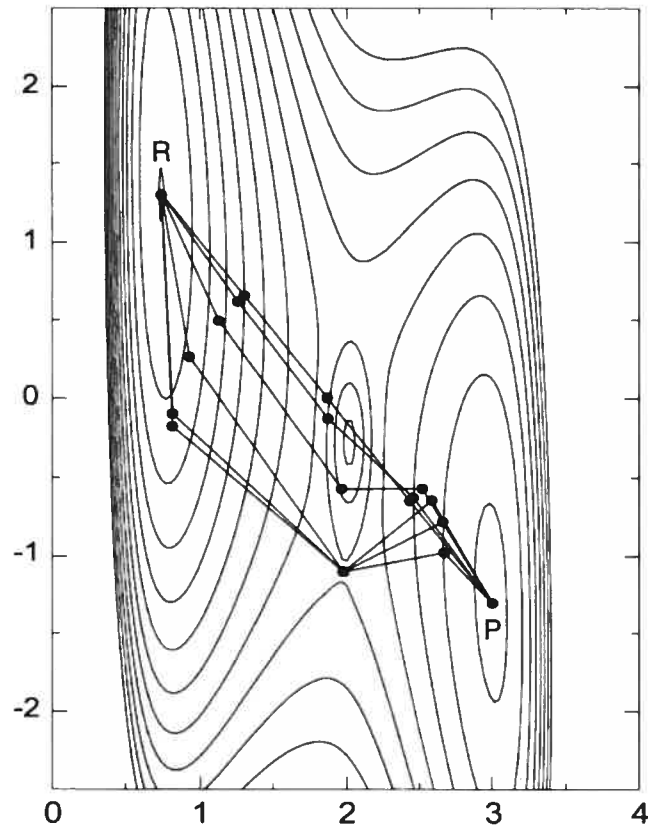


FIG. 3.4 – Le chemin de migration tel que reconstitué par la méthode de la bande élastique (Tiré de la référence [7]). Ici cinq images (les points noirs) dont trois images intermédiaires sont utilisées. Elles sont initialement placées sur la ligne droite qui relie l'état initial (R) et l'état final (P). Après quelques étapes d'optimisation, les images relaxent vers le chemin de migration minimal. La troisième image allant de la droite converge vers le point de selle.

de selle ou des configurations métastables apparaissent le long du chemin de transition.

Le groupe de David Wales [81, 82] a récemment apporté une modification à la méthode de la bande élastique supportée qui améliore jusqu'à deux fois son efficacité. Elle a reçu le nom de la méthode de la bande élastique "doublement supportée" (DNEBM).

3.3.4 Technique d'activation et de relaxation

La technique d'activation et de relaxation (TAR) est utilisée pour explorer la surface d'énergie potentielle à partir d'un minimum initial et sans la moindre connaissance

de l'état final ou des états intermédiaires. TAR a été implémentée par Mousseau et Barkema [83] en 1996. Elle a subi quelques modifications pour améliorer son rendement et accélérer le processus de recherche de point de selle. Cette nouvelle version est connue sous le nom "TAR-nouveau" [84]. La recherche de la trajectoire activée avec TAR-nouveau nécessite trois étapes :

1. Quitter le bassin harmonique ;
2. Converger vers le point de selle ;
3. Relaxer au prochain minimum.

Durant la première phase de l'algorithme, la configuration doit quitter le bassin harmonique. Cette partie est très sensible aux détails du système étudié. Il est souvent nécessaire d'essayer plusieurs moyens pour quitter le bassin. Les semiconducteurs amorphes et les protéines ont un bassin harmonique très étroit qui est facile à quitter. Cependant, les matériaux plus compacts comme les verres métalliques et les matériaux flexibles comme les verres de silice ont un bassin harmonique large et très difficile à quitter.

La matrice dynamique au minimum ne contient aucune information sur les points de selle qui entourent le minimum. Un événement commence à partir d'un minimum local où la force totale est nulle ; une direction aléatoire qui pointe loin du minimum est choisie. Pour éviter de créer plusieurs événements compétitifs, ce déplacement aléatoire est localisé dans l'espace (à titre d'exemple seulement un atome et quelques uns de ses voisins peuvent bouger aléatoirement). La configuration est poussée lentement le long de cette direction loin du minimum initial. À chaque étape, (1) l'énergie totale est relaxée dans le plan perpendiculaire au déplacement aléatoire (2) la valeur propre la plus basse de la matrice dynamique est évaluée. Puisque nous avons besoin de connaître seulement la valeur propre la plus basse, il n'est donc pas nécessaire de diagonaliser la matrice dynamique. Cette valeur propre est calculée à partir de l'algorithme de Lanczòs [85]— présenté dans l'annexe II— nécessitant des opérations d'ordre N . La direction aléatoire est suivie jusqu'à ce que la courbure de la surface d'énergie devienne négative le long d'une seule direction. En d'autres termes, on continue à suivre cette direction jusqu'à ce que la valeur propre la plus basse dépasse une certaine valeur critique, indiquant que

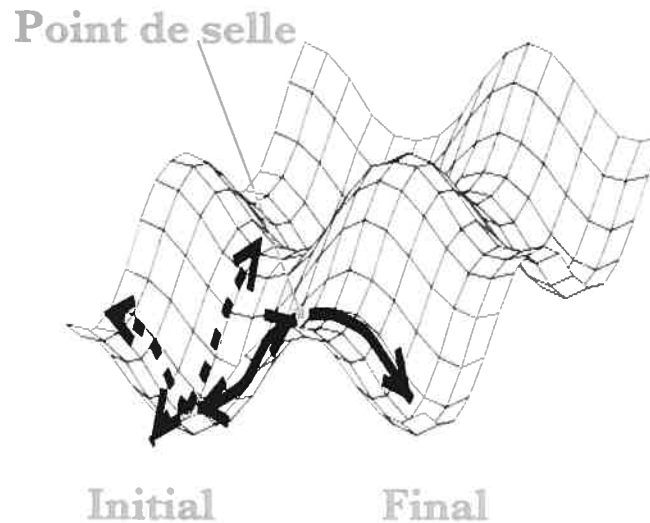


FIG. 3.5 – Exploration de la surface d'énergie par la technique d'activation et de relaxation

la configuration a quitté le bassin harmonique (voir la figure 3.5). Plusieurs tentatives peuvent être nécessaires avant d'arriver à la valeur critique de la valeur propre négative et trouver la bonne direction qui permet de quitter le bassin harmonique.

La seconde étape se charge de converger vers le point de selle. Une fois que la configuration est poussée loin du bassin harmonique, une valeur propre négative qui correspond à la direction de courbure négative est déjà apparue. L'activation continue suivant cette valeur propre négative et son vecteur propre. À chaque étape, l'énergie est minimisée dans l'espace à $3N - 1$ dimensions. La convergence au point de selle est atteinte quand les forces deviennent égales à zéro. La barrière de transition $E_{Barrière}$ peut être calculée en prenant la différence d'énergie entre l'état initial et le point de selle.

L'étape de relaxation du point de selle trouvé vers le nouveau minimum peut commencer. Pour faire la minimisation globale, on peut utiliser plusieurs méthodes comme la méthode du gradient conjugué. Un événement est accepté, selon l'algorithme de Metropolis [70], si la probabilité d'acceptation de chaque événement correspond au facteur de Boltzman $P_{accepté} = \min[1, \exp(-\Delta E_{Barrière}/k_B T)]$, où T est une température fictive.

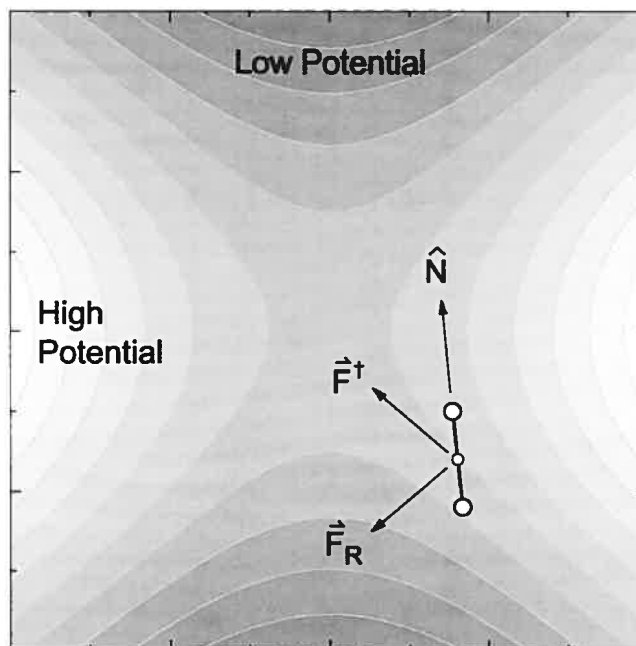


FIG. 3.6 – Une représentation du processus itératif dans la méthode du dimère (Tiré de la référence [7]). À chaque pas, le dimère formé par deux images du système change son orientation progressivement jusqu'à ce qu'il se stabilise suivant le vecteur propre associé à la valeur propre la plus basse. Après la stabilisation, la configuration des images est incrémentée et l'énergie est minimisée de nouveau en variant l'orientation.

La technique d'activation et de relaxation est une méthode puissante qui permet de voyager dans la surface d'énergie potentielle. Elle est utilisée avec succès pour identifier les événements complexes et peu attendus dans de nombreux systèmes comme les semiconducteurs amorphes [86], les verres, les protéines [75, 87–89] etc.

3.3.5 Méthode du dimère

La première version de la méthode du dimère a été proposée par Henkelman et Jónsson [90] en 1999. Cette méthode est capable de prédire le point de selle de la transition sans aucune connaissance de l'état final. La valeur propre la plus basse est déterminée par une méthode différente de celle utilisée dans TAR-nouveau. Elle se base sur une idée proposée par Voter [66] avec laquelle la valeur propre la plus basse est trouvée en utilisant la première dérivée de l'énergie seulement.

Cette méthode utilise un dimère composé de deux images du système. Les configurations de ces images sont $X_l^1 = X_l + \delta X_D \hat{N}$ et $X_l^2 = X_l - \delta X_D \hat{N}$. Où, X_l est la configuration actuelle du système, δX_D représente le déplacement des deux images le long d'un vecteur unitaire \hat{N} . Ce vecteur est choisi aléatoirement lors du premier cycle. Pour les cycles suivants, il correspond au vecteur propre associé à la valeur propre la plus basse.

L'énergie du dimère est définie comme étant la somme des énergies des images ($V_D = V_1 + V_2$). Cette énergie est minimisée par la méthode du gradient conjugué² en gardant δX_D et X_l fixes. Le seul paramètre qui change est l'orientation du dimère ou en d'autre terme l'angle de rotation du dimère autour de son centre. À la fin du cycle de minimisation, la force qui agit sur le centre du dimère est modifiée et le dimère s'aligne suivant le vecteur propre associé à la valeur propre la plus basse. En changeant progressivement l'orientation du dimère, la valeur propre la plus basse qui conduit vers le point de selle est identifiée. La figure 3.6 montre une configuration du dimère dans l'espace des phases. Durant chaque cycle itératif, un nouveau dimère est créé en suivant la direction du vecteur propre associé à la valeur propre la plus basse et la configuration avance vers le point de selle de la transition. La méthode du dimère est utilisée avec succès en conjonction avec le code *ab-initio* VASP pour simuler les transitions dans plusieurs systèmes. Citons les simulations les plus récentes qui étudient l'agrégation des dopants dans le Si [91], l'autodiffusion dans le GaSb [40] et l'influence de l'hydrogène sur la migration des lacunes d'azote dans le GaN [41].

²La méthode du gradient conjugué est la plus efficace, mais d'autres méthodes peuvent être utilisées aussi.

CHAPITRE 4

RÉSOLUTION DU PROBLÈME DE LA STRUCTURE ÉLECTRONIQUE

La théorie fondamentale qui décrit les matériaux en matière condensée se base sur les interactions Coulombiennes entre les électrons et les noyaux. Ces interactions sont décrites par l'Hamiltonien non-relativiste à N-corps dont la forme générale est la suivante :

$$\hat{H} = - \sum_{I=1}^{N_n} \frac{\hbar^2}{2M_I} \nabla_I^2 - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m} \nabla_i^2 + \frac{e^2}{2} \sum_{I=1}^{N_n} \sum_{J \neq I}^{N_n} \frac{Z_I Z_J}{|R_I - R_J|} + \frac{e^2}{2} \sum_{i=1}^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|r_i - r_j|} - e^2 \sum_{I=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_I}{|R_I - r_i|} \quad (4.1)$$

Où R et M représentent la position et la masse des noyaux alors que r et m sont la position et la masse des électrons. Les deux premiers termes sont les énergies cinétiques des noyaux et des électrons respectivement. Les autres termes décrivent les interactions noyau-noyau, électron-électron et électron-noyau respectivement. Toutes les propriétés de ce système à N-corps peuvent être obtenues après la résolution de l'équation de Schrödinger :

$$\hat{H}\psi_i(r, R) = E_i\psi_i(r, R) \quad (4.2)$$

Cette équation est très complexe et très difficile à résoudre analytiquement car la fonction d'onde $\psi_i(r, R)$ contient l'information sur toutes les particules (électrons et noyaux) qui interagissent mutuellement. La solution analytique n'est possible que pour les systèmes à un électron [92]. Dès que le nombre d'électrons dans le problème augmente (comme l'exemple de l'atome de He), l'équation 4.2 ne pourrait plus être résolue exactement.

Or, de nos jours, la simulation numérique nous permet d'estimer les énergies totales pour des systèmes réalistes traitant des cellules unitaires comportant des centaines d'atomes. Ceci n'est devenu possible qu'après la simplification de l'Hamiltonien d'interaction à N-corps en introduisant une série d'approximations discutées dans l'ordre

suivant : l'approximation de Born-Oppenheimer est introduite en premier, elle permet de décomposer la fonction d'onde en fonction d'onde électronique et fonction d'onde du noyau. En second lieu, la théorie de la fonctionnelle de la densité est discutée brièvement. Puis, la méthode du pseudopotentiel et les méthodes de construction de la fonction d'onde sont présentées. Enfin, la méthode SIEST-A-RT est expliquée en mettant l'emphasis sur le choix de la méthode SIESTA pour le calcul des énergies totales et des forces.

4.1 Approximation de Born-Oppenheimer

Appelée également l'approximation adiabatique [93], cette approximation est toujours introduite en premier. Elle se base sur le fait que la vitesse typique des électrons est plus importante que la vitesse des ions qui forment l'atome. La vitesse électronique étant de l'ordre de 10^8 cm/sec, alors que la vitesse typique des ions est de l'ordre de 10^5 cm/sec. Les ions sont donc beaucoup plus lents que les électrons car la masse des ions est nettement supérieure à la masse des électrons.

Dans l'approximation adiabatique, on décompose le problème en traitant les électrons et les noyaux séparément. Durant les calculs avec la dynamique moléculaire quantique ou durant la minimisation, on peut traiter les ions comme des particules classiques. Pour chaque position instantanée des ions, on peut supposer que les électrons occupent l'état fondamental du Hamiltonien électronique.

On peut donc décomposer la fonction d'onde en la fonction d'onde des ions $\Theta(R, t)$ et la fonction d'onde électronique $\phi_n(R, r)$:

$$\psi_{BO}(R, r, t) = \Theta(R, t)\phi_n(R, r) \quad (4.3)$$

En supposant que le système est à l'état fondamental, les états électroniques sont la solution de l'équation de Schrödinger indépendante du temps :

$$\hat{h}_e\phi_n(R, r) = \epsilon_n(R)\phi_n(R, r) \quad (4.4)$$

Où, \hat{h}_e est le Hamiltonien électronique pour une configuration particulière des ions.

$$\hat{h}_e = - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m} \nabla_i^2 + \frac{e^2}{2} \sum_{i=1}^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|r_i - r_j|} - e^2 \sum_{I=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_I}{|R_I - r_i|} \quad (4.5)$$

On résout l'équation 4.4 pour avoir la fonction d'onde électronique $\phi_n(R, r)$ correspondant à une position donnée des ions. Les ions sont soit traités comme des particules classiques et bougent suivant les équations de mouvement de Newton (DM) ou suivant un algorithme préétabli (NEBM, ART, etc.). Ils peuvent aussi être traités quantiquement en résolvant l'équation de Schrödinger dépendante du temps pour le système ions + électrons :

$$\frac{\partial \psi_{BO}(R, r, t)}{\partial t} = \hat{H} \psi_{BO}(R, r, t) \quad (4.6)$$

4.2 Théorie de la fonctionnelle de la densité (DFT)

La théorie de la fonctionnelle de la densité est l'approximation la plus utilisée de nos jours en matière condensée, sa première formulation en 1964 revient à Hohenberg et Kohn [94]. Elle permet de comprendre la physique de base des systèmes à plusieurs électrons en simplifiant la fonction d'onde complexe. Cette théorie a été largement discutée dans plusieurs ouvrages [95, 96] et plusieurs thèses. Pour cette raison, nous avons choisi de faire un survol de cette approximation en passant par les idées principales qui la régissent. Nous discuterons, par la suite, les différents types de fonctionnelles largement utilisées en matière condensée et qui apportent une simplification considérable au problème.

La DFT simplifie le problème à N-corps vers un problème à un seul corps. La valeur clé, dans la théorie de Hohenberg et Kohn, est la densité électronique $n(r)$. *Pour un système d'électrons interagissant sous l'effet d'un potentiel externe $V_{ext}(r)$, il est possible d'évaluer ce potentiel connaissant la densité électronique à l'état fondamental $n(r)$.* Si cette densité est connue, il sera possible d'obtenir la fonction d'onde, le potentiel ainsi que les différentes observables.

L'énergie de l'état fondamental est déterminée uniquement par la densité $n(r)$. C'est

une fonctionnelle de la densité $E[n]$. Elle est réécrite en séparant l'interaction des électrons avec le potentiel externe et le reste de l'énergie :

$$E[n] = \int V(r)n(r)dr + F[n] \quad (4.7)$$

Où $F[n]$ est la somme de l'énergie cinétique des électrons interagissants et leur énergie d'interaction Coulombienne mutuelle :

$$F[n] = U[n] + T[n] = \langle \Psi_0 | \widehat{V}_{ee} + \widehat{T} | \Psi_0 \rangle \quad (4.8)$$

À ce niveau, nous introduisons le second point dans le théorème de Hohenberg et Kohn : *Si le potentiel externe $V(r)$ est maintenu fixe et que la densité électronique commence à varier, la valeur de $n(r)$ qui minimise l'énergie totale $E[n]$ est la vraie densité électronique du système à l'état fondamental associée au potentiel externe $V(r)$. Donc, si nous connaissons la forme de la fonctionnelle $F[n]$, il serait possible d'obtenir l'énergie de l'état fondamental en minimisant l'énergie par rapport à la variation de la densité électronique ¹.*

Cette théorie simplifie considérablement le problème, mais elle ne peut pas être directement appliquée dans les codes *ab-initio* tant que la forme de la fonctionnelle $F[n]$ n'est pas assez simple et bien connue.

Kohn et Sham [17] ont travaillé pour donner une forme plus simple au terme $F[n]$ qui comprend les potentiels électroniques. Au lieu de considérer l'interaction électron-électron individuellement, on suppose que chaque électron situé à la position r ressent l'effet des autres électrons par la distribution de leur densité au point r' . On définit un potentiel d'interaction classique entre les électrons, appelé potentiel de Hartree, qui décrit le potentiel exercé par la distribution de la densité électronique $n(r')$ sur un électron situé à r :

¹ $n(r)$ peut varier dans la limite inférieure $n[r] \geq 0$ et la limite supérieure $\int n(r)dr = N_e$ où N_e est le nombre total des électrons.

$$V_H(r) = \int \frac{n(r')}{|r - r'|} dr' \quad (4.9)$$

L'énergie de Hartree est une énergie purement classique qui provient de l'interaction de la densité électronique avec elle même.

$$E_H[n] = \int V_H(r)n(r)dr = \frac{1}{2} \int \frac{n(r)n(r')}{|r - r'|} dr dr' \quad (4.10)$$

Après avoir approximé le potentiel d'interaction électron-électron par le potentiel de Hartree, il reste le terme d'énergie cinétique.

L'énergie cinétique des électrons (second terme dans l'équation 4.8) est approximée par l'énergie cinétique d'un système d'électrons non-interagissants. Toutes les autres formes d'interactions entre les électrons qui ne sont pas prises en compte dans le terme du potentiel de Hartree et dans le terme énergie cinétique sont représentés par un terme appelé *énergie d'échange et de corrélation*. L'énergie d'échange et de corrélation contient la différence d'énergie cinétique et potentielle entre les électrons réels interagissants et les électrons non interagissants.

L'énergie totale finale contiendra le potentiel externe, l'énergie de Hartree, l'énergie cinétique des électrons indépendants et finalement le terme connu sous le nom de l'énergie d'échange et de corrélation :

$$E[n] = \int V_{ext.}(r)n(r) + E_H[n] + T_e[n] + E_{XC}[n] \quad (4.11)$$

Maintenant, l'Hamiltonien électronique est simplifié et il dépend de la densité électronique. Les états électroniques des électrons non interagissants Ψ_i sont orthogonaux. Ils peuvent être calculés à partir de l'équation de Khon-Sham :

$$\hat{h}^{KS}\Psi_i(r) = -\frac{\hbar^2}{2m}\nabla^2 + V_{eff}\Psi_i(r) = \epsilon_i\Psi_i(r) \quad (4.12)$$

Où V_{eff} est le potentiel effectif que ressentent les électrons indépendants. Il comprend le potentiel externe provenant des noyaux, le potentiel de Hartree ainsi que le

potentiel d'échange et de corrélation :

$$V_{eff} = V_{ext}(r) + V_H[n] + V_{XC}[n] \quad (4.13)$$

Nous sommes, enfin, arrivés à la forme finale de l'équation qui devrait être résolue durant les calculs *ab-initio*. Cette équation est résolue par autocoherence selon les étapes suivantes :

1. Faire une estimation initiale de la densité électronique $n(r)$;
2. construire l'Hamiltonien de Kohn Sham ;
3. résoudre le problème à un corps et extraire les états Ψ_i ;
4. générer la nouvelle densité de charge à partir des états occupés ($n(r) = \sum_{Ne}^{i=1} |\Psi_i(r)|^2$) ;
5. réinjecter la nouvelle densité et itérer jusqu'à ce que l'autocoherence soit atteinte.

En d'autre terme, jusqu'à ce que les densités de deux itérations successives convergent l'une vers l'autre.

La seule quantité qui n'a pas été définie explicitement est la fonctionnelle d'échange et de corrélation. Celle-ci ne peut pas être connue exactement. Encore une fois, des approximations sont utilisées pour donner une forme satisfaisante de cette fonctionnelle. La toute première approximation, connue sous le nom de l'approximation de la densité locale (nous utiliserons souvent l'abréviation LDA tirée de l'anglais *Local Density Approximation*), a été introduite par Kohn et Sham [17]. À chaque point, la densité électronique dans le matériau étudié par la simulation, possède une contribution au potentiel d'échange et de corrélation qui est la même que pour un gaz d'électrons homogène ayant la même densité.

$$E_{XC}^{LDA} = \int \epsilon_{xc}(n(r))n(r)dr \quad (4.14)$$

Où, ϵ_{xc} est l'énergie d'échange et de corrélation d'un gaz uniforme d'électrons de densité n .

La LDA a connu beaucoup de succès et demeure largement utilisée car elle permet de traiter des systèmes assez larges et donne des résultats satisfaisants malgré sa simplicité. Elle a des limitations qui sont bien connues, à savoir :

- Les atomes sont trop fortement liés (les liens sont donc plus courts que l'expérience) parce que les électrons ne sont pas assez localisés et que la densité électronique est généralement trop faible.
- Les résultats sont de qualité moyenne pour les systèmes avec des interactions longue portée (les liens hydrogène, les liaisons Van derWaals, etc).
- Les résultats sont médiocres pour les systèmes à électrons fortement corrélés.

Il existe d'autres approximations plus avancées que la LDA comme l'approximation du gradient généralisé [95] (GGA pour *Generalized Gradient Approximation*). Dans la GGA, l'énergie d'échange et de corrélation contient un premier terme semblable à celui dans la formulation de la LDA plus un second terme qui comprend une fonctionnelle de la densité $n(r)$ et du gradient de la densité $\nabla n(r)$:

$$E_{XC}^{GGA} = \int \epsilon_{xc}(n(r))n(r)dr + \int f(n(r), |\nabla n(r)|)n(r)dr \quad (4.15)$$

La GGA est a des problèmes bien connus comme :

- Elle surestime parfois la longueur des liens atomiques, $\approx 0-2\%$ plus que les valeurs expérimentales et les énergies de cohésion sont plus petites $\approx 10-20\%$.
- Les résultats s'améliorent par rapport à la LDA pour les liens hydrogène alors que pour les liaisons Van derWaals, il n'y a pas une nette amélioration.
- Elle est plus sollicitée pour l'étude des atomes, des molécules et des surfaces car elle décrit les effets d'échange et de corrélation dans des régions de densité réduite par rapport à la LDA.
- Les résultats restent toujours mauvais pour les systèmes à électrons fortement corrélés.

4.3 Les pseudopotentiels

Une des préoccupations principales dans les calculs de principes premiers est la représentation de la fonction d'onde $\psi_i(r)$. Quand l'électron se trouve dans les couches internes de l'atome et proche du noyau, il s'agit d'un électron du coeur. Son énergie cinétique est importante et sa fonction d'onde oscille fortement. Alors que, si l'électron se

trouve dans les couches externes de l'atome, il s'agit d'un électron de valence. À mesure que sa fonction d'onde s'éloigne de la région du coeur, son énergie cinétique diminue et sa fonction d'onde varie plus lentement et de manière plus homogène.

Les orbitales de valence sont responsables des liaisons chimiques entre les atomes, alors que les orbitales du coeur sont plus confinées autour du noyau. Toute perturbation introduite à l'environnement chimique de l'atome affecte considérablement les orbitales de valence alors que les orbitales du coeur ressentent très peu cette perturbation. Pour plusieurs systèmes, il est très utile de séparer les orbitales du coeur des orbitales de valence.

Une approximation peut être introduite en supposant que le noyau et les électrons dans les couches internes de l'atome forment le coeur de l'atome. Ce coeur est entouré par les électrons de valence. Dans la méthode du pseudopotentiel, le potentiel du coeur est remplacé par un pseudopotentiel ionique construit de manière judicieuse qui simule l'effet du coeur ionique et qui laisse les énergies de Kohn-Sham inchangées. Pour les atomes de la même espèce, on utilise donc un pseudopotentiel unique pour simuler l'effet du coeur sans se soucier de construire la fonction d'onde pour les électrons du coeur.

La méthode du pseudopotentiel est largement utilisée dans les calculs en matière condensée car elle diminue nettement le temps de calcul et permet de traiter des systèmes de plus en plus larges².

4.4 Construction de la fonction d'onde

Nous discutons dans cette partie l'efficacité de la base utilisés pour construire la fonction d'onde. En matière condensée, il est plus commun d'utiliser les orbitales atomiques localisées obtenues par une combinaison linéaire d'orbitales atomiques (LCAO) ou les ondes planes. Dans ce qui suit, nous présenterons les avantages et les inconvénients de chacune de ces méthodes.

²Seule la fonction d'onde des électrons de valence est construite.

4.4.1 Les ondes planes

La fonction d'onde est exprimée par une superposition d'ondes planes :

$$\psi_{nk}(r) = \sum_{G=0}^{G_{max}} C_{nk}(G) e^{iGr} \quad (4.16)$$

La somme est sur les vecteurs de l'espace réciproque. G_{max} représente la valeur limite de l'expansion en ondes planes.

Les avantages liés à l'usage des ondes planes sont :

- La qualité de la base peut être améliorée en ajustant un seul paramètre qui est l'énergie de coupure (G_{max}) qui doit être assez grande pour assurer la convergence ;
- Tous les points dans l'espace sont traités de la même manière ;
- Le calcul des éléments de la matrice Hamiltonienne est simple et peu coûteux.

Alors que ses inconvénients se résument en :

- Un grand nombre d'ondes planes est nécessaire pour décrire correctement chacune des orbitales ;
- Les conditions aux limites périodiques doivent être appliquées ;
- Les espaces vides qu'on trouve dans les supercellules contenant les agrégats ou les surfaces doivent aussi être remplis d'ondes planes ; ce qui alourdit considérablement les calculs.

4.4.2 Les orbitales atomiques

La fonction d'onde est exprimée par une combinaison linéaire d'orbitales atomiques ϕ_{μ} :

$$\psi_i(r) = \sum_{\mu} C_{i\mu} \phi_{\mu} \quad (4.17)$$

Les orbitales atomiques possèdent des avantages qui rendent leur utilisation en matière condensée de plus en plus populaire :

- Un petit nombre d'orbitales est nécessaire pour avoir une base efficace, ce qui

allège considérablement le calcul et le rend moins coûteux comparé avec les ondes planes ;

- Les fonctions de la base ont un caractère atomique. Donc, il est facile d’extraire l’information sur le transfert de charge, les populations de Mulliken ³ etc ;
- Les orbitales atomiques sont localisées, elles sont très utiles pour l’application d’algorithmes d’ordre N ;
- Les espaces vides qu’on trouve dans les supercellules contenant les molécules, les agrégats ou les surfaces n’ont pas besoin d’être représentés par les orbitales atomiques, ce qui allège considérablement les calculs.

Cependant, elles ont les inconvénients suivants :

- Manque de flexibilité lors de l’optimisation de la base localisée : Il y a un grand nombre de paramètres à ajuster pour obtenir la meilleure performance comme le nombre d’orbitales, leur forme, comment elles sont confinées et leur longueur, etc ;
- La base n’est pas orthogonale ; ce qui fait apparaître des erreurs lors de l’utilisation d’un grand nombre d’orbitales ;
- Le calcul de la matrice Hamiltonienne reste compliqué et coûteux.

4.5 Fonctionnement du code SIEST-A-RT

SIESTA (Spanish Initiative for Electronic Simulations with Thousands of Atoms) est l’implémentation d’une méthode de simulation numérique développée par un groupe espagnol dès 1996 [97]. Elle permet de faire des calculs de la structure électronique aussi bien que de la dynamique moléculaire quantique pour les molécules et les solides. C’est un code qui possède plusieurs caractéristiques qui font de lui un excellent candidat pour l’intégration avec la technique d’activation et de relaxation (TAR). Les principales caractéristiques de la méthode sont les suivantes :

- Utilise l’approximation de Born-Oppenheimer ;
- Utilise la méthode de la fonctionnelle de la densité de Kohn-Sham dans ses différentes formes comme l’approximation de la densité locale dépendante du spin

³Ce type d’analyse dépend fortement de la qualité de la base utilisée.

- (LSD) ou indépendante du spin (LDA) ainsi que l'approximation du gradient généralisé (GGA) et la fonctionnelle B3LYP ;
- Utilise les pseudopotentiels qui conservent la norme dans la forme non-locale de Kleinman-Bylander ;
 - Utilise une base générale et flexible constituée d'une combinaison linéaire d'orbitales atomiques (LCAO). La qualité de ces orbitales peut être améliorée en ajoutant la polarisation et en augmentant le nombre de fonctions radiales. Cette base allège beaucoup les calculs dans SIESTA et assure une rapidité dans le temps de calcul comparé à d'autres codes qui utilisent les bases à ondes planes ;
 - Utilise des schémas efficaces pour échantillonner des points spéciaux dans la zone de Brillouin (l'espace réciproque) ;
 - Projette les fonctions d'ondes et les densités électroniques sur une grille dans l'espace réel à 3 dimensions. Cette projection permet d'évaluer numériquement les potentiels de Hartree et d'échange et de corrélation ainsi que leurs éléments de matrice ;
 - Permet de faire la relaxation des géométries avec des cellules variables ou fixes (par la méthode du gradient conjugué)
 - Permet de faire de la dynamique moléculaire à température constante ou la dynamique en variant la cellule du type Parrinello-Rahman ;
 - Donne l'information sur la structure de bande électronique et les densités d'états électroniques locales ou projetées sur les orbitales ;
 - Pratique pour l'utilisation avec d'autres codes, comme ART, car il n'est pas nécessaire de recompiler le code SIESTA quand la dimension du problème change. Puisqu'il est écrit en Fortran90, la mémoire est allouée dynamiquement ;
 - Efficace en séquentiel et en parallèle (utilise MPI) et permet de traiter de larges systèmes.

Les calculs avec SIESTA peuvent donner les informations suivantes :

1. Les énergies totales et partielles ;
2. Les forces atomiques ;

3. Le tenseur de contrainte ;
4. Le moment dipolaire électrique ;
5. Les populations de Mulliken associées à chaque atome, orbitale ou à chaque lien atomique ;
6. La densité électronique.

Durant les différentes étapes de TAR, le code SIESTA est utilisé pour calculer les énergies totales, les forces et pour faire les minimisations. La figure 4.1 montre comment se passe l'interaction entre les codes TAR et SIESTA :

1. Avant de procéder à la recherche des chemins de diffusion, il est important de vérifier si les paramètres de SIESTA sont bien optimisés et si par exemple l'énergie de formation calculée est assez convergée. Il est, cependant, très important d'étudier, au préalable, la qualité des paramètres de SIESTA comme les pseudopotentiels, la densité des points spéciaux dans l'espace réciproque, la fonctionnelle d'échange et de corrélation, la taille de la boîte de simulation, la tolérance de convergence des cycles autoconsistants, la qualité de la base atomique et la finesse de la grille dans l'espace réel.
2. Durant l'étape d'activation, plusieurs appels de calcul de forces par SIESTA sont effectués. Par conséquent, la stabilité des forces calculées par SIESTA est cruciale pour la réussite de l'activation. Si la grille utilisée pour subdiviser l'espace réel n'est pas assez fine, un petit déplacement de la configuration engendre une déstabilisation de la force exercée.
3. Une fois que TAR converge vers le point de selle, la configuration est légèrement poussée dans la direction la plus négative puis l'énergie totale est minimisée par SIESTA en utilisant la méthode du gradient conjugué.

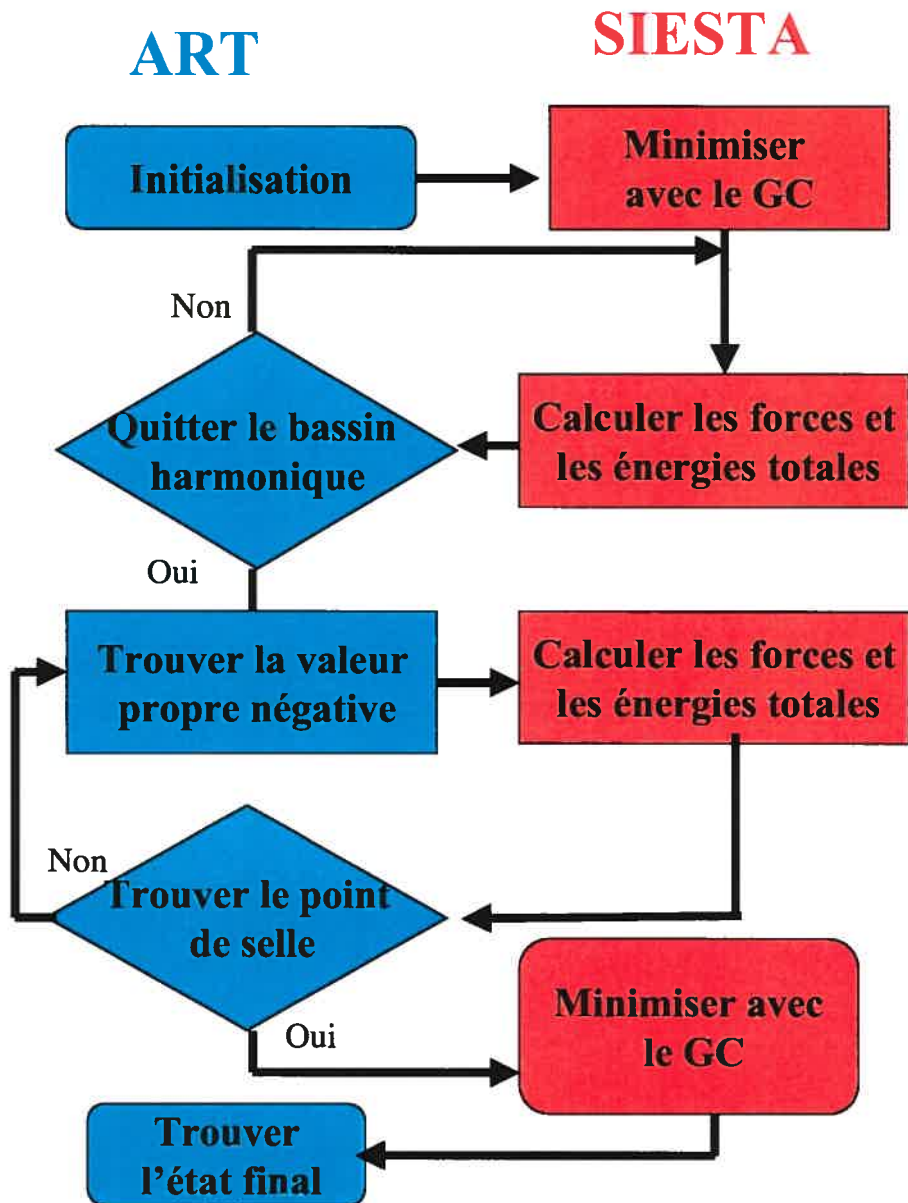


FIG. 4.1 – Organigramme qui montre l'interaction entre TAR et SIESTA durant la recherche des trajectoires activées.

CHAPITRE 5

L'AUTODIFFUSION DES LACUNES DANS LES SEMICONDUCTEURS ÉLÉMENTAIRES : SILICIUM

Ce chapitre présente de nombreux aspects reliés aux lacunes dans le silicium. Ces aspects sont traités par l'article intitulé : *Sampling the diffusion paths of a neutral vacancy in Silicon with SIEST-A-RT* et publié dans le journal Physical Review B, volume 70, pages 205202 :1-9 en date du 12 Mai 2004.

Cet article discute les facteurs qui influencent le calcul des énergies de formation des lacunes neutres dans le silicium et définit clairement les limites minimales qui donnent des énergies de formation assez convergées. L'influence de ces facteurs est générale et s'applique pour tous les semiconducteurs . L'article traite aussi les divers mécanismes associés à la lacune neutre dans le silicium. Les mécanismes de diffusion que nous avons identifiés montrent des sauts au premier voisin seulement. Nous avons aussi trouvé des mécanismes peu attendus tel que le réarrangement de la structure autour du défaut et des mécanismes d'échange de liens du type WWW (Wooten-Winer-Weaire). Nous avons reconstruit les chemins de migration détaillés à l'aide de la méthode NEBM et nous avons calculé l'entropie de migration et les préfacteurs pour l'autodiffusion. L'énergie de formation et la barrière de diffusion calculées pour la lacune silicium sont en bon accord avec les mesures expérimentales et d'autres simulations récentes.

Sampling the diffusion paths of a neutral vacancy in Silicon with SIEST-A-RT

Fedwa El-Mellouhi¹, Normand Mousseau¹ and Pablo Ordejón²

¹ Département de physique and Regroupement québécois sur les matériaux de pointe, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal (Québec) H3C 3J7, Canada

² Institut de Ciència de Materials de Barcelona (CSIC), Campus de la Universitat Autònoma de Barcelona, Bellaterra, E-08193, Barcelona, Spain

PACS numbers: 61.72.Ji, 71.15.Mb, 71.15.Pd, 71.55.Cn

Abstract

We report a first-principles study of vacancy-induced self-diffusion in crystalline silicon. Starting from a fully relaxed configuration with a neutral vacancy, we proceed to search for local diffusion paths. The diffusion of the vacancy proceeds by hops to first nearest neighbor with an energy barrier of 0.40 eV in agreement with experimental results. Competing mechanisms are identified, like the reorientation, and the recombination of dangling bonds by Wooten-Winer-Weaire process.

5.1 Introduction

Self-diffusion in homogeneous solids is a fundamental process of mass transport, in addition, it is responsible of the annealing of implantation damage, crystal to amorphous transition and the nucleation of extended defects. The native defects, such as vacancies and self-interstitial, have been identified as the prime entities controlling the self-diffusion, as well as the diffusion of impurities in semiconductor lattices. For example, in silicon, antimony impurities are vacancy diffusers and phosphorus are interstitial

diffusers [1]. Therefore, self-diffusion in Si has been extensively investigated by experiments [1–5] and simulations [6–33].

In covalently bonded silicon, the neutral vacancy is the most stable charge state near room temperature [3,34]. While the diffusion mechanism and the barrier for silicon have been experimentally determined decades ago [4, 5], we still lack a detailed knowledge of the various activated mechanisms associated with the neutral vacancy. With standard *ab initio* molecular dynamics methods, the number of events generated is always insufficient to give reliable data for migration energies. Accordingly, most of the focus has been on formation energies, distortion and reorientation geometries [31]. Few papers reported a calculated activation energy and a migration path that are in agreement with the experimental data [29, 35]. Until now, the only studies available for exploring the energy landscape were that of Munro and Wales [24] who examined the dynamics of a vacancy, among others, using a tight-binding approach, and the work of Kumeda *et al.* [14] using first-principles calculations. They identified the migration barrier and the underlying mechanism. However, a complete study for the topology of the landscape around a stable minimum is still needed.

The principal goal of this study is to give a complete description of the diffusion mechanism in elementary semiconductors. As such, the SIEST-A-RT method is built up to develop an entire methodology for quantum mechanically precise and reasonably fast energy landscape exploration method. Our simulations are performed on supercells containing 215 atoms. This supercell size is computationally tractable and constitutes a reasonable size to simulate the host crystal. We generate the diffusion paths using the activation-relaxation technique (ART) [36, 37], which can sample efficiently the energy landscape of complex systems. The forces and energies are evaluated using SIESTA [38, 39], a self-consistent density functional method using standard norm-conserving pseudo-potentials and a flexible numerical linear combination of atomic orbital basis set. Combining these two methods allows us to identify very efficiently diffusion paths of various energy height.

This paper is constructed as follows. In the next section, we discuss the SIEST-A-RT method; simulation details are given in section 5.3. We apply the method to the vacancy

in silicon, the corresponding activated dynamics are detailed and discussed in section 5.4. Finally we emphasize the capability of the method in identifying competing events in silicon and its possible extension for future applications.

5.2 Method: SIEST-A-RT

In order to sample the diffusion paths associated with a vacancy in bulk Si, we integrate the activation-relaxation technique (ART nouveau) [37,40] to the *ab-initio* program SIESTA. This approach is similar to the integration of the hybrid eigenvector-following method to a tight-binding and an *ab-initio* code as performed by Munro, Kumeda and Wales [14, 24] as well as the dimer-method [41] which was coupled with VASP [42] by Henkelman and Jönsson [23].

As both ART nouveau and SIESTA are described elsewhere in the literature [39,40], we review these methods only briefly.

5.2.1 ART nouveau

The activation-relaxation technique is a method for sampling efficiently the energy landscape of activated systems. [36, 37] Instead of following in details the thermal fluctuations, it searches directly for transition states; these are characterized as first-order saddle points in the energy landscape. This method was refined recently to ensure a controlled convergence to the saddle point (ART nouveau). [40] An ART event can be decomposed into the following steps. *Activation*: (1) Starting from a local minimum, push the configuration in a direction chosen randomly in the $3N$ -dimensional space; stop when the lowest eigenvalue becomes negative or when the configuration is far enough from the initial configuration. (2) Follow the eigendirection corresponding to this negative eigenvalue while minimizing the energy in the perpendicular directions; stop when the total force (parallel and perpendicular to the eigendirection) falls below a given threshold (0.1 eV/\AA). This is the transition state. *Relaxation*: After pushing over the saddle point, minimize the energy using any standard minimization algorithm.

In order to sample only the activated mechanisms associated with the vacancy, we

restrict the initial random projection for each event to the atoms surrounding the defect. Except for this initial selection, all atoms are allowed to move during the ART event.

ART uses the Lanczòs [43] algorithm to extract the lowest eigenvalue and its corresponding eigenvector; 16 to 20 force evaluations are sufficient, irrespective of the size of the system, to extract a stable eigenvalue value and eigenvector. Recent unpublished tests indicate that the number of force evaluations needed to reach a saddle point is about the same for the dimer method and ART nouveau [41], indicating no significant difference between the various activation-relaxation methods.

5.2.2 Siesta

Forces and energies are evaluated using SIESTA [38, 39], a self-consistent density functional method using standard norm-conserving pseudo-potentials of Troullier-Martins [44] factorized in the Kleiman-Baylander form [45]. Matrix elements are evaluated on a 3D grid in real space. The one-particle problem is solved using linear combination of pseudo-atomic orbitals (PAO) basis set of finite range. The quality of the basis can be improved by *radial* flexibilization, add more than one radial function within the same angular momentum (Multiple- ζ), and *angular* flexibilization or polarization, add shells of different atomic symmetry (different l).

Because ART requires a numerical derivative of the force to compute the curvature of the energy landscape, it is essential to obtain highly converged forces. The various parameters used are discussed below.

5.2.3 Optimization of the parameters

The introduction of a vacancy in the crystal affects considerably the relaxation of the simulation cell. Previous studies of the silicon vacancy using supercells ranging from 32 to 216 atoms with various Brillouin zone sampling techniques [11, 14, 15, 19, 22, 25–27, 33] show a broad spread of formation energies (from 2.6 to 4.6 eV) and structural relaxation with (D_{2d} , C_{2v} , C_{3v} , T_d) symmetries. This scattering is mainly due to various convergence problems like basis set size, Brillouin zone sampling and supercell size. A

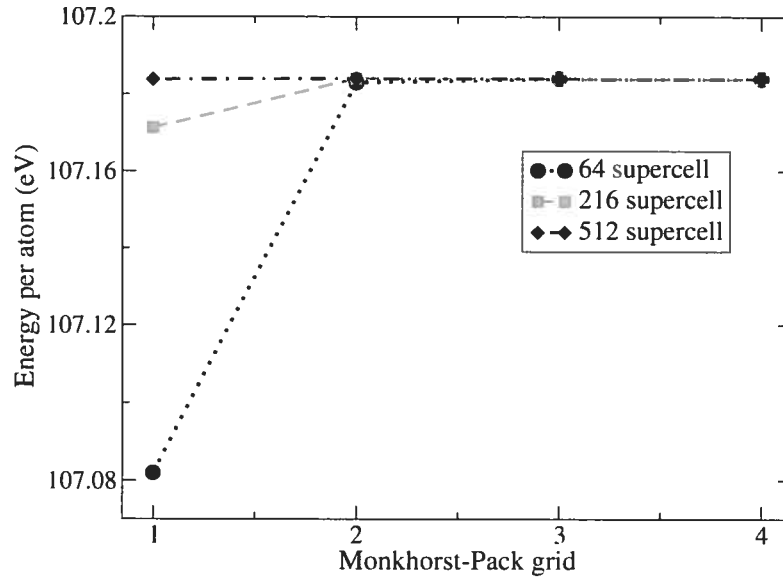


Figure 5.1: Convergence tests for the energy per atom as a function of Brillouin zone sampling for 64, 216 and 512 cubic supercells. The lines are guide to the eye. (see the text).

recent work of Probert and Payne [27] propose a systematic methodology for accurate calculations of defect structure in supercells. We proceed in a similar way in using supercells with up to 512-atoms.

5.2.3.1 Mesh cut-off and the basis set

With the presence of a defect, the single- ζ basis (SZ), i.e., a single orbital per occupied state, is unsatisfactory: both the formation energy and reconstruction surrounding the vacancy are far from experimental value. Moreover, we find that the lack of overlap between the PAO induced discontinuities in the force, causing serious problems to the application of ART. It is possible to address these problems by using the optimized basis generated using the procedure of Anglada *et al.* [46] (SZ-optimized). Although much improved, the discontinuity in the force remains due to the presence of an underlying grid for the integration in real space which causes breaking of the translational symmetry. While the energy is converged with a real-space grid of about 0.03 \AA (a mesh cut-off of 35-40 Ry), we find that we need a grid space of at most 0.02 \AA (mesh cut-off of 50

Ry) to stabilize the force derivative; this spacing is comparable to the typical atomic displacement during the dynamics monitored by ART.

5.2.3.2 Brillouin zone sampling

Because the defect is highly localized even in large supercells, we need to have fully converged forces and total energies for the system. In order to isolate Brillouin zone sampling effects, we study force and energy convergence with respect to the density of k -points. We use a Monkhorst and Pack [47] mesh to sample the Brillouin zone. Uniform meshes of q^3 shifted by (k_0, k_0, k_0) are used, with q going from 1 to 4 and $k_0 = 0$ or 0.5. For $q = 1$ no offsets are used ($k_0 = 0$) to sample the Γ point. Offsets of $k_0 = 0.5$ are used for $q = 2, 3, 4$ for each cubic supercell size.

Figure 5.1 displays the total energy per atom embedded in differently sized supercells with respect to q . The total energy is converged to 1 meV/atom at a mesh of 3^3 for the 64 supercell, this corresponds to a density of 0.031 \AA^{-3} calculated using our LDA lattice parameter 5.39 \AA . As discussed by Probert and Payne [27] this density is sufficient to converge the total energy for larger systems. It is equivalent to a mesh of 2^3 for 216 and 512 cells with densities of 0.031 \AA^{-3} and 0.023 \AA^{-3} respectively. Comparing the forces for unrelaxed vacancy using different meshes and supercells of 63, 215 and 511 atoms, we find that atomic forces converge faster than the total energy as a function of k -point sampling: convergence to 1 meV/ \AA per degree of freedom is reached with a 2^3 mesh for the 63 cell and with the Γ point for larger cells.

5.2.3.3 Supercell calculations

As reported by Puska et al. [25], supercell method has obvious drawback because of the interaction of the defect and its periodic replicas. If the defect-defect distance is not large enough, the electronic structure of an isolated defect is distorted because the deep levels in the band gap form energy bands with a finite dispersion. The size of the supercell restricts also the ionic relaxation. The relaxation pattern is truncated midway between the defect and its nearest image. In order to identify the smallest cell-

size acceptable for the simulation of diffusion mechanisms with negligible size effects, we study the structural relaxation and the formation energies with respect to supercell size by employing supercells as large as 511 ionic sites.

Formation energies are calculated for fully converged supercells with respect to the basis set and Brillouin zone sampling as detailed in sections 5.2.3.1, 5.2.3.2. Results for 63 and 215 atoms supercells are overestimated compared to the 511-cell results by 9 % and 2% respectively as shown in table 5.1 . This error on vacancy formation energy for small supercell was previously reported by Puska *et al.* [11] as due to the interaction of the vacancy with its periodic images and reflects a repulsive component of the defect-defect interaction, this component tends to vanish for the 215 supercell.

Structural relaxation is also of a major importance, it informs us about the correctness of the atomic forces. The 63-atom supercell shows a wrong relaxation pattern. Tetrahedral (T_d) symmetry is a metastable minimum irrespective of the density of k points used with a formation energy of 3.95 eV (see the second line on in table 5.1). Symmetry need to be broken artificially to get a Jahn-Teller distortion leading to a D_{2d} symmetry with lower formation energy (3.59 eV). This is misleading and shows that size effects are important for supercells of 63 ionic sites and smaller. This size is insufficient to represent the system and the diffusion pattern correctly as discussed in more details in section 5.4.4.

For larger system, size effects tend to have minor effects on the relaxation geometry (see table 5.2). Already at 215 cells, D_{2d} symmetry is obtained even for low density of k-points (Γ only). The formation energy using Γ point for the 215 supercell is underestimated by 6% compared to the best converged results we got with the 511 supercell, it tends to converge to the formation energy of larger models relaxed with Γ -point only.

We conclude, in agreement with previous works [11, 25, 27] that a supercell of 215-atoms with Γ -point is the minimal size and sampling we can use in order to ensure that both energy and forces are sufficiently accurate to provide results of a quality at least equal to experimental precision.

Table 5.1: Converged formation energies for (E_f) for different supercells and k-points meshes compared with various *ab-initio* calculations. For each work we specify the DFT functionals used: local density approximation (LDA), screened-exchange LDA (sx-LDA) and conjugate gradient approximation (GGA) with the corresponding plane-wave (PW) energy cutoff when applicable. (TB-EF) denotes tight binding calculations with eigenvalue following method. Experimentally, Watkins *et al.* [4] measured a formation energy of 3.6 ± 0.5 eV and Dannefaer *et al.* [5] 3.6 ± 0.2 eV

Authors	Method	Cell-size	k-points	E_f (eV)
This work	LDA	63- D_{2d}	3^3	3.6
		63- T_d	3^3	3.95
		215	Γ	3.1
			2^3	3.36
		511	2^3	3.29
Lento and Nieminen [33]	LDA (PW 15 Ry)	32	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	
		256	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	3.6
	sx-LDA	256	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	3.515
Probert <i>et al.</i> [27]	GGA (PW 12Ry)	256	2^3	3.17
Puska <i>et al.</i> [11]	LDA (PW 15 Ry)	63	Γ	2.86
			2^3	3.41
		215	Γ	3.27
Mercer <i>et al.</i> [26]	LDA (PW 12-20 Ry)	127	2^3	3.63
Kumeda <i>et al.</i> [14]	LDA (PW12 Ry)	63	Γ	3.32
		215	Γ	3.902
	GGA	215	Γ	3.336
Seong and Lewis [10]	LDA (PW 10 Ry)	63	Γ	3.65
		511	Γ	4.12
Munro and Wales [24]	TB-EF	63	Γ	3.73
		215	Γ	3.902
Antonelli <i>et al.</i> [12]	LDA	63	2^3	3.49

Table 5.2: Relaxation geometries for different supercells and k-points meshes compared with various *ab-initio* calculations. In the D_{2d} symmetry or the pairing mode, the distorted tetrahedron formed by atoms around the vacant site has two short bonds for atoms belonging to the same pair (atom-atom) and four long bonds between atoms from different pairs (pair-pair). The resulting symmetries of the defect are reported as well.

Authors	Cell-size	k-points	Distances (Å)		Symmetry
			pair - pair	atom - atom	
This work	63	3^3	3.34	2.79	D_{2d}
		3^3	3.19	3.19	T_d
	215	Γ	3.34	2.89	D_{2d}
		2^3	3.35	2.9	D_{2d}
	511	2^3	3.35	2.9	D_{2d}
Lento and Nieminen [33]	32	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$		3.084	D_{2d}
	256	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	3.490	2.950	D_{2d}
	256	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$			D_{2d} -out
Probert <i>et al.</i> [27]	256	2^3	3.72	3.1	D_{2d}
Puska <i>et al.</i> [11]	63	Γ			C_{2v}
		2^3	3.42	3.40	D_{2d}
	215	Γ	3.38	2.89	D_{2d}
Mercer <i>et al.</i> [26]	127	2^3	3.50	2.89	D_{2d}
Seong and Lewis [10]	63	Γ	3.53	3.00	D_{2d}
Antonelli <i>et al.</i> [12]	63	2^3	3.53	3.01	D_{2d}
			3.52	3.40	D_{2d}

5.3 Details of the simulation

We perform first-principles electronic structure calculations based on the density functional theory within the local-density approximation (LDA) [48,49] for the exchange-correlation energy. The static zero-temperature calculations for energies and forces are computed from SIESTA. Forces on the atoms are computed using the Hellman-Feynman theorem. The nuclear positions are optimized by using a conjugate gradient algorithm (CG) to minimize the total energy. Using the unit cell of two atoms and increasing the number of k points the equilibrium lattice constant of bulk Si converges to 5.39 Å. The cell volume and shape are kept fixed during the simulation. The crystalline 216 supercell is relaxed until the largest force component is of about 0.002 eV/Å. One conjugate gradient (CG) step is necessary to relax the structure with vanishingly small force components. To create the starting configuration with a defect, one atom is removed from the ideal crystalline cell. The obtained structure contains a mono-vacancy characterized by a vacancy-vacancy distance of 16.17 Å, with tetrahedral symmetry (T_d). The structure is then allowed to relax at constant volume. All the events are generated starting from a fully relaxed 215-atom unit-cell with a single vacancy displaying an expected Jahn-Teller distortion: the four atoms around the vacancy are paired two by two with D_{2d} symmetry, as described in section 5.4.1. More than 120 events were started from this geometry, each launched in a different random direction, providing an extensive sampling of the energy landscape around the vacancy. The results from this search confirm *a posteriori* that this is indeed the minimum-energy configuration for the isolated vacancy. Using the set of parameters detailed in this section, one event with SIEST-A-RT requires about 700 force evaluations and takes about 3 days of CPU times on a single IBM Power 4 processor.

5.4 Results

5.4.1 Neutral vacancy relaxation

The details of the relaxation around the vacancy depend on the local charge [3, 50]. Since there is one electron per dangling bond for the neutral charge state (V^0), it is expected to undergo a Jahn-Teller distortion [3, 50] by forming two pairs out of the four dangling bonds of the vacancy. This is what we see. During the relaxation of the 215-atom cell, the total energy of the system is reduced by about 1.71 eV, due to the distortion of the lattice and the relaxation of the atoms around the vacancy. This relaxation energy is larger than the 0.36 eV as obtained from a study on 63-atom cell by Seong et Lewis [10] and the 0.9 eV calculated on a HF cluster calculation [30] and compared to values from a more recent study [27] 1.186 eV using GGA on a 255 supercell.

The relaxation can be described as follows. The vacancy nearest neighbors (NN) first move away from each other by pair along the $\langle 110 \rangle$ axis [10]. The net local displacement of 0.39 Å per atom result into a tetragonal symmetry configuration (D_{2d}). The distorted tetrahedron formed by the atoms has four long bonds of 3.35 Å and two short bonds of 2.9 Å as shown in Fig. 5.3(A). The formation of bonds between the two pairs weakens the back bonds which are then significantly stretched [12], up to 2.49 Å. The range of perturbation introduced by the defect in the 215 supercell affects up to 5 surrounding shells: the average amplitude of relaxation falls progressively from 0.39 Å for the first shell to 0.074 Å for the fifth shell which is of the order of atomic vibrations at room temperature.

The amplitude and the type of displacement compare very well with the previous results using DFT [11, 14, 15, 19, 22, 25–27]. The LDA results appear to be at odds with recent screened-exchange LDA calculations on partially relaxed neutral Si vacancy, it indicates that the relaxation could be outward [33]. More work remains to establish if LDA really fails here. Our results on the pairing mode are compared with some of these results in table 5.2.

The formation energy E_f is calculated by taking into account the distortion of the

lattice, using the expression:

$$E_f = E_{N-1} - \frac{N-1}{N} E_N \quad (5.1)$$

Where N is the number of atoms and E_N is the relaxed total energy of the ideal structure. E_{N-1} is the total energy of the distorted system containing the defect. Within SIESTA, the formation energy of the mono-vacancy is 3.36 eV, which is in agreement with previous LDA and experimental studies. On the simulation side, Kelly *et al.* [51] calculated 3.5 eV, Blöchl *et al.* [52] 4.1 eV, Mercer *et al.* [26] 3.63 eV and Kumeda *et al.* [14] 3.32 eV. Experimentally, Watkins *et al.* [4] measured 3.6 ± 0.5 eV and Dannefaer *et al.* [5] 3.6 ± 0.2 eV. A complete comparison is reported in table 5.1.

The energy of 511 supercell, E_f^{511} , is found to be 3.29 eV after a full convergence of the energy. Previous works by Wang *et al.* [19] and Seong and Lewis [10] reported $E_f^{511} = 4.12$ eV and 4.10, respectively. A recent study by Probert *et al.* give GGA results of a fully converged calculation for a vacancy in BCC silicon lattice with 256 atoms, they obtain a formation energy of 3.17 eV.

Even a defect as simple as a neutral vacancy can be associated with many activated mechanisms, some of which leading to migration: among the generated events, 60 % are simple diffusion mechanisms, while 15 % are reorientation and the remaining are more complex events involving a pair or more of active atoms.

5.4.1.1 Nearest neighbor diffusion

The simplest migration process involves hopping of the vacancy to a nearest-neighbor site as shown in Fig. 5.2(a): a nearest-neighbor of the vacancy moves along the $\langle 111 \rangle$ direction toward the vacant site (shown in Fig. 5.3A) and passes the metastable split-vacancy site before falling in the tetrahedral site previously formed by the vacancy. The energy barrier for this mechanism is 0.40 ± 0.02 eV. The error estimate comes from the convergence criterion to the saddle point as explained in section 5.2.1. This result for the migration energy is in good agreement with the experimental findings of Watkins *et al.* [3] who measured 0.45 ± 0.04 eV. In a work similar to this one, but using the CPMD

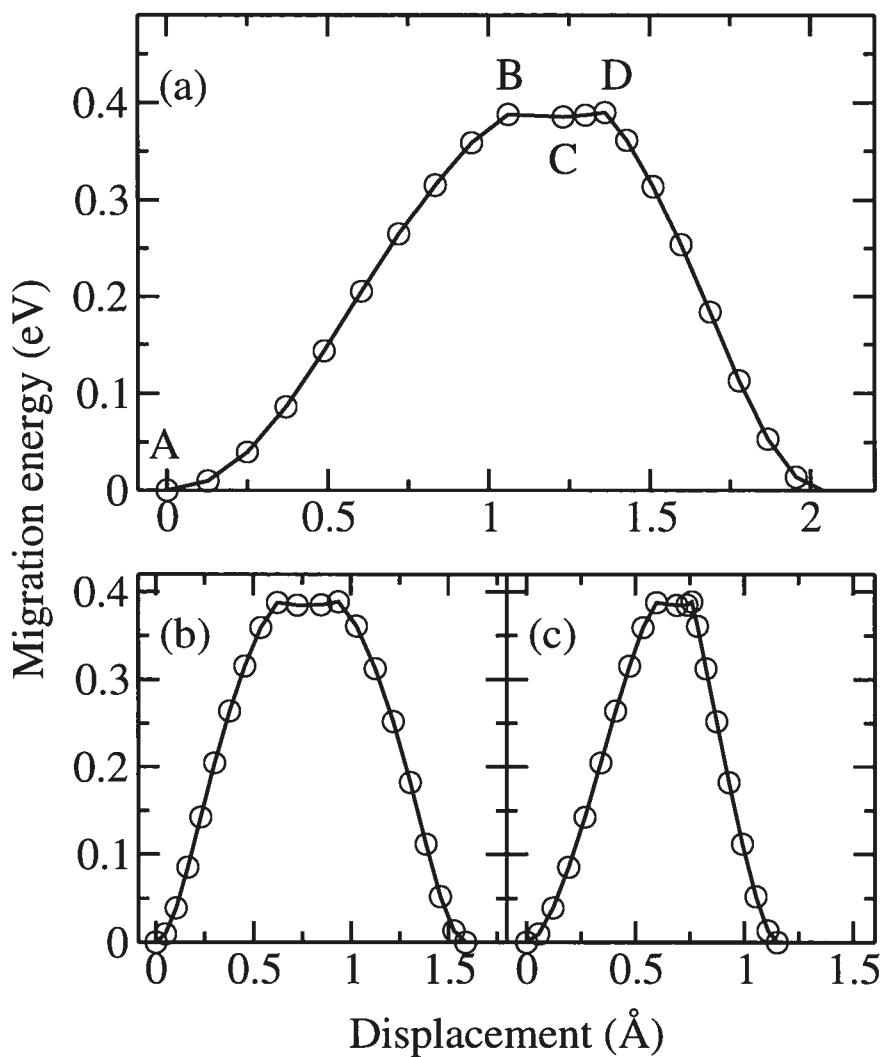


Figure 5.2: Top: (a) The total minimum-energy path for the simple diffusion. The saddle points identified in various ART events are indicated by the same labels (B, C and D) as in Table II. The path is generated by previous knowledge of the initial, the saddle point and final states, then the overall configuration is relaxed using the nudged-elastic-band method [28]. The contributions from different moving atoms are decomposed in two: (b) The path followed by the diffusing atom; (c) Diffusion of the other atoms around the defect. Here The assymetry is due to reconstruction.

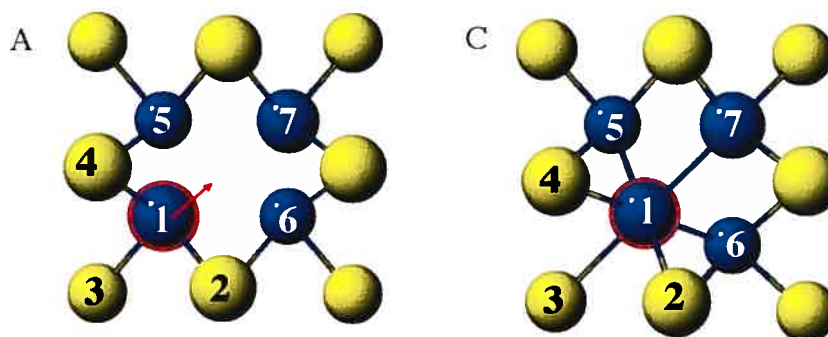


Figure 5.3: The vacant site is visible in the snapshots A and C. A is the initial minimum and C is the ideal split vacancy site, the arrow shows the direction of the diffusion.

code with GGA (BLYP) functional, Kumeda *et al.* [14] found a barrier of 0.58 eV.

The diffusion path is reconstructed in detail using the nudged-elastic band method [28] and relaxed until the total force is of $0.2 \text{ eV}/\text{\AA}$. This path shows an unusually long transition plateau —ranging from $d_r = 0.7$ to 1.2 \AA — with a slight energy minimum at the split-vacancy site. At this point, the 3 back bonds are stretched and the diffusing atom forms weak bonds (length 2.685 \AA) with the surrounding atoms of the six membered ring. From this site, a small displacement of 0.11 \AA is sufficient to overcome the barrier of 0.04 eV and complete the jump to the new minimum (see table 5.3). This value is of the order of the precision of the method; in addition, at room temperature this minimum would be smeared out by thermal vibrations. Because of a lack of a better approximation, we use Vineyard’s harmonic approximations [53] to get a rough estimate of the attempt frequency (see section 5.4.2.2), approximating the transition plateau by a single barrier.

5.4.2 Simple diffusion paths

Although the total diffusion path in Fig. 5.2(a) is asymmetric, after separating the contribution coming from different moving atoms, we find that diffusion path is symmetric for the diffusing atom as shown in figure (Fig. 5.2(b)). The asymmetry in the path shown in Fig. 5.2(c) is due to the reconstruction: with four atoms reconstructing two-by-two, there are three possible orientations. Here, the orientation is different for the initial and final state. This should average out as the vacancy moves around.

Table 5.3: Interatomic distances along the minimum energy path for a simple diffusion. The positions correspond to the labels in Fig. 5.2(a). A corresponds to the initial minimum with a relaxed vacancy: the active atom (1) is bonded to four neighbors, B, C and D are transition states. During the migration bonds are stretched progressively until all the distances becomes equal at the configuration C.

Distance(Å)	d_{1-2}	d_{1-3}	d_{1-4}	d_{1-5}	d_{1-6}	d_{1-7}
A	2.49	2.39	2.39	2.89	3.34	3.34
B	2.67	2.67	2.65	2.65	2.76	2.73
C	2.68	2.68	2.68	2.68	2.68	2.68
D	2.68	2.87	2.61	2.69	2.61	2.69

A comparison of the formation energies using the Stillinger-Weber [35] and the MEAM [54] potentials for the vacancy and the split vacancy gives a migration energy of 0.43 and 0.37 eV, respectively. Although the formation energies for the vacancy are not well represented by empirical potentials the energy difference is in agreement with experiment and with our results. The minimum-energy path, however, is not well described by these potential and the local minima identified by Maroudas *et al.* [35] during the migration are absent in our case.

5.4.2.1 Reorientation process

In addition to finding diffusion mechanisms associated with the vacancy, we also identify the trajectories responsible for atomic rearrangement in the reconstructed configuration. As there are three pairing possibilities for the D_{2d} symmetry, it is possible for the configuration to hop from one of these states to either of the two others.

SIEST-A-RT generates easily this reorientation of the paired atoms. As illustrated in Fig. 5.4, starting with the reconstructed pairs formed initially — 1-4 and 2-3—, atom “3” moves away from its neighbor along the $\langle 111 \rangle$ direction, breaks the weak bond with “2” and reaches the transition state, which is in the trigonal symmetry (C_{3v}). The

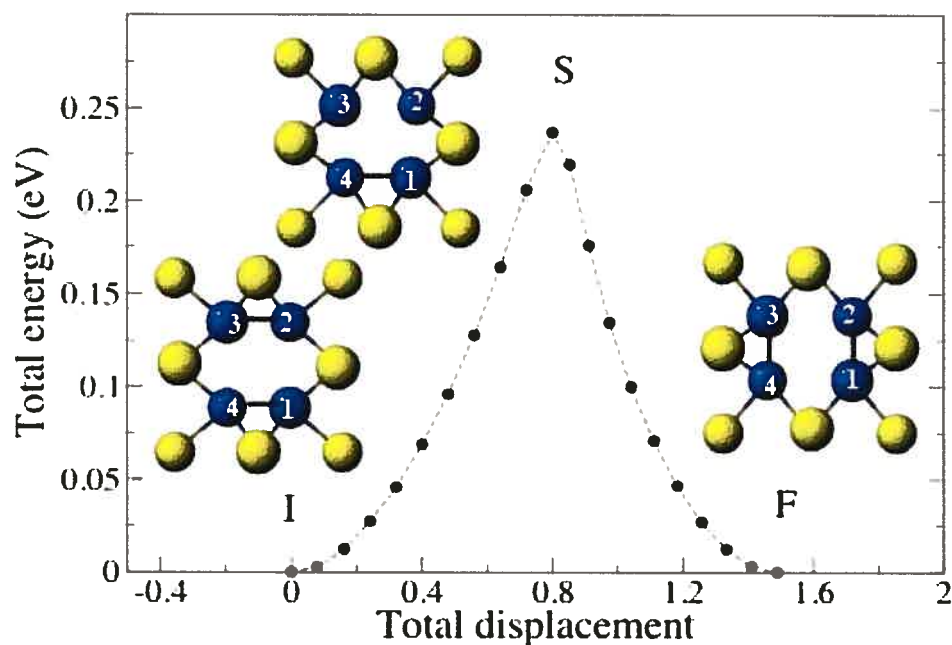


Figure 5.4: The reorientation path passing over the saddle point configuration. (I→S→F), the tetragonal axis is fixed throughout this process. At the saddle point the local atomic coordinates of the distorted lattice shows a trigonal symmetry. The pairing mode changes to one of the equivalent modes.

relaxation continues until new bonds form between the pairs 1-2 and 3-4. The activation energy for the reorientation of the neutral vacancy between two tetragonal orientations is 0.2 eV. Roberson *et al.* [29] in a HF calculation scheme performed calculation on small clusters (44 silicon atoms) and found 0.37 eV; another cluster calculation, by Ögüt *et al.* [30] found 0.32 eV.

The deformation to the C_{3v} can be observed experimentally. The experiment consists of applying a stress at high temperature in the dark, then cooling down to about 20K, relieving the stress, illuminating the sample and monitoring the resulting vacancy species. The barrier measured from this alignment experiment is at 0.23 eV, in excellent agreement with our calculations [3].

Our simulations also show that the pairing mode can change *during* the migration

of the vacancy. For the simple diffusion plus reorientation, we find an overall energy barrier of 0.47 eV and a total displacement of 0.68 Å at the saddle point.

5.4.2.2 Diffusion constants and diffusion rate

The diffusion is a two step process, consisting of the creation and the subsequent migration of a vacancy. Assuming that both processes are thermally activated, the activation energy E_a for diffusion is the sum of the formation energy E_f and the migration energy E_m . By collecting our data from the previous sections we get an activation energy of $E_a = 3.5$ eV, which is in good agreement with $E_a = 4.14$ eV recently measured by Bracht *et al.* [55], and with the activation enthalpy for self-diffusion $H_V = 4.07 \pm 0.2$ eV obtained by Tang *et al.* [20] using tight-binding MD.

It is well known that there is a strong competition between the interstitial and the vacancy-mediated mechanisms of self-diffusion. These two contributions were separated in a metal-diffusion experiment [55] and the vacancy contribution to the self-diffusion was measured to be

$$C_V^{eq}(T)D_v(T) = 0.92 \exp \frac{(-4.14\text{eV})}{k_B T} \text{cm}^2 \text{s}^{-1} \quad (5.2)$$

Other experimental data [56] shows that

$$C_V^{eq}(T)D_v(T) = 0.6 \exp \frac{(-4.03\text{eV})}{k_B T} \text{cm}^2 \text{s}^{-1} \quad (5.3)$$

Both data sets predict a self-diffusion dominated by vacancies at low temperature, where $C_V^{eq}(T)$ is the concentration of vacancies at a certain temperature.

Experiments under thermal equilibrium concentrations of native point defects have shown that the Si self-diffusion coefficient D_{Si}^{eq} exhibits Arrhenius behavior,

$$D_{Si}^{eq}(T) = D_0 e^{-H_a/k_B T}, \quad (5.4)$$

where H_a is a single activation enthalpy, D_0 is the prefactor in a certain temperature range, T denotes the absolute temperature, and k_B is the Boltzman's constant. Un-

der the assumption that the diffusion proceeds through discrete jumps of equal length r —the distance between nearest neighbor equilibrium sites, in diamond cubic lattices $r = a\sqrt{3}/4$, where a is the lattice parameter — the prefactor is written as [34]

$$D_0 = \frac{z}{6}\Gamma_0 r^2 \quad (5.5)$$

with Γ_0 the jump, or attempt, frequency of the point defect from one equilibrium site to another, z denotes the number of neighbor sites, in the diamond lattice a vacancy can jump to one of the four neighbors. In the harmonic approximation, this attempt frequency is defined as [53, 57]

$$\Gamma_0 = \frac{\prod_{i=1}^{3N} \nu_i^{(i)}}{\prod_{i=1}^{3N-1} \nu_i^{(s)}} \quad (5.6)$$

where $\nu_i^{(i)}$ and $\nu_i^{(s)}$ are the normal mode frequencies at the minimum and the saddle point respectively and the product does not include the imaginary frequency at the saddle point. The eigenvalues are computed by diagonalizing the Hessian, obtained by numerical derivation with a step of 0.01 Å.

Accordingly, the migration entropy defined as the entropy difference between the saddle point and the equilibrium point configuration at low temperature can be written as

$$\Delta S = S_V^m = k_B \ln \left[\frac{\prod_{i=1}^{3N-1} \nu_i^{(i)}}{\prod_{i=1}^{3N-1} \nu_i^{(s)}} \right] \quad (5.7)$$

The phonon frequency corresponding the direction of the jump is removed from the numerator. In addition the self diffusion entropy evaluated experimentally is the sum of the diffusion entropy and the formation entropy $S_V^{SD} = S_V^f + S_V^m$. Using the experimental $S_V^{SD} = 5.5k_B$ by Bracht *et al.* [55] and a $S_V^f = (5\pm 2)k_B$ as calculated by Blöchl *et al.* [52] and Dobson *et al.* [9], migration entropy is expected to lie within $1k_B$ and $2.5k_B$,

our calculated data $\Delta S = 3.036k_B$.

The corresponding attempt frequency $\Gamma_0 = 8.65 \times 10^{12} s^{-1}$ is similar to the result of Maroudas et al. [35] ($1.0539 \times 10^{13} s^{-1}$), obtained with the Stillinger-Weber potential. The corresponding diffusion prefactor is $D_0 = 3.141 \times 10^{-3} cm^2 s^{-1}$. Our diffusion constant can be compared with results from Maroudas *et al.* [35] by taking into account the z factor that had been omitted, the resulting prefactor $D_0 = 3.88 \times 10^{-3} cm^2 s^{-1}$ is in agreement with our results. Tang *et al.* [8,20] report $D_0 = 1.18 \times 10^{-4} cm^2 s^{-1}$ using TB-MD and thermodynamical integration, their migration energy (E_m) is of 0.1 eV.

Our numerical estimate of the diffusion rate $\Gamma = \Gamma_0 e^{E_m/k_B T}$ at room temperature gives a characteristic time for diffusion ($\tau = 1/\Gamma$) of the order of microseconds. This timescale is not directly accessible to *ab-initio* molecular dynamics, so simulations must generally be run at temperatures close to the melting point. With SIEST-A-RT, we can study the activated mechanisms from the local energy minimum, allowing us to identify accurately subtle mechanisms such as that associated with the split-vacancy site.

5.4.3 Complex diffusion paths

The SIEST-A-RT sampling of the energy landscape around the vacancy also identifies a number of high-energy barrier events which involve a complex atomistic rearrangements. Two types of such events occur. The first one can be described as a diffusion event mediated by bond exchange. In this case, the overall result is a simple vacancy hop but with a more complex transition state and a barrier of 2.4 eV.

The second type of events involves what has been dubbed *spectator* events [14] (see Fig. 5.5) where the diamond network surrounding the vacancy is altered. These spectator events all correspond basically to the bond-switching mechanisms proposed many years ago by Wooten, Winer and Weaire as mechanism for amorphization [58]. In recent years, the WWW mechanism has also been identified as the bond defect [59] and the four fold coordinated defect (FFCD) [16].

Two types of *spectator* events are identified. In one of these events, two of the direct neighbors of the vacancy are pushed toward each other, leading to the saturation of the dangling bonds; as a consequence one of the long covalent bonds between the

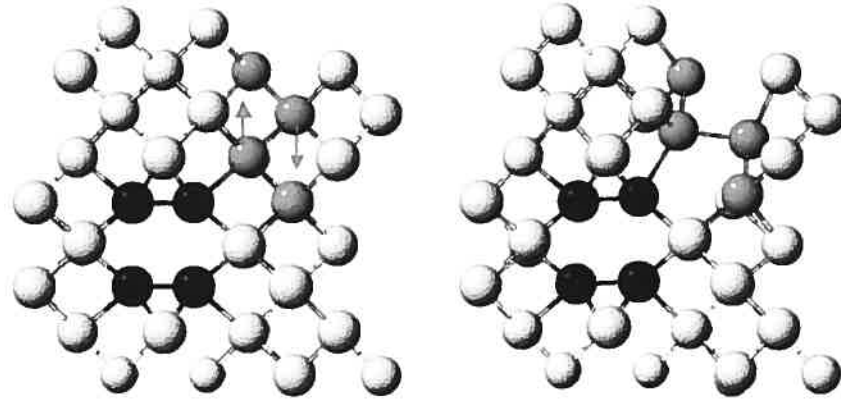


Figure 5.5: A *spectator* WWW process involving atoms far from the vacant site. Left panel: atoms in grey are the atoms involved in the exchange, the direction of the exchange is illustrated by the arrows. Right panel: the final configuration is a metastable minimum by which a disorder in the crystalline lattice is introduced.

NN of the vacancy transform to a real bond in a form very similar to the bond defect complex identified by Marqués *et al* [17]. The lowest barrier we found is 2.63 eV for a process involving a direct neighbor of the vacancy (one of the bonds is already weak), the final configuration is 2.16 eV higher than the initial minimum. Kumeda *et al.* found the single *spectator* WWW barrier of 1.719 eV using LDA and 1.867 eV using BLYP functional [14].

The second type, atoms far from the vacant site are involved. The effect of the vacancy on the WWW event barrier is less strong for complex events located far from the vacancy without involving a direct neighbor. As the lattice distortion increases, so does the energy barrier. For this mechanism, for example, the barrier ranges from 3.31-3.8 eV, a value which can be compared to the I-V recombination process identified to be responsible for the amorphization of crystalline silicon of Marqués *et al* [17] the barrier is estimated to be between 4.58 and 4.75 eV using *ab initio* calculation [59]. One can estimate the effect of the vacancy to lower the barrier by about 1.27 eV. Since the barrier

is very high, this kind of events is unlikely to play any significant role in crystals at room temperature.

5.4.4 Size effects

As discussed, previously, it is necessary to use a large simulation cell in order to minimize the interaction between the images of the vacancy. This is particularly noticeable in a 63-atom supercell, even though the defect-defect separation is larger than 4 inter-atomic distances. It is this interaction that cause the unrealistic relaxation pattern identified in previous studies [15, 26], allowing the vacancy to relax both with a T_d and a D_{2d} symmetries.

Relaxing a 63-atom supercell with SIESTA and the parameters described above, we find that two stable minima: the D_{2d} (tetragonal) relaxation described previously as well as a T_d (tetrahedral) configuration, a symmetry which is *unstable* in cells of 215 and 511 atoms. Within the later symmetry, all atoms relax inward, conserving the tetrahedral symmetry of the system at equal distance of 3.4 Å. The T_d structure can be reached with a single jump from the D_{2d} site during a reorientation mechanism similar to that described in section 5.4.2.1, crossing a small barrier of of 94 meV with an overall displacement of 0.64 Å. The energy difference between the two symmetries is estimated to be 83 meV compared to 20 meV as obtained by Mercer *et al.* in a simulation also on a 63-atom cell [26]. We obtain qualitatively similar results with 2^3 k-points mesh.

The interaction between images also affects the barrier for the simple diffusion and the reorientation, which is underestimated by 50 % compared to the 215 cell results. Interestingly, this underestimation is compensated by a higher formation energy, leading to very similar activation energies for the two size systems.

The underestimation of the diffusion barrier in the 63-atom cell makes it easier for the vacancy to jump to the second-neighbor position. Although we saw many such events in the small cell, this mechanism was not generated in the 215-atom box; there does not appear to be any direct path for second-neighbor diffusion in silicon.

5.5 Discussion and Conclusions

In this paper, we report on the study of the activated dynamics associated with a neutral vacancy in bulk crystalline silicon. Following the work of Munro, Kumeda and Wales [14,24], we couple an *ab initio* code, SIESTA, with the activation-relaxation technique, ART nouveau, in order to sample the various activated mechanisms taking place around the vacancy accurately and efficiently. Defining moves directly in the energy landscape, it is as easy to generate the diffusion trajectory responsible for the high-energy bond-switching mechanisms than it is to find the path associated with the jump reorientation of the reconstructed state ($D_{2d} \rightarrow D_{2d}$).

Simulating more than 120 activated trajectories, we find a number of different diffusion mechanisms associated with the vacancy. In particular, we recover the basic vacancy hopping diffusion mechanism and show that the path associated with it includes a very metastable state at the split-vacancy site. Moreover, we can reproduce without difficulty the jump between the symmetrically equivalent reconstructed states. We also generate a number of higher-energy diffusion mechanisms which reminiscent of those found in amorphous silicon as well as at the crystalline-amorphous interface.

Because we obtain, with good accuracy, the transition state associated with each of these mechanisms, we also compute the transition rate in the harmonic approximation. The result obtained is in good agreement with experiment as well as previous high-temperature TB-MD simulation [20].

Overall, various barrier and relaxation energies computed here are in agreement with results obtained previously both experimentally and numerically. Using a large simulation cell and an extensive sampling of the energy landscape around the vacancy, we could also show that some of the previously obtained mechanisms were artifact due to the small size of the cell used. These results establish the validity of the SIESTA-ART approach as a unique tool for the study of the diffusion and relaxation mechanisms of defects in crystalline materials. This method can be easily extended to study semiconductor compounds, adatom diffusion, interlayer diffusion, all systems with a landscape too complicated for high-temperature molecular dynamics. Even in system as simple as

a vacancy in *c*-Si, we found a number of activated mechanism associated with non-trivial transition paths. The situation promises to be more complex in binary semiconductors, such as GaAs and GaN, where the bonding character and chemical composition are much more complex than for Si.

5.6 Acknowledgments

This work is funded in part by NSERC and NATEQ. NM is a Cottrell Scholar of the Research Corporation. Most of the simulations were run on the computers of the Réseau québécois de calcul de haute performance (RQCHP) whose support is gratefully acknowledged.

BIBLIOGRAPHY

- [1] A. Ural, P. B. Griffin, and J. D. Plummer, *Phys. Rev. Lett.* **83**, 3454 (1999).
- [2] A. L. Lagna and S. Coffa, *Phys. Rev. Lett.* **82**, 1720 (1999).
- [3] G. D. Watkins, *Deep Centers in semiconductors*, edited by S. T. Pantelides (Gordon and Breach, New York, 1992), chap. 3.
- [4] G. Watkins and J. Corbett, *Phys. Rev.* **134**, A1389 (1964).
- [5] S. Dannefaer, P. Mascher, and D. Kerr, *Phys. Rev. Lett.* **56**, 2195 (1986).
- [6] S. V. Ghaisas, *Phys. Rev. B* **43**, 1808 (1991).
- [7] S. Ögüt and J. R. Chelikowsky, *Phys. Rev. Lett.* **83**, 3852 (1999).
- [8] A. Jääskeläinen, L. Colombo, and R. Nieminen, *Phys. Rev. B* **64**, 233203 (2001).
- [9] T. W. Dobson, J. F. Wager, and J. A. Van Vechten, *Phys. Rev. B* **40**, 2962 (1989).
- [10] H. Seong and L. J. Lewis, *Phys. Rev. B* **53**, 9791 (1996).
- [11] M. J. Puska, S. Pöykkö, M. Pesola, and R. M. Nieminen, *Phys. Rev. B* **58**, 1318 (1998).
- [12] A. Antonelli, E. Kaxiras, and D. J. Chadi, *Phys. Rev. Lett.* **81**, 2088 (1998).
- [13] Y. Bar-Yam and J. D. Joannopoulos, *Phys. Rev. B* **30**, 2216 (1984).
- [14] Y. Kumeda, D. J. Wales, and L. J. Munro, *Chem. Phys. Lett.* **341**, 185 (2001).
- [15] R. Virkkunen, M. Alatalo, M. Puska, and R. M. Nieminen, *Comp. Mat. Sc.* **1**, 151 (1993).
- [16] S. Goedecker, L. Billard, and T. Deutch, *Phys. Rev. Lett.* **88**, 235501 (2002).

- [17] L. Marques, L. Pelaz, J. Hernandez, J. Barbolla, and G. H. Gilmer, *Phys. Rev. B* **64**, 045214 (2001).
- [18] J. Kim, F. Kirchhoff, J. Watkins, and F. S. Khan, *Phys. Rev. Lett.* **84**, 503 (2000).
- [19] C. Z. Wang, C. T. Chan, and K. M. Ho, *Phys. Rev. Lett.* **66**, 189 (1991).
- [20] M. Tang, L. Colombo, J. Zhu, and T. D. de la Rubia, *Phys. Rev. Lett.* **55**, 14279 (1997).
- [21] W. K. Leung, R. J. Needs, and G. Rajogopal, *Phys. Rev. Lett.* **83**, 2351 (1999).
- [22] S. K. Estreicher, *Phys. Stat. sol. (b)* **217**, 513 (2000).
- [23] G. Henkelman, B. Uberuaga, S. Dunham, and H. Jónsson, *Phys. Stat. sol. (b)* **233**, 24 (2002).
- [24] L. J. Munro and D. Wales, *Phys. Rev. B* **59**, 3969 (1999).
- [25] M. J. Puska, *Comp. Mat. Sc.* **17**, 365 (2000).
- [26] J. L. Mercer, J. S. Nelson, A. F. Wright, and E. B. Stechel, *Modelling. Simul. Mater. Sci. Eng.* **6**, 1 (1998).
- [27] M. I. J. Probert and M. C. Payne, *Phys. Rev. B* **67**, 075204 (2003).
- [28] G. Henkelman, B. P. Uberuaga, and H. Jónsson, *J. Chem. Phys.* **113**, 9901 (2000).
- [29] M. A. Roberson and S. K. Estreicher, *Phys. Rev. B* **49**, 17040 (1994).
- [30] S. Ögüt, H. Kim, and J. Chelikowsky, *Phys. Rev. B* **56**, R11 353 (1997).
- [31] F. Anderson, F. S. Ham, and G. Grossmann, *Phys. Rev. B* **53**, 7205 (1996).
- [32] S. J. Clarck and G. J. Ackland, *Phys. Rev. B* **56**, 47 (1997).
- [33] J. Lento and R. M. Nieminen, *J. Phys.: Condens. Matter* **15**, 4387 (2003).

- [34] F. Agullo-Lopez, C. Catlow, and P. Townsend, *Point Defects in Materials* (Academic Press, 1988).
- [35] D. Maroudas and R. A. Brown, *Phys. Rev. B* **47**, 15562 (1993).
- [36] N. Mousseau and G. T. Barkema, *Phys. Rev. E* **57**, 2419 (1998).
- [37] G. T. Barkema and N. Mousseau, *Phys. Rev. Lett.* **77**, 4358 (1996).
- [38] D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, *Int. J. Quant. Chem.* **65**, 453 (1997).
- [39] J. M. Soler, E. Artacho, J. Gale, A. García, J. Junquera, and P. Ordejón, *J. Phys. Condens. Matter* **14**, 2745 (2002).
- [40] R. Malek and N. Mousseau, *Phys. Rev. E* **62**, 7723 (2000).
- [41] G. Henkelman and J. Jónsson, *J. Chem. Phys.* **111**, 7010 (1999).
- [42] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- [43] C. Lanczós, *Applied Analysis* (Dover, New York, 1988).
- [44] N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- [45] L. Kleiman and D. M. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).
- [46] E. Anglada, J. M. Soler, J. Junquera, and E. Artacho, *Phys. Rev. B* **66**, 205101 (2002).
- [47] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [48] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- [49] J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [50] J. Bourgoin and M. Lanoo, *Point defects in Semiconductors I and II* (Springer, Berlin, 1981).

- [51] P. Kelly and R. Car, Phys. Rev. B **45**, 6543 (1992).
- [52] P. Blöchl, E. Smargiassi, R. Car, D. B. Laks, W. Andreoni, and S. Pantelides, Phys. Rev. Lett. **70**, 2435 (1993).
- [53] H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).
- [54] M. I. Baskes, J. S. Nelson, and A. F. Wright, Phys. Rev. B **40**, 6085 (1989).
- [55] H. Bracht, E. E. Haller, and R. Clark-Phelps, Phys. Rev. Lett. **81**, 393 (1998).
- [56] T. Y. Tan and U. Gösele, Appl. Phys. A: Solids Surf. **37**, 1 (1985).
- [57] R. Phillips, *crystals, Defects and Microstructures, Modelling a cross Scales* (Cambridge university press, 2001).
- [58] F. Wooten, K. Winer, and D. Weaire, Phys. Rev. Lett. **54**, 1392 (1985).
- [59] F. Cargnoni, C. Gatti, and L. Colombo, Phys. Rev. B **57**, 170 (1998).

CHAPITRE 6

STABILITÉ DES LACUNES DANS LE GAAS EN FONCTION DE LEUR ÉTAT DE CHARGE

Ce chapitre présente les propriétés statiques des lacunes dans le composé semi-conducteur binaire GaAs. Les propriétés de ces lacunes sont traitées par l'article intitulé *Self-vacancies in Gallium Arsenide : an ab initio calculation* et publié dans le journal *Physical Review B*, volume 71, pages 125207 :1-12 en date du 18 Mars 2005.

Cet article réexamine les propriétés statiques des lacunes dans le GaAs en utilisant des calculs de principes premiers basés sur la théorie de la fonctionnelle de la densité. Les énergies de formation calculées sont en bon accord avec les expériences et les calculs *ab-initio* récents. Cet article comporte une description complète de la géométrie de relaxation et des énergies d'ionisation associées aux différents états de charge des lacunes d'As et de Ga. Les lacunes de Ga sont stables dans les états de charge 0, -1, -2 et -3 mais la lacune V_{Ga}^{-3} demeure le défaut qui domine l'autodiffusion dans le GaAs intrinsèque ou de type N.

Les lacunes d'As présentent des aspects très intéressants. En fait, la transition d'un état de charge vers un autre a lieu en transférant deux électrons à la fois. Il s'agit de l'effet U-négatif. Par conséquent, les états de charge +1, -1 et -3 sont stables alors que les états intermédiaires sont métastables. Chacune de ces transitions induit une déformation structurelle autour de la lacune d'As. La seconde transition (-/-3) donne lieu à une relaxation peu observée jusqu'à date. Il s'agit de la relaxation en mode de lien résonant pour V_{As}^{-3} qui a déjà été observée pour les bilacunes dans le silicium et le GaAs.

Self-vacancies in Gallium Arsenide: an *ab-initio* calculation

Fedwa El-Mellouhi and Normand Mousseau

Département de physique and Regroupement québécois sur les matériaux de pointe, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal (Québec) H3C 3J7, Canada

PACS numbers: 61.72.Ji, 71.15.Mb, 71.15.Pd

Abstract

We report here a reexamination of the static properties of vacancies in GaAs by means of first-principles density-functional calculations using localized basis sets. Our calculated formation energies yields results that are in good agreement with recent experimental and *ab-initio* calculation and provide a complete description of the relaxation geometry and energetic for various charge state of vacancies from both sublattices. Gallium vacancies are stable in the 0, -1 , -2 , -3 charge state, but V_{Ga}^{-3} remains the dominant charge state for intrinsic and n-type GaAs, confirming results from positron annihilation. Interestingly, Arsenic vacancies show two successive negative-U transitions making only $+1$, -1 and -3 charge states stable, while the intermediate defects are metastable. The second transition (-1 / -3) brings a resonant bond relaxation for V_{As}^{-3} similar to the one identified for silicon and GaAs divacancies.

6.1 Introduction

Native point defects are involved in virtually every process during which an atom incorporated in the lattice of a semiconductor migrates toward another lattice site. This

diffusion mediated by point defects is responsible for a number of important effects, for instance, those encountered during fabrication of microelectronic devices. It is not surprising therefore that point defects in semiconductors have been extensively studied using *ab initio* techniques. Some work still remains to be done, however, especially in the case of alloyed semiconductors.

If Ga vacancies in GaAs are relatively well understood — because most dopants used in technology (Si donor, Zn, Be, and Mg acceptors) occupy the Ga sites— much less is known about the As vacancy. The recent introduction of carbon as a prospective As-site acceptor has raised significantly the interest for this defect [1], however. There is therefore an abundant literature on calculations [1–10] of formation energies of point defects in GaAs which is nicely reviewed and summarized in a paper of Deepak *et al.* [11]. Since formation energies are difficult to measure, calculations are the primary method for obtaining these values. However because of the assumptions and approximations taken into account, the reported values in the literature during the last two decades differ greatly from paper to paper.

Most calculations were done for small supercells. These calculations take advantage of error cancellations to obtain energy differences that are more precise than the total energies themselves. Here we repeat these calculations for all possible charge states for both Ga and As vacancies using the strict convergence criteria and a large simulation cell.

The outline of the paper is as follows. We first describe the method for defect calculation in section 6.2, then we turn to the convergence tests made to set up the methodology. We focus on the effect of the k -point sampling, system-size and completeness of the basis on the formation energy and structural properties of the defects. Section 6.3 deals with these effects on the most important charge states like V_{As}^{-1} and V_{Ga}^{-3} . In section 6.4 we present and discuss the most converged results using localized basis sets compared to previous results obtained from theory and experiment.

6.2 Simulation details and methodology

6.2.1 Total energy calculations

The total energies for this work are evaluated using SIESTA [12, 13], a self-consistent density functional method (DFT) within local-density approximation (LDA). Core electrons are represented by the standard norm-conserving Troullier-Martins pseudo-potentials [14] factorized in the Kleiman-Bylander form [15] and the one-particle problem is solved using linear combination of pseudo-atomic orbitals (PAO) basis set of finite range. These orbitals are strictly localized and represent well the local electronic densities; few of them are therefore needed, decreasing considerably the computational costs by comparison with standard plane-waves calculations. The main drawback of this approach, however, is the lack of a systematic procedure to ensure a rapid variational convergence with respect to the number of basis orbitals and to the range and shape of each orbital. Consequently, while extending plane-wave basis sets is trivial, some efforts are needed to prepare unbiased pseudo-atomic basis sets (see, for example, Ref [16, 17]).

In this work, we use the following sequence to test the convergence of the basis set. Starting with the simplest scheme, a single ζ (SZ) basis, a second group of valence orbitals is added for flexibility, forming the double- ζ (DZ) basis. For completeness, we also add polarization orbitals to both valence sets, generating SZP and DZP bases.

Finally, it is possible to optimize the localization radius in order to increase the accuracy for a given basis set. While the computational efficiency is slightly reduced, as the optimized orbitals have generally a longer tail, it is often a good alternative to increase the size of the basis set (for details see Ref. [12, 13, 16]). We also test the accuracy of these basis sets optimized with respect to the amount of overlap between atomic orbitals around the defect using the optimizing procedure of Anglada et al. [17] at 0.0 GPa for the SZ basis set. We find that the efficiency of these orbitals with (SZP-O) and with (SZ-O) is comparable to those of DZP and DZ, respectively.

Table 6.1 reports the values of a number of structural and thermodynamical quantities for bulk GaAs as computed using these various bases with a k -point sampling density of 0.03 \AA^{-1} , corresponding to that for a 216-atom unit cell with a $2 \times 2 \times 2$ k -point sampling.

Table 6.1: Comparison between converged basic parameters for bulk GaAs. a , B and E_{total} represent the lattice constant, the bulk modulus, and the total energy per GaAs pair, respectively. E_g denotes the energy gap and ΔH is the heat of formation of GaAs calculated using Eq. 6.4. See the text for description of the basis set label. Results are compared to a recent plane-wave calculation (PW) [18] and to experimental values from Ref. [19] at 0K, unless other references are cited.

	SZ	SZ-O	SZP-O	DZ	DZP	PW	Expt.
$a(\text{\AA})$	5.68	5.66	5.60	5.64	5.60	5.55	5.65
$B(\text{GPa})$	59.85	68.9	78.8	67.7	70.4		75.3
$E_{total}(\text{eV}/\text{pair})$	235.2	235.5	235.9	235.7	236.0		
$E_g(\text{eV})$	0.61	0.78	0.98	0.66	0.82	0.7 ^b	1.52
$\Delta H(\text{eV})$	0.66	0.99	0.78	0.81	0.72	0.67 ^c 0.83 ^d	0.73 ^e

^a Reference [18] ^b Reference [3] ^c Reference [8]

^d Reference [1] ^e Reference [20]

For SZ, the lattice constant at zero pressure is found to be 5.68 \AA , overestimating the experimental value by only 0.03 \AA . The density of the GaAs crystal increases with the size of the basis set and the lattice constant for DZP is found to be too small by 0.05 \AA with respect to experiment. The relatively contracted structure obtained with DZP is characteristic of LDA; plane-wave calculations also using LDA give 5.55 \AA . [18] Increasing the basis set leads to a significant improvement on the calculated value of the bulk modulus as it goes from 59.85 GPa for SZ to 70.4 GPa for DZP, close to the experimental value of 75.3 GPa.

The LDA band gap is found to be 0.61 eV, 0.66 eV and 0.82 eV for SZ, DZ and DZP basis sets respectively, underestimating, as usual with this approximation, the experimental gap of 1.52 eV. The DZP band gap lies well within the energy gap obtained from PW calculation (around 0.7 eV [3]) and can be considered converged.

No such systematic problem is found for the total energy and the heat of formation. In particular, the heat of formation obtained with DZ and DZP is comparable to PW results [1, 8] and agree well with the experimental value [20].

Overall, therefore, we see a well-defined trend in the structural and thermodynamical values shown in Table 6.1: most quantities converge rapidly as the basis goes from SZ to DZ to DZP, with DZP providing an excellent agreement with experiment. Moreover, it appears that the optimal basis sets, SZ-O and SZP-O, compare very well with DZP, suggesting that they could be used when computational costs are an issue. The application of these optimized bases to study the diffusion of vacancies in GaAs will be reported somewhere else. [21]

6.2.2 Defects formation energies in supercell calculations

When computing structural and energetic properties of defects using *ab-initio* methods, it is important to ensure that the size of the basis set is complete enough but also that the simulation cell is sufficiently large to avoid self-interaction between the defect and its images. We have shown previously, [22] in a study of the neutral vacancy in silicon, that a cubic supercell of at least 216 atomic sites can be necessary in order to reduce the elastic and electronic self-interaction and obtain the right symmetry around the defect. [23] As discussed in Sect. 6.4.2.1, we find a similar behavior for As vacancy; unless indicated, therefore, we use a 215-atom cell for all our calculations of defects.

In all calculations, this initial 215-atom configuration cell is randomly distorted, to avoid imposing spurious symmetry in the fully relaxed defect state. All atoms are allowed to relax without any constraint until every force component falls below 0.04 eV/\AA . The energy minimization takes place at a constant volume, using the optimal lattice constant obtained with DZP, 5.6\AA (see Tab. 6.1), 1% denser than the experimental value.

The formation energy can be evaluated directly from total energies obtained from electronic structure calculations. For binary compounds it is current to use the formalism of Zhang and Northrup [2] (see Ref. [24] for intermediate steps). The formation energy

of a defect of charge state q is defined as:

$$E_f = E'_f + q(E_V + \mu_e) - \frac{1}{2}(n_{As} - n_{Ga})\Delta\mu \quad (6.1)$$

where E'_f is independent of $\Delta\mu$ and μ_e , and is represented by

$$E'_f = E_{tot}(q) - \frac{1}{2}(n_{As} + n_{Ga})\mu_{GaAs}^{bulk} - \frac{1}{2}(n_{As} - n_{Ga})(\mu_{As}^{bulk} - \mu_{Ga}^{bulk}) \quad (6.2)$$

where n_{As} and n_{Ga} are the number of As and Ga ions present in the sample, q denotes the net number of electrons or holes supported by the vacancy, μ_e is the electron chemical potential or the Fermi energy E_F , and E_V is the energy at the valence band maximum. Errors in E_V due to the finite supercell are corrected by aligning the vacuum levels of the defective supercell and undefected supercell [25].

If $\Delta\mu$ is defined as the chemical potential difference:

$$\Delta\mu = (\mu_{As} - \mu_{Ga}) - (\mu_{As}^{bulk} - \mu_{Ga}^{bulk}), \quad (6.3)$$

the restriction on the chemical potentials becomes: $0 \leq \mu_e \leq E_g$ and $-\Delta H \leq \Delta\mu \leq \Delta H$, where E_g is the energy gap, and the heat of formation ΔH of bulk GaAs is defined as the difference between the chemical potential of bulk As and bulk Ga crystals and that of bulk GaAs. This latter quantity represents the energy necessary to dissociate GaAs into its individual components:

$$\Delta H = \mu_{As}^{bulk} + \mu_{Ga}^{bulk} - \mu_{GaAs}^{bulk} \quad (6.4)$$

For Ga vacancies of charge q the Eq. 6.1 reduces to :

$$E_f = E'_f + q(E_V + \mu_e) - \frac{1}{2}\Delta\mu \quad (6.5)$$

with $E'_f = E_{tot}(q) - \frac{215}{2}\mu_{GaAs}^{bulk} + \frac{1}{2}(\mu_{As}^{bulk} - \mu_{Ga}^{bulk})$; for As vacancies, it becomes:

$$E_f = E'_f + q(E_V + \mu_e) + \frac{1}{2}\Delta\mu \quad (6.6)$$

and $E'_f = E_{tot}(q) - \frac{215}{2}\mu_{GaAs}^{bulk} - \frac{1}{2}(\mu_{As}^{bulk} - \mu_{Ga}^{bulk})$.

6.2.3 Computing the ionization energy of charged defects

The concentration of charged defects is controlled by the position of the Fermi level which is determined by the local concentration of carriers. Since GaAs is used in a doped state in devices, it is important to assess the possible charged states of defects.

Charges can affect strongly the formation energy as well as the structure of a defect, changing the symmetry of the relaxed state and altering considerably the local electronic properties. For charged defects, the effects of finite-size supercell will be even more marked due to long-ranged nature of the Coulomb interaction; the use of a sufficiently large supercell is therefore even more important.

To account for the electrostatic interaction of periodically arranged defects of charge q as well as their interaction with the compensating background, we follow the approximate procedure of Makov and Payne. [26] The correction to the total energy of a charged system is handled by SIESTA, and it consists of a monopole correction only ($q^2\alpha/2\epsilon L$), where α is the Madelung constant of the simple cubic lattice, L is the defect-defect distance (16.8 Å) and ϵ is the experimental static dielectric constant. The monopole correction is found to be 0.094 eV, 0.37 eV and 0.84 eV for the charge states ± 1 , ± 2 , and ± 3 , respectively. The quadrupole correction, which we evaluated by hand, is proportional to $1/L^3$. For the 215-atom supercell, it is $2.42 \times 10^{-6} * qQ$ eV (where Q is the quadrupole moment), and can therefore be neglected (see also Ref. [27, 28]).

Because of the limitations of LDA, localized DFT eigenvalues are not equivalent to the measured electronic levels. Thus ionization energy is obtained from the difference between q_1 and q_2 electron total energy calculations ($\epsilon(q_2/q_1) = E_{tot}^{q_1} - E_{tot}^{q_2} - (q_2 - q_1)E_V$), rather than the difference of q_2 and q_1 electron eigenvalues of a single calculation. Usually only one electron is transferred between the electron reservoir and the

defect levels. When two electrons are transferred at the same time the electron-electron repulsion is compensated by a relaxation of the structure around the defect that arises from a strong electron-phonon coupling. This so-called *negative-U* effect is found when the ionization level $\epsilon((q-1)/q)$ appears above $\epsilon(q/(q+1))$, thus a direct transition $((q-1)/(q+1))$ is energetically more favorable. [29]

6.3 Converging defect formation and ionization energies

In this section, we study the effects of the basis set, the k -sampling and the simulation cell size on the formation energy and the relaxed geometry of neutral and charged Ga and As vacancies. A special emphasis is put on V_{Ga}^{-3} and V_{As}^{+1} for the analysis of convergence. These are the dominant charge states as is discussed in Section 6.4

6.3.1 Local basis set effect

We first study the effect of the choice of the local basis set on the defects. In order to cancel out all other effects, we use supercells of 215 atomic sites and a Monkhorst-Pack grid [30] of $2 \times 2 \times 2$ corresponding to a density of 0.03 \AA^{-1} of k -points.

Cohesive energies as well as bulk moduli studied in Section 6.2.1 for the different basis sets used show that atomic bonding is strengthened progressively as we go from SZ basis to DZP.

Structural relaxation is directly related to the interatomic forces acting on atoms around the vacancy and on the strength of atomic bonding. Atoms around the vacancy form initially an ideal tetrahedron with six equal distances labeled $d_1 - d_6$ with tetrahedral symmetry T_d . After the full relaxation (see Sect. 6.2.1), distances and angles can be altered and the symmetry is either conserved or broken.

Relaxations around the vacancies are given for different charge states in Tables 6.2 and 6.3 for Ga and As vacancies respectively. In both cases the corresponding formation energy is reported as well as the relative change in the volume of the tetrahedron with respect to the ideal one. [23, 24]

Due to the finite precision in the relaxation, there is some imprecision in the iden-

tification of the defect symmetry. Here, if the highest relative difference between two bonds is lower than 1 % (equivalent to a precision of 0.04 Å), the structure is assigned to highest symmetry group. The last column lists the symmetry groups for the different defects in the DZP basis set. Unless specified, the symmetry group for all bases is the same as that of DZP.

6.3.1.1 V_{Ga}

The Ga vacancy maintains the same symmetry for all charge states irrespective of the basis set used. The structural relaxation is most important for the smallest basis set, SZ, decreasing progressively by about one third as the number of orbitals is increased, but the T_d symmetry is maintained in all cases, with the atoms moving inwards systematically. Moreover, for each basis set the degree of structural relaxation, is almost independent of the charge state, the variation of the structural relaxation from the neutral to the -3 charge state goes from 2% (SZ) to 8% (DZP). As can be seen in Table 6.2, the formation energy is also rather well converged with the minimal basis set (SZ), by comparison with the more accurate DZP: the difference between the two bases is at most 0.2 eV.

Heavily charged defects like V_{Ga}^{-3} , which is the most likely charge state in a heavily doped material, are more sensitive to the completeness of the basis set. This effect explains the fact that the formation energy decreases with an improved basis, contrary to the other defects.

Because of its technological importance, we must ensure that the orbital overlap around the vacant site is sufficient to accommodate the extra electrons in V_{Ga}^{-3} . We can do so by placing a *ghost* Ga atom at the defect site. For that purpose, we generate a set of orbitals and place them on the crystalline site, without adding any pseudo-potential or extra electrons. The system is then relaxed using the same convergence criterion as before. This ghost atom does not have any significant effect on the total energy of the defect when the DZ and the DZP basis sets are used. In contrast, SZ basis total energies are corrected by 0.34 eV, suggesting that the SZ orbital are too short. This additional set of orbitals is sufficient to correct for the overestimated formation energy with SZ basis, decreasing its value from 4.06 to 3.72 eV, following the general trend observed for other

Table 6.2: Convergence of the formation energy E'_f in eV with respect to the basis set for the Ga vacancy. Relaxation around the vacancy are given in % compared to the ideal tetrahedral distance between As nearest neighbors. The distances are labeled $d_1 - d_6$, the negative sign indicates an inward relaxation. The tetrahedron volume change is also given in % of the ideal volume ($\Delta V = 100 * (V - V_0)/V_0$). The last column displays the symmetry group of the defect (see the text for more details).

Basis	E'_f (eV)	Distances in %						ΔV	Symmetry
		d_1	d_2	d_3	d_4	d_5	d_6		
V_{Ga}^0									
SZ	2.7	-19.3	-19.2	-19.3	-19.3	-19.2	-19.3	-47.3	
DZ	2.8	-14.5	-14.4	-14.5	-14.4	-14.5	-14.4	-37.4	
DZP	2.9	-13.5	-13.4	-13.5	-13.4	-13.4	-13.5	-35.2	T_d
V_{Ga}^{-1}									
SZ	2.9	-19.0	-19.1	-19.0	-19.0	-19.0	-19.0	-46.9	
DZ	2.9	-14.7	-14.7	-14.7	-14.7	-14.7	-14.7	-37.9	
DZP	3.0	-14.2	-14.2	-14.2	-14.2	-14.2	-14.2	-36.9	T_d
V_{Ga}^{-2}									
SZ	3.4	-19.5	-19.5	-19.5	-19.5	-19.5	-19.5	-47.8	
DZ	3.2	-14.4	-14.4	-14.4	-14.4	-14.4	-14.4	-37.3	
DZP	3.4	-14.0	-14.1	-14.0	-14.0	-14.1	-14.0	-36.5	T_d
V_{Ga}^{-3}									
SZ	4.1	-19.7	-19.8	-19.7	-19.7	-19.8	-19.7	-48.3	
DZ	3.8	-15.1	-15.2	-15.1	-15.1	-15.2	-15.1	-38.9	
DZP	3.9	-14.6	-14.6	-14.6	-14.6	-14.7	-14.6	-37.7	T_d

charge states (see table 6.2).

6.3.1.2 V_{As}

The situation is very different for the As vacancy: the local symmetry is broken for most charge states and the completeness of the basis set impacts strongly on the reconstruction around the defect. Except for the positively charged vacancy, the bonds are stretched considerably to form pairs, leading to volume deformation by as much as 60%. Because of this strong deformation, we relax the threshold condition on the relative difference between two bonds used to determine the symmetry. To allow the reader to judge the impact of a relaxed classification, the distances $d_1 - d_6$ are also indicated (in ascendent order) in Table 6.3.

As a general trend, the SZ basis is less efficient for As than for Ga vacancies: the formation energy is underestimated by as much as 0.6 eV (21%) for charge +1 and 1.0 eV (17%) for charges -3 as compared with DZP. The underestimation drops to 2-8 % with DZ, a considerable improvement, for all charge states. The improvement in formation energy can be directly correlated with the prediction quality of the local relaxation. For example, while T_d symmetry is conserved for the three basis sets for V_{As}^{+1} , the change in the volume around the V_{As}^{+1} is highly overestimated by 131 % using SZ basis compared to the DZP results. The overestimation drop to 33 % with DZ, leading to an error of less than 0.1 eV compared with DZP.

As more electrons are added to the defect level this trend tends to diminish, the symmetry and the relaxation of the defect can be described with reasonable accuracy using the DZ basis. Applied to the singly negative As vacancy V_{As}^{-1} , the inclusion of the ghost atom at the vacant site has a smaller impact on the energy level and relaxation using SZ than for V_{Ga}^{-3} , and next to none with DZ and DZP. For the minimal basis set, the correction is negligible and accounts only for 0.067 eV.

Table 6.3: Convergence of the formation energy E'_f in eV with respect to the basis set for the As vacancy. The relaxation around the vacancy is given in % compared to the ideal tetrahedral distance between **Ga** nearest neighbors. The distances are labeled $d_1 - d_6$, the negative sign indicates an inward relaxation. The volume change around the vacancy is also given in % of the ideal volume ($\Delta V = 100 * (V - V_0)/V_0$). The last column displays the symmetry group of the defect (see the text for more details).

Basis	E'_f (eV)	Distances in %						ΔV	Symmetry
		d_1	d_2	d_3	d_4	d_5	d_6		
V_{As}^{+1}									
SZ	2.2	-14.5	-14.9	-15.1	-15.7	-15.9	-16.6	-39.5	
DZ	2.7	-7.9	-8.0	-8.1	-8.4	-8.4	-8.7	-22.7	
DZP	2.8	-5.6	-5.9	-6.1	-6.3	-6.2	-6.4	-17.1	T_d
V_{As}^0									
SZ	2.4	-13.8	-14.3	-15.2	-15.2	-31.3	-31.7	-53.4	C_{1h}
DZ	3.1	-12.2	-12.2	-12.7	-13.0	-26.9	-27.0	-46.7	$\sim D_{2d}$
DZP	3.2	-11.1	-11.2	-11.5	-11.6	-24.6	-24.5	-43.0	D_{2d}
V_{As}^{-1}									
SZ	2.5	-14.3	-14.7	-15.3	-15.4	-33.3	-33.5	-55.6	
DZ	3.2	-12.5	-12.7	-13.2	-13.4	-31.8	-32.0	-52.4	
DZP	3.3	-11.6	-11.9	-12.5	-12.6	-31.2	-31.5	-51.2	$\sim D_{2d}$
V_{As}^{-2}									
SZ	3.8	-17.8	-18.5	-18.5	-19.3	-33.6	-33.7	-58.6	$\sim C_{3v}$
DZ	4.4	-16.5	-17.1	-17.5	-18.2	-32.0	-32.3	-56.2	$\sim C_{3v}$
DZP	4.6	-16.5	-16.8	-17.0	-17.6	-31.6	-31.7	-55.4	C_{3v}/D_{2d}
V_{As}^{-3}									
SZ	4.9	-18.4	-18.4	-31.4	-31.6	-31.7	-31.8	-65.5	
DZ	5.4	-18.3	-18.2	-30.4	-30.6	-30.7	-31.0	-63.8	
DZP	5.9	-18.1	-18.2	-30.4	-30.2	-30.2	-30.2	-62.9	D_{2d} -resonant

6.3.1.3 Convergence of the ionization energy with the basis sets

As defined in section 6.2.3, ionization energies are taken as the difference between the total energy in different charge states, eliminating or decreasing many errors present in the formation energy. For example, errors coming from LDA are canceled out and those coming from different chemical potentials are eliminated. The main remaining sources of error are those coming from the basis set convergence, the evaluation of the valence band edge energy and the Madelung correction.

Figure 6.1 displays the convergence of the ionization energies as a function of the basis set used for V_{As} and V_{Ga} . For both types of defects, the preliminary results obtained using SZ basis set give a rough estimate of the location of ionization energies in the band gap. These energies converge with increasing basis sets, but slower than the formation energies. For V_{Ga} , while the formation energies are reasonably converged already with the minimal basis set, the ionization levels found using the SZ basis are noticeably overestimated compared to DZP. For V_{As} , we found in section 6.3.1.2 that formation energies are underestimated by less than 1eV with SZ for most charge states. Once ionization energies are calculated, errors on the formation energies coming from the relaxation around the defect cancel out since defects in all charge states suffer from this effect. SZ basis gives a correct qualitative description of the nature of the electronic transition (double negative-U effect, to be discussed in section 6.4.2.1). The location of the levels in the band gap as well as the distance between them are also reasonably converged. Ionization energies with SZ are slightly underestimated compared to the DZP ionization levels.

By studying the effect of the choice of the basis set on the structural relaxation, the formation and the ionization energies, we conclude that the SZ basis is significantly less efficient when the local symmetry is broken. In these cases, the use of a second radial function (DZ) is necessary to obtain reasonable numbers. Moreover, both DZ and DZP are complete enough to represent the properties associated with a defect without the need for a ghost atom.

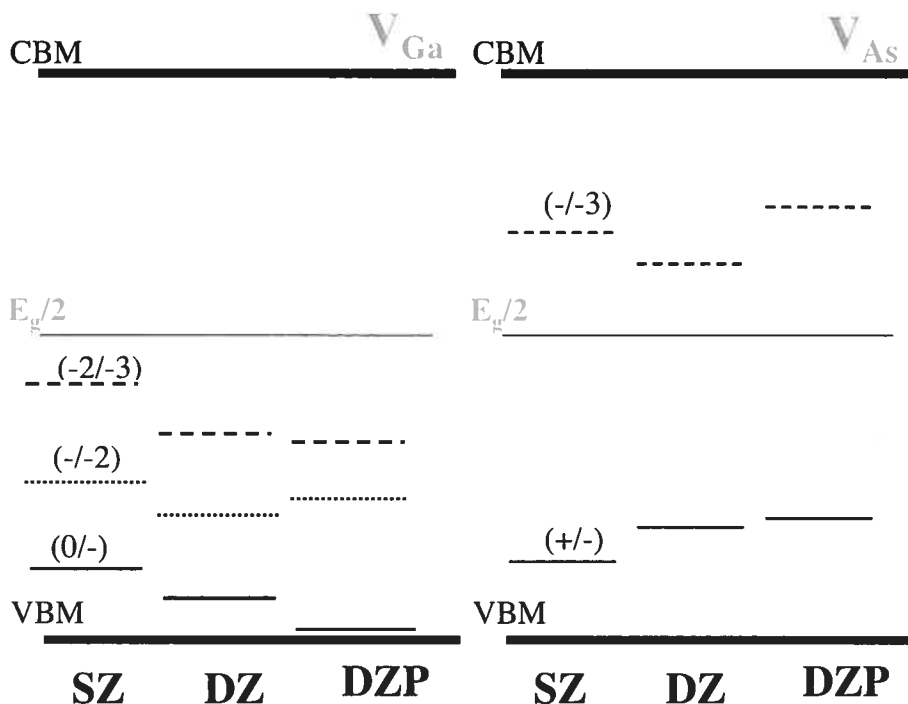


Figure 6.1: Schematic representation of the convergence of ionization energies as function of the basis set. Due to the underestimation of the gap, as a consequence of LDA, it is usual to align the conduction band minimum (CBM) with the experimental value, at 1.52 eV from the valence band maximum (VBM). The left panel displays the ionization energies as function of the basis set for the Ga vacancy that are all below the experimental mid gap. The right panel shows the two ionization levels for the As vacancy located below and above the mid gap and their convergence with respect to the different basis sets used. Refer to the text for the basis set description.

6.3.2 K-points effects

The effects of Brillouin zone sampling are studied by comparing the formation energies and the relaxation volume of both As and Ga vacancies for all charge states. We consider two k-point sampling, a Γ point sampling and a $2 \times 2 \times 2$ Monkhorst-Pack mesh [30], corresponding, for the 215-atom cell used here, to a density of 0.06 \AA^{-3} and 0.03 \AA^{-3} , respectively. To isolate sampling effects, all calculations are done using the DZP basis set. Results are shown in Table 6.4.

The use of Γ point only gives formation energies that are reasonably converged for Ga vacancies. Similarly, the bond lengths around the vacancy and the relaxation volume

Table 6.4: Convergence of the formation energies E'_f (in eV) for Ga and As vacancies with respect to the Brillouin zone sampling. Percentages refer to the corresponding relative difference in formation energies between the two sampling schemes

	+1	Neutral	-1	-2	-3
	$E'_f(V_{Ga})$				
Γ		2.51	2.73	3.17	3.81
$2 \times 2 \times 2$		2.94	3.00	3.40	3.94
		15%	9%	7%	3%
	$E'_f(V_{As})$				
Γ	2.73	3.21	3.52	4.70	6.08
$2 \times 2 \times 2$	2.79	3.25	3.33	4.52	5.86
	-2%	-1%	6%	4%	4%

(Fig. 6.2) are near those of the $2 \times 2 \times 2$, preserving the T_d relaxation symmetry for all charge states.

The use of Γ -point only produces less reliable results in the case of As vacancies. Formation energies are well converged but the relaxation symmetry around the defect is not correctly predicted for all charge states. With Γ -point, the distance between the atoms forming the tetrahedron as well as its associated volume are already converged for charge states +1, 0, and -1, as shown in Fig. 6.2.

For V_{As}^{-2} , Γ -point sampling overestimates the volume contraction around the defect by 7.5%. The long bonds in the pairing mode associated with the D_{2d} structure are shortened from 3.46 Å for this sampling to 3.26 Å for $2 \times 2 \times 2$ sampling. The k-point sampling effect becomes even stronger for the highly charged V_{As}^{-3} . The resonant bond geometry present with high-density sampling (discussed further in Sect. 6.4.2.1) is not found with the Γ -point sampling and the defects relaxes into the usual pairing mode with two short bonds and four long bonds.

Comparing the two samplings for V_{As}^{-3} , we find that: (1) with DZP + 2^3 mesh, the

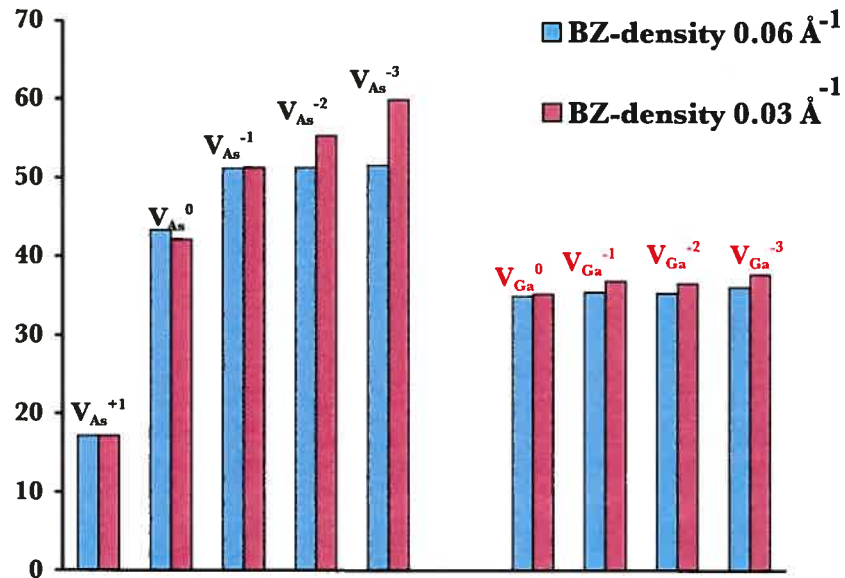


Figure 6.2: Histogram of the change in the volume of the relaxed tetrahedron formed by atoms surrounding the vacancy in % of the ideal one for two densities of k-points in the Brillouin zone (refer to the text for details). For comparison, data for Ga and As vacancies in different charge states are plotted in the same figure.

resonant-bond configuration has the lowest energy; (2) the pairing mode configuration is unstable using DZP + 2³ mesh, starting from this configuration, the defect relaxes back into the resonant-bond state; (3) for Γ -point sampling the resonant bond configuration is found to be metastable but the pairing mode is favored with an energy difference of 0.6 eV.

Overall, the use of Γ -point sampling gives a reasonable description of the energetic and structural properties of the Ga vacancy as well as most charge states for As. It fails, however, for heavily charged defects, when many electrons are involved in the bonding.

6.3.3 Supercell size effects

Size effects on defect formation energies have been widely discussed for a number of systems including the silicon vacancy (see Ref. [22] and references therein) and GaAs. [7] Size effects are found to be strong for cubic supercells smaller than 216±1

atoms for both materials and becomes negligible for larger systems. For this reason, we will study in this section size effects on charged Ga and As vacancies for supercells of 216 ionic sites and smaller.

To characterize the size effects, we consider the dominant charge state for each vacancy type, V_{As}^{-1} and V_{Ga}^{-3} , with DZP basis set. We simulate two cell-size with the same k-point density: a 63-atom unit cell with a $3 \times 3 \times 3$ sampling and a 215-atom cell with $2 \times 2 \times 2$ sampling. [7]

We find a formation energy of 3.83 eV compared to 3.33 eV obtained for the 215 cell for V_{As}^{-1} . In this case size effects are strong and overestimate the formation energy by 0.5 eV. More important, the symmetry of the relaxed defect is different for the two sizes: in the 63-atom cell, the defects relax to a C_{2v} symmetry while the defects 215-atom cell adopts a D_{2d} symmetry (Table 6.3). We check that the 63-atom cell is not caught into a metastable state by relaxing the defect starting in a D_{2d} symmetric state. After full relaxation, the cell relaxes into a C_{2v} , confirming that this is the lowest-energy symmetry for this cell. The different symmetry also impacts the change in the volume surrounding the defect. As a consequence of the strong defect-defect interaction in the smaller cell, the structural relaxation is hindered and the volume decreases by 43.8% compared with 51.2% for the 215-atom cell.

As could be expected from the previous sections, size effects are less important for V_{Ga}^{-3} . In particular, the initial tetrahedral symmetry (T_d), with atoms equidistant, is maintained around the defect for the relaxed 215-atom supercell. The same symmetry is found in the 63-atom supercell. Intercell defect-defect interaction rigidify the lattice, however, and the change in volume is only 34.8 % for the 63-atom cell compared to 37.7 % for the larger supercell. As with V_{As}^{-1} , the formation energy is overestimated with the small cell: $E_f^{63} = 4.24$ eV while $E_f^{215} = 3.94$ eV. This difference is considerable as it is on the order of the ionization energies.

6.3.4 Summary

In this section, we have studied in details the effects of the choice of basis set, k-point sampling and simulation cell on the properties of charged defects. In summary, we find

that: (1) The DZP basis set is well-converged and ensures reliable results for all charge states. For a number of charged states, it is also possible to use cheaper optimized basis sets for a similar accuracy. This is not always the case, however, and the applicability of these basis sets must be evaluated on a case by case basis. (2) SZ is less efficient than DZP for defects where symmetry is broken, but it gives a satisfactory estimation of the location of ionization levels in the band gap. (3) For supercell of 216 atoms or more, the density of k-points has only a minor effect on the defect relaxation. The use of the Γ point only gives a relaxation and a symmetry that are satisfactory, in most cases. However, this reduced sampling must be used with care for highly charged defects such as V_{As}^{-3} . (4) The errors arising from size effects are much more important than those coming from the density of k-points. In particular, size effects can be the source of errors in estimating the ionization energies, especially when the transition from a charge state to another induces breaking of the symmetry.

6.4 Discussion

We discuss here the results reported in the previous sections; we concentrate on the highly converged results for the Ga and As vacancies, obtained using the DZP basis set, a 215-supercell and a Monkhorst-pack grid of $2 \times 2 \times 2$ in the reciprocal space. We invite the reader to refer to the DZP results reported, in bold, in Tables 6.2 and 6.3. We first deal with the stoichiometric case where $n_{Ga} = n_{As}$ ($\Delta\mu = 0$), then we study the formation energies for each kind of defects separately under ideal growing conditions as a function of the doping level. Next, the dominant vacancy type defects in real GaAs crystals are identified by taking into account growing conditions. The ionic chemical potentials varies from Ga-rich conditions ($\Delta\mu = -\Delta H$) to As-rich conditions ($\Delta\mu = +\Delta H$) as the Fermi level is changed progressively.

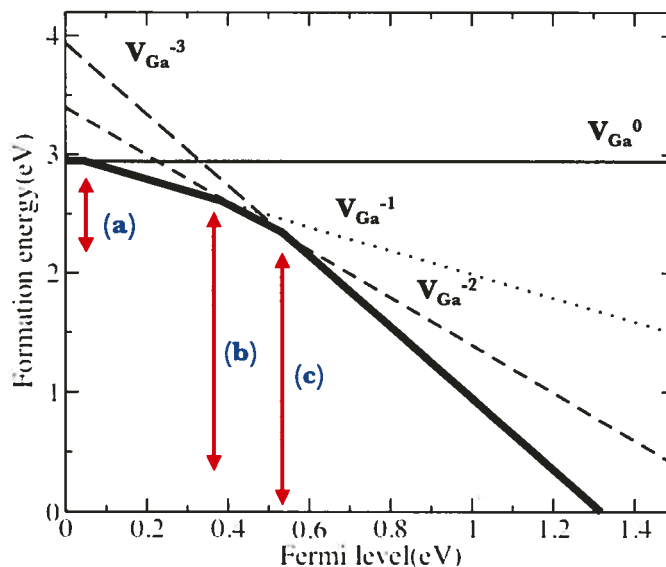


Figure 6.3: Formation energies as function of Fermi level in various charge states of Ga vacancies at 0 K calculated for stoichiometric GaAs ($\Delta\mu = 0$). The Fermi level is defined by reference to the valence band maximum. Ionization levels are defined as the intersection between the formation energies of different defects. Defect with the lowest formation energy is dominant. Arrows point to the location of the ionization levels labeled (a) for (0/-), (b) for (-/-2) and (c) for (-2/-3)

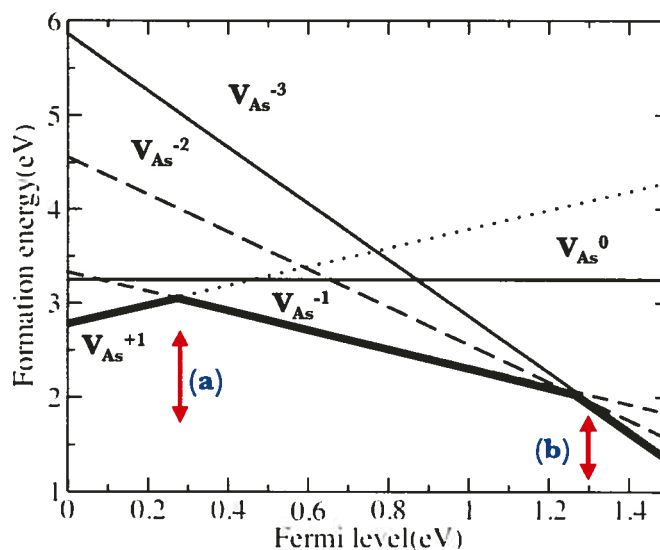


Figure 6.4: Formation energies as a function of the Fermi energy of various charge states of As vacancies at 0 K. The Fermi level is calculated with respect to the valence band maximum. Arrows point to the location of the ionization levels labeled (a) for (+/-), (b) for (-/-3)

6.4.1 Gallium vacancies

6.4.1.1 Relaxation geometry

Table 6.2 describes the fully relaxed geometry of the defect. The structural deformation obtained is well localized around the vacant site; the magnitude of the relaxations is listed for the nearest neighbor As atoms in all relevant charge states, reaching 15% of the bulk bond distances. The tetrahedral symmetry T_d is always conserved for this defect, irrespective of the charge state: only the breathing mode matters here. The As dangling bonds do not form pairs in any charge state, but the back bonds formed with Ga atoms are clearly weakened; this is in agreement with the observation that the pairing mode is generically not energetically favorable for cation vacancies (Ga). [32] We also observe that all As atoms relax inward but the amount of the relaxation does not increase significantly as more electrons are added to the vacancy levels but remains almost stable (13.5-14.5%). Our results are in good agreement with the *ab-initio* calculations of Laasonen *et al.* [3] and the empirical tight-binding simulations of Seong and Lewis [4] who found a systematic inward relaxation with tetrahedral symmetry for Ga vacancies in (0,-1,-2) charge state.

6.4.1.2 Energetics

Experimentally, the Ga vacancy is found to exist in the 0,-1,-2,-3 charge states. Until recently, the preferred charge state for the Ga vacancy in GaAs was the subject of a hot debate: most calculations [1, 4-6] find that where GaAs is either semi-insulating or n-type and Fermi energy is away from the valence band edge, the gallium vacancy is in the triply negative charge, while diffusion experiments suggest a charge of -2 or -1. [33]

Using positron annihilation to determine the Gibbs free energy of formation for Ga vacancies in GaAs, Gebauer *et al.* [34] could finally resolve this debate, giving a quantitative estimation of the formation enthalpy for V_{Ga}^0 and V_{Ga}^{-3} . The vacancy concentration is directly probed with positron annihilation in Te doped GaAs as function of doping concentration, temperature and chemical potential. Our estimate of the formation energies

Table 6.5: Comparison between ionization energies (measured from the valence band edge) of the Ga vacancy in GaAs. Results are grouped following the three classes discussed in the text.

Authors	Ionization levels (eV)				
	Negative U				
	0/-1	-1/-2	-2/-3	+1/-1	-1/-3
Seong and Lewis [4]				0.035	0.078
Northrup and Zhang [5]	0.19	0.2	0.32		
Pöykkö et al. [6]	0.11	0.22	0.33		
Schick et al. [7]	0.09	0.13	0.20		
Janoti et al. [8]	0.13	0.15	0.18		
Jansen and Sankey [9]	0.1	0.35	0.50		
Baraff and Schlüter [10]	0.2	0.5	0.7		
Cheong and Chang [1]		0.49	0.69		
Gorczyca et al [31]	0.39	0.52	0.78		
This work	0.05	0.4	0.55		

reported in the first column of Table 6.2 are in good agreement with recent experimental and theoretical data. For the neutral vacancy, we get $E'_f(V_{Ga}^0) = 2.94$ eV, a value that agrees with $E'_f(V_{Ga}^0) = 2.8$ eV from Bockstedte and Scheffler [35] (earlier first-principles calculations [5] predict 3.0 eV) and the experimental results of Gebauer *et al.*, [34] $H'_f(V_{Ga}^0) = 3.2 \pm 0.5$ eV. A formation enthalpy of 1.8 ± 0.5 eV was also measured by Mitev *et al.* [36] using inter-diffusion experiment on AlGaAs/GaAs heterostructure. However, the charge state of the associated defect is unknown, rendering the comparison with our results difficult.

For the triply negative charge state, we find that $E'_f(V_{Ga}^{-3}) = 3.9$ eV. This value compares well with recent experiments [34] which give $H'_f(V_{Ga}^{-3}) = 3.6$ eV as well as with recent theoretical study by Janoti *et al.* [8] (3.6 ± 0.2 eV).

Considering the stoichiometric case where $n_{Ga} = n_{As}$ ($\Delta\mu = 0$), we can study the formation energies of GaAs under ideal growing conditions as a function of the doping level. Figure 6.3 displays the formation energies as function of Fermi energy for various charge states of Ga vacancies at 0K. As E_f depends linearly on the electronic chemical potential, μ_e , the slope of each of the lines represents the net charge for the system, q . Intersections determine the ionization levels where one electron is transferred from the electron reservoir to the defect level. We can see that at each transition only one electron is transferred at a time. Moreover, the ionization levels labeled a, b, and c favor the stability of the -3 charge states for intrinsic and n-type GaAs.

In order to compare with earlier theoretical work, we summarize the results into three types of behavior reported in Table 6.5: (1) Sole among all calculations, Seong et Lewis [4] find a negative-U effect for the Ga vacancy using TB-MD method, the transition levels identified are very shallow and favor the triply negative state in almost the entire range of the Fermi level. (2) The second category of levels reported in Ref. [5–8] are shallow and lie well below the mid gap with no negative-U effect detected. (3) Defects level can also lie deeper below the mid gap. This is the case for levels a, b and c in Fig. 6.3, which compare well with the results of Cheong and Chang [1], Baraff and Schlüter [10]. Also with a recent study from Gorczyca et al [31] where the -3 charge of V_{Ga} is relevant only if the Fermi energy is above 0.55 eV for arsenic-rich conditions.

Unfortunately, ionization levels cannot be directly measured experimentally, only their sum is obtainable. Nevertheless, experimental values obtained by electron irradiation of GaAs [37] support the assignment of deep-lying levels in GaAs, in agreement with the third category. Gebauer *et al.* [34] confirm this assignment using a model to fit their experimental data in order to identify the charge state of the vacancy in GaAs from the location of the ionization levels. This calculation shows that the -3 charge state is the most stable charge state if deep ionization energies coming from a calculation on unrelaxed Ga vacancies of Baraff and Schlüter, [10] are used, which are in agreement with our more precise calculations.

Table 6.6: Comparison between ionization energies (measured from the valence band edge) for the As vacancy in GaAs. The data are grouped according to the two categories discussed in the text.

Authors	Ionization levels (eV)			
	+1/0	0/-1	Negative U	
			+1/-1	-1/-3
Seong and Lewis [4]	1.41	1.54		
Jansen and Sankey [9]	1.30	1.40		
Cheong and Chang [1]			0.785	
Pöykkö et al [6]			0.86	
This work			0.27	1.27

6.4.2 Arsenic vacancies

6.4.2.1 Relaxation geometry

As vacancies in various charge states have been studied by a number of authors [3,4, 38] taking into consideration ionic relaxation and relaxation geometry. Except for V_{As}^{+1} , breathing mode displacement breaks the vacancy local symmetry. [3,4] Contrary to Ga vacancy, the volume of the As vacancy increases as electrons are added and shrinks when electrons are removed.

Our results confirm part of these findings. From Table 6.3, the volume of the tetrahedron shrinks from -17% to -60% of the initial volume as electron are added to the vacancy. However, all charge states, even positive ones, display inward relaxation with respect to the unrelaxed volume.

Comparing the bond lengths obtained after a full relaxation of the structure (Table 6.3), we find that for the positively charged vacancy (V_{As}^{+1}) there is no electron in the localized states and all four atoms relax inward by about -6% conserving the tetrahedral symmetry, with no Jahn-Teller distortion. Although this inwards relaxation is more im-

portant, in absolute value, than the outward relaxation reported in references [3,4], they agree in term of the conservation of the symmetry.

For the neutral As vacancy in GaAs (V_{As}^0), there is just one electron in the localized state formed by the dangling bonds. The volume reduction is more than twice as large as for the positively charged defect. This change in volume is associated with a Jahn-Teller distortion with the formation of two dimers, leading to two short (-24.5%) and four long Ga-Ga bonds (-11.5%) arranged in a D_{2d} symmetry. This stretches and weakens the back bonds but it allows all atoms to recover a fourfold coordination.

Using *ab initio* molecular dynamics, Laasonen *et al.* [3] see a small (2-3%) outward relaxation, and an even smaller (0.6%) pairing-mode relaxation, leading to a weak tetragonal distortion with D_{2d} symmetry. This calculation was found to suffer from band dispersion for the localized defect states due to the artificial interaction between unit cells; as a consequence atoms surrounding the vacancy are not allowed to relax properly. Using tight-binding molecular dynamics, a larger breathing-mode displacement was obtained by Seong and Lewis. [4] They find that local tetrahedral symmetry was broken, as one neighbor atom of the defect relaxes inward while the other three relax outward. The pairing mode relaxation is also found to be very small. Feng *et al.*, [38] using a similar method with a 64-atom supercell, found similar results leading to trigonal distortion with C_{3v} symmetry.

The difference between our results and previous calculations come from the use of a better converged potential as well as of a larger unit cell. Moreover, as stated previously, we have started the relaxation from various random geometries, always converging to the same final state: the lowest energy configuration has D_{2d} symmetry for neutral V_{As} .

The extra electron added to get V_{As}^{-1} can be accommodated in the same localized level as the previous one. A stronger pairing mode relaxation appears and the two short bonds become stronger (from -24.5% to -31.5%) atoms of the dimer get closer, while the long bond are almost kept fixed. In agreement with our results reported in Table 6.3, Chadi, [32] using LDA and 32-atom supercell, finds that the -1 charge state arises from a direct transition from the $+1$ state due to a pairing of the neighboring Ga atoms and has C_{2v} symmetry (pairing mode). The Ga atoms would then move by 0.8 \AA ($\sim -20\%$)

from their ideal position to form two sets of paired bonds.

The next electron added to the localized levels, occupy a different state. The arrangement of the atoms around the doubly negative vacancy V_{As}^{-2} is directly affected: the short bonds remain unchanged but the long bond become stronger passing from -12.6% to -17.6% with a slight change in the relaxation volume. The dimers are therefore brought closer without affecting the intradimer distance.

Most interestingly, the relaxation geometry is modified when a third electron is added. In particular, there is an inversion in the Jahn-Teller distortion and the pairs of atoms forming the two dimers get closer to each other and form new weak Ga-Ga bonds with a length equal to the intradimer distance. Finally a tetramer is formed where the four Ga atoms are equidistant and fivefold coordinated (three covalent Ga-As bonds and two weaker Ga-Ga bonds). The tetrahedron formed by the vacancy's first neighbor has 4 short bonds and 2 long bond, as can be clearly seen from Table 6.3.

This type of relaxation, “resonant bond” model, was first seen in calculations for the singly negative divacancy in silicon, [39, 40] then for divacancies in GaAs. [6] More recently, this relaxation pattern have been observed experimentally then confirmed using *ab-initio* cluster calculations [41] for the $As_{Si} - V_{Si}$ pair in silicon.

6.4.2.2 Energetics

We calculate the formation energies for all possible charge states of the As vacancy; the most relevant defects are reported in the first row of Table 6.4. Most of the earlier calculations deal with the positively charged vacancy and do not go beyond the -1 charge state; there are no recent calculations that report formation energies for the -2 and -3 charge states.

In a 32-atom supercell LDA calculation for the formation energy of charged defects, Northrup and Zhang find that the outwards relaxed V_{As}^{+3} shows an sp²-like bonding. It is stable for both intrinsic and p-type materials, [42] with a formation energy 1.7 eV lower than the (+1) charge state.

Our results are in complete disagreement with this calculation. In the stable geometry, atoms around the defect relax inwards conserving the T_d symmetry. The local

volume is decreased by -16.21%. In addition, the corresponding formation energy we obtain does not favor the (+3) charge state under any doping or growing conditions: $E'_f(V_{As}^{+3}) = 3.5$ eV compared to $E'_f(V_{As}^{+1}) = 2.8$ eV (in agreement with other calculations: $E'_f(V_{As}^{+1}) = 2.97$ eV [5] using LDA and 3.09 eV [43] using self-consistent-charge density-based tight-binding [44]). This confirms that +1 has the lowest energy among all other defect.

Figure 6.4 displays the formation energies as a function of Fermi level for various charge states of As vacancies at 0 K. For p-type GaAs (Fermi level close to the valence band maximum) the +1 charge state has the lowest energy, in agreement with recent results from Chadi. [32]

We note, moreover, that lines for the neutral and the negative charge state intersect before those for the positive and neutral state; the (0/−) ionization level is located well below the (+/0) level and represents a net signature of level inversion or of the so-called negative-U effect. [29] It is therefore energetically more favorable to transfer two electrons at the same time to the defect level from the Fermi level with the reaction $V_{As}^+ + 2e \rightarrow V_{As}^-$.

Such transfer is associated to the strong Jahn-Teller distortion discussed in Section 6.4.2.1 as the systems goes from the +1 to the −1 state. The negatively charged vacancy remains stable for intrinsic and n-type GaAs but is superseded by the triply negative vacancy in the heavily n-doped GaAs corresponding to $E_F = 1.27$ eV (level *b* in figure 6.4): the (−/−2) and (−2/−3) defect levels almost collapse, thus the direct (−/−3) transition is favored. Interestingly, the transition (−/−3) is associated with a structural change from the pairing mode relaxation to the resonant mode relaxation as discussed in Section 6.4.2.1. The four added electron are paired two by two as the energy gained from the structural relaxation overcomes the Coulombic repulsion for each of the two electrons, supplying a net effective attractive interaction between the electrons.

Only the +1, 0 and −1 charge states seem to have been studied previously by *ab initio* calculations. While Northrup et Zhang [5] predict that arsenic vacancy in GaAs exist in the + charge state only, other calculated ionization levels split into two main categories classified in Table 6.6: (i) A direct transition from a charge state to another

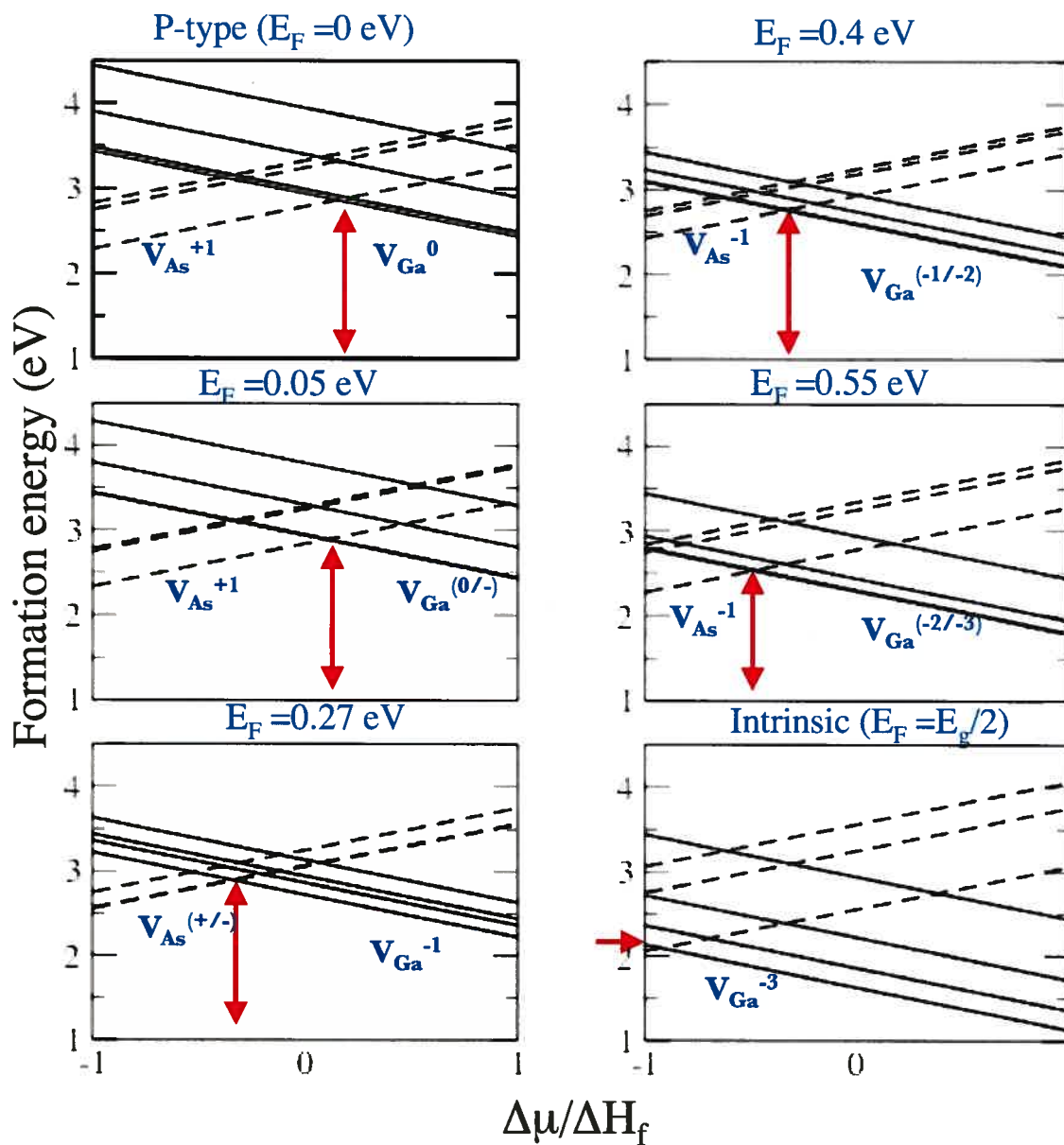


Figure 6.5: Formation energies of Ga (solid line) and As (dashed line) vacancies in GaAs as a function of the growth conditions ($\Delta\mu$). Different panels are for different critical values of the Fermi level or ionization levels identified earlier.

is possible with only one electron transferred at time, and (ii) a negative-U effect for the (+/−) transition. The first type of transition is found by Seong and Lewis, [4] who predict V_{As} to exist only in the + charge state, at 1.41 eV above the valence band this charge state changes to neutral then to negative charge state at the limit of the conduction band maximum. In an earlier study for unrelaxed As vacancies reported by Jansen and Sankey [9] ionization energies are located in the range of experimental band gap and located near the conduction band maximum. A negative-U effect is reported by both Cheong and Chang [1] and Pöykkö *et al* [6] for the (+/−) transition, the transfer of the two electron occurs above the middle of the band gap. This negative-U behavior is confirmed by a recent calculation from Chadi [32] which finds that the direct transition $V_{As}^+ + 2e \rightarrow V_{As}^-$ is favored after a Jahn-Teller distortion and that −1 charge state is the most stable for Fermi levels above midgap.

Our results agree partially with this results. There is a negative-U effect but the level for the (+/−) transition is shallower. This might be due to the important structural relaxation that affects the neutral and − charged vacancies. Moreover, as discussed above, we also find a second negative-U transition level at $E_F = 1.27$ eV with the reaction $V_{As}^- + 2e \rightarrow V_{As}^{-3}$.

Real GaAs crystals are far from being perfectly stoichiometric, during growth there will be excess of Ga or As ions. A more general study concerns the effect of the growing conditions on the stability of the defect under certain doping conditions. In Figure 6.5, the various panels show the progressive doping of the GaAs sample and how the stability of the defect gets affected consequently. The Fermi level ranges between the valence band maximum and the midgap where most of the ionization levels computed in this work have been identified. Each of the panels shows the transition between two charge states for As and Ga vacancies at the critical values of the ionization levels. For example at $E_f=0.05$ eV the first transition for gallium vacancies takes place, the two lines collapse and are indicated as $V_{Ga}^{(0/-1)}$.

For p-type GaAs, at the Ga-rich limit in GaAs, corresponding to $\Delta\mu/\Delta H_f$ close to −1, the dominant charge state are the As vacancies that probably compete with As antisites, while for the As-rich limit Ga vacancies have lower formation energies. For a

Fermi energy at midgap and for n-type GaAs, regardless of the growing condition, the triply negative charge state is the most stable among others and have the lowest formation energy.

6.5 Conclusions

We have presented a complete description of the energetic and the relaxation geometry for relevant vacancy type defects in GaAs using the SIESTA *ab initio* program. Various convergence tests show that size effects, the completeness of the basis sets and the sampling of the Brillouin zone can become very important when the symmetry of the defect is broken or when the defect is highly charged.

Using the DZP basis set, with a 216-atom unit cell and a $2 \times 2 \times 2$ k -point sampling, we find that Ga vacancies have shallow ionization levels below midgap in agreement with experiment, and do not show any recombination of the dangling bonds as was shown in earlier calculation [3,4]. For the less studied As vacancy, we find that the ionization level (+/-) of As is located in the lower half of the band gap and lies near the valence band while the second negative-U level (-/- 3) is located above midgap with a significant difference in the relaxation pattern reported earlier. The triply negative charge state for As vacancy reconstruct in the resonant bond mode, in a similar fashion as divacancies in Si and GaAs reported earlier. Finally, we find that only a few vacancy types can act as vehicles for self-diffusion of dopants in real GaAs devices under different doping and growing conditions, suvy as the triply negative Ga vacancy for intrinsic and n-type GaAs.

These results will be used as a starting point for a detailed study of self-diffusion using the SIESTA-A-RT method presented elsewhere. [21]

6.6 Acknowledgments

We thank Pablo Ordejón for his help in the initial optimization of the basis sets. This work is funded in part by NSERC, FQRNT and the Research Canada Chair program. NM was supported by a Cottrell Research Corporation Scholarship. Most of the simula-

tions were run on the computers of the Réseau québécois de calcul de haute performance (RQCHP) whose support is gratefully acknowledged.

BIBLIOGRAPHY

- [1] B.-H. Cheong and K. J. Chang, Phys. Rev. B **49**, 17436 (1994).
- [2] S. B. Zhang and J. E. Northrup, Phys. Rev. Lett. **67**, 2339 (1991).
- [3] K. Laasonen, R. M. Nieminen, and M. J. Puska, Phys. Rev. B **45**, 4122 (1992).
- [4] H. Seong and L. J. Lewis, Phys. Rev. B **52**, 5675 (1995).
- [5] J. E. Northrup and S. B. Zhang, Phys. Rev. B **47**, 6791 (1993).
- [6] S. Pöykkö, M. J. Puska, and R. M. Nieminen, Phys. Rev. B **53**, 3813 (1996).
- [7] J. T. Schick, C. G. Morgan, and P. Papoulias, Phys. Rev. B **66**, 195302 (2002).
- [8] A. Janotti, S.-H. Wei, S. B. Zhang, and S. Kurtz, Phys. Rev. B **67**, 161201 (2003).
- [9] R. W. Jansen and O. F. Sankey, Phys. Rev. B **39**, 3192 (1989).
- [10] G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **55**, 1327 (1985).
- [11] Deepak, D. Balamurugan, and K. Nandi, Bull. Mater. Sci. **1**, 169 (2003).
- [12] D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, Int. J. Quant. Chem. **65**, 453 (1997).
- [13] J. M. Soler, E. Artacho, J. Gale, A. García, J. Junquera, and P. Ordejón, J. Phys. Condens. Matter **14**, 2745 (2002).
- [14] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
- [15] L. Kleiman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
- [16] J. Junquera, O. Paz, D. Sánchez-Portal, and E. Artacho, Phys. Rev. B **64**, 235111 (2001).

- [17] E. Anglada, J. M. Soler, J. Junquera, and E. Artacho, *Phys. Rev. B* **66**, 205101 (2002).
- [18] G. Zollo and R. M. Nieminen, *J. Phy. Condens. Matter* **15**, 843 (2003).
- [19] M. Levinshtein, S. Rumyantsev, and M. Shur, eds., *Handbook Series on Semiconductor Parameters*, vol. 1 (World Scientific, London, 1996).
- [20] C. Weast, ed., *Handbook of Chemistry and Physics* (CRC, Boca Raton, FL, 1992), 73rd ed.
- [21] F. El-Mellouhi and N. Mousseau, in preparation.
- [22] F. El-Mellouhi, N. Mousseau, and P. Ordejón, *Phys. Rev. B* **70**, 205202 (2004).
- [23] M. I. J. Probert and M. C. Payne, *Phys. Rev. B* **67**, 075204 (2003).
- [24] L. Torpo, M. Marlo, T. E. M. Staab, and R. M. Nieminen, *J. Phys. Condens. Matter* **13**, 6203 (2001).
- [25] S. B. Zhang and A. Zunger, *Phys. Rev. Lett.* **77**, 119 (1996).
- [26] G. Makov and M. C. Payne, *Phys. Rev. B* **51**, 4014 (1995).
- [27] M. Bockstedte, A. Mattausch, and O. Pankratov, *Phys. Rev. B* **68**, 205201 (2003).
- [28] J. Lento, J. L. Mozos, and R. M. Nieminen, *J. Phys.: Condens. Matter* **14**, 2637 (2002).
- [29] P. W. Anderson, *Phys. Rev. Lett.* **34**, 953 (1975).
- [30] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [31] I. Gorczyca, N. E. Christensen, and A. Svane, *Phys. Rev. B* **66**, 075210 (2002).
- [32] D. J. Chadi, *Materials Science in Semiconductors Processing* **6**, 281 (2003).
- [33] B. Bracht, E. E. Haller, K. Eberl, and M. Cardona, *App. Phys. Lett.* **74**, 49 (1999).

- [34] J. Gebauer, M. Lausmann, F. Redmann, R. Krause-Rehberg, H. S. Leipner, E. R. Weber, and P. Ebert, *Phys. Rev. B* **67**, 235207 (2003).
- [35] M. Bockstedte and M. Scheffler, *Z. Phy. Chem. (munich)* **200**, 195 (1997).
- [36] P. Mitev, S. Seshadri, L. J. Guido, D. T. Schaafsma, and D. H. Christensen, *Appl. Phys. Lett.* **73**, 3718 (1998).
- [37] J. V. Jia, Y Q nad Bardeleben, D. Stienvard, and C. Delerue, *Phys. Rev. B* **45**, 1645 (1992).
- [38] Y. P. Feng, C. K. Ong, H. C. Poon, and D. Tománek, *J. Phys. Condens. Matter* **9**, 4345 (1997).
- [39] M. Saito and A. Oshiyama, *Phys. Rev. Lett.* **73**, 866 (1994).
- [40] H. Seong and L. J. Lewis, *Phys. Rev. B* **53**, 9791 (1996).
- [41] S. Ögüt and J. R. Chelikowsky, *Phys. Rev. Lett.* **91**, 235503 (2003).
- [42] J. E. Northrup and S. B. Zhang, *Phys. Rev. B* **50**, 4962 (1994).
- [43] T. E. M. Staab, M. Haugk, T. Frauenheim, and H. S. Leipner, *Phys. Rev. Lett.* **83**, 5519 (1999).
- [44] Self-consistent-charge Density-functional based tight-binding.

CHAPITRE 7

L'AUTODIFFUSION DÉPENDANTE DE LA CHARGE POUR LES LACUNES DE GA DANS LE GAAS

Ce chapitre traite l'effet de la charge sur le chemin de migration des lacunes de gallium dans le GaAs. Les mécanismes de diffusion sont étudiés par l'article intitulé *Charge-dependent migration pathways for the Ga vacancy in GaAs* qui est prêt à la soumission.

La méthode SIEST-A-RT a été utilisée pour étudier la diffusion des lacunes de Ga chargées dans le semiconducteur binaire GaAs. La diffusion au second voisin se fait principalement par le chemin le plus simple pour tous les états de charge avec des barrières variant entre 1.7 et 2.0 eV. Les lacunes négatives possèdent un chemin de migration additionnel leur permettant de diffuser au second voisin. En plus, V_{Ga} diffuse au premier voisin si $q = 0, -1, -2$ en formant des complexes du type $\text{As}_{\text{Ga}}-\text{X}$ qui sont déformés si la lacune est négativement chargée. Finalement, les lacunes négativement chargées peuvent diffuser directement vers le quatrième voisin en suivant la direction 100 en franchissant des barrières supérieures à 4 eV.

Charge-dependent migration pathways for the Ga vacancy in GaAs

Fedwa El-Mellouhi¹ and Normand Mousseau^{1,2}

¹ Département de physique and Regroupement québécois sur les matériaux de pointe, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal (Québec) H3C 3J7, Canada

² Service de Recherches de Métallurgie Physique, Commissariat à l'énergie atomique-Saclay, 91191 Gif-sur-Yvette, France.

PACS numbers: 61.72.Ji, 71.15.Mb, 71.55.Eq, 71.20.Nr

Abstract

We use SIEST-ART, a combination of the local-basis *ab-initio* program SIESTA and the activation-relaxation technique (ART nouveau) to study the diffusion mechanisms of the gallium vacancy in GaAs. Vacancies are found to diffuse to the second neighbor using two different mechanisms, as well as to the first and fourth neighbors following various mechanisms. We find that the height of the energy barrier is sensitive to the Fermi-level and generally increases with the charge state. Migration pathways themselves can be strongly charge-dependent and may appear or disappear as a function of the charge state. These differences in transition state and migration barrier are explained by the charge transfer that takes place during the vacancy migration.

7.1 Introduction

Self-diffusion is one of the basic mass-transport mechanisms in materials and one of the most powerful tools used in the preparation of nanostructures [1–4]. Yet, many questions remain regarding the microscopic details of self-diffusion and recent studies have shown that even in the simplest cases, the mechanisms can be much more complicated than was initially thought. Diffusion on simple metallic surfaces, for example, was found to take place through a range of mechanisms involving from one to at least seven atoms [5, 6]. Similarly, recent studies on self-interstitial clusters [7, 8] and disordered systems [9, 10] have highlighted the importance of collective moves in easing atomic motion even in bulk systems and underlined the importance for a detailed characterization of these mechanisms in materials of technological interests. This is the case for semiconductors, for example, that are at the heart of the electronic industry. While one predicts, using symmetric considerations, the self-diffusion pathways in elemental materials, such as silicon, this approach becomes impossible when impurities are added or multi-compound systems are considered. Thus, one must turn to experiments or numerical simulations to provide a direct and comprehensive study of diffusion mechanisms in semiconductors. The difficulty to extract precise information on the diffusion mechanisms in these materials is compounded by the nature of bonding and the importance of charged defects that complicate seriously both experimental and theoretical studies.

Previous numerical studies of the migration pathways of intrinsic defects in binary semiconductors have focused on GaAs [11], SiC [12, 13] and GaN [14]. Except in one case [11], where high-temperature molecular dynamics was used, these works focus on optimizing preselected pathways using algorithms such as the drag [11] or the ridge method [15] that rely on the knowledge of the initial and final states in addition to a decent guess of the overall diffusion mechanism. While these approaches work efficiently to identify with precision the migration energy of previously known diffusion trajectories, they cannot help to identify complex or unexpected mechanisms that could also play an important role in the diffusion process.

Here, we present the application of the activation-relaxation technique in its more re-

cent implementation (ART nouveau) to explore systematically the diffusion mechanisms of complex systems. More precisely, we focus on the diffusion of V_{Ga} in GaAs because of its apparent simplicity but also because of its technological interest and its role in affecting the properties of bulk materials and nanostructures. Gallium vacancies are found to be mobile at typical growth and annealing temperatures [16]— with a dominant charge state strongly depending on growth conditions, temperature, dopants, etc.— playing the main role in dopant diffusion. For example, Bracht *et al.* [17] showed recently that the contribution of V_{Ga} to Ga self-diffusion in GaAs is even more important than earlier estimation giving an important contribution to the total diffusion profile. Furthermore, Tuomisto *et al.* [18] found that Ga vacancies play a central role in the migration of Mn in $Ga_{1-x}Mn_xAs$ alloys. Finally, the energy blueshift in PL spectra of InGaAs/GaAs [19] and InAs/GaInP quantum dots [20] has also been recently attributed to V_{Ga} diffusion.

Focusing on a simple defect, a gallium vacancy (V_{Ga}), in the weakly ionic GaAs, we show here that diffusion in bulk semiconductors is a complex phenomenon that depends not only on the geometry of the defect and the surrounding lattice but also on the charge of the defect. In particular, we identify a new mechanism for the diffusion to the second neighbor in addition to the one already found by Bockstedte and Scheffler [11], plus a number of other jumps to the first and fourth neighbors. Not all these pathways are likely to occur in a normal range of temperature, and some exist only for a subset of charge states, but their existence underlines the under-estimated richness of diffusion mechanisms in bulk materials.

This paper is organized as follows, in Sec. 7.2, we present the details of activated event generation using SIEST-A-RT. A description of initial and different transition states in various charge states are presented in Sec. 7.3. In Sec 7.4 we discuss the origin of charge dependent migration barrier and we compare our results with experimental and theoretical works.

7.2 Details of the simulation

Our simulations are performed using SIEST-A-RT, a method combining a self-consistent density functional method (SIESTA) [21] with the most recent version of activation relaxation technique (ART nouveau) [22]. Integrating various empirical potentials, ART nouveau was shown to sample efficiently the energy landscape of amorphous semiconductors [23], glasses [22], and proteins [24], for example.

SIEST-A-RT was used for the first time to study vacancy diffusion in Si and details of its implementation can be found in Ref. [25]. Forces and energies are evaluated using density-functional theory (DFT) with the local-density approximation (LDA) using standard norm-conserving pseudo-potentials of Troullier-Martins [26] factorized in the Kleiman-Baylander form [27]. Matrix elements are evaluated on a 3D grid in real space. The one-particle problem is solved using linear combination of pseudo-atomic orbitals (PAO) basis set of finite range. Here, we use the extended local basis set [28], which was shown to reproduce closely the best solution while minimizing computational costs. Calculations are performed on a 215-atom GaAs supercell sampled at the Γ special point. All relevant charge states (0, -1 , -2 and -3) are fully relaxed until the residual force falls below 0.002 eV/\AA , then we proceed to search for local diffusion pathways by assuming that the charge state of the defect is preserved during any transformation. Unless mentioned, all events start from fully relaxed gallium vacancy geometries.

In order to break the initial local symmetry, activated events are started by displacing in a random direction a region of the cell centered around a chosen atom in the first, second or fourth-shell around the vacancy, involving typically between 10 and 30 atoms. The structure around the vacancy is then deformed along this random direction until the lowest curvature, corresponding to the lowest eigenvalue of the Hessian matrix, becomes negative, falling below a preset threshold value. The system is then pushed along the corresponding eigenvector, while minimizing the energy in the perpendicular hyperplane, until the total force falls below 0.1 eV/\AA , indicating that the transition state has been reached.

About 60 events were generated in total for all charge states, with 20 events for

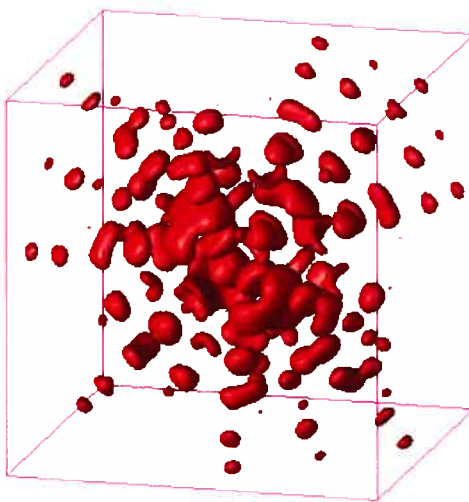


Figure 7.1: Constant surface at $0.0004 \text{ electron}/\text{\AA}^3$ of the difference in electronic charge density between the neutral and the -3 charged gallium vacancy. The gallium vacancy is located initially at the center of the box.

jumps to the first nearest neighbor and 30 for jumps to the second neighbor. In both cases the structure is deformed by at least 0.9 \AA before a sufficiently negative eigenvalue appears. Diffusion to the fourth neighbor was more difficult to complete (we generated 10 such events in total) and we set up a threshold displacement of 1.4 \AA to allow the configuration to escape from the harmonic well.

7.3 Results

7.3.1 Gallium vacancies at the initial state

We first examine the vacancy at the initial state before moving to the analysis of transition states. The structure of V_{Ga} for all charge states, for a number of basis sets within SIESTA, is presented in Ref. [28] and agrees well with other simulations and experiments. Relaxed at its energy minimum, V_{Ga} conserves the T_d symmetry for all charges, while the open volume associated with the vacancy decreases with increasing charge. Spin-polarized LDA relaxations for V_{Ga} also lead to a T_d symmetry, confirming that the symmetry conservation of V_{Ga}^q is not a drawback of LDA but a behavior proper

Table 7.1: Nearest neighbor distances in 111 direction relevant for first nearest neighbor diffusion from the initial \rightarrow to the final state. Distances are calculated by taking the initial position of V_{Ga} as reference. The last column describes the geometry of the final state.

	$V_{Ga} - As^{1st}$	$As^{1st} - Ga^{2nd}$	$Ga^{2nd} - As^{3rd}$	Final geometry
V_{Ga}^0	2.08 \rightarrow 0.44	2.43 \rightarrow 3.40	2.45 \rightarrow 2.46	$(As_{Ga} + V_{As}^{1st})^0$
V_{Ga}^{-1}	2.08 \rightarrow 0.90	2.42 \rightarrow 2.65	2.45 \rightarrow 2.58	$(V_{Ga} + I_{As} + V_{As}^{1st})^{-1}$
V_{Ga}^{-2}	2.06 \rightarrow 0.53	2.41 \rightarrow 2.69	2.45 \rightarrow 2.64	$(As_{Ga} + V_{As}^{1st} + I_{Ga} + V_{Ga}^{2nd})^{-2}$
V_{Ga}^{-3}	<i>Do not diffuse to the first neighbor</i>			

to cation vacancies, as pointed out previously by Chadi [29]. This behavior has been further confirmed by a recent calculation [30] on V_{Ga} in GaAs using two-component density-functional theory-LDA applied to a cubic 64 atom supercell together with a 4^3 k-point mesh Brillouin zone sampling .

A detailed analysis of the charge densities plots (see Fig. 7.1) and the Mulliken populations reveals that the electrons added to the neutral vacancy are delocalized and distribute themselves on the 111 axes passing by the As dangling bonds in a similar way for all studied charge states. Less than 4% of the charge is localized on the four As neighboring the vacancy (As^{1st}) while the remaining 96 % is spread over more distant neighbors on the 111 axes. This suggests that these predominantly covalent As–Ga bonds are progressively stiffened upon the addition of extra electrons, thus making them more difficult to break compared to the remaining bonds and affecting directly the height of the diffusion barriers.

7.3.2 Diffusion path to first neighbor

The activated events we generated using SIEST-A-RT show that diffusion to the first nearest neighbor is not possible for all sequential charge states as proposed by Van Vechten [31] nor impossible as found by Bockstedte and Scheffler [11]. We find rather that diffusion to the first nearest neighbor is very much charge dependent.

For $q = 0$ the first neighbor of the vacancy (As^{1st}) diffuses, along the 111 direction, toward the vacant site via a split vacancy configuration by optimizing its bonds with the close neighbors. This mechanism is similar to the diffusion mechanism of a neutral silicon vacancy in silicon [25, 32, 33]. First, each back bond of the diffusing atom ($As^{1st} - Ga^{2nd}$) stretches until it breaks during the migration of As^{1st} toward the vacancy. As^{1st} proceeds in its migration until a metastable vacancy-antisite structure ($As_{Ga} + V_{As}^{1st}$) forms. The $As^{1st} - Ga^{2nd}$ bond evolves from 2.43 Å at the initial state to 3.40 Å at the final state. This metastable complex is 0.67 eV higher than V_{Ga} and the recorded migration barrier for the first neighbor diffusion is 0.84 eV. The next first neighbor jump ($V_{As}^{1st} \rightarrow Ga^{2nd}$) leading to the formation of $As_{Ga} + Ga_{As}^{1st} + V_{Ga}^{2nd}$ complex is also possible by crossing a barrier of 1.55 eV. Starting from this last complex we did not find any mechanism for V_{Ga}^{2nd} to diffuse further to the first neighbor, suggesting that this jump is unfavorable.

For $q = -1$ the vacancy follows the same path to the saddle point as for $q = 0$: it crosses a barrier of 0.9 eV passing by a split vacancy site, but $As^{1st} - Ga^{2nd}$ bond is less stretched. As^{1st} relaxes then in a split vacancy site, forming a $(V_{Ga} + I_{As} + V_{As}^{1st})^{-1}$ complex. $As^{1st} - V_{Ga}$ distance is reduced to half (2.08 Å \rightarrow 0.90 Å) suggesting that As^{1st} is at half way between the two vacancies. This metastable configuration is 0.81 eV higher than the initial minimum. We confirmed that this final metastable state is not only a local minimum along the diffusion path to the first neighbor by starting from an ideal antisite (As_{Ga}) structure, then relaxing until the residual force becomes lower than 0.002 eV/Å. This antisite structure is found to be unstable in -1 charge state because As^{1st} leaves the ideal antisite structure and prefers to relax at a metastable state half way between the two vacancies.

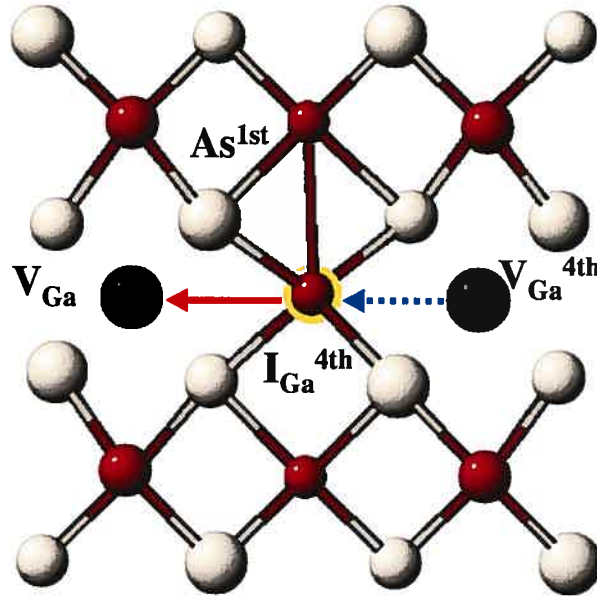


Figure 7.2: Migration path of V_{Ga}^q to the fourth neighbor along (100) direction (see the text).

For $q = -2$, we find only a collective motion of a $As^{1st} - Ga^{2nd}$ pair toward the vacancy along 111 direction. $As^{1st} - Ga^{2nd}$ bond stretches slightly, while As^{1st} approaches V_{Ga}^{-2} as close as 0.53 \AA . Consequently, Ga^{2nd} is forced to stretch its back bonds and stabilizes in an interstitial position. Finally, As^{1st} atom occupies V_{Ga} while Ga^{2nd} is located at a split interstitial position between V_{As}^{1st} and V_{Ga}^{2nd} . The resulting metastable complex $(As_{Ga} + V_{As}^{1st} + I_{Ga} + V_{Ga}^{2nd})^{-2}$ is 1.74 eV higher in energy than V_{Ga}^{-2} and can be obtained by crossing a barrier of 1.86 eV . Relevant distances for successful first nearest neighbor diffusion in $q = 0, -1, -2$ are summarized in Table 7.2.

For $q = -3$, all attempts for first nearest neighbor diffusion failed and the configuration always relaxes back to the initial minimum. We do not find any way to make the vacancy diffuse to the first neighbor even by forcing: using the transition state at the neutral charge state as starting point for a convergence of the $q = 3$ to its saddle point, the vacancy systematically returns to its original state.

7.3.3 Diffusion path to fourth neighbor

The situation is almost opposite for the diffusion to the fourth neighbour: all negatively charged vacancies succeed in diffusing along the pathway, only the neutral vacancy refuses to go this way.

The fourth neighbor of the vacancy (Ga^{4th}) approaches the interstitial region near the vacancy by diffusing along the 100 direction. Figure 7.2 illustrates this configuration: the dashed arrow shows the direction of the jump from the initial state to the saddle point, while the full arrow shows the path from the saddle to the final. Ga^{4th} diffuses to an unstable interstitial position close to the vacancy. In the cubic zincblende structure of GaAs, with lattice constant a and a vacancy located initially at $(0,0,0)$, the fourth neighbor diffuses first from $(a,0,0)$ to the interstitial position at $(a/2, 0,0)$. The transition state of this mechanism can be described as a gallium interstitial placed between two distant vacancies ($V_{Ga} - I_{Ga}^{4th} - V_{Ga}^{4th}$). This interstitial atom is not a direct neighbor of either of the two vacancies because the $Ga^{4th} - V_{Ga}$ distance shortens from 5.6 \AA at the initial minimum to 2.88 \AA at the transition state. This diffusion mechanism requires an elevated barrier (around 4.24 eV) since the Ga^{4th} needs to break two bonds initially with As^{5th} farthest away from the vacancy and to twist the remaining As^{3rd} bonds. Once it arrives at the saddle point two new bonds with As^{1st} are formed. It is interesting to note, as is shown in Table 7.3 that the energy barrier is almost independent of the charge for the defects that manage to diffuse to the fourth neighbour; this weak dependence can probably be attributed to the constant electronic charge distribution along 100 direction for different negative charging.

7.3.4 Diffusion path to second neighbor

By exploring the energy landscape of the system, we find that vacancy migration to the second neighbor occurs by two mechanisms. The simplest diffusion pathway is already well known [11] and is considered to mediate self-diffusion in binary semiconductors, while the second pathway, which has not been reported, to our knowledge, for this system, is more complex and involves the correlated motion of many atoms neigh-

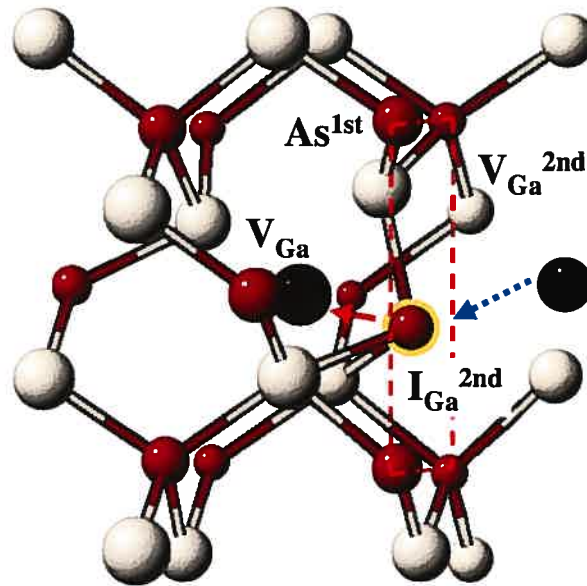


Figure 7.3: Diffusion to second neighbor by plane-passing mechanism (refer to the text).

bouring the vacancy.

7.3.4.1 Plane-passing mechanism

The most intuitive diffusion pathway, which we call the *plane-passing mechanism*, brings one Ga^{2nd} to the interstitial region, joining the diffusing atom and the vacancy, along the 110 direction. The diffusing Ga^{2nd} atom must go through the diffusion plane perpendicular to the 110 direction during its way to the vacancy site causing three As^{1st} atoms to move away from the vacancy and opening the cage. The transition state of this mechanism, first identified by Bockstedte and Scheffler [11], can be described as a gallium interstitial placed between two close vacancies ($V_{Ga} + I_{Ga}^{2nd} + V_{Ga}^{2nd}$). Figure 7.3 illustrates this configuration: the dashed arrow shows the direction of the jump from the initial state to the saddle point, while the full arrow shows the path from the saddle to the final; the diffusion plane formed by the second nearest neighbors of the vacancy is also shown.

From Table 7.2, one can notice that the structural details of the jumps are charge-dependent. The position of the transition state, in particular, changes by 20 % as the

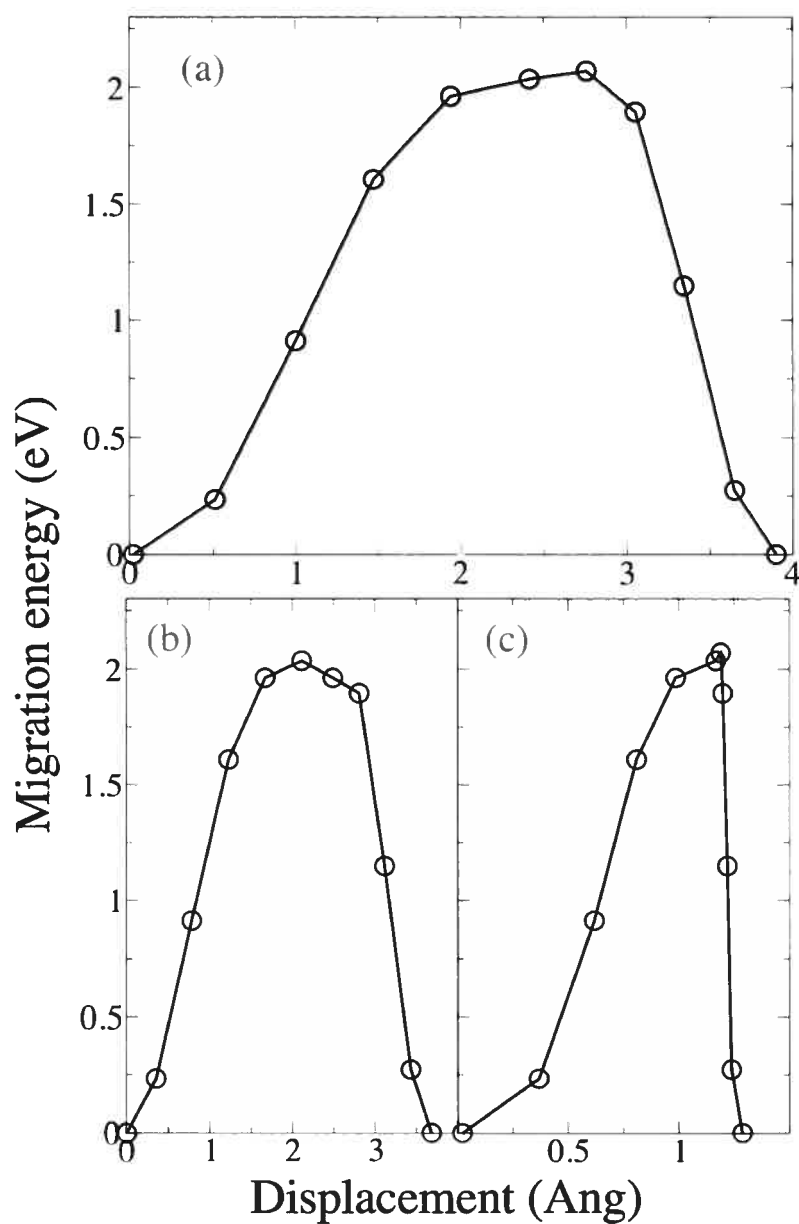


Figure 7.4: (a) Migration trajectory to the second neighbor by the plane-passing mechanism for -3 charge state. 11 images (circles) are relaxed with force tolerance of $1\text{eV}/\text{\AA}$, lines are guide to the eye. The contributions from different moving atoms are decomposed in two: (b) The path followed by the diffusing atom; (c) Arrangement of the other atoms around the defect.

Table 7.2: Evolution of the distance (in Å) between the initial position of the vacancy and the diffusing atom ($V_{Ga}^q - Ga^{2nd}$) in the initial state and the transition state for both diffusion mechanisms to the second neighbour.

	Plane-passing		Cluster-assisted	
	Initial	Saddle	Initial	Saddle
V_{Ga}^0	3.89	1.89		
V_{Ga}^{-1}	3.88	2.09	3.88	2.5
V_{Ga}^{-2}	3.82	2.19	3.82	2.78
V_{Ga}^{-3}	3.80	2.35	3.80	2.89

vacancy goes from neutral to a charge of -3 . As can be expected, the displacement of the unstable interstitial position as the diffusing atom becomes a nearest neighbour of the vacancy has noticeable impact on the energy barrier, which goes from 1.7 eV for neutral and -1 charge state to 1.84 and 2.0 eV for -2 and -3 charges respectively.

More precisely, for V_{Ga}^0 and V_{Ga}^{-1} , the saddle point is a site close the hexagonal interstitial, while it is closer to the tetrahedral interstitial configuration for the more negative V_{Ga}^{-2} and V_{Ga}^{-3} . At the transition state, as the charge increases, the displacement of the gallium atom out of the plane becomes less pronounced; for the -3 charge state, it almost vanishes, leaving the moving atom on the plane.

This can be seen by looking at the the full migration trajectory for -3 charge state in Figure 7.4. This path is generated by initially interpolating between the ART-generated initial, saddle and final states, generating 11 images. These image configurations are then relaxed using the CI-nudged-elastic-band method [34] until the total force becomes lower than $1\text{eV}/\text{Å}$. After decomposing the total path into contributions coming from different moving atoms one can notice that the path followed by the diffusing atom is nearly symmetric, while most asymmetry in the total path comes from arrangement of the other atoms around the defect.

Differences in the migration barriers are mainly due to the diffusion of the electronic

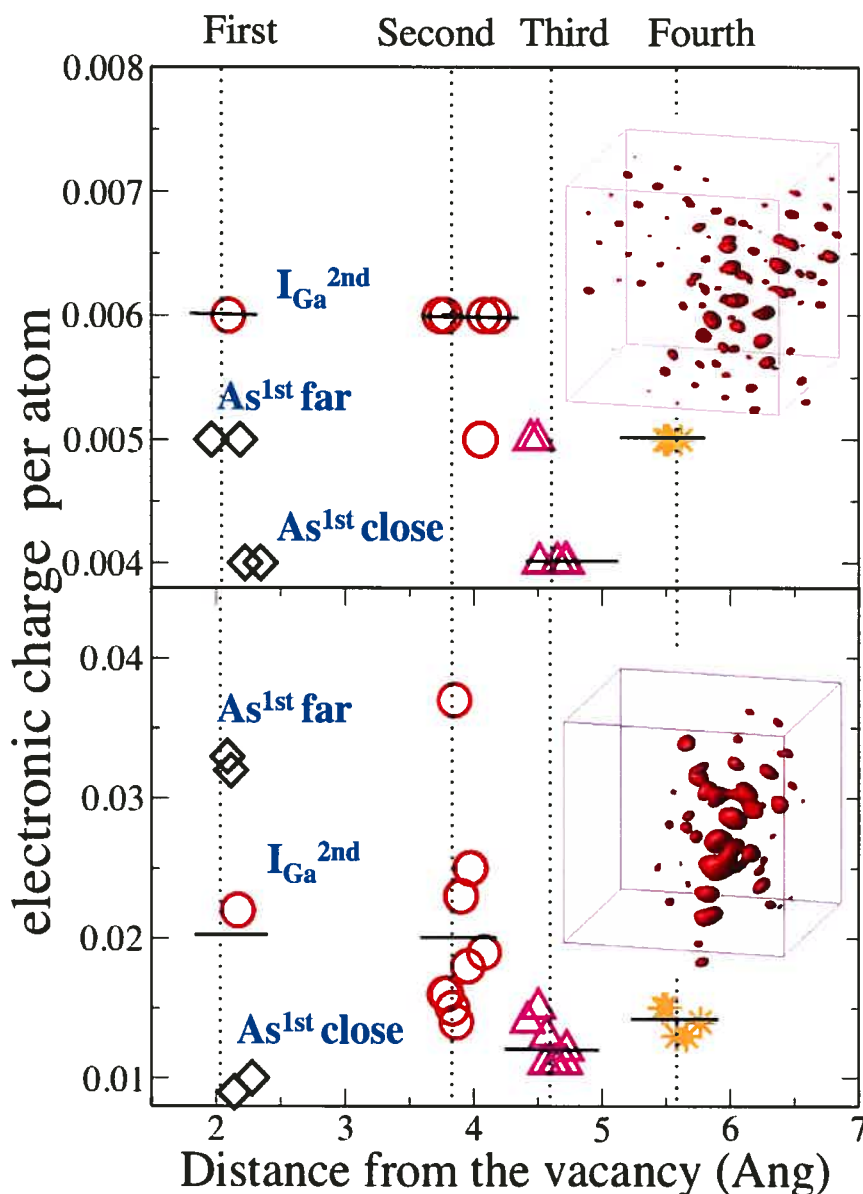


Figure 7.5: Mulliken population analysis of the difference in electronic charge between the neutral and the -1 (top panel), and -3 (low panel) charges at the saddle point of plane-passing mechanism as function of the distance from the centre of the vacancy. The insets show these charge densities in 3D at 0.00018 and 0.0008 electron/ \AA^3 near V_{Ga}^{-1} and V_{Ga}^{-3} respectively. The charge transfert during the plane-passing path could be deduced by comparing the initial and the saddle point populations: Mulliken populations at the initial state are indicated by horizontal solid lines, vertical dotted lines indicate the position of the first, second, third and fourth neighboring shell of the vacancy. At the saddle point, populations for each of the atoms originating from the first arsenic (\diamond), second gallium (\circ), third arsenic (\triangle) and fourth gallium shells ($*$) are shown.

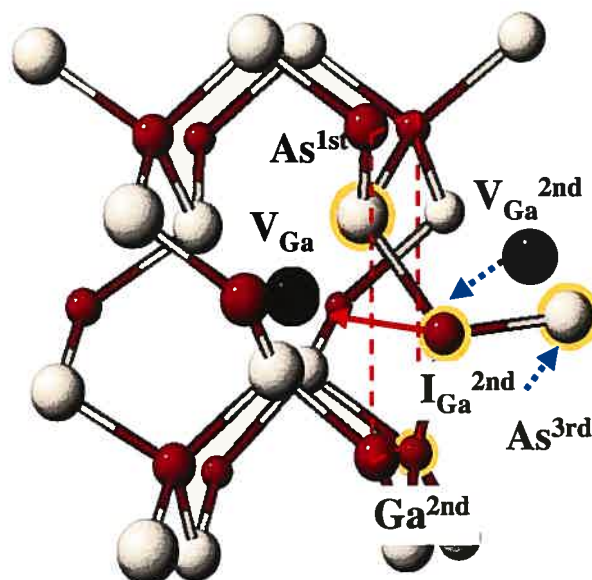


Figure 7.6: Diffusion to second neighbor by cluster-assisted mechanism (refer to the text).

charge during the jump. For the singly negative vacancy, Mulliken population analysis at the saddle point (top panel in Figure 7.5) is compared to the initial configuration. It shows the diffusion of the electronic charge during the migration of the Ga^{2nd} atom. The electronic charge at Ga^{2nd} diffuses with it to the hexagonal interstitial site thus saturating partially the As^{1st} dangling bonds. The charge is consequently suppressed from the neighborhood of the diffusing atom (I_{Ga}^{2nd}) and spreads uniformly over more distant shells as can be seen from the 3D charge density (top inset of Figure 7.5).

For $V_{Ga}^{-2,-3}$ the diffusing atom is less engaged toward the vacancy: circles in the lower panel in Figure 7.5 are shifted to the right as a signature of volume opening that affects the second neighbor of the vacancy as well. However, the 3D charge densities for V_{Ga}^{-2} and V_{Ga}^{-3} are different from V_{Ga}^{-1} . After the jump, the charge becomes highly localized around the dangling bonds belonging to As atoms farthest away from I_{Ga}^{2nd} labeled As_{far}^{1st} . Some of the charge carried by two closer As_{close}^{1st} to I_{Ga}^{2nd} get transferred to the other As_{far}^{1st} ; these have non saturated bonds that are still pending.

7.3.4.2 Cluster-assisted mechanism

Negatively charged V_{Ga} can also diffuse on the Ga sublattice through a mechanism that we call the *cluster-assisted path* that, as far as we know, had not been reported until now. Instead of crossing directly the diffusion plane, Ga^{2nd} approaches the interstitial region far from the plane, being assisted by two As atoms— respectively first (As^{1st}) and third neighbors (As^{3rd}) of the vacancy— and one gallium atom second neighbor of the vacancy Ga^{2nd} . From the initial to the transition state (dashed arrows on Figure 7.6), an incomplete bond exchange mechanism of type Wooten-Winer-Weaire [35] —which we also find in Si— occurs between Ga^{2nd} and As^{3rd} , then Ga^{2nd} is pushed into the interstitial region. No cluster-assisted path was found at the neutral state. We tried to generate this event by activating with SIEST-A-RT starting from the transition state at -1 charge, but, the neutral vacancy relaxes back to the initial minimum rather than to diffuse to the second neighbor site demonstrating that this path is impossible at the neutral charge state.

The cluster formed by $As^{1st} + I_{Ga}^{2nd} + As^{3rd} + Ga^{2nd}$ plays the major role for diffusion since the bond distances $I_{Ga}^{2nd} - As^{1st}$ and $I_{Ga}^{2nd} - As^{3rd}$ remain unchanged (2.4 Å) regardless of the charge state, while the $V_{Ga}^q - I_{Ga}^{2nd}$ distance increases by adding extra electrons (see Table 7.2). This assumes that the whole cluster becomes less engaged toward the vacancy when passing from -1 to -3 charge states. During the relaxation from the saddle point to the final state (full arrow on Figure 7.6) only the gallium atom at the interstitial position (I_{Ga}^{2nd}) continues its motion toward the vacancy leaving the remaining constituents of the cluster close to their initial positions. This is confirmed by looking at the full migration trajectory for -3 charge state for the cluster-assisted mechanism plotted in Figure 7.7. From the initial to the saddle point many atoms are experiencing rearrangements and displacements, but the main contribution comes from the diffusing atom which is less engaged toward the vacancy than for plane-passing mechanism, while during the relaxation from the saddle point to the final state the main contribution to the total path comes from the relaxation of I_{Ga}^{2nd} atom (Figure 7.7(b)).

Mulliken population analysis shows that the electronic charge diffusion for cluster-

Table 7.3: Calculated diffusion barriers (in eV) for $V_{Ga}^{0,-1,-2,-3}$ in GaAs for all possible migration paths identified. The empty cell means that the migration was not possible with this mechanism.

	V_{Ga}^0	V_{Ga}^{-1}	V_{Ga}^{-2}	V_{Ga}^{-3}
First neighbor	0.84	0.90	1.86	
Plane-passing	1.7	1.7	1.85	2.0
Cluster-assisted		2.44	2.89	3.24
Fourth neighbor		4.24	4.24	4.3

assisted mechanism is completely different from the plane-passing mechanism. By increasing the charge state of the vacancy from -1 to -3 , positions of the cluster atoms are all shifted to the right confirming that the whole cluster is experiencing a displacement far away from the vacancy. In addition, the electronic charge becomes more and more concentrated around the cluster as the charge on the defect increases. Consequently, the migration barrier increases in significant way with the charge state reaching 2.44, 2.89 to 3.24 eV for -1 , -2 , -3 charges respectively.

The summary of the calculated migration barriers is presented in Table 7.3: First neighbor diffusion has the lowest barrier for 0 and -1 charges, but cannot be considered to be dominant for V_{Ga} since the complete diffusion process is impossible to achieve even for the neutral vacancy. However, this mechanism induces the formation of defects belonging to the As_{Ga} family of great importance since they are responsible of EL type defects in GaAs (see the discussion below).

Self-diffusion will rather be dominated to jumps to the same sublattice. Gallium vacancies can diffuse to the second neighbour through the plane-passing mechanism for all charge states by crossing barriers lower than 2 eV. The impact of the charge state on the trajectory is reflected on the charge-dependent barrier-height energy, for example, the barrier height for V_{Ga}^{-3} is 18 % higher than for V_{Ga}^0 , making the crossing 30 times

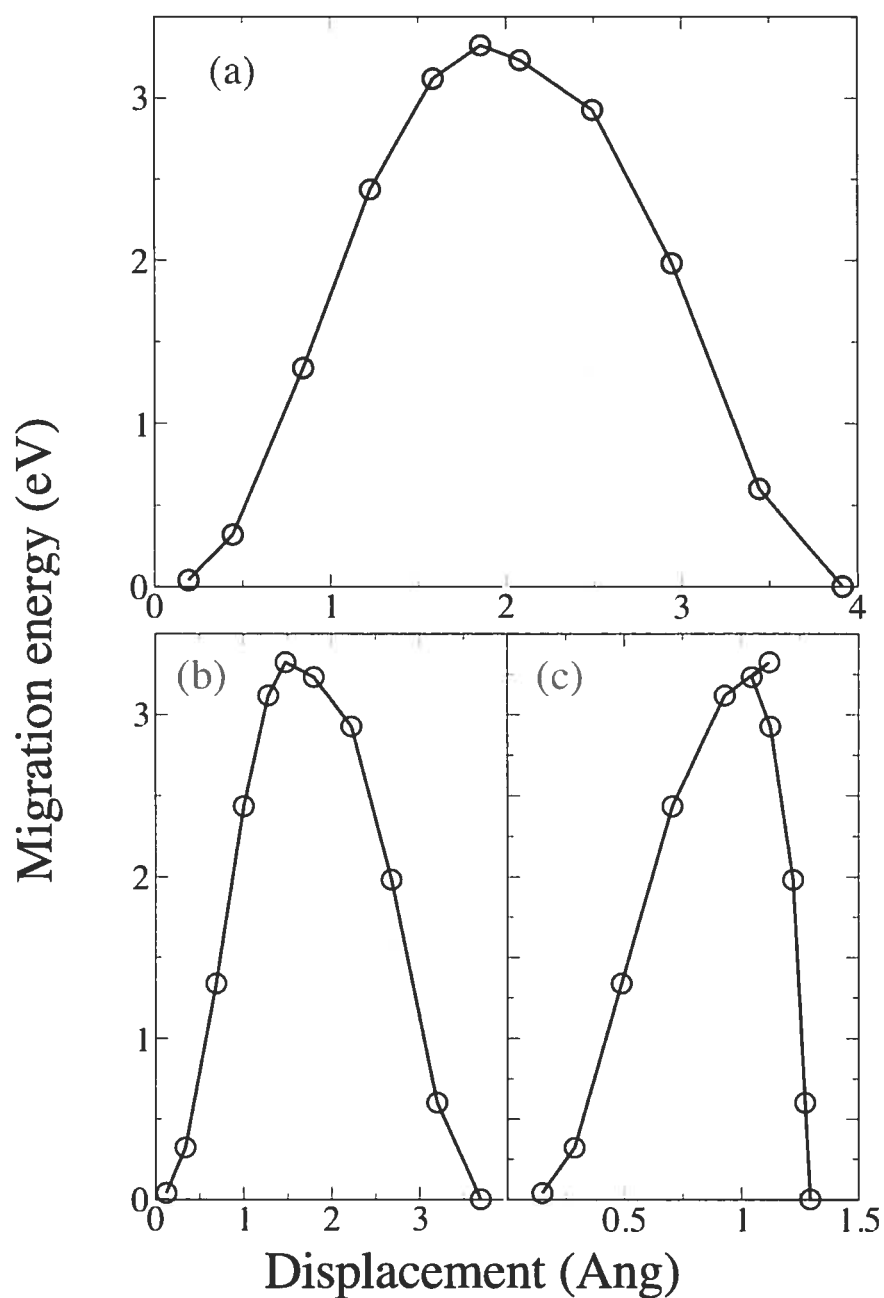


Figure 7.7: (a) Migration trajectory to the second neighbor by the cluster-assisted mechanism for -3 charge state. 11 images (circles) are relaxed with force tolerance of $1\text{eV}/\text{\AA}$, lines are guide to the eye. The contributions from different moving atoms are decomposed in two: (b) The path followed by the diffusing atom; (c) Arrangement of the other atoms around the defect.

less likely at 1000 K. The simplicity of this path as well as the relatively low migration barrier make the plane-passing mechanism the potential candidate to mediate vacancy self-diffusion in GaAs.

Negatively charged vacancies might also follow the cluster-assisted pathway to diffuse to the second neighbor or diffuse directly to the fourth neighbor with energy-barrier which is, by about 50 % and 100 % respectively, higher than that of the plane-passing mechanism. Although these two mechanisms have elevated barriers, their existence is by itself a very interesting finding showing the underestimated richness of self-diffusion phenomena in bulk semiconductors.

7.4 Discussion and comparisons

7.4.1 Self-diffusion to first-neighboring sites

In a binary lattice, vacancy hops to the first-neighbor sites require a complex sequence of moves to preserve chemical ordering on the long run. In 1984, Van Vechten [31] proposed a model leading to vacancy diffusion to the second neighbor in GaAs. His model describes a 11 first-neighbor hop process on a six membered ring and is based on two assumptions: (i) jumps to the first neighbors are always possible, and (ii) vacancy-antisite complexes are always stable. During successive hops the vacancy should proceed leaving behind a chain of antisites, which is energetically unfavorable. This is avoided by the vacancy passing twice through the same sixfold ring. At the final stage all the antisite defects are removed. Otherwise such a mechanism would create an unfavorable excess of antisite defects beyond the equilibrium concentration. Bockstedte and Schefler [11] studied the validity of these assumptions —possible diffusion via first neighbor hops— using the drag method and LDA. They found that $As_{Ga} + V_{As}^{1st}$ is metastable in the neutral state while it is unstable for other negative charge states for a 64-GaAs atoms system. They concluded that diffusion by first nearest neighbor hops was impossible.

Our study shows that the complex is rather metastable in -1 charge state too. The first nearest neighbor diffusion of V_{Ga} in -1 and -2 charges is found to be possible via *deformed* structures belonging to As_{Ga} family. While V_{Ga} diffusion to the first neighbor

lead to a symmetric $As_{Ga} - V_{As}$ complex in the neutral state, this configuration is distorted in the case of -1 charge states. In the deformed structure for -1 charge state, As atom do not occupy the tetrahedral gallium vacancy site, it is rather is displaced from the Ga vacancy site by about 0.9 eV toward V_{As}^{1st} . Interestingly, this deformed structure is no longer metastable for -2 , the transition state originates from a $As_{Ga} + Ga_{As} + V_{Ga}$ complex, with Ga_{As} pushed toward V_{Ga} and occupying a split vacancy site.

We attribute this charge-dependent first-neighbor diffusion to two competing factors: (1) the progressive increase in the strength of $As^{1st} - Ga^{2nd}$ bonds as electrons are added make them more difficult to break, and (2) the electron density in the region surrounding the vacancy —especially on As dangling bonds— increase as electrons are added to the relaxed system. Thus, Coulomb repulsion between three As^{1st} atoms and the diffusing As^{1st} atom can become so strong that this atom cannot approach further the vacant site. This picture causes the As^{1st} atom to relax at the split vacancy configuration in -1 charge state. If the system is charged -2 , the $As^{1st} - Ga^{2nd}$ bond is so strong that As^{1st} pulls Ga^{2nd} with it during its diffusion to the saddle point position leaving behind V_{As}^{1st} and V_{Ga}^{2nd} . The electronic density is partially transferred from the initial vacancy to these two new vacancies allowing As^{1st} to relax on the initial vacant site and leaving the Ga^{2nd} atom stacked between V_{As}^{1st} and V_{Ga}^{2nd} .

Similarly deformed structures for negatively charged defects have been recently reported in the literature. Chadi [36] found that the isolated As antisite structure (As_{Ga} is generally accepted to be the basic structure of EL2 defect) exists in charge states $+2$, $+1$ and 0 occupying the tetrahedrally symmetric position and in -1 , -2 charge states when it undergoes a small displacement that causes a deviation from T_d symmetry. The main difference between different charge states is the degree of relaxation that the antisite and its direct neighbor undergo. This possible high negative charging of As_{Ga} inducing a structural relaxation is similar to the deformed negatively charged $As_{Ga} - X$ observed in this work.

Moreover, the electronic structure of $As_{Ga} - V_{As}$ has been studied by total-energy Green's-function calculations treating many-body effects within LSDA-DF [37] and shows that the electronic levels allow at most -1 charging for this complex which is

in agreement with our results.

On the experimental side, negatively charged $V_{As} - X$ complexes have been recently detected by positron annihilation experiments [39] between 20–330 K in weakly p-type GaAs under arsenic rich condition. The detection of $V_{As} - X$ complexes in these samples was surprising and have never been reported before since under these conditions [28] isolated arsenic vacancies are unlikely to form. V_{Ga} diffusion to the first neighbor after annealing could explain these finding assuming that the metastable complex $As_{Ga} - V_{As}$ could be detected at sufficiently low temperatures.

In addition, $As_{Ga} - V_{As}$ structure was proposed by Steinegger *et al.* [38] to be a potential candidate for EL6 defect. Measuring the relative concentration of EL6 by photon-induced current-transient spectroscopy at room temperature, they observed an increase of As_{Ga} concentration with annealing time by roughly a factor of 2 corresponding to a decrease of V_{Ga} concentration from about 10^{16} to zero. While the first nearest-neighbor diffusion cannot play a role in self-diffusion, we suggest that it is dominant in these conditions, leading to the transformation of a large number of V_{Ga} into $As_{Ga} - V_{As}$ complexes.

7.4.2 Diffusion to second neighbor

Nevertheless, diffusion to the second neighbor is more interesting since it conserves the equilibrium concentration of vacancies. The plane-passing paths of V_{Ga} simulated with SIEST-A-RT confirm that the transition state does cross the diffusion plane in all charge states. For all our simulations the As^{1st} atom crosses the plane before Ga^{2nd} atom; we did not record any event where As^{1st} crossed out the plane or where it was located on the plane as suggested by Bockstedte and Scheffler [11]. However, the presence of the extra charge affects the diffusion trajectory by increasing the distance between the moving atom (Ga^{2nd}) and the three As^{1st} dangling bonds. For V_{Ga}^{-2} and V_{Ga}^{-3} the charge on As^{1st} is so strong that it scatters the moving atom and pushes it away from the vacancy toward the plane. Consequently, the distance between the plane and the atom is lowered as the charge of the vacancy increases, it almost vanishes for -3 charge state.

The diffusion barrier of plane-passing mechanism is found to *increases* moderately

as extra electrons are added. Recent theoretical works reported a charge-dependent migration to the second neighbor by plane-passing mechanism for vacancies in SiC and GaN binary semiconductors. Bockstedte *et al.* [12] found that the migration barrier for diffusion V_C and V_{Si} in 4C-Si *decreases* when the vacancy charges are *increasing* progressively. A similar trend was observed for V_N in GaN [14] and recently for V_{Ga} in GaN [40]. Our most recent results with SIEST-A-RT concerning V_{As} in GaAs [41] suggest that the migration barrier by plane-passing mechanism *decreases* by *increasing* progressively vacancy charging. Thus, V_{Ga} in GaAs shows an opposite trend compared to previously studied vacancies. This trend was observed previously by Bockstedte and Scheffler [11] as they found a migration barrier of 1.7 eV for neutral and 1.9 eV for -3 charge state. This behavior cannot be a drawback of SIEST-A-RT, since the total path relaxed using CI-NEBM lead to the same barrier, it is rather correlated to the electronic charge diffusion of V_{Ga} in GaAs.

Experimentally, negatively charged gallium vacancy migration on GaAs(110) surface have been found to be stimulated by the tip during STM experiments [42] only when *holes* are injected on the surface reducing consequently the negative charging of the vacancy. This supports our results suggesting that the migration barrier is *lowered* as the vacancy becomes less negatively charged.

7.4.3 Diffusion in experimental systems

Bliss *et al.* [43] identified the migration barriers for V_{Ga} in LT-GaAs by positron annihilation technique. They found a migration enthalpy of 1.5 ± 0.3 eV for the vacancy diffusion to the second neighbor and 1.1 ± 0.3 eV for diffusion to the first neighbor. Within the experimental error bars, these results agree with our calculated barriers summarized in Table 7.3. Our results are also consistent with the widely accepted value of 1.7-1.8 eV for GaAs vacancy migration in bulk and GaAs-based QW materials [44–48]. This value has been recently reported for inter-diffusion in InGaAs/GaAs quantum dots [19] as well. Thus our calculated diffusion barriers are accurate enough and could be extended to analyze experimental data in bulk and even in nanostructured materials.

Our extensive study is very useful for interpreting more accurately experimental

data, especially if taking into account the fact that most experiments for gallium self-diffusion [49, 50] have been performed in n-type or intrinsic material where negatively charged vacancies are more abundant. The most relevant pathway for self-diffusion is the plane-passing mechanism since it has the lowest barrier for diffusion to the second neighbor, while at sufficiently low temperatures diffusion to the first neighbor could be observed for -1 and -2 charge states. By combining information coming from our calculated first and second neighbor diffusion barrier it becomes possible to identify the charge state of the diffusing vacancy in some experiments. For example, in the experiment of Bliss *et al.* [43], V_{Ga} diffuses to the first neighbor by crossing a barrier of 1.1 ± 0.3 eV, which is close to our calculated barrier of 0.84 eV for V_{Ga}^{-1} . In addition, measured diffusion barrier to the second neighbor is 1.5 ± 0.3 eV which is comparable to 1.7 eV we calculated for -1 charge state. Thus, we can state that diffusing vacancies in this experiment are more probably singly negative instead of the generally accepted triply negative. Recent research works support more and more the hypothesis of -1 charge state. This was the case, for example, in a recent theoretical study comparing simultaneously the relaxations and lifetimes obtained from simulation and Doppler spectra of positron annihilation experiments support the possibility of -1 charge state for V_{Ga} rather than the -3 charge state [30].

7.5 Conclusions

The SIEST-A-RT approach has been used to study various diffusion pathways for charged gallium vacancies in GaAs demonstrating that diffusion in bulk semiconductors is a rich and complex phenomenon closely related to the charge state. Novel diffusion pathways are identified for negatively charged vacancies, showing that diffusion process could be non intuitive and more complex than initially thought. V_{Ga} can diffuse to the second neighbor using two different mechanisms in addition to diffusion to the first and fourth neighbor. Diffusion to the second neighbor is possible by the *plane-passing mechanism* at all charge states as well as by the newly found *cluster-assisted mechanism* which becomes possible only in the presence of negatively charged vacancies.

In addition, gallium vacancy diffusion to the first nearest neighbor is possible only for $q = 0, -1, -2$ charges passing by regular or distorted $As_{Ga} - X$ defects. However this mechanism cannot be considered to be dominant for V_{Ga} since the complete diffusion process was impossible to achieve. Finally, the highest diffusion barrier was recorded for the direct diffusion to the fourth neighbor along the 100 direction for negatively charged vacancies. All barriers for the migration pathways of V_{Ga} are found to increase with increasing the number of electron, a behaviour opposite to what was recently found in the case of vacancy-mediated self-diffusion in SiC [12, 13] and GaN [14, 40].

Acknowledgements. We thank Prof. Michel Côté and Vladimir Timochevski for fruitful discussions. NM acknowledges partial support from FQRNT (Québec), NSERC (Canada) and the Canada Chair program. We thank the RQCHP for generous allocation of computer time.

BIBLIOGRAPHY

- [1] Z. M. Wang, S. Seydmohamadi, J. H. Lee, and G. J. Salamo, *Appl. Phys. Lett.* **85**, 5031 (2004).
- [2] Z. M. Wang, K. Holmes, J. L. Shultz, and G. J. Salamo, *Phys. stat. sol. (a)* **202**, R85 (2005).
- [3] A. Rastelli, S. Stufler, A. Schliwa, R. Songmuang, C. Manzano, G. Costantini, K. Kern, A. Zrenner, D. Bimberg, and O. G. Schmidt, *Phys. Rev. Lett.* **92**, 166104 (2004).
- [4] H. Yu and W. Lu, *Acta Materialia* **53**, 1799 (2005).
- [5] P. J. Feibelman, *Phys. Rev. Lett.* **65**, 729 (1990).
- [6] G. Henkelman, G. Jóhannesson, and H. Jónsson, in *Progress on Theoretical Chemistry and Physics*, edited by S. D. Schwartz (Kluwer Academic Publishers, 2000), pp. 269–300.
- [7] M. Cogoni, B. P. Uberuaga, A. F. Voter, and L. Colombo, *Phys. Rev. B* **71**, 121203(R) (2005).
- [8] Y. A. Du, S. A. Barr, K. R. A. Hazzard, T. J. Lenosky, R. G. Hennig, and J. W. Wilkins, *Phys. Rev. B* **72**, 241306(R) (2005).
- [9] N. Mousseau and G. T. Barkema, *Phys. Rev. E* **57**, 2419 (1998).
- [10] T. F. Middleton and D. J. Wales, *Phys. Rev. B* **64**, 024205 (2001).
- [11] M. Bockstedte and M. Scheffler, *Z. Phy. Chem. (munich)* **200**, 195 (1997).
- [12] M. Bockstedte, A. Mattausch, and O. Pankratov, *Phys. Rev. B* **68**, 205201 (2003).
- [13] R. Rurali, E. Hernández, P. Godignon, J. Rebollo, and P. Ordejón, *Comp. Mat. Sc.* **27**, 36 (2003).

- [14] S. Limpijumnong and C. G. Van de Walle, *Phys. Rev. B* **69**, 035207 (2004).
- [15] I. V. Ionova and E. A. Carter, *J. Chem. Phys.* **98**, 6377 (1993).
- [16] C. Corbel, F. Pierre, K. Saarinen, , P. Hautojärvi, and P. Moser, *Phys. Rev. B* **45**, 3386 (1992).
- [17] H. Bracht and S. Brotzmann, *Phys. Rev. B* **71**, 115216 (2005).
- [18] F. Tuomisto, K. Pennanen, and K. Saarinen, J. Sadowski, *Phys. Rev. Lett.* **93**, 055505 (2004).
- [19] H. S. Djie, O. Gunawan, D.-N. Wang, B. S. Ooi, and J. C. M. Hwang, *Phys. Rev. B* **73**, 155324 (2006).
- [20] T. Raz, N. Shuall, G. Bahir, D. Rittera, D. Gershoni, and S. N. G. Chu, *Appl. Phys. Lett.* **85**, 3578 (2004).
- [21] D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, *Int. J. Quant. Chem.* **65**, 453 (1997).
- [22] R. Malek and N. Mousseau, *Phys. Rev. E* **62**, 7723 (2000).
- [23] F. Valiquette and N. Mousseau, *Phys. Rev. B* **68**, 125209:1 (2003).
- [24] G. Wei, N. Mousseau, and P. Derreumaux, *J. Chem. Phys.* **117**, 11379 (2002).
- [25] F. El-Mellouhi, N. Mousseau, and P. Ordejón, *Phys. Rev. B* **70**, 205202 (2004).
- [26] N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- [27] L. Kleinman and D. M. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).
- [28] F. El-Mellouhi and N. Mousseau, *Phys. Rev. B* **71**, 125207 (2005).
- [29] D. J. Chadi, *Materials Science in Semiconductors Processing* **6**, 281 (2003).
- [30] I. Makkonen, M. Hakala, and M. J. Puska, *Phys. Rev. B* **73**, 035103 (2006).

- [31] J. A. Van Vechten, *J. Phys. C: Solid State Phys.* **17**, L933 (1984).
- [32] P. E. Blöchl, E. Smargiassi, R. Car, D. B. Laks, W. Andreoni, and S. T. Pantelides, *Phys. Rev. Lett.* **70**, 2435 (1993).
- [33] J. Dabrowski, *Solid State phenomena* **71**, 23 (2000).
- [34] G. Henkelman, B. P. Uberuaga, and H. Jónsson, *J. Chem. Phys.* **113**, 9901 (2000).
- [35] F. Wooten, K. Winer, and D. Weaire, *Phys. Rev. Lett.* **54**, 1392 (1985).
- [36] D. J. Chadi, *Phys. Rev. B* **68**, 193204 (2003).
- [37] H. Overhof and J.-M. Spaeth, *Phys. Rev. B* **72**, 115205 (2005).
- [38] T. Steinegger, B. Gründig-Wendrock, M. Jurisch, and J. R. Niklas, *Physica B* **308-310**, 745 (2001).
- [39] V. Bondarenko, J. Gebauer, F. Redman, and R. Krause-Rehberg, *Appl. Phys. Lett.* **87**, 161906 (2005).
- [40] M. G. Ganchenkova and R. M. Nieminen, *Phys. Rev. Lett.* **96**, 196402 (2006).
- [41] F. El-Mellouhi and N. Mousseau, in preparation.
- [42] G. Lengel, J. Harper, and M. Weimer, *Phys. Rev. Lett.* **76**, 4725 (1996).
- [43] D. E. Bliss, W. Walukiewicz, and E. E. Haller, *Journal of Electronic Materials* **22**, 1401 (1993).
- [44] J. L. Rouviere, Y. Kim, J. Cunningham, J. A. Rentschler, A. Bourret, and A. Ourmazd, *Phys. Rev. Lett.* **68**, 2798 (1992).
- [45] R. Geursen, I. Lahiri, M. Dinu, M. R. Melloch, and D. D. Nolte, *Phys. Rev. B* **60**, 10926 (1999).
- [46] D. E. Bliss, W. Walukiewicz, J. W. Ager, E. E. Haller, K. T. Chan, and S. Tanigawa, *J. Appl. Phys.* **71**, 1699 (1992).

- [47] I. Lahiri, D. D. Nolte, M. R. Melloch, J. Woodall, and W. Walukiewicz, *Appl. Phys. Lett.* **69**, 239 (1996).
- [48] S. Balasubramanian, S. W. Mansour, M. R. Melloch, and D. D. Nolte, *Phys. Rev. B* **63**, 033305 (2000).
- [49] B. Bracht, M. Norseng, E. E. Haller, K. Eberl, and M. Cardona, *Sol. Stat. com.* **112**, 301 (1999).
- [50] J. Gebauer, M. Lausmann, F. Redmann, R. Krause-Rehberg, H. S. Leipner, E. R. Weber, and P. Ebert, *Phys. Rev. B* **67**, 235207 (2003).

CHAPITRE 8

EFFET DE LA TEMPÉRATURE SUR LE RENVERSEMENT DE L'ÉTAT DE CHARGE DES LACUNES DE GALLIUM DANS LE GAAS

Ce chapitre traite l'effet de température sur la stabilité des lacunes de gallium chargées dans le GaAs. Les effets de température sont traités par l'article intitulé *Thermally-activated charge reversibility of gallium vacancies in GaAs* et accepté pour publication dans *Journal of Applied Physics* en date du 2 Mai 2006.

Il y a un long débat concernant l'état de charge de la lacune dominante dans le GaAs. La simulation numérique propose que la lacune V_{Ga}^{-3} soit dominante pour le GaAs de type N ou intrinsèque. Certaines expériences, faites à haute température, proposent la charge -3 alors que d'autres en proposent les charges -2 et -1 . Ce débat, entre les différents travaux expérimentaux, a persisté jusqu'à récemment. Nous avons résolu ce problème en introduisant l'effet de la température sur l'énergie libre de Gibbs en plus de l'effet de température sur le niveau de Fermi. Nous avons trouvé que ces effets de température sont si importants qu'ils renversent et changent l'état de charge de la lacune dominante. V_{Ga} préfère diminuer progressivement son état de charge au fur et à mesure que la température augmente. Nos profils de concentration calculés à l'aide du modèle de concentration peuvent bien ajuster les courbes expérimentales de Gebauer *et al.* [2] en même temps que celles de Bracht *et al.* [1]. Ce résultat montre que même si les effets de température sur les niveaux d'ionisation sont faibles ; une fois combinés avec l'effet du niveau de Fermi, ils deviennent pertinents afin de bien ajuster les données expérimentales.

Dans ce qui suit, nous expliquerons la méthode utilisée par les deux groupes afin d'ajuster leurs résultats expérimentaux. En second lieu, nous analyserons les ajustements de Gebauer *et al.* et de Bracht *et al.* pour trouver la source du débat sur l'état de charge de V_{Ga} .

Thermally-activated charge reversibility of gallium vacancies in GaAs

Fedwa El-Mellouhi¹ and Normand Mousseau^{1,2}

¹ Département de physique and Regroupement québécois sur les matériaux de pointe, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal (Québec) H3C 3J7, Canada

² Service de Recherches de Métallurgie Physique, Commissariat à l'énergie atomique-Saclay, 91191 Gif-sur-Yvette, France.

PACS numbers: 65.40.Gr, 61.72.Ji, 71.15.Mb

Abstract

The dominant charge state for the Ga vacancy in GaAs has been the subject of a long debate, with experiments proposing -1 , -2 or -3 as the best answer. We revisit this problem using *ab initio* calculations to compute the effects of temperature on the Gibbs free energy of formation, and we find that the thermal dependence of the Fermi level and of the ionization levels lead to a reversal of the preferred charge state as the temperature increases. Calculating the concentrations of gallium vacancies based on these results, we reproduce two conflicting experimental measurements, showing that these can be understood from a single set of coherent LDA results when thermal effects are included.

8.1 Introduction

There has been a long-standing controversy about the nature of the charge state of the gallium vacancy (V_{Ga}^q) mediating self-diffusion in GaAs. While experiments have deduced a preferred charge state ranging from -1 and -2 [1] to -3 [2, 3] under N-type or intrinsic conditions, most (zero temperature) *ab-initio* calculations favor the -3

charge state for V_{Ga} , see Ref. [4–14]. Extracting the charge state from experiments is challenging: these are usually performed at high temperature, in a regime where the formation entropy becomes a significant part of the Gibbs free energy. These entropic contributions can therefore alter the charge of the dominant species as calculated at $T = 0$, leading to a strong competition between the charge states. Moreover, the separate ionization energy for each state is not directly accessible experimentally; only the sum over the ionization free energies can be obtained by fitting the experimental data over a range of temperatures, using a number of assumptions typically inspired from *ab-initio* data [2, 3] or set as free parameters in the fit [1].

Since the ionization free energies are key for fitting experimental data, it is important to evaluate them from *ab-initio* calculations. Because of the high computational cost associated with computing vibrational spectra [15], however, various assumptions are made typically to simplify the problem. For example, it is usually assumed that the free energy of ionization is independent of temperature and that ionization levels are constant with respect to the valence band edge so that the formation entropy of all charge defects is taken as equal to that of V_{Ga}^0 [16]. Often, it is also supposed that ionization energies follow the thermal behavior of band gap (or the conduction band minimum) [17], which results typically into shallower ionization levels with increasing temperature (within this approximation, for example, heavily charged defects in Si become favored at high T [18]).

In this article, we revisit the question of the dominant charge state of V_{Ga} in GaAs, calculating directly, using LDA, the ionization Gibbs free energies ($F^{q1/q2}$) for all possible charge state transitions, using an approach applied to SiC [15] and Si [19]. We show that, in spite of well known LDA limitations, apparently conflicting data coming from two experiments performed under different conditions [1, 2] are well fitted. This satisfactory fit is obtained for both experimental data by using the same set of inputs like the concentration model, the Fermi level effect and the computed temperature dependence of $F^{q1/q2}$.

This article is organized as follows: calculation details are presented in section 8.2, next we analyze the source of conflict that arose between different experiments in sec-

tion 8.3, presenting at the same time the common concentration and diffusion models used. Section 8.4 contains a discussion of phonon calculation results followed by our fitting procedure and its application on two experimental data. The main conclusions are presented in section 8.5.

8.2 Computational details

From the computational point of view, the stability of V_{Ga}^q is determined by comparing total energies of supercells with and without defects. In a previous ground state *ab-initio* calculation using SIESTA, a local-basis LDA code [4], we computed the ionization levels of gallium vacancies for 0, -1 , -2 , -3 charges.

Here, we include the thermal contributions by computing the Gibbs free energies in the harmonic limit. Note that entropy changes in bulk Ga and As metals is accounted for (see Refs. [15, 20] as well as Ref. [4] for more details):

$$G = E_{tot} + U_{vib} - TS_{vib} + p.V + \sum_i \mu_i N_i \quad (8.1)$$

where, at constant number of particles and volume, the vibrational energy and entropy are given, respectively, by

$$U_{vib} = \sum_{i=1}^{3N} \left\{ \frac{\hbar\omega_i}{\exp(\hbar\omega_i/k_B T) - 1} + \frac{1}{2}\hbar\omega_i \right\} \quad (8.2)$$

and

$$S_{vib} = k_B \sum_{i=1}^{3N} \left\{ \frac{\hbar\omega_i}{k_B T} \left[\exp\left(\frac{\hbar\omega_i}{k_B T}\right) - 1 \right]^{-1} - \ln \left[1 - \exp\left(\frac{-\hbar\omega_i}{k_B T}\right) \right] \right\}. \quad (8.3)$$

To obtain the Gibbs free energy, the phonon spectrum (ω_i) is computed for all stable charge states of V_{Ga}^q using a numerical derivative to construct the dynamical matrix.

The calculation of the dynamical matrix of cells with 215 sites is a very time-consuming task. Since many atomic shells around the vacancy are distorted it is not possible to reduce the number of calculations by symmetry operations. We thus use the full dynamical matrix for the complete supercell. A mesh cutoff of 120 Ry is used and

the Brillouin zone of the 215-atom cubic supercell is sampled at the Γ -point. The eigenfrequencies are then extracted at 146 k-points in the Brillouin zone [19].

Except for the splitting of the optical branches at the Γ -point, which cannot be obtained in the harmonic approximation, the total phonon density of states and the dispersion curve obtained by this procedure for the GaAs crystal are in good agreement with experiment and calculations [21, 22]. Since we are interested here in the entropy difference and not its absolute value, these discrepancies should not play an important role.

8.3 Analysis of the experimental results

Before discussing our analysis, we first focus on the source of discrepancy that exists in explaining different experiments. Although the diffusion model and the experimental Fermi level dependence used in different experiments [1, 2] are the same, conclusions about the dominant charge state of V_{Ga} are contradictory. This contradiction can therefore be due to experimental errors or the additional assumptions about the temperature dependence of the ionization levels used in the experimental fits. The following analysis shows that the choice of the ionization levels is an important factor that suffices to explain the conflict. Some calculation details are skipped in the manuscript, we refer the reader to Refs. [1, 3] for further details on the model used.

8.3.1 Concentration model

Both fitting procedures used by Gebauer *et al.* [2] and Bracht *et al.* [1] are based on the concentration model used previously by Tan and coworkers [3]. The Ga self-diffusion coefficient D_{Ga} is given by the sum of the transport coefficients of vacancies in various charge states q :

$$D_{Ga} = \frac{1}{C_0} \sum_{q=0}^3 f_q C_{V_{Ga}^{-q}}^{eq} D_{V_{Ga}^{-q}} \quad (8.4)$$

where, C_0 represents the Ga atom density in GaAs ($C_0 = 2.215 \times 10^{22} \text{cm}^{-3}$), f_q is the diffusion correlation factor that contains information about the microscopic jump mechanism, $C_{V_{Ga}^{-q}}^{eq}$ is the thermal equilibrium concentration and $D_{V_{Ga}^{-q}}$ is the diffusion coefficient of the vacancy V_{Ga}^{-q} for $q = 0, 1, 2, 3$.

The thermal equilibrium concentration of V_{Ga}^{-q} in GaAs is defined as

$$C_{V_{Ga}^{-q}}^{eq} = C^* P_{As_4}^{1/4} T^{-5/8} \exp\left(-\frac{G_{V_{Ga}^{-q}}^f}{k_B T}\right), \quad (8.5)$$

where C^* is a pre-exponential factor, $G_{V_{Ga}^{-q}}^f$ is the Gibbs free energy of formation of V_{Ga}^{-q} , T is temperature and k_B is the Boltzmann constant. This equation takes into account the influence of As_4 over-pressure (P_{As_4}), often used in experiments, on vacancy concentrations. If the arsenic over-pressure increases, it affects the stoichiometry of gallium arsenide by introducing an excess of As atoms. With this excess of As atoms, GaAs sample can become As-rich and the concentrations of Ga vacancies in GaAs are enhanced.

Different vacancies can introduce energy levels within the energy band gap of GaAs. The occupation of these energy states depends on the position of the Fermi level. Under extrinsic conditions, when the hole or the electron concentration introduced by doping exceeds the intrinsic carrier concentration, the Fermi level deviates from its intrinsic position. As a consequence, the ratio of the charged to neutral vacancy concentrations is changed. Vacancies can introduce acceptor levels, i.e., V_{Ga}^{-1} introduces one level located at $E_{V_{Ga}^{-1}}$ above the valence band edge, V_{Ga}^{-2} introduces one at $E_{V_{Ga}^{-2}}$ and V_{Ga}^{-3} introduces one at $E_{V_{Ga}^{-3}}$ above the valence band edge.

The concentration profiles measured experimentally contain contributions from all active vacancy charge states that cannot be separated. We take advantage of the fact that the Gibbs free energy of formation — and thus the equilibrium concentration $C_{V_{Ga}^0}^{eq}$ — of neutral vacancies do not depend on the position of the extrinsic Fermi level. Therefore, we can express the total concentration as function of the ratio of the charged to neutral vacancy concentrations:

$$C_{V_{Ga}}^{eq} = \sum_{q=0}^3 C_{V_{Ga}^{-q}}^{eq} = C_{V_{Ga}^0}^{eq} \sum_{q=0}^3 \frac{C_{V_{Ga}^{-q}}^{eq}}{C_{V_{Ga}^0}^{eq}} \quad (8.6)$$

The ratios are given by:

$$\frac{C_{V_{Ga}^{-1}}^{eq}}{C_{V_{Ga}^0}^{eq}} = g_{V_{Ga}^{-1}} \exp\left(\frac{E_f - E_{V_{Ga}^{-1}}}{k_B T}\right) \quad (8.7)$$

$$\frac{C_{V_{Ga}^{-2}}^{eq}}{C_{V_{Ga}^0}^{eq}} = g_{V_{Ga}^{-2}} \exp\left(\frac{2E_f - E_{V_{Ga}^{-2}} - E_{V_{Ga}^{-1}}}{k_B T}\right) \quad (8.8)$$

$$\frac{C_{V_{Ga}^{-3}}^{eq}}{C_{V_{Ga}^0}^{eq}} = g_{V_{Ga}^{-3}} \exp\left(\frac{3E_f - E_{V_{Ga}^{-3}} - E_{V_{Ga}^{-2}} - E_{V_{Ga}^{-1}}}{k_B T}\right) \quad (8.9)$$

where $g_{V_{Ga}^{-q}}$ is approximated to unity (further explanations and intermediate steps used in this work can be found in Ref. [1]).

The concentration of negatively charged vacancies can be deduced if we know the position of the Fermi level E_f and the ionization energies $E_{V_{Ga}^{-q}}$. Ionization levels do not cause much problems as they are derived from available *ab-initio* calculation or included in the fit. However, the Fermi level depends strongly on intrinsic (n_i) and extrinsic carrier concentrations (n) by the relation:

$$\frac{n}{n_i} = \exp\left(\frac{E_f - E_f^i}{k_B T}\right) \quad (8.10)$$

Intrinsic properties of GaAs as function of temperature are taken from empirical data [23], like intrinsic band gap E_G , electron and hole concentrations (n_i and p_i) and intrinsic Fermi level E_f^i . The band gap of GaAs is supposed to vary with temperature [24], following Varshni's [25] empirical expression for GaAs $E_G(T) = 1.519 - 5.0405 * 10^{-4} T^2 / (T + 204)$ deduced from optical experiments [23]. The Fermi energy is the important quantity governing the equilibrium concentrations of defects. However, it changes if any electrically active impurities or dopants are present in the material as well as under the presence of charged vacancies. The concentrations of these impurities or dopants (C_{Donor} , $C_{Acceptor}$) and vacancy concentration in all charge states ($C_{V_{Ga}^{-q}}^{eq}$)

must therefore be taken into account. The charge balance equation, requiring that the free electron and hole concentrations n and p (which also depend on the Fermi level) must cancel out any net charge resulting from the concentrations of all positively and negatively charged defects and impurities:

$$n = \frac{1}{2} \left(C_{Donor} - \sum_{q=0}^3 q C_{V_{Ga}^{-q}}^{eq} \right) + \sqrt{n_i^2 + \frac{1}{4} \left(C_{Donor} - \sum_{q=0}^3 q C_{V_{Ga}^{-q}}^{eq} \right)^2} \quad (8.11)$$

$$\frac{n_i^2}{n} = p = \frac{1}{2} \left(C_{Acceptor} + \sum_{m=0}^3 m C_{V_{Ga}^{-m}}^{eq} \right) + \sqrt{n_i^2 + \frac{1}{4} \left(C_{Acceptor} + \sum_{m=0}^3 m C_{V_{Ga}^{-m}}^{eq} \right)^2} \quad (8.12)$$

The extrinsic Fermi level, for its part, is obtained by solving simultaneously Equations 8.10, 8.11, 8.7, 8.8 and 8.9 self-consistently using the Newton Raphson method [26]. Once the Fermi level is known, it is used to determine the final Gibbs free energy of formation and the corresponding equilibrium concentrations of all the defects present in the structure. Up to this point, the procedure is sufficient to reproduce the fit of Gebauer *et al.* and Bracht *et al.* shown in Fig. 7 of Ref. [2] and Fig. 9 of Ref. [1], respectively.

8.3.2 Diffusion model

Few steps are left to fit the measured diffusion coefficients of Bracht *et al.* (see the detailed discussion in Section 5.2 of Ref. [1]). Based on the theoretical work of Bockstedte and Scheffler [16], they assume that the entropy change associated with the V_{Ga}^{-q} migration and the correlation factor f_q are similar for all charge states, i.e., $D_{V_{Ga}^{-q}} \approx D_{V_{Ga}^{-1}}$ and ($f_q \approx f = 1$). In this approximation, Eq. 8.4 simplifies to:

$$D_{Ga} = \frac{1}{C_0} C_{V_{Ga}^0}^{eq} D_{V_{Ga}^0} \sum_{q=0}^3 \frac{C_{V_{Ga}^{-q}}^{eq}}{C_{V_{Ga}^0}^{eq}} \quad (8.13)$$

While the extrinsic diffusion coefficient ($D_{V_{Ga}}$) is unknown, it is, however, related to the intrinsic diffusion coefficient by the relation:

$$\frac{1}{C_0} C_{V_{Ga}}^{eq} D_{V_{Ga}} = \frac{D_{Ga}(n_i)}{1 + \sum_{q=1}^3 \exp \frac{qE_f^i - \sum_{q=1}^3 E_{V_{Ga}}^{-q}}{k_B T}} \quad (8.14)$$

The missing quantity in this equation is $D_{Ga}(n_i)$. Bracht *et al.* use the Ga diffusion coefficient they measured in **intrinsic** GaAs and find D_{Ga} , using Eqns 8.13, 8.14 and :

$$D_{Ga}(n_i) = 0.64 \exp \left(\frac{-3.71}{k_B T} \right) \text{cm}^2 \text{s}^{-1} \quad (8.15)$$

8.3.3 Fits of positron-annihilation measurements

From positron-annihilation experiments, Gebauer *et al.* [2] deduce the Gibbs free energy of formation and the dominant charge state of V_{Ga} by measuring directly vacancy concentrations. Measures were done at different high temperatures (900-1450K) by changing dopants concentrations and the chemical potential (stoichiometry).

The resulting concentration profiles were fitted using the model presented above [3] using *ab-initio* ionization levels calculated by Baraff et Schlüter [13] at 0K corresponding to $F^{0/-1} = 0.13E_G$, $F^{-1/-2} = 0.35E_G$ and $F^{-2/-3} = 0.49E_G$. These ionization levels are supposed to vary with temperature like the band gap following Varshni's [25] empirical relation.

Gebauer *et al.* [2] fitted the concentration profiles for N-GaAs with $[Te] = 2 \times 10^{18} \text{cm}^{-3}$ at arsenic over pressure of $P_{As_4} = 5.6 \text{ atm}$ for temperature ranging between 900 and 1450K. The resulting fit (corresponding to Fig. 7 of their article) is in reasonable agreement with measured concentration profiles especially at high temperature. It shows a negative temperature dependence that cannot be reproduced unless the -3 charge state dominates all other charge states.

Fig. 8.1 shows the variation of ionization levels used as function of temperature, the extrinsic Fermi level corresponding to the the dopant concentration used in this ex-

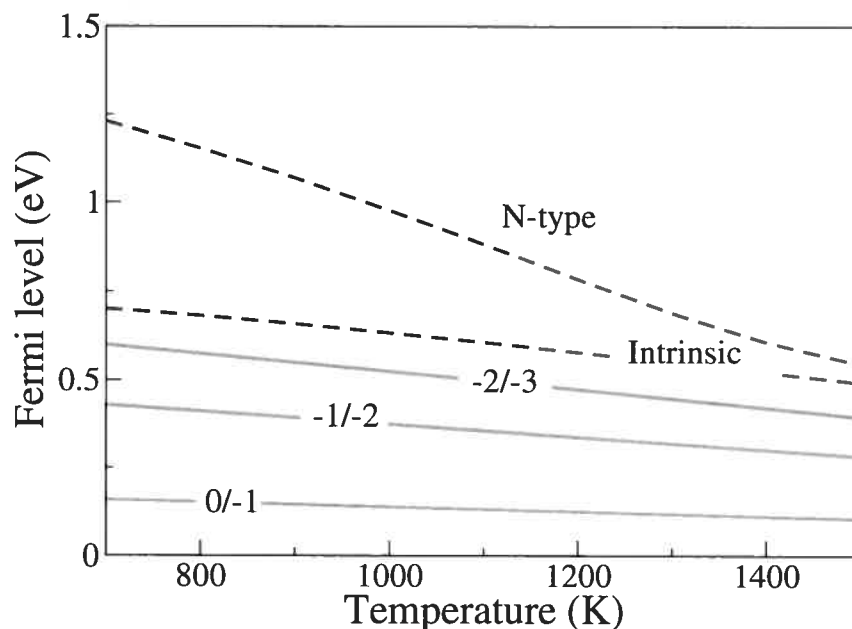


Figure 8.1: Ionization levels (solid lines) used by Gebauer *et al.* [2], they are taken to have a negative temperature dependence and vary like the experimental band gap. The position of the extrinsic Fermi level is also shown (dash line).

periment is also shown. We can see that this choice of the ionization levels favors the stability of the -3 charge state for N-type and intrinsic GaAs over a broad temperature range.

8.3.4 Fits of interdiffusion measurements

For their part, Bracht *et al.* [1], measured interdiffusion in isotope heterostructure $^{71}\text{GaAs}/^{\text{nat}}\text{GaAs}$ exposed to arsenic over pressure of $P_{\text{As}} = 1$ atm under intrinsic and extrinsic conditions (N-type with $[Si] = 3 \times 10^{18} \text{cm}^{-3}$ and P-type with $[Be] = 3 \times 10^{18} \text{cm}^{-3}$). They measured the temperature dependence of the Ga self-diffusion coefficient (D_{Ga}) at high temperature ranging between 1032K and 1350K.

The simultaneous fit of the three curves (N-type, P-type and intrinsic curves in Fig. 10 of Ref. [1]) is complicated and lengthy, especially that the ionization levels of Baraff

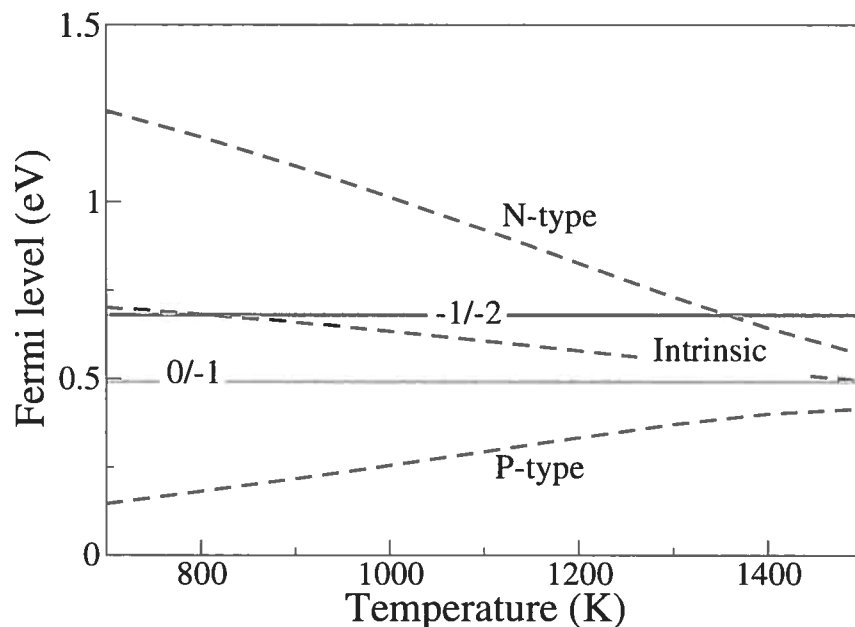


Figure 8.2: Ionization levels (solid lines) as obtained by Bracht *et al.* [1] from a fit of the experimental data in the temperature range 1032K-1350K and are assumed to remain constant. The position of the extrinsic Fermi level is also shown (dash line).

and Schlüter did not give a satisfactory fit. Therefore they include the ionization levels in the fitting procedure. The best fit obtained corresponds to $H_f = 1.9$ eV and includes two ionization levels only ($0/-1$ and $-1/-2$) located at $F^{0/-1} = 0.42$ eV and $F^{-1/-2} = 0.6$ eV from the VBM that must remain unchanged with temperature.

Figure 8.2 shows the variation of ionization levels used as function of temperature. The Fermi level corresponding to intrinsic, N-type and P-type GaAs is also shown. By analyzing this figure, one can notice that the -2 charge state dominates at low temperature, in intrinsic and N-type GaAs, then it is reversed to -1 and 0 as temperature increases. In this temperature range there is a competition between different charge states for intrinsic and N-type GaAs with no contribution coming from the -3 charge state. This choice of ionization level favors the stability of less charged vacancies as temperature increases. For P-type GaAs, the neutral vacancy dominates self-diffusion

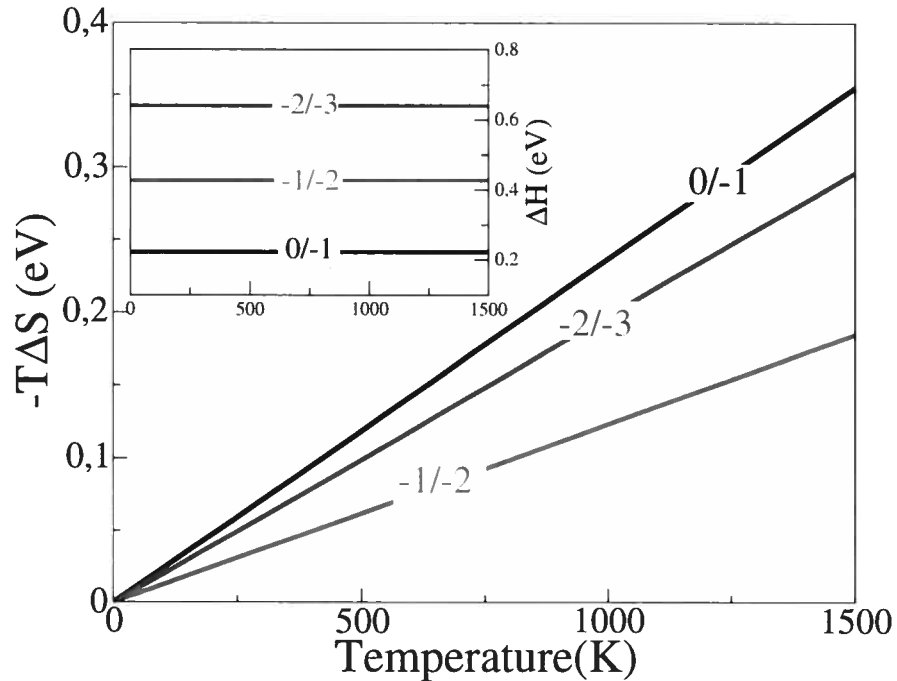


Figure 8.3: Entropy and enthalpy (inset) differences obtained after progressive ionization of V_{Ga}^q . While $\Delta H^{q1/q2}$ shows little temperature dependence, the non-vanishing entropy difference between charge states lead to a linearly increasing entropic term, $T\Delta S^{q1/q2} = T(S^{q2} - S^{q1})$, in the Gibbs free energy.

over the entire temperature range.

In both experiments the dominant part of the temperature dependence is due to the change of the chemical potential (i.e., the Fermi-level effect), treated extensively by Tan and coworkers [3]. However, the above analysis shows unequivocally that only including the experimentally-measured thermal dependence of the Fermi level alone is not sufficient to explain the two conflicting experiments at the same time. Figures 8.1 and 8.2 show the crucial role played by the ionization level behaviour for determining the charge state of the vacancy and that they are the main source of conflict.

8.4 Results and discussion

The calculated thermal effect on the enthalpy $\Delta H^{(q1/q2)}$ and the entropy differences $T\Delta S^{(q1/q2)}$ between the relevant sequential charge states are shown in Fig. 8.3. For all

differences, the ionization enthalpy is independent of the temperature. In contrast, since the entropy difference between these states is finite and mostly independent of T , the ionization entropy term increases linearly with temperature. Moreover, contrary to what could be expected, the neutral vacancy has the highest entropy so that the formation entropy of V_{Ga} should decrease with an increase in the number of electrons. Important changes with temperature occur in the band gap, and, depending on the doping type, in the Fermi level. This temperature dependence of the Fermi level affects the relative energy of neighboring charge states by the same amount. It is this effect that is responsible for the interesting effect of *negative temperature dependence* of the V_{Ga} concentration discussed in Refs. [2, 3].

While the change in the charge transition level due to entropy term is smaller than the shift in the Fermi level — the $(-2/-3)$ transition is moved by about 0.3 eV — its effect is charge dependant, affecting the relation between the various levels. At 0K, V_{Ga} ionization levels in GaAs are located in the lower part of the band gap and distant by 0.1–0.2 eV from each other. Hence, an increase of 0.3 eV coupled with the Fermi level effect cause important changes in the dominant vacancy charge.

Similar entropic effects are known to affect the properties of defects in semiconductors. For example, the entropy term of amplitude similar to ours appears to be responsible for the lowering of the migration barrier in SiC with increasing temperature [15]. It is also associated with the destabilization of a metastable configuration in GaAs [27, 28] and the thermal behaviour of the binding free energy of acceptor-oxygen complexes in Si [19].

These entropic effects are directly associated with changes in the phonon density of states (DOS) as electrons are added to the system, modifying the bonding strength in the region surrounding the defect. In the case of V_{Ga} , the phonon DOS for the neutral vacancy shows the highest deviation from the crystalline value. The effective impact of the vacancy on the lattice is equivalent to the application of a local negative pressure: the transverse acoustic (TA) modes move from the crystalline value by 3 cm^{-1} to higher frequencies (mode hardening), while the LA, LO, TO modes move to lower frequency (mode softening) by 2 cm^{-1} . As the defect is progressively charged, the DOS tends to

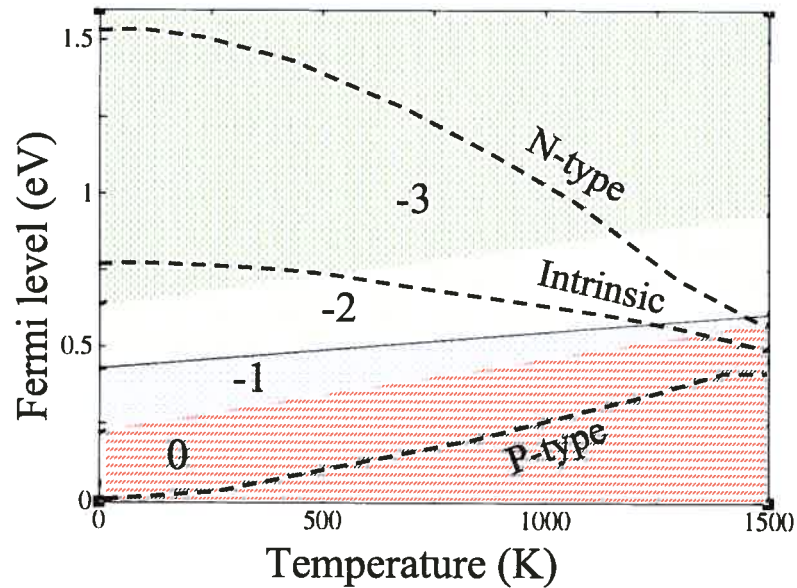


Figure 8.4: Stability domains of V_{Ga}^q in GaAs as function of temperature, indicated by the charge state. Each region denotes the temperature and the doping range where a charge state dominates self diffusion. The dashed lines show the typical position of the chemical potential as function of temperature for N, P and intrinsic conditions. At temperatures above 1000K the competition between different charges becomes significant.

recover the frequencies of the crystal DOS: TA modes move to lower frequencies while LA, LO and TO move to higher frequencies leading to an overall decrease in the entropy with the charge state.

Structurally, this behaviour is associated with the relaxation of the arsenic atoms surrounding the empty site. In an unrelaxed model, As atoms form a tetrahedron around the empty site with an interatomic distance of 3.95 \AA . As electron are added to the relaxed system, the symmetry is conserved but volume of the vacancy decreases. A detailed structural analysis reveals that the addition of the first electron does not affect considerably the bond length around V_{Ga} . As-As bonds decrease progressively to 3.37 \AA for 0 and -1 charges and 3.36 \AA and 3.35 \AA for -2 and -3 charge states. To compensate, since the simulation is performed at fixed volume, the Ga-As bonds, for the atoms surrounding the vacancy, stretch out with the charge; this effect propagates in the direction of the link until the borders of the supercell especially on the high symmetry axes. The pressure increases locally around the vacancy but it is released farthest away from it. The

phonon density of states suffers from a strong competition between the compression and the expansion of bonds, and the electron-phonon interaction.

Finally, it appears that the DOS is under the effects of a small progressive compression as electrons are added (TA softening and LA, LO, TO hardening) that clearly overcomes the proposed phonon mode softening induced by electron-phonon interaction studied earlier for GaAs in Ref. [29]. Instead, our data show a brutal decrease of the phonon vibrational modes as a neutral vacancy is introduced in a perfect crystal and a progressive strengthening as electrons are added to the defect, agreeing with the experimental observation [30] that LO(Γ) phonon peak is shifted toward the low frequency side (mode softening) and broadens its width depending on the concentration of vacancies.

Results for our calculated formation entropies and enthalpies are summarized in Table 8.1. Formation entropy for the neutral state, the only one available in the literature, compare well with the *ab-initio* values using anharmonic approximation by Bockstedte and Scheffler [16] and a recent fit of experimental data by Gebauer *et al.* [2] (further comparisons with earlier works can be found in these two references). The formation enthalpy we calculated in a previous work [4] is in good agreement with recent LDA calculation by Zollo *et al.* [31] using 64 and 216 supercells, differences do not exceed 0.2 eV for V_{Ga}^0 . We relaxed V_{Ga} in different charge states using spin-polarized LDA calculation but no Jahn-Teller distortion has been observed. This confirms that the symmetry conservation of V_{Ga}^q is not a drawback of LDA but a behaviour proper to cation vacancies [32]. To our knowledge no calculation using GGA has been performed on this defect at this point.

Combining the results presented above, we obtain the free energies difference between the various charge states: $F^{0/-1} = 0.22 + 2.74k_B T$, $F^{-1/-2} = 0.42 + 1.42k_B T$ and $F^{-2/-3} = 0.63 + 2.28k_B T$. While the formation entropies are known to suffer from finite-size effects, entropy differences should be much more reliable. Formation entropy errors due to the finite-size of the supercell are expected to be of similar magnitude for all charge states and thus cancel out when only differences are considered [15]. Contrary to earlier assumptions [17], the actual temperature dependence of the ionization energies

Table 8.1: Calculated formation entropies S_f and formation enthalpies H_f for $V_{Ga}^{0,-1,-2,-3}$ in GaAs are compared to earlier calculation and experimental fits (see the text)

	Neutral	-1	-2	-3
S_f/k_B	7.23 ^a , 7.3 ^b , 9.6±1 ^c	4.46	3.04	0.755
H_f (eV)	3.2 ^a , 2.8 ^b , 3.2±0.5 ^c	3.35	3.78	4.423

^a Present work

^b *Ab-initio* calculation by Bockstedte and Scheffler [16]

^c Fit of concentration profiles by Gebauer *et al.* [2]

favors the formation of less charged point defects at high temperature. Figure 8.4 shows the stability domains within the experimental band gap for the dominant charge states as function of temperature.

The transition from a charge state to another takes place as soon as the Fermi level (under certain doping conditions) crosses the ionization energy line, falling into a new stability domain. For P-type, the 0/-1 level is always higher than the extrinsic Fermi energy suggesting that the neutral vacancy remains stable at all temperatures. For intrinsic and N-type conditions, however, the charge state of the vacancy is lowered as the temperature increases. The critical temperature associated with these changes is located in the range of 520-1500K where V_{Ga}^{-2} , V_{Ga}^{-1} and V_{Ga}^0 start to compete with V_{Ga}^{-3} .

8.4.1 Uncertainties on the calculated Gibbs free energy

The formation enthalpies and transition levels calculated with LDA involve uncertainties that can exceed the vibrational effects, but that are more likely to cancel out. Thus, one must be aware of possible source of errors affecting the Gibbs free energy

of formation. Formation entropies suffers from errors associated with the calculation of vibrational density of states [33]. For example, the microscopic structure of the defect itself, which is possibly not obtained accurately. This may be a result of the defect-defect interaction of the supercells, but also errors from the construction of the pseudopotentials or the LDA may contribute. In general, these errors only affect slightly the total energy calculation but are more significant for vibrational modes. Moreover, uncertainty in the lattice constant which causes the defect to exert pressure on its environment and can modify the equilibrium lattice constant of the defective supercell compared to the LDA bulk value. Since we are only interested in Gibbs free energy differences we may expect a significant amount of error cancellation. Recent tests on phonon frequencies of silicon vacancy in bulk silicon [34] showed that a cancellation of errors does tend to occur because free energy differences depend only weakly on the choice of lattice constant.

Uncertainties on formation enthalpy can come mostly from two sources. First, the LDA band gap problem, since it is well-known that LDA usually underestimates the band gap of materials. The simplest way, used here, for dealing with the band gap problem [4] is to shift the conduction band states up uniformly by the amount needed to reproduce the experimental gap [35] (align the conduction band minimum (CBM) with the experimental value, at 1.52 eV from the valence band maximum (VBM)). Since the LDA can produce similarly large errors in the energies of the deep defect states, it is also important to correct these errors. Therefore, defect states with predominantly conduction band character experience the same upward shift as the the conduction band states themselves, while leaving the defect states with predominantly valence band character fixed relative to the valence band edge. In our case, V_{Ga}^q ionization levels are calculated by total energy differences, the correction we use here leaves these ionization levels unchanged since they have a valence band character. Many other correction schemes can be used with good efficiency as discussed in the work of Castleton and coworkers [36], we discuss below the impact of such corrections on the fitting procedure. With current computational power limitations, a full GW calculation, which would correct LDA errors, is unfortunately not possible for the large supercells needed for defect studies. Therefore, LDA calculation is the best choice currently since it provides a good estimate of the

electronic ionisation levels that enables us to fit efficiently experimental data. Second, supercell corrections could also introduce additional errors on the calculated formation enthalpy. In this work, total energies of charged defect are corrected using the first order Makov-Payne method [37]. Many articles [36, 38, 39] have recently shown that Makov and Payne correction can be overestimated for highly charged defects.

Our calculated formation energies for 216 supercell corrected using the first order Makov and Payne term lead to deep-lying ionization energies [4] which are in agreement with experimental values obtained by electron irradiation of GaAs [40]. This correction is 0.094 for V_{Ga}^{-1} , and 0.37 and 0.84 eV for V_{Ga}^{-2} , V_{Ga}^{-3} respectively. To evaluate errors coming from this supercell correction we take advantage of a recently published finite-size scaling study [36] of V_{In}^{-3} in InP. A direct comparison is possible here since the two compounds have close lattice and dielectric constants. Considering the largest correction, for the -3 charge defect, Castleton *et al.* [36] find that the formation energy of V_{In}^{-3} in a 216 supercell is underestimated by 0.7 eV with respect to the infinite system while the Makov Payne approximation suggests a correction of 0.84 eV, leading to an error of ± 0.13 eV. Given the similarities between InP and GaAs, we expect that the maximum error due to size effect should be similar, after the Makov Payne corrections, for the V_{Ga}^{-3} in GaAs.

8.4.2 Fitting experimental measurements

We revisit the analysis of two apparently conflicting results by looking at diffusion data of Bracht *et al.* and concentration profiles of Gebauer *et al.* We first reproduce exactly the fits of Gebauer *et al.* and Bracht *et al.* using exactly the same conditions described in the respective original papers (dash lines in Fig. 8.5 and circles in Fig. 8.6) and explained above. We then use the Gibbs free energy of formation and all resulting ionisation levels calculated by LDA to generate our fits. The contributions coming from each charge state is shown and the total concentration and diffusion coefficient are plotted with solid lines in Figs. 8.5 and 8.6.

Using our calculated values, with only one free parameter to adjust the prefactor (C^* in equation 8.5), and -2 and -3 charges, we reproduce the experimental data of

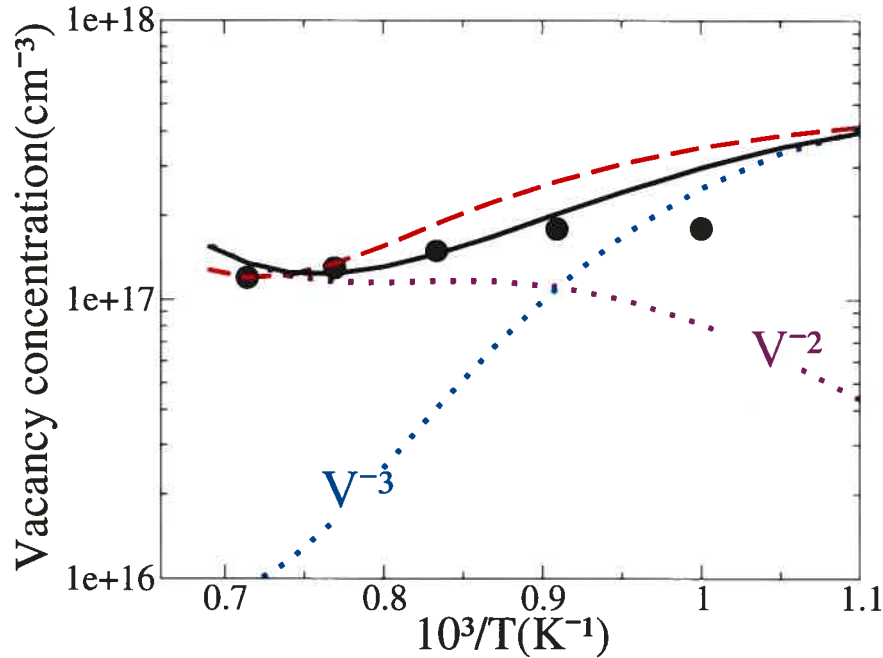


Figure 8.5: Experimental measurement (filled circle), and fit (dashed line) for temperature dependence of the vacancy concentration, as reported by Gebauer *et al.* [2]. Our calculation including -2 and -3 contributions (solid line) is shown together as well as separately (dotted line).

Gebauer *et al.* with a good accuracy (Fig. 8.5) showing that the -2 charge state is the main contributor to self-diffusion at high T.

Self-diffusion coefficient measured by of Bracht *et al.* are also efficiently fitted. Comparing our calculated Ga self-diffusion parameters with experimental data and the corresponding fits [1], with an adjustable prefactor C^* we notice (see Fig. 8.6(b)) a strong competition between -2 , -1 and 0 charge states for intrinsic GaAs while -3 does not contribute to the total concentration. Moreover, a progressive charge reversibility occurs: V_{Ga}^{-2} dominates at high temperature until $T_I^{(-2/-1)} = 1230\text{K}$, then V_{Ga}^{-1} dominates until $T_I^{(-1/0)} = 1345\text{K}$. Even if details about the location of transition temperatures are different, the total diffusion coefficient is very well reproduced with our model too. Contrary to what was proposed by Tan *et al.* [3], our results support the experiment, indicating that under intrinsic conditions V_{Ga}^{-3} plays very little role.

For Si-doped GaAs, while Bracht *et al.* find that the concentration decreases with

temperature for V_{Ga}^{-2} but shows no contribution from the -3 charge states, we see a more complex situation (Fig. 8.6(a)): self-diffusion is mediated by V_{Ga}^{-3} at temperature below $T_N^{(-3/-2)} = 1150\text{K}$, while above this transition temperature -2 charge state is dominant until $T_N^{(-2/-1)} = 1450\text{K}$. Finally, for P-type GaAs (Fig. 8.6(c)), we obtain results that are very similar to Bracht: the neutral vacancy is dominant over the whole temperature range.

As is seen in Figs. 8.5 and 8.6, our model reproduces very well all data at high temperature for different doping condition. The agreement is weaker for N-type GaAs below 1100K , however. In this regime both experimental data sets lie below our predictions. This is likely due to the formation of vacancy complexes and slow Ga-self-diffusion which appear to delay the establishment of equilibrium condition, as noted by Gebauer *et al.* [2]. Thus, the measured concentration is a lower limit of the calculated equilibrium concentration.

A gap or supercell correction that applies the same rigid shift to the ionization levels would not change the observed sequential charge reversibility, caused by entropic effects, but only the location of transition temperature. In this particular case, if the ionization levels are shifted upwards, the contribution of -2 charge state increases. Thus, the remaining discrepancies that appears between experiment and theory visible in Fig. 8.6(a) for the low temperature limit can be significantly attenuated. At a first insight, we propose to use the simplest LDA corrections, if entropy effects alone give unsatisfactory results, one can go beyond by using more advanced gap correction schemes.

8.5 Conclusion

The present work shows that the deformation potential and the electron-phonon interaction induced by charged vacancies affect the phonon spectrum in a much more complicated way than was thought previously [16, 17]. Consequently, the picture about ionization levels obtained at 0K [4] is not applicable in the experimental range as ionization free energies are temperature dependent. We have found that the entropy of formation calculated with LDA contributes to lower the free energy of formation at high

temperature when associated with empirical thermal effects on the Fermi level. Thus, the formation of less charged Ga vacancies mediating self-diffusion in GaAs becomes much more favored at high temperature. In spite of LDA limitations, the resulting coherent Gibbs free energies of formation are a possible solution to the apparently contradictory experimental results about the dominant charge state for vacancy diffusion in GaAs. In addition to resolving this controversy, our results underline the need to include thermal effects on the Fermi level together with ionization entropies and ionization enthalpies. Even though the Fermi level effect dominates the correction term, the smaller contribution coming from ionization entropy is crucial for a correct interpretation of experimental data.

Acknowledgements. We are thankful to Dr Hartmut Bracht for sending us his data and a number of discussions. NM acknowledges partial support from FQRNT (Québec), NSERC (Canada) and the Canada Research Chair program. We thank the RQCHP for generous allocation of computer time.

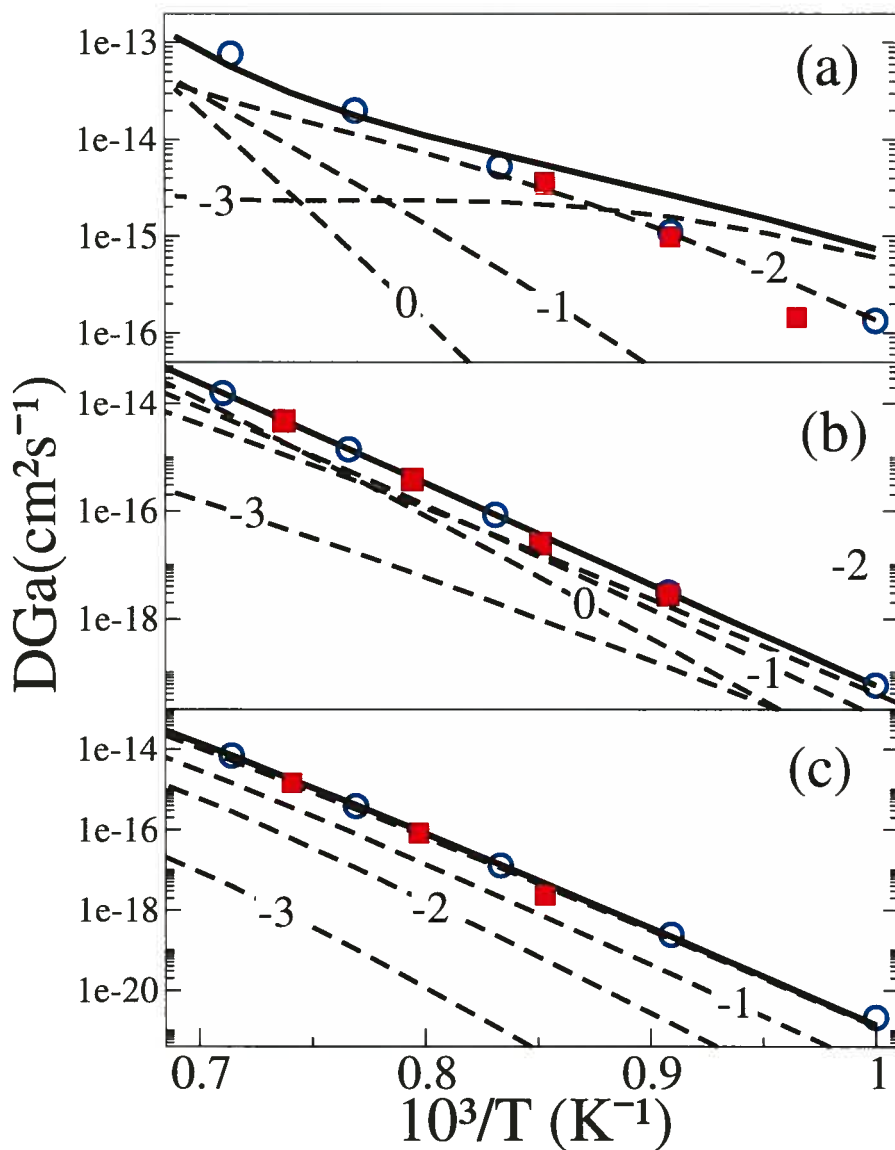


Figure 8.6: Gallium self diffusion constant versus inverse temperature for (a) N-type (b) intrinsic (c) P-type. Theoretical results for D_{Ga}^q (dashed lines) show the contribution of each charge state. The transition temperature lie at the crossing of the curves. The total D_{Ga} (solid) is in good agreement with the measured interdiffusion data (square) and fit by Bracht *et al.* [1](circle).

BIBLIOGRAPHY

- [1] B. Bracht, M. Norseng, E. E. Haller, K. Eberl, and M. Cardona, *Sol. Stat. com.* **112**, 301 (1999).
- [2] J. Gebauer, M. Lausmann, F. Redmann, R. Krause-Rehberg, H. S. Leipner, E. R. Weber, and P. Ebert, *Phys. Rev. B* **67**, 235207 (2003).
- [3] T. Y. Tan, H. M. You, and U. M. Gösele, *Appl. Phys. A* **56**, 249 (1993).
- [4] F. El-Mellouhi and N. Mousseau, *Phys. Rev. B* **71**, 125207 (2005).
- [5] S. B. Zhang and J. E. Northrup, *Phys. Rev. Lett.* **67**, 2339 (1991).
- [6] K. Laasonen, R. M. Nieminen, and M. J. Puska, *Phys. Rev. B* **45**, 4122 (1992).
- [7] H. Seong and L. J. Lewis, *Phys. Rev. B* **52**, 5675 (1995).
- [8] J. E. Northrup and S. B. Zhang, *Phys. Rev. B* **47**, 6791 (1993).
- [9] S. Pöykkö, M. J. Puska, and R. M. Nieminen, *Phys. Rev. B* **53**, 3813 (1996).
- [10] J. T. Schick, C. G. Morgan, and P. Papoulias, *Phys. Rev. B* **66**, 195302 (2002).
- [11] A. Janotti, S.-H. Wei, S. B. Zhang, and S. Kurtz, *Phys. Rev. B* **67**, 161201 (2003).
- [12] R. W. Jansen and O. F. Sankey, *Phys. Rev. B* **39**, 3192 (1989).
- [13] G. A. Baraff and M. Schlüter, *Phys. Rev. Lett.* **55**, 1327 (1985).
- [14] B.-H. Cheong and K. J. Chang, *Phys. Rev. B* **49**, 17436 (1994).
- [15] E. Rauls and T. Frauenheim, *Phys. Rev. B* **69**, 155213 (2004).
- [16] M. Bockstedte and M. Scheffler, *Z. Phy. Chem. (munich)* **200**, 195 (1997).
- [17] J. A. Van Vechten and C. D. Thurmond, *Phys. Rev. B* **14**, 3539 (1976).

- [18] K. Dev and E. G. Seebauer, Phys. Rev. B **67**, 035312 (2003).
- [19] M. Sanati and S. K. Estreicher, Phys. Rev. B **72**, 165206 (2005).
- [20] E. Rauls, T. Frauenheim, A. Gali, and P. Deák, Phys. Rev. B **68**, 155208 (2003).
- [21] K. Petzke, C. Göbel, C. Schrepel, and U. Scherz, Mater. Sci. Forum **258-263**, 861 (1997).
- [22] D. Strauch and B. Dorner, J. Phys. Chem. **2**, 1457 (1990).
- [23] J. S. Blakemore, J. Appl. Phys. **53**, R123 (1982).
- [24] C. D. Thurmond, J. Electrochem. Soc. **122**, 1133 (1975).
- [25] Y. P. Varshni, Physica **34**, 149 (1967).
- [26] Numerical Recipes F77, section 9.4 <http://www.library.cornell.edu/nr/cbookfpdf.html>.
- [27] B. Hamilton, A. R. Peaker, and S. T. Pantelides, Phys. Rev. Lett. **61**, 1627 (1988).
- [28] J. W. Fang, Z Q an Kim and P. W. Yu, Appl. Phys. Lett. **78**, 2506 (2001).
- [29] V. Heine and J. A. Van Vechten, Phys. Rev. B **13**, 1622 (1976).
- [30] K. Ishioka, K. G. Nakamura, and M. Kitajima, Phys. Rev. B **52**, 2539 (1995).
- [31] G. Zollo, J. Tarus, and R. M. Nieminen, J. Phys. Cond. Matt. **16**, 3923–3932 (2004).
- [32] D. J. Chadi, Materials Science in Semiconductors Processing **6**, 281 (2003).
- [33] A. Mattausch, M. Bockstedte, and O. Pankratov, Phys. Rev. B **70**, 235211 (2004).
- [34] O. K. Al-Murshadani and R. J. Needs, Phys. Rev. B **68**, 235205 (2003).
- [35] A. Baraff and M. Schlüter, Phys. Rev. B **30**, 1853 (1984).
- [36] C. W. M. Castleton, A. Höglund, and S. Mirbt, Phys. Rev. B **73**, 035215 (2006).
- [37] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).

- [38] J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, *Phys. Rev. B* **71**, 035206 (2005).
- [39] U. Gerstmann, P. Deák, R. Rurali, B. Aradi, T. Frauenheim, and H. Overhof, *Physica B cond. Matter* **340**, 190 (2003).
- [40] Y. Q. Jia, J. V. Bardeleben, D. Stienvard, and C. Delerue, *Phys. Rev. B* **45**, 1645 (1992).

SYNTHÈSE ET CONCLUSION

Les mécanismes de diffusion des lacunes dans deux semiconducteurs cristallins, le silicium et l'arséniure de gallium, ont été étudiés. Plusieurs détails concernant l'auto-diffusion dans ces matériaux simples et composés de grande importance technologique ont été révélés. Le programme *ab-initio* à base d'orbitales localisées (SIESTA) combiné avec la technique d'activation et de relaxation (ART) ont été utilisés pour générer directement des dynamiques activées. Au lieu de suivre en détail les fluctuations thermiques, les états de transition menant à la migration d'un défaut sont directement identifiés. Cette approche est efficace pour étudier les mécanismes de diffusion des lacunes dans différents types de semiconducteurs allant des plus simples comme le silicium jusqu'aux composés tel que le GaAs.

Cette thèse démontre que la diffusion des lacunes dans les semiconducteurs est un phénomène beaucoup plus riche et plus complexe qu'on ne le pensait initialement. La méthode SIESTA-ART nous a permis d'identifier des mécanismes subtils qui peuvent être observés expérimentalement à basse température comme la réorientation dans le Si et les différentes structures des défauts de la famille de As_{Ga} dans le GaAs. En même temps, elle offre la possibilité d'identifier les mécanismes de diffusion simples qui dominent l'autodiffusion des défauts dans les semiconducteurs et des mécanismes plus complexes, rares et peu attendus comme les sauts au quatrième voisin pour $V_{Ga}^{-1,-2,-3}$.

Pour un défaut aussi simple que V_{Si} dans le silicium, nous avons trouvé que le mécanisme de migration dominant est régi par des sauts successifs aux premiers voisins en traversant une barrière de 0.4 eV. Toutefois, il peut se produire un réarrangement de la structure autour du défaut, un mécanisme qui est rarement observé par les autres simulations mais qui a été observé expérimentalement à 20 K par Watkins *et al.* [1]. Il y a également introduction de désordre dans la structure, accompagné par la saturation des liens pendants nécessitant des barrières beaucoup plus élevées. Ce désordre conduit à la transformation du défaut ponctuel en un défaut plus complexe. Les propriétés thermodynamiques comme les barrières et les constantes de diffusion calculées sont dans l'ensemble comparables à ceux qui sont obtenus par des travaux théoriques et expéri-

mentaux précédents. Nos résultats ont déjà été cités par d'autres groupes de recherche pour comparer le mode de relaxation de V_{Si} [99] et sa barrière de diffusion [100].

En étudiant les lacunes de Ga chargées dans le semiconducteur binaire le GaAs des aspects encore plus intéressants nous sont révélées car de nouveaux mécanismes de diffusion apparaissent si les lacunes sont chargées négativement. La diffusion au second voisin se fait principalement par le chemin le plus simple appelé *plane-passing mechanism* pour tous les états de charge avec des barrières variant entre 1.7 et 2.0 eV. Les lacunes négatives possèdent un chemin de migration additionnel leur permettant de diffuser au second voisin que nous avons appelé *cluster-assisted mechanism*. En plus, V_{Ga} diffuse au premier voisin si $q = 0, -1, -2$ en formant des complexes du type $As_{Ga}-X$ qui sont déformés si la lacune est négativement chargée. Cependant, ce mécanisme ne peut pas être considéré comme dominant pour l'autodiffusion de V_{Ga} car le chemin de migration n'est pas complet. Finalement, les lacunes négativement chargées peuvent diffuser directement vers le quatrième voisin en suivant la direction 100 en franchissant des barrières supérieures à 4 eV.

Nous avons aussi trouvé que les différentes barrières de diffusion de V_{Ga} dans le GaAs augmentent avec l'augmentation de l'état de charge. Cette dépendance de la charge est confirmée par des calculs précédents de Bockstedte et Scheffler [6] et par l'expérience [101], mais elle va l'encontre de la tendance habituellement observée par les calculs *ab initio* [52–54, 102, 103] sur les lacunes dans les semiconducteurs binaire tel que le SiC et le GaN.

En plus de ces conclusions majeures, d'autres aspects importants reliés à ces défauts ont été découverts. Les limitations imposées par l'utilisation des calculs *ab-initio* dans l'étude des défauts ont été clairement identifiées. Les erreurs introduites sont principalement dues aux effets de taille, à la densité des points-k, à l'usage de la DFT et finalement à la qualité de la base utilisée. Ces conclusions sont standards et les tests proposés pour améliorer les résultats peuvent être généralisés pour différents types de semiconducteurs.

Les résultats les mieux convergés révèlent que la symétrie tétraédrique T_d est toujours conservée pour V_{Ga} . Les liens pendants d'As ne se recombinent pas en laissant le volume de relaxation presque inchangé quand la charge augmente. À chaque transition

(intersection), un électron est transféré à la fois du niveau de Fermi au niveau du défaut faisant de V_{Ga}^{-3} l'état de charge favorisé pour le GaAs intrinsèque ou de type N à basse température. Les niveaux d'ionisation calculés sont profonds et en bon accord avec les expériences de radiation par électron [104].

Contrairement à V_{Ga} , le volume de V_{As} diminue progressivement quand la charge du défaut augmente. La symétrie locale autour de la lacune se trouve brisée sauf pour V_{As}^{+1} . À chaque transition électronique, deux électrons sont transférés en même temps (effet U-négatif) faisant passer directement de la charge + à la charge -. Il en résulte une distorsion du type Jahn-Teller ; ce qui est en bon accord avec les calculs récents de Chadi [105]. Un comportement très similaire vient d'être découvert en 2006 pour les lacunes d'azote dans le GaN [103] en utilisant des calculs *ab-initio* basés sur la LDA. Une nouvelle transition du type U-Négatif est trouvée lors de la transition (-/-3) à $E_F = 1.27$ eV accompagnée d'une déformation structurale peu habituelle. Cette déformation mène à une configuration avec des liens résonnants. Ce type de relaxation n'a été observé jusqu'à date que pour les bilacunes d'As dans le GaAs et les bilacunes de Si dans le silicium.

La stabilité de V_{As} et V_{Ga} est affectée par le dopage progressif de l'échantillon dont la croissance est sous des conditions riches en As ou en Ga. La compétition entre V_{As} et V_{Ga} à des valeurs critiques des niveaux de Fermi révèle que pour le GaAs de type N ou intrinsèque, V_{Ga}^{-3} est le défaut le plus stable par rapport à tous les autres indépendamment des conditions de croissance.

L'étude de la stabilité de V_{Ga}^{-3} à haute température révèle qu'il peut y avoir un renversement de la charge vers des états moins chargés ce qui produit un excellent ajustement des résultats expérimentaux. Nous avons résolu le débat concernant l'état de charge de la lacune dominante dans le GaAs en introduisant l'effet de la température sur l'énergie libre de Gibbs en plus de l'effet de température sur le niveau de Fermi. Ces effets de température sont si importants qu'ils renversent et changent l'état de charge de la lacune dominante. V_{Ga} préfère diminuer progressivement son état de charge au fur et à mesure que la température augmente. Nos profils de concentration calculés à l'aide d'un

modèle de concentration peuvent bien ajuster les courbes expérimentales de Gebauer *et al.* [16] en même temps que celles de Bracht *et al.* [49]. Ce résultat montre que même si les effets de températures sur les niveaux d'ionisation sont faibles, une fois combinés avec l'effet du niveau de Fermi, ils deviennent pertinents afin de bien ajuster les données expérimentales.

Bien que les résultats de cette thèse se comparent bien avec l'expérience, leur précision est limitée par les erreurs introduites lors des calculs *ab-initio* (discutés dans la section 2.1), comme les effets de taille finie, les erreurs causées par la DFT et la qualité de la base utilisée. La plupart de ces erreurs sont corrigées ou atténuées lorsque nous prenons des différences. Toutefois, la comparaison avec certaines expériences est peu évidente du fait que les énergies de formation et les barrières de diffusion que nous avons calculé ne prennent en considération ni les effets de température ni l'influence des dopants ou d'autres défauts se trouvant au voisinage des lacunes.

Cette étude peut être étendue pour traiter des systèmes plus complexes comme les composés GaN et GaAsN dopés, où on s'attend à des mécanismes encore plus complexes, en menant une étude sur la diffusion simultanée des autodéfauts et des dopants. La compréhension deviendra encore plus complète concernant les mécanismes de diffusion possibles et les propriétés des défauts ponctuels qui existent dans ces matériaux, leur nature, leur état de charge ainsi que leur contribution à l'autodiffusion. Avec les outils et les analyses faites durant ce travail, il sera aussi possible de :

- Mener une étude similaire pour les chemins de diffusion de V_{As} afin d'avoir une idée complète sur sa compétition avec V_{Ga} . Nous avons utilisé SIEST-A-RT avec succès pour générer les chemins de diffusion de V_{As} . Ces travaux de simulation sont bien avancés, mais l'analyse des résultats n'est pas encore achevée, ils feront objet d'un article une fois que ce sera terminé.
- Étudier aussi les chemins de diffusion des interstitiels de Ga dans le GaAs. L'article de Bracht *et al.* [42] a suscité notre intérêt pour cette question car il n'y a pas encore de travaux de simulation qui ont été réalisés dans cette direction. Nous pouvons imaginer la richesse des événements activés qui peuvent être générés pour ce défaut. Un étudiant d'été, Mr Marc-André Malouin avait commencé à

travailler sur ce sujet durant l'été 2005 en collaboration avec l'auteur et Prof. Normand Mousseau et continuera son travail afin de finaliser les calculs et les analyses nécessaires. À ces deux points s'ajoutent d'autres études intéressantes :

- Étudier l'effet de l'état de charge et de la température sur le chemin de migration de V_{Si} ;
- Utiliser des méthodes plus précises que la LDA pour les calculs de relaxation (voir par exemple l'étude menée sur les interstitiels de Si [106]), qui est en voie de développement mais encore limitée par le nombre d'atomes utilisés ;
- Faire des études plus avancées, qui comparent la diffusion des lacunes à haute et basse température.

CONTRIBUTIONS

L'ensemble des travaux de simulation par le code *ab-initio* Siesta [107, 108] ainsi que les analyses présentées dans cette thèse ont été effectués par l'auteur. La première intégration des codes ART [83, 109] et Siesta a été faite par Prof. Normand Mousseau et Prof. Pablo Ordejón en 2001. L'auteur a longuement travaillé sur l'amélioration du rendement du code SIEST-A-RT et l'étude détaillée des paramètres de simulation jusqu'à ce qu'ils deviennent fonctionnels et appropriés pour l'étude des semiconducteurs. Le premier article présenté dans cette thèse, au chapitre 5, constitue la toute première utilisation fructueuse de cette méthode.

Plusieurs programmes permettant l'analyse et le raffinement des résultats ont été écrits et utilisés par l'auteur. À titre d'exemple, dans le cadre du projet d'été 2005 de Mr Marc-André Malouin, l'auteur a suggéré, supervisé et analysé les besoins et les spécifications pour le développement d'un outil de conversion de données permettant la visualisation des sorties de Siesta avec OpenDx appelé *Siesta2Dx*. Cet outil fait maintenant partie de la version 2.0 de SIESTA comme contribution externe.

L'auteur a écrit la première version des quatre articles présentés dans cette thèse. Prof. Normand Mousseau les a, par la suite, revu et révisé. Quelques itérations étaient suffisantes pour aboutir aux versions soumises à la publication. Quelques suggestions de la part des arbitres de la revue *Physical Review B* ont permis d'obtenir les versions finales des deux premiers articles qui sont déjà publiés. Mis à part quelques figures, la plupart des figures utilisées dans cette thèse ont été générées par l'auteur.

BIBLIOGRAPHIE

- [1] G. D. Watkins, *Deep Centers in semiconductors*, edited by S. T. Pantelides (Gordon and Breach, New York, 1992), chap. 3.
- [2] T. Raz, N. Shuall, G. Bahir, D. Rittera, D. Gershoni, and S. N. G. Chu, *Appl. Phys. Lett.* **85**, 3578 (2004).
- [3] Z. M. Wang, K. Holmes, J. L. Shultz, and G. J. Salamo, *Phys. stat. sol. (a)* **202**, R85 (2005).
- [4] Deepak, D. Balamurugan, and K. Nandi, *Bull. Mater. Sci.* **1**, 169 (2003).
- [5] J. A. Van Vechten, *J. Phys. C : Solid State Phys.* **17**, L933 (1984).
- [6] M. Bockstedte and M. Scheffler, *Z. Phy. Chem. (munich)* **200**, 195 (1997).
- [7] G. Henkelman, G. Jóhannesson, and H. Jónsson, in *Progress on Theoretical Chemistry and Physics*, edited by S. D. Schwartz (Kluwer Academic Publishers, 2000), pp. 269–300.
- [8] E. R. Weber, *Physica B* **340-342**, 1 (2003).
- [9] Z. M. Wang, S. Seydmohamadi, J. H. Lee, and G. J. Salamo, *Appl. Phys. Lett.* **85**, 5031 (2004).
- [10] S. Adachi, ed., *Properties of Aluminium Gallium Arsenide, Data Review Series*, vol. 7 (The Institute of Electrical Engineers, London, 1993).
- [11] A. Rastelli, S. Stufler, A. Schliwa, R. Songmuang, C. Manzano, G. Costantini, K. Kern, A. Zrenner, D. Bimberg, and O. G. Schmidt, *Phys. Rev. Lett.* **92**, 166104 (2004).
- [12] H. H. Evans, *Nature* **229**, 403 (1971).
- [13] H. Yu and W. Lu, *Acta Materialia* **53**, 1799 (2005).
- [14] C. Woo and W. Frank, *J. Nucl. Mater.* **137**, 7 (1985).
- [15] B. Bracht, M. Norseng, E. E. Haller, K. Eberl, and M. Cardona, *Sol. Stat. com.* **112**, 301 (1999).

- [16] J. Gebauer, M. Lausmann, F. Redmann, R. Krause-Rehberg, H. S. Leipner, E. R. Weber, and P. Ebert, *Phys. Rev. B* **67**, 235207 (2003).
- [17] W. Kohn and L. J. Sham, *Physical Review* **140**, A1133 (1965).
- [18] J. Bourgoin and M. Lanoo, *Point defects in Semiconductors I and II* (Springer, Berlin, 1981).
- [19] F. Wooten, K. Winer, and D. Weaire, *Phys. Rev. Lett.* **54**, 1392 (1985).
- [20] H. J. Queisser and E. E. Haller, *Control and use of defects in materials* **281**, 945 (1998).
- [21] J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, *Phys. Rev. B* **71**, 035206 (2005).
- [22] H. Bracht, E. E. Haller, and R. Clark-Phelps, *Phys. Rev. Lett.* **81**, 393 (1998).
- [23] G. A. Baraff, E. O. Kane, and M. A. Schlüter, *Phys. Rev. B* **21**, 5662 (1980).
- [24] O. Sugino and A. Oshiyama, *Phys. Rev. Lett.* **68**, 1858 (1992).
- [25] J. Dabrowski, *Solid State phenomena* **71**, 23 (2000).
- [26] P. W. Anderson, *Phys. Rev. Lett.* **34**, 953 (1975).
- [27] F. Anderson, F. S. Ham, and G. Grossmann, *Phys. Rev. B* **53**, 7205 (1996).
- [28] C. S. Nicols, C. G. Van der Walle, and S. T. Pantelides, *Phys. Rev. B* **40**, 5484 (1989).
- [29] C. S. Nicols, C. G. Van der Walle, and S. T. Pantelides, *Phys. Rev. Lett.* **62**, 1049 (1989).
- [30] P. Blöchl, E. Smargiassi, R. Car, D. B. Laks, W. Andreoni, and S. Pantelides, *Phys. Rev. Lett.* **70**, 2435 (1993).
- [31] M. J. Puska, S. Pöykkö, M. Pesola, and R. M. Nieminen, *Phys. Rev. B* **58**, 1318 (1998).
- [32] M. I. J. Probert and M. C. Payne, *Phys. Rev. B* **67**, 075204 (2003).
- [33] J. Lento and R. M. Nieminen, *J. Phys. : Condens. Matter* **15**, 4387 (2003).
- [34] V. Ranki and K. Saarinen, *Phys. Rev. Lett.* **93**, 255502 (2004).

- [35] H. Bracht, J. F. Pedersen, N. Zangenberg, A. Nylandsted Larsen, E. E. Haller, G. Lulli, and M. Posselt, *Phys. Rev. Lett.* **91**, 245502 (2003).
- [36] R. Car, P. J. Kelly, A. Oshiyama, and S. T. Pantelides, *Phys. Rev. Lett.* **52**, 1814 (1984).
- [37] M. M. De Souza and E. M. Sankara Narayanan, *Def. Diff. Forum* **153**, 69 (1998).
- [38] M. Tang, L. Colombo, J. Zhu, and T. D. de la Rubia, *Phys. Rev. Lett.* **55**, 14279 (1997).
- [39] T. Kitashima, K. Kakimoto, and H. Ozoe, *Journal of The Electrochemical Society* **150**, G198 (2003).
- [40] M. Hakala, M. J. Puska, , and R. M. Nieminen, *J. Appl. Phys.* **91**, 4998 (2002).
- [41] R. R. Wixoma and A. F. Wright, *Appl. Phys. Lett.* **87**, 201901 (2005).
- [42] H. A. Bracht and S. Brotzman, *Phys. Rev. B* **71**, 115216 (2005).
- [43] G. A. Baraff and M. Schlüter, *Phys. Rev. Lett.* **55**, 1327 (1985).
- [44] R. W. Jansen and O. F. Sankey, *Phys. Rev. B* **39**, 3192 (1989).
- [45] H. Seong and L. J. Lewis, *Phys. Rev. B* **53**, 9791 (1996).
- [46] B.-H. Cheong and K. J. Chang, *Phys. Rev. B* **49**, 17436 (1994).
- [47] S. Pöykkö, M. J. Puska, and R. M. Nieminen, *Phys. Rev. B* **53**, 3813 (1996).
- [48] S. B. Zhang and J. E. Northrup, *Phys. Rev. Lett.* **67**, 2339 (1991).
- [49] B. Bracht, E. E. Haller, K. Eberl, and M. Cardona, *App. Phys. Lett.* **74**, 49 (1999).
- [50] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [51] L. Torpo, M. Marlo, T. E. M. Staab, and R. M. Nieminen, *J. Phys. Condens. Matter* **13**, 6203 (2001).
- [52] M. Bockstedte, A. Mattausch, and O. Pankratov, *Phys. Rev. B* **68**, 205201 (2003).
- [53] R. Rurali, E. Hernández, P. Godignon, J. Rebollo, and P. Ordejón, *Comp. Mat. Sc.* **27**, 36 (2003).
- [54] S. Limpiumnong and C. G. Van de Walle, *Phys. Rev. B* **69**, 035207 (2004).

- [55] K. H. Chow, G. D. Watkins, A. Usui, and M. Mizuta, *Phys. Rev. Lett.* **85**, 2761 (2000).
- [56] P. J. Feibelman, *Phys. Rev. Lett.* **65**, 729 (1990).
- [57] J. Tersoff, *Phys. Rev. B* **39**, 5566 (1989).
- [58] F. Sillinger and T. Weber, *Phys. Rev. B* **36**, 1208 (1987).
- [59] M. J. Caturla, T. Diaz de la Rubia, L. A. Marqués, and G. H. Gilmer, *Phys. Rev. B* **54**, 16683 (1996).
- [60] M. Jaraiz, G. H. Gilmer, J. M. Poate, and T. Diaz de la Rubia, *Appl. Phys. Lett.* **68**, 409 (1996).
- [61] A. Jääskeläinen., L. Colombo, and R. Nieminen, *Phys. Rev. B* **64**, 233203 (2001).
- [62] M. Parrinello and A. Rahman, *Phys. Rev. B* **45**, 1196 (1980).
- [63] A. F. Voter, F. Montalenti, and T. C. Germann, *Annu. Rev. Mater. Res.* **34**, 321 (2002).
- [64] A. F. Voter, *J. Chem. Phys.* **106**, 4665 (1997).
- [65] M. Cogoni, B. P. Uberuaga, A. F. Voter, and L. Colombo, *Phys. Rev. B* **71**, 121203(R) (2005).
- [66] A. F. Voter, *Phys. Rev. Lett.* **78**, 3908 (1997).
- [67] A. F. Voter, *Phys. Rev. B* **57**, R13985 (1998).
- [68] X. Wu and S. Wang, *J. Chem. Phys.* **110**, 9401 (1999).
- [69] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12562 (2002).
- [70] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [71] D. Chandler, *J. Chem. Phys.* **68**, 2959 (1978).
- [72] J. A. Montgomery, D. Chandler, and B. J. Berne, *J. Chem. Phys.* **70**, 4056 (1979).
- [73] A. F. Voter and J. D. Doll, *J. Chem. Phys.* **82**, 80 (1985).
- [74] H. Vineyard, *J. Phys. Chem. Solids* **3**, 121 (1957).

- [75] G. Wei, N. Mousseau, and P. Derreumaux, *Biophysical J.* **87**, 3648 (2004).
- [76] R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jónsson, *J. Phys. Chem.* **121**, 9776 (2004).
- [77] I. V. Ionova and E. A. Carter, *J. Chem. Phys.* **98**, 6377 (1993).
- [78] G. Mills and H. Jónsson, *Phys. Rev. Lett.* **72** (1994).
- [79] H. Jónsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, Singapore, 1998), p. 385.
- [80] G. Henkelman, B. P. Uberuaga, and H. Jónsson, *J. Chem. Phys.* **113**, 9901 (2000).
- [81] S. A. Trygubenko and D. J. Wales, *J. Chem. Phys.* **120**, 2082 (2004).
- [82] W. David, *International Journal of Modern Physics B* **19**, 2877 (2005).
- [83] G. T. Barkema and N. Mousseau, *Phys. Rev. Lett.* **77**, 4358 (1996).
- [84] R. Malek and N. Mousseau, *Phys. Rev. E* **62**, 7723 (2000).
- [85] C. Lanczòs, *Applied Analysis* (Dover, New York, 1988).
- [86] F. Valiquette and N. Mousseau, *Phys. Rev. B* **68**, 125209 (2003).
- [87] G. Wei, N. Mousseau, and P. Derreumaux, *J. Chem. Phys.* **117**, 11379 (2002).
- [88] S. Santini, G. Wei, N. Mousseau, and P. Derreumaux, *Internet Electron. J. Mol. Des.* **2**, 564 (2003).
- [89] A. Melquiond, N. Mousseau, G. Boucher, and P. Derreumaux, *J. Chem. Phys.* **122**, 174904 (2005).
- [90] G. Henkelman and H. Jónsson, *J. Chem. Phys.* **111**, 7010 (1999).
- [91] G. Henkelman, B. Uberuaga, S. Dunham, and H. Jónsson, *Phys. Stat. sol. (b)* **233**, 24 (2002).
- [92] A. Szabo and N. S. Ostlund, eds., *Modern Quantum Chemistry : introduction to advanced electronic structure theory* (McGraw-Hill, New York, 1989).
- [93] M. Born and J. R. Oppenheimer, *Ann. Phys. (Leipzig)* **84**, 457 (1927).

- [94] P. Hohenberg and W. Kohn, *Physical Review* **136**, B864 (1964).
- [95] J. P. Perdew and S. Kurth, eds., *Density functional theory of atoms and molecules* (Oxford, 1989).
- [96] C. Klaus (2004), cond-mat/0211443 v3.
- [97] P. Ordejón, E. Artacho, and J. M. Soler, *Phys. Rev. B* **53**, R10 :441 (1996).
- [98] M. Levinshtein, S. Rumyantsev, and M. Shur, eds., *Handbook Series on Semiconductor Parameters*, vol. 1 (World Scientific, London, 1996).
- [99] A. Satta, E. Albertazzi, G. Lulli, and L. Colombo, *Phys. Rev. B* **72**, 235206 (2005).
- [100] H. A. Bracht, *Physica B : Condensed Matter* **376-377**, 11 (2006).
- [101] G. Lengel, J. Harper, and M. Weimer, *Phys. Rev. Lett.* **76**, 4725 (1996).
- [102] F. El-Mellouhi and N. Mousseau, in preparation.
- [103] M. G. Ganchenkova and R. M. Nieminen, *Phys. Rev. Lett.* **96**, 196402 (2006).
- [104] Y. Q. Jia, J. V. Bardeleben, D. Stienvard, and C. Delerue, *Phys. Rev. B* **45**, 1645 (1992).
- [105] D. J. Chadi, *Materials Science in Semiconductors Processing* **6**, 281 (2003).
- [106] M. Lopez, G and V. Fiorentini, *Phys. Rev. B* **69**, 155206 (2004).
- [107] D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler, *Int. J. Quant. Chem.* **65**, 453 (1997).
- [108] J. M. Soler, E. Artacho, J. Gale, A. García, J. Junquera, and P. Ordejón, *J. Phys. Condens. Matter* **14**, 2745 (2002).
- [109] N. Mousseau and G. T. Barkema, *Phys. Rev. E* **57**, 2419 (1998).
- [110] M. Leslie and J. Gillian, *J. Phys. C* **18**, 973 (1985).
- [111] G. Makov and M. C. Payne, *Phys. Rev. B* **51**, 4014 (1995).
- [112] U. Gerstmann, P. Deák, R. Rurali, B. Aradi, T. Frauenheim, and H. Overhof, *Physica B cond. Matter* **340**, 190 (2003).
- [113] C. W. M. Castleton, A. Höglund, and S. Mirbt, *Phys. Rev. B* **73**, 035215 (2006).

[114] P. A. Schultz, *Phys. Rev. Lett.* **84**, 1942 (2000).

[115] N. Mousseau, P. Derreumaux, G. T. Barkema, and Malek, *J. Mol. Graph. and Modelling* **19**, 78 (2001).

Annexe I

Conditions aux limites périodiques

Les conditions aux limites périodiques sont largement utilisées pour simuler les propriétés des solides cristallins. Un réseau de Bravais infini peut être simulé par un nombre fini d'atomes répétés infiniment dans l'espace grâce aux conditions aux limites périodiques. Le terme supercellule est généralement utilisé pour désigner la portion du cristal qui est répétée infiniment. Les conditions aux limites périodiques sont particulièrement utiles pour étudier les matériaux réels contenant des défauts ou les liquides et les matériaux amorphes. Néanmoins, l'utilisation des supercellules induit des différences importantes entre les calculs avec des systèmes périodiques (cristaux parfaits) et apériodiques (cristaux comprenant quelques imperfections)¹.

Pour un système comprenant un défaut, les bordures de la supercellule ressemblent à celles d'un matériau volumique mais le défaut interagit avec un nombre infini de ces images. L'énergie totale du système est influencée et devient dépendante de la taille de la supercellule. Les "vraies" propriétés du défaut sont retrouvées à la limite d'une supercellule infiniment large. Le problème est particulièrement important dans le cas des défauts chargés, où l'énergie de Madelung devient infinie si la charge n'est pas neutralisée par un plan compensateur uniforme [110]. Les énergies de formation calculées peuvent contenir des erreurs de quelques électronvolts pour des cellules contenant des dizaines ou des centaines d'atome. Les effets de taille peuvent aussi affecter les défauts neutres comme nous l'avons montré pour la lacune de Si dans la chapitre 5. Pour les systèmes apériodiques, les valeurs calculées sont totalement convergées si et seulement si on utilise des supercellules infiniment grandes ($L \rightarrow \infty$ où L est la dimension de la supercellule). La taille de la supercellule est jugée suffisante si l'énergie totale du système commence à converger en augmentant la taille de la supercellule (L). Cette convergence est régie par l'atténuation des forces électrostatiques et des forces élastiques à longue portée.

¹Connus comme *les effets de taille finie*.

Les tests sur le taux de convergence de l'énergie en fonction de L peuvent être très longs surtout pour les systèmes chargés, où l'énergie converge comme L^{-1} . L'estimation du taux de convergence est très importante, vu que le temps de calcul par les méthodes *ab-initio* est généralement proportionnel au volume de la supercellule (L^3). Il est donc important de trouver la taille de la supercellule optimale qui offre un compromis entre un temps de calcul raisonnable et des énergies assez convergées. Si la dépendance exacte du taux de convergence par rapport à la taille de la supercellule est connue, l'énergie calculée pour un L fini peut être corrigée pour estimer l'énergie à $L \rightarrow \infty$. En augmentant L , l'énergie totale calculée converge une fois que les forces électrostatiques et élastiques de longue portée deviennent négligeables et que le défaut ne ressent plus la présence de ses images.

Le calcul de l'énergie d'un système aperiodique, en utilisant les conditions aux limites périodiques, devrait donner l'énergie $E_{L \rightarrow \infty}$ à la limite $L \rightarrow \infty$. Pour un système localement aperiodique, l'énergie calculée pour une supercellule finie $E(L)$ est différente de $E_{L \rightarrow \infty}$ à cause des interactions entre la densité de charge aperiodique et ses images dans les supercellules avoisinantes (voir la figure I.1). Ces interactions modifient la densité de charge aperiodique avec un facteur qui dépend de L .

Afin d'estimer $E_{L \rightarrow \infty}$, nous avons besoin de connaître la dépendance asymptotique de l'énergie par rapport à L .

Pour un défaut chargé dans un solide cristallin, la densité de charge totale est la somme de la densité de charge du solide cristallin $\rho_p(r)$ et celle du défaut aperiodique $\rho_{ap}(r)$:

$$\begin{aligned}\rho(r) &= \rho_p(r) + \rho_{ap}(r) \\ \rho_p(r + L) &= \rho_p(r)\end{aligned}\tag{I.1}$$

La densité de charge aperiodique $\rho_{ap}(r)$ est supposée être localisée en analogie avec une densité de charge moléculaire. L'énergie électrostatique est la somme des interactions entre :

1. La densité de charge périodique et elle-même (indépendante de L) ;

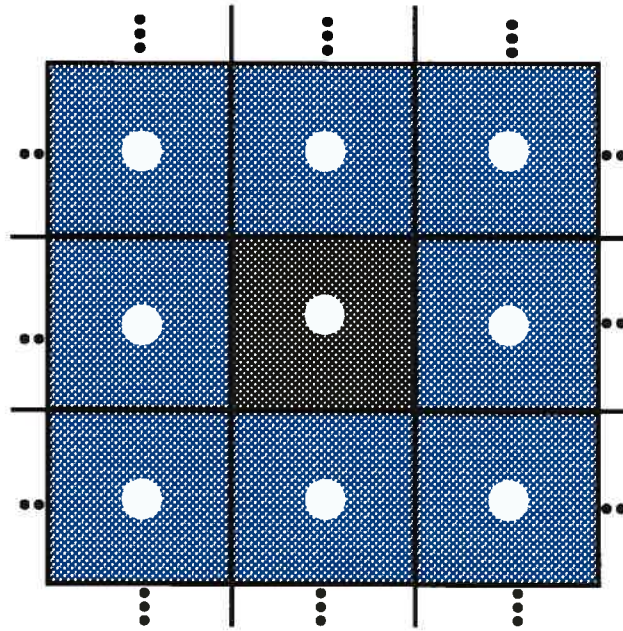


FIG. I.1 – Les conditions aux limites périodiques génèrent un système constitué de la cellule de base et un nombre infini d’images.

2. Les densités de charge périodiques et aperiodiques (indépendante de L);
3. Les densités aperiodiques dans différentes supercellules (dépend de L).

La densité de charge périodique ne doit pas contribuer à la charge nette de la supercellule. Généralement, la densité de charge périodique est neutralisée par un plan compensateur ayant une densité de charge uniforme et de signe opposé $n_0 = -q/V$ (où q est la charge nette et V est le volume de la supercellule).

Les deux premières contributions peuvent donc être remplacées par la réponse diélectrique du milieu périodique à la densité aperiodique. Le défaut chargé est supposé d’être noyé dans un milieu diélectrique (le volume) et toutes les interactions indésirables sont remplacées par la réponse diélectrique du matériau. En général, la constante diélectrique (ϵ) expérimentale du milieu est utilisée. Il est, cependant, plus correct d’utiliser la constante diélectrique calculée par le code *ab-initio* lui-même.

Il reste à trouver la relation entre la densité aperiodique et ses multipôles et la dimension de la supercellule L . Pour une supercellule de taille L , la correction apportée à l’énergie $E(L)$ est exacte dans la limite de L large. Makov et Payne [111] proposent une

correction de la forme² :

$$E_{L \rightarrow \infty} = E(L) - \frac{q^2 \alpha}{2L\epsilon} - \frac{2\pi q Q \alpha}{3L^3 \epsilon} + O(L^{-5}) \quad (I.2)$$

Le premier terme correspond à l'interaction monopole ayant la forme d'une énergie de Madelung. q est la charge nette, ϵ est la constante diélectrique du milieu et α est la constante de Madelung (2.83373, 2.8883 et 2.885 pour les supercellules cubiques : simple, centrée et à faces centrées respectivement). Le second terme est le terme quadripôle. $Q = \int r^2 \Delta n(r) d^3r$ est le quadripôle ou le second moment radial de la densité électronique $n(r)$ du système a périodique.

Il est très courant de corriger l'énergie totale avec le premier terme seulement [52]. Ce terme varie comme le carré de la charge. Dans certains cas, la correction monopole est surestimée et dans d'autres cas les termes d'ordres plus élevés doivent être considérés. Il existe trois cas critiques où la correction totale de Makov et Payne pourrait devenir très importante [21] :

1. Les semiconducteurs qui contiennent des éléments de la première ligne du tableau périodique comme le carbone pour lequel la constante du réseau et la distance entre les défauts voisins est petite. Par conséquent, le terme $1/L^3$ peut devenir important ;
2. Les semiconducteurs avec une constante diélectrique ϵ faible ;
3. Si la charge nette est élevée comme +3, +4, -3, -4.

Si ces trois critères sont satisfaits en même temps, la correction peut atteindre plusieurs électronvolts. À titre d'exemple, la correction du premier ordre calculée par Bockstedt *et al.* [6] pour les défauts de charge +4 dans le SiC est de ≈ 4 eV. Pour les défauts doublement chargés dans le SiC, la correction totale est de l'ordre de 1 eV. Une telle correction peut même inverser l'ordre de transition des niveaux du défaut . Une transition, qui induit un seul électron avant la correction, peut devenir de type U-négatif après la correction [53].

²La correction totale n'est possible que pour les cellules **cubiques**. Sinon, il faut utiliser des cellules aussi larges que possible.

Un article récent de Gerstmann *et al.* [112] utilise l'exemple des défauts ponctuels dans le SiC pour tester la validité de la correction de Makov et Payne. En comparant les énergies des supercellules corrigées et les énergies obtenues par les calculs avec les fonctions de Green, ils ont trouvé que :

1. Le terme quadrupôle dépend fortement de la distribution spatiale de la densité de charge ;
2. L'utilisation de l'énergie de Madelung seule donne une bonne correction pour les défauts localisés ;
3. Le terme quadrupôle peut être complètement négligé pour les molécules ou les défauts localisés ;
4. L'énergie de Madelung surestime la correction des défauts délocalisés, et doit être supprimée de la correction ;
5. L'utilisation du terme quadrupôle seul est une bonne correction pour les défauts délocalisés.

La correction complète de Makov et Payne est raisonnable si et seulement si la perturbation induite par la densité de charge est strictement localisée autour du défaut. Avant d'appliquer cette correction, il est impératif de vérifier la localité de la différence de charge entre les densités de charge du défaut dans la cellule perturbée et la densité de charge idéale non perturbée [21].

en résumé, la correction de Makov et Payne a été très largement utilisée dans la littérature pour corriger les énergies de formation des défauts. Elle est d'ailleurs utilisée dans les travaux de cette thèse dans les chapitres 6 et 8. Nous avons continué à utiliser cette méthode malgré les critiques publiées récemment à ce sujet à cause de l'absence d'une meilleure alternative et du fait que nous diminuons les effets de taille en utilisant des supercellules d'au moins 216 atomes. L'article de Castleton *et al.* [113], publié en 2006, analyse la correction de Makov et Payne ainsi que les autres corrections, comme la méthode d'alignement de potentiel introduite par Schultz [114]. Ils proposent une méthode simple et efficace pour bien corriger les effets de taille. Dans cette méthode, ils procèdent au calcul de l'énergie totale d'une série de cellules de tailles différentes ayant la même

géométrie (par exemple des cellules comportant 8, 64, 216, 512 atomes) et observent la tendance pour ensuite déduire la correction nécessaire. Les auteurs concluent que la meilleure alternative serait d'utiliser des supercellules aussi grandes que les ressources et le temps de calcul le permettent avec la méthode d'alignement de potentiel. Ces conclusions n'ont malheureusement pas été exploitées dans les travaux de cette thèse, mais nous les recommandons fortement pour les travaux futurs dans ce domaine.

En conclusion, les problèmes liés aux effets de taille sont pour le moment atténués en faisant des corrections. Nous pensons qu'ils seront considérablement réduits grâce à l'amélioration des ressources de calcul qui permettraient l'utilisation de supercellules de l'ordre des milliers d'atomes.

Annexe II

Algorithme de Lanczòs

L'algorithme de Lanczòs [85] permet d'extraire de manière efficace un nombre limité de valeurs et de vecteurs propres. Pour les systèmes de haute dimension, par exemple un système comprenant N atomes, quelques valeurs et vecteurs propres sont obtenus sans évaluer ni diagonaliser la matrice dynamique de dimension $3N \times 3N$.

Cet algorithme est basé sur une méthode récursive. La matrice dynamique locale est appliquée plusieurs fois à un vecteur de déplacement aléatoire. À partir d'un certain point dans l'espace des phases, une configuration située à \vec{x} subit un léger déplacement suivant le vecteur propre aléatoire $|\vec{x}_0\rangle$. Le résultat de l'application de la matrice H sur ce vecteur revient à l'estimation de la différence des forces exercées sur les configurations situées à $|x\rangle$ et $|x\rangle + |x_0\rangle$. La matrice dynamique n'est pas évaluée durant ce processus ; seule la valeur propre la plus basse est extraite. L'application de la matrice dynamique H sur les vecteurs propres x_i avec $i = 0, \dots, l$ donne un résultat sous forme d'une combinaison linéaire de $|x_i\rangle$ et deux autres vecteurs $|x_{i+1}\rangle$ et $|x_{i-1}\rangle$ perpendiculaires à x_i . Chaque application de la matrice dynamique est une opération d'ordre N car elle requiert une seule évaluation de la force.

$$\begin{aligned} H|x_0\rangle &= a_0|x_0\rangle + b_1|x_1\rangle \\ H|x_1\rangle &= a_1|x_1\rangle + b_1|x_0\rangle + b_2|x_2\rangle \\ H|x_2\rangle &= a_2|x_2\rangle + b_2|x_1\rangle + b_3|x_3\rangle \\ &\cdot \\ &\cdot \\ H|x_{l-1}\rangle &= a_{l-1}|x_{l-1}\rangle + b_{l-1}|x_{l-2}\rangle + b_l|x_l\rangle \\ H|x_l\rangle &= a_l|x_l\rangle + b_l|l-1\rangle \end{aligned}$$

La construction de la matrice L de dimension $l \times l$ nécessite $3lN$ opérations. L'avantage de cet algorithme est que le nombre de vecteurs utilisés l est nettement inférieur à la dimension du problème ($l \ll 3N$). Une grande partie du temps de calcul est consacrée à la construction de la matrice tridiagonale L . La diagonalisation de cette matrice est peu coûteuse comparée à la matrice dynamique entière.

$$\begin{pmatrix} a_0 & b_1 & & & & \\ b_1 & a_1 & b_2 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & b_{l-1} & a_{l-1} & b_l \\ & & & & & a_l & b_l \end{pmatrix} \quad (\text{II.1})$$

Il y a diverses méthodes pour accélérer l'algorithme de Lanczós dans TAR [115]. En général 10 à 30 vecteurs sont suffisants pour donner une bonne estimation de la valeur propre la plus basse. On gagne aussi à utiliser le vecteur propre le plus bas qui découle de l'étape d'activation comme germe pour la nouvelle opération récursive.

Annexe III

Construction de la matrice dynamique

Les éléments de la matrice dynamique (ou matrice des constantes de force) sont générés à partir des déplacements finis dans chaque direction cartésienne de chacun des atomes ($i = 1 \dots N$) du système. Chaque coordonnée atomique x_i est incrémentée d'un petit déplacement Δx_i alors que le reste des coordonnées x_j du système restent fixes. Ce déplacement a pour but de simuler les vibrations atomiques. La réponse linéaire de chacun des atomes, suite à ce déplacement, est alors enregistrée.

$$H_{i,j} = \frac{\partial^2 E}{\partial x_i \partial x_j} \quad (\text{III.1})$$

Numériquement, la seconde dérivée est obtenue par la méthode des différences finies et les éléments de la matrice sont calculés à partir des différences des forces exercées sur les atomes j :

$$H_{i,j} = \frac{\left| \frac{\partial E}{\partial x_j} \right|_{x_i + \Delta x_i} - \left| \frac{\partial E}{\partial x_j} \right|_{x_i - \Delta x_i}}{2\Delta x_i} \quad (\text{III.2})$$

On refait le déplacement Δx_i pour les $3N$ coordonnées cartésiennes. Puisque la matrice dynamique est symétrique $H_{i,j} = H_{j,i}$, les erreurs aléatoires dans le calcul du gradient peuvent être réduites en recalculant la matrice dynamique :

$$H_{i,j} = \frac{1}{2} \frac{\left| \frac{\partial E}{\partial x_i} \right|_{x_j + \Delta x_j} - \left| \frac{\partial E}{\partial x_i} \right|_{x_j - \Delta x_j}}{2\Delta x_j} + \frac{\left(\frac{\partial E}{\partial x_j} \right|_{x_i + \Delta x_i} - \left| \frac{\partial E}{\partial x_j} \right|_{x_i - \Delta x_i}}{2\Delta x_i} \quad (\text{III.3})$$

La diagonalisation de la matrice dynamique donne des valeurs propres qui correspondent aux fréquences de vibration et des vecteurs propres qui correspondent aux modes de vibration.