

### Theory of X-ray circular dichroism and application to materials under pressure

Nadejda Bouldi

#### ▶ To cite this version:

Nadejda Bouldi. Theory of X-ray circular dichroism and application to materials under pressure. Chemical Physics [physics.chem-ph]. Université Pierre et Marie Curie - Paris VI, 2017. English. NNT: 2017PA066491. tel-01803912

### HAL Id: tel-01803912 https://theses.hal.science/tel-01803912

Submitted on 31 May 2018  $\,$ 

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



### Université Pierre et Marie Curie

ED 397 IMPMC - SOLEIL synchrotron

# Theory of X-ray circular dichroism and application to materials under pressure

par Nadejda BOULDI née MAS

Thèse de doctorat de Physique et Chimie des matériaux

Présentée et soutenue publiquement le 11 décembre 2017 devant un jury composé de:

Sakura PascarelliRapporteurLucia ReiningRapporteurChristel GervaisExaminateurAndrei RogalevExaminateurHubert EbertExaminateurAmélie JuhinCo-directeur de thèsePhilippe SainctavitCo-directeur de thèse

Membres invités:

François Baudelet Co-encadrant Christian Brouder Co-encadrant

#### Acknowledgments

This PhD has been a more than three years journey during which numerous people contributed, directly or indirectly, to my work. It is with pleasure that I take the opportunity of this page to thank them.

First, and above all, I want to express my gratitude to my supervisors: Amélie Juhin, Christian Brouder and Philippe Sainctavit at IMPMC and François Baudelet at SOLEIL. All four of them never spared a constructive comment and I learned a lot from them. People from the field know that we could hardly put a better team together to work on XMCD at K-edge and I am grateful to Philippe Sainctavit for giving a fresh momentum to this subject. I thank my supervisors for the ideas and advices they gave me, for the numerous corrections of my work (including this thesis) and also for their great kindness. In particular, Amélie Juhin perfectly managed to reconcile her great scientific rigor with a very human supervision of my PhD.

I would also like to extend my thanks to the members of the jury, Lucia Reining, Sakura Pascarelli, Christel Gervais, Andrei Rogalev and Hubert Ebert for accepting to consider this work.

The first chapter of this thesis was written in close collaboration with Christian Brouder and I am particularly grateful for his help with the theoretical problems that arose during my PhD and which seemed insurmountable at first sight. The numerical part of this thesis is the continuance of Christos Gougoussis' PhD work. Christos developed XSpectra in Quantum-Espresso and his codes served as a basis to continue implementing XMCD in XSpectra. Uwe Gerstmann implemented collinear spin-orbit-coupling in the code pw. I thank him for his numerous comments on my numerical work, in particular when I went to Paderborn in October 2015. The cross examinations of the code for XMCD by Nora Vollmers, a fellow PhD candidate in Paderborn, were useful to exclude coding mistakes. I also acknowledge the participation of Matteo Calandra in the whole process of code development. Lorenzo Paulatto was of great help for learning to use Quantum-Espresso and for understanding the subtleties of the code pw. Delphine Cabaret was very present to answer my questions and share ideas on spectroscopy calculations and I thank her for that. I would also like to thank Yves Joly for implementing the D-SP term in FDMNES which has been very useful when I had doubts on its sign and order magnitude. I also thank him for fruitful discussions on various topics all along my PhD and I know that the projects that we started together will continue to yield interesting results. The experiments of XMCD under pressure on ODE beamline at SOLEIL were all performed with Lucie Nataf and François Baudelet. The experiment on iron hydride was done before my PhD even started. I thank them both for their great and efficient involvement, for their numerous explanations on the experiment, and for interesting scientific discussions. I also thank Jessica Brest and Ludovic Delbes for their help in the preparation and DRX of the  $CrO_2$  powder.

I thank all of my colleagues and fellow PhD candidates, who made IMPMC a pleasant work environment. In particular, Maud Jusot, Manoj Ducher, Henri Hay and Florent Caste with whom I shared an office and also many friendly moments.

I gratefully acknowledge the funding sources that made my PhD work possible. This work was supported by SOLEIL synchrotron and by the Cluster of Excellence MATISSE (Investissements d'Avenir programme Reference No. ANR-11-IDEX-0004-02).

Finally, most of all, I owe a lot to my family. My parents instilled in me a great intellectual curiosity which was the fuel of all my studies. The unfailing support from my husband, Lazhare, was more than precious all along my PhD and even more since a little treasure joined the family.

ii

# Contents

| In | Introduction en français 1  |           |  |  |  |  |
|----|---|-----------|--|--|--|--|
| In | ntroduction   | 7         |  |  |  |  |
| 1  | Gauge invariance and relativistic effects in photon absorption and scattering | r<br>S    |  |  |  |  |
|    | by matter   | 13        |  |  |  |  |
|    | 1.1 Introduction  | 14        |  |  |  |  |
|    | 1.2 Time-dependent perturbation theory  | 15        |  |  |  |  |
|    | 1.3 One-body Foldy-Wouthuysen Hamiltonian and transition rate                 | 16        |  |  |  |  |
|    | 1.4 Gauge Invariance  | 22        |  |  |  |  |
|    | 1.5 Relativistic absorption and scattering cross sections                     | 25        |  |  |  |  |
|    | 1.6 Many-body Foldy-Wouthuysen transformation                                 | 29        |  |  |  |  |
|    | 1.7 Semi-relativistic absorption cross sections                               | 32        |  |  |  |  |
|    | 1.8 Semi-relativistic scattering cross section                                | 35        |  |  |  |  |
|    | 1.9 Conclusion  | 38        |  |  |  |  |
| 2  | Method for ab initio calculation of XAS, XMCD and XNCD spectra at K-          |           |  |  |  |  |
|    | $\mathbf{edge}$   | <b>43</b> |  |  |  |  |
|    | 2.1 Introduction  | 44        |  |  |  |  |
|    | 2.2 X-ray Circular Dichroism cross section                                    | 46        |  |  |  |  |
|    | 2.3 Method  | 49        |  |  |  |  |
|    | 2.4 Convergence with calculation parameters                                   | 59        |  |  |  |  |
|    | 2.5 Inclusion of the core hole  | 69        |  |  |  |  |
|    | 2.6 Conclusion  | 71        |  |  |  |  |
| 3  | Calculation of XMCD and XNCD: application to the $3d$ ferromagnetic metals    |           |  |  |  |  |
|    | and to $\alpha$ -LiIO <sub>3</sub>  | 77        |  |  |  |  |
|    | 3.1 Introduction  | 78        |  |  |  |  |
|    | 3.2 XMCD at $K$ -edge in Fe, Co and Ni $\ldots$                               | 78        |  |  |  |  |
|    | 3.3 XNCD in $\alpha$ -LiIO <sub>3</sub>                                       | 94        |  |  |  |  |
|    | 3.4 Sum-rules at $K$ - and $L_1$ -edge  | 101       |  |  |  |  |
|    | 3.5 Conclusion  | 116       |  |  |  |  |
| 4  | XAS and XMCD to study matter under pressure                                   |           |  |  |  |  |
|    | 4.1 Introduction  | 124       |  |  |  |  |
|    | 4.2 Experimental method: XMCD under pressure on a dispersive beamline         | 125       |  |  |  |  |
|    | 4.3 Pressure in the calculation: Fe   | 132       |  |  |  |  |
|    | 4.4 Iron hydride  | 138       |  |  |  |  |
|    | 4.5 Chromium Dioxide  | 152       |  |  |  |  |

| Conclusion   |  |  |
|--|--|--|
| Appendix A Abbreviations and Notations   |  |  |
| Appendix B Definitions and useful formulaB.1 Pauli MatricesB.2 Angular Momentum matrix for $l=1$ B.3 Dirac MatricesB.4 Commutators   |  |  |
| Appendix C Point groups  | 179  |  |
| Appendix D Operators in XSpectra code         D.1 Spherical harmonics          D.2 Operators          D.3 Initial vector $ \tilde{\varphi}_{\mathbf{R}_0}\rangle$ D.4 Prefactors | <b>180</b><br>. 180<br>. 181<br>. 182<br>. 184 |  |
| Appendix E Technical details of the pseudopotentials         E.1 Norm conserving (NC) pseudopotentials         E.2 Ultrasoft pseudopotentials                                    | <b>185</b><br>. 185<br>. 186                   |  |
| Appendix F Articles 1  |  |  |

# List of Figures

| 1    | Principe du XAS  | 1  |
|------|--|----|
| 2    | Effet photoélectrique  | 2  |
| 3    | Régions spectre d'absorption X   | 2  |
| 4    | Expérience de diffusion  | 3  |
| 5    | Principe des mesures de XMCD   | 4  |
| 6    | Principle of XAS   | 7  |
| 7    | Photoelectric effect   | 8  |
| 8    | X-ray absorption regions   | 8  |
| 9    | Scattering experiment  | 9  |
| 10   | Principle of XMCD experiment   | 10 |
| 2.1  | Illustration of muffin-tin potentials  | 45 |
| 2.2  | Illustration on pseudopotentials   | 51 |
| 2.3  | Augmentation regions in the PAW framework  | 53 |
| 2.4  | Contributions of the different terms of the PAW SOC in XMCD  | 54 |
| 2.5  | Comparison of the XMCD spectra obtained by including, or not, SOC in the   |    |
|      | relaxation.  | 58 |
| 2.6  | $p\mbox{-}{\rm projected}$ density of states computed with the fully-relativistic version of QE $% p\mbox{-}{\rm projected}$ . | 59 |
| 2.7  | Magnetization and total energy $vs$ size of the $k$ -point grid  | 60 |
| 2.8  | Convergence of the spectra with the number of $k$ -points in pwscf $\ldots \ldots \ldots$                                      | 61 |
| 2.9  | Magnetization and total energy $vs$ cut-off energy. $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$                               | 62 |
| 2.10 | Convergence of the spectra with the cut-off energy   | 63 |
| 2.11 | Magnetization and total energy $vs$ spreading  | 64 |
| 2.12 | Convergence of the spectra with the spreading  | 65 |
| 2.13 | Schematic representation of the scf loop   | 66 |
| 2.14 | Convergence of the spectra with the convergence threshold  | 66 |
| 2.15 | Convergence of the spectra with the number of $k$ -points in XSpectra  | 67 |
| 2.16 | Convergence of the spectra with the error in XSpectra  | 68 |
| 2.17 | Illustration of a supercell  | 69 |
| 2.18 | Convergence of the spectra with the distance between core-holes  | 70 |
| 3.1  | Illustration on Zeeman effect.   | 80 |
| 3.2  | Illustration on spin-orbit coupling  | 80 |
| 3.3  | Table of Clebsch-Gordan coefficients for $l = 1$ and $s = \frac{1}{2}$ .   | 81 |
| 3.4  | Applied broadening for the $3d$ transition metals  | 82 |
| 3.5  | Calculated contributions to the K-edge XAS for <i>bcc</i> -Fe, <i>hcp</i> -Co and <i>fcc</i> -Ni                               | 83 |
| 3.6  | Calculated contributions to the K-edge XMCD for bcc-Fe, hcp-Co and fcc-Ni  | 84 |
| 3.7  | Comparison between the calculated and the experimental spectra at $K$ -edge for  |    |
|      | <i>bcc</i> -Fe, <i>hcp</i> -Co and <i>fcc</i> -Ni  | 85 |

| 3.8  | Effect of the core-hole at Fe and Co K-edge  | 87  |
|------|--|-----|
| 3.9  | XMCD at Ni K-edge calculated with FDMNES   | 88  |
| 3.10 | Easy and hard magnetization directions for Fe, Ni and Co   | 89  |
| 3.11 | Experimental spectra for Co: applied magnetic field and temperature  | 90  |
| 3.12 | Calculated spectra for Co: two different directions  | 91  |
| 3.13 | Quasiparticle self-energy for Cu   | 92  |
| 3.14 | Calculated <i>bcc</i> -Fe <i>K</i> -edge XMCD spectra using PBE and LDA functionals  | 93  |
| 3.15 | Calculations for $\alpha$ -LiIO <sub>3</sub> from Eur. Phys. J. B 4, 1-11 (1998)   | 95  |
| 3.16 | $\alpha$ -LiIO <sub>3</sub> structure  | 95  |
| 3.17 | Calculated XAS at I $L_1$ -edge in $\alpha$ -LiIO <sub>3</sub> .   | 96  |
| 3.18 | XAS and XNCD spectra at the I $L_1$ -edge in LiIO <sub>2</sub>   | 97  |
| 3.19 | Angular dependence of the calculated XAS and XNCD at the I $L_1$ -edge in LiIO <sub>3</sub>  | 98  |
| 3.20 | Contributions to XAS at the L $L_1$ -edge in LiIO <sub>2</sub> with and without core-hole  | 99  |
| 3.21 | XAS and XNCD at the L $L_1$ -edge in LiIO <sub>2</sub> computed with fractions of 2s core-hole   | 99  |
| 3 22 | Effect of the core-hole in XAS and XNCD at the L $L_1$ - and K-edge in LiIO <sub>2</sub>   | 100 |
| 3 23 | Badial part of the wave functions for an hydrogen-like atom  | 100 |
| 3 24 | <i>n</i> -projected density of states and D-D spectrum in Fe and Ni  | 115 |
| 3 25 | $L_{z_{i}}(E)$ and $S_{z^{1,-1}}(E)$ compared to D-D and D-SP XMCD spectra in Fe and Ni  | 116 |
| 0.20 | D = p(D) and $D = p$ (D) compared to $D = D$ and $D = D$ . Three speeric in relation of the probability of | 110 |
| 4.1  | Diamond anvil cell.  | 125 |
| 4.2  | Ruby fluorescence spectra under several pressures  | 126 |
| 4.3  | SOLEIL synchrotron   | 127 |
| 4.4  | Bending magnet   | 127 |
| 4.5  | Dispersive setup   | 128 |
| 4.6  | Baseline removal XAS   | 130 |
| 4.7  | Baseline removal XMCD  | 131 |
| 4.8  | Comparison of the pressure vs lattice parameter obtained by calculation and  |     |
|      | experimentally   | 133 |
| 4.9  | Experimental XAS and XMCD spectra for Fe at ambient pressure and 10 GPa.   | 134 |
| 4.10 | Calculated total magnetic moment and amplitude of the experimental XMCD .  | 134 |
| 4.11 | Calculated XMCD and XAS spectra for Fe at ambient pressure and 10 GPa.   | 135 |
| 4.12 | Illustration for band ferromagnetism   | 136 |
| 4.13 | DOS of <i>bcc</i> -Fe under pressure   | 137 |
| 4.14 | Experimental XAS spectra for Fe under H <sub>2</sub> atmosphere with increasing pressure   |     |
|      | up to 4 GPa.   | 138 |
| 4.15 | Experimental XMCD spectra for Fe under H <sub>2</sub> atmosphere of increasing pressure  |     |
|      | up to 4 GPa.   | 139 |
| 4.16 | Experimental XAS and XMCD spectra at the Fe $K$ -edge for Fe under H <sub>2</sub> atmo-  |     |
|      | sphere of increasing pressure.   | 140 |
| 4.17 | Crystal structure of $dhcp$ -FeH $\ldots$  | 140 |
| 4.18 | Volume as a function of pressure FeH   | 141 |
| 4.19 | Calculated contributions to the XAS and XMCD spectra in FeH.   | 143 |
| 4.20 | Comparison experiment/calculation FeH  | 143 |
| 4.21 | Spectra for FeH with and without core hole   | 144 |
| 4.22 | Relaxed position of H atoms in FeH   | 145 |
| 4.23 | Effect of the position of H on the XMCD spectra.   | 147 |
| 4.24 | Effect of the position of H on the <i>d</i> DOS  | 148 |
| 4.25 | 2D spin density in FeH   | 149 |

| 4.26 | Total energy vs volume for FM FeH and NM FeH   | 150 |
|------|--|-----|
| 4.27 | Calculated XAS and XMCD spectra at the Fe $K$ -edge in FeH at several pressures.   | 151 |
| 4.28 | Left: CrO <sub>2</sub> structure (Figure from [Yamasaki et al., 2006]). We define the local  |     |
|      | coordinate systems with $\mathbf{z}_{1,2}$ pointing exactly to an O atom and $\mathbf{x}_{1,2}$ almost   |     |
|      | pointing to an O atom. Right: the unit cell in the local coordinates of the  |     |
|      | octahedron around Cr1.   | 153 |
| 4.29 | Energy levels of $Cr^{4+}$ in $CrO_2$  | 154 |
| 4.30 | Projected DOS in $CrO_2$   | 155 |
| 4.31 | XAS of $CrO_2$ at the Cr K-edge $\ldots \ldots \ldots$ | 156 |
| 4.32 | Experimental absorption spectra and DOS with and without core hole 1   | 157 |
| 4.33 | Interpretation of the core-hole effect on $K$ pre-edge in $CrO_2$  | 158 |
| 4.34 | XRD pattern of the $CrO_2$ powder  | 159 |
| 4.35 | Experimental XAS and XMCD under pressure in CrO2 at room temperature   |     |
|      | and at $T=15K$ .   | 161 |
| 4.36 | Experimental and calculated XAS and XMCD in $CrO_2$  | 162 |
| 4.37 | Contributions to XMCD at the Cr $K$ -edge in $CrO_2$   | 162 |
| С.1  | Stereographic projections of the crystal classes.  | 179 |

# List of Tables

| 2.1 | Orders of magnitude of the operators in the absorption cross section  |
|-----|---|
| 3.1 | Selection rules of the operators  |
| 3.2 | Fe, Co and Ni crystal structure   |
| 3.3 | Electric dipole XAS sum-rule  |
| 3.4 | Electric quadrupole XAS sum-rule  |
| 3.5 | Cross term electric dipole-electric quadrupole XAS sum-rule   |
| 3.6 | Cross term electric dipole-spin-position XAS sum-rule   |
| 4.1 | <i>dhcp</i> -FeH lattice parameters and atomic positions  |
| 4.2 | Integral of the XMCD spectrum in Fe and FeH   |
| 4.3 | $CrO_2$ lattice parameters and atomic positions $\dots \dots \dots$ |
| D.1 | Correspondence between variables in XS<br>pectra and corresponding quantities $\therefore$ . 183  |
| E.1 | Technical details of the norm-conserving pseudopotentials   |
| E.2 | Technical details of the ultrasoft pseudopotentials   |

Nasreddine ! Je suis surpris de voir chaque jour les gens courir de tous côtés.
Il faut s'en réjouir, mon ami, car si tous allaient du même côté, la terre perdrait l'équilibre.

Contes des sages et facétieux Djeha et Nasreddine Hodja – Jean Muzi

### Introduction en français

Le mot **spectroscopie** est composé d'une première partie d'origine latine, *spectrum* (apparition, vision) et d'une seconde partie d'origine grecque  $\sigma \times \sigma \pi \epsilon \omega$  (skop $\epsilon \bar{o}$ , voir). Comme ce nom l'indique, les techniques de spectroscopie consistent à étudier des fantômes.

C'est, par exemple, le cas de la spectroscopie d'absorption X (XAS) dont il sera beaucoup question dans cette thèse. Le principe de cette expérience est très simple: il consiste à éclairer l'échantillon que l'on souhaite étudier avec des rayons X d'énergie choisie et de mesurer l'intensité lumineuse transmise. Par comparaison avec l'intensité lumineuse incidente, il est possible de savoir quelle quantité de lumière a été absorbée par la matière. On mesure donc ce que l'échantillon a retiré à la lumière : il s'agit du spectre du matériau. Plus précisément, si, conformément à la Figure 1, on note  $I_0$  l'intensité incidente et I l'intensité transmise, la mesure d'absorption consiste à évaluer une quantité proportionnelle à  $\ln\left(\frac{I_0}{I}\right)$ . Cette absorption est généralement tracée en fonction de l'énergie E des rayons X incidents (elle pourrait de manière équivalente être tracée en fonction de la longueur d'onde ou de la fréquence des rayons X car ces quantités sont liées par des relations très simples). Lorsque l'intensité est suffisamment faible, on constate qu'elle décroit exponentiellement quand la lumière se propage selon z dans l'échantillon suivant la relation  $I(z) = I_0 e^{-\mu z}$  où  $\mu$  est appelé coefficient d'absorption linéaire des rayons X.<sup>1</sup>





A l'échelle microscopique, l'absorption des rayons X par le matériau est décrite par la section efficace d'absorption qui est définie comme le nombre de photons absorbés par centre absorbeur divisé par le nombre de photons incidents par unité de surface perpendiculaire à la direction de propagation. Il s'agit en fait de la surface opaque équivalente à un centre absorbeur qui serait nécessaire pour obtenir la même absorption (d'où le nom de section efficace). L'intensité absorbée dans une lamelle dz d'échantillon est  $I(z)\mu dz$ , le nombre de centres absorbeurs est  $\rho_a Sdz$  où  $\rho_a$  est la densité de centres absorbeurs (qui s'exprime en atomes/volume)

<sup>&</sup>lt;sup>1</sup>S. Ravy (2013) Cours "Structure de la matière condensée" (https://www.lps.u-psud.fr/spip.php? article531)

et l'intensité par unité de surface est I(z)/S. La section efficace s'exprime donc en fonction du coefficient d'absorption linéaire comme  $\sigma = \frac{\mu}{\rho_a}$ .

La section efficace est habituellement exprimée en barns (1 barn =  $10^{-28}$  m<sup>2</sup>). Elle dépend de la nature des atomes absorbeurs mais aussi de leur environnement. Dans la gamme d'énergie des rayons X (100 eV à 100 keV), le processus responsable de l'absorption est l'effet photoélectrique dont le principe est illustré Figure 2.



**Fig. 2:** Illustration de l'effet photoélectrique. Le photon est absorbé par l'atome si son énergie est supérieure à l'énergie de liaison d'un électron. L'électron excité est appelé *photoélectron* et l'état vacant dans son niveau d'origine est appelé *trou de coeur*.

Quand l'énergie des photons incidents correspond à l'énergie de liaison d'un électron, il y a une augmentation soudaine de l'absorption. Cette résonance est appelée *seuil*. Par exemple, le seuil K correspond à l'excitation d'un électron 1s.



Fig. 3: Courbe d'absorption au seuil K du fer dans Fe<sub>2</sub>O<sub>3</sub>. La partie non oscillante correspond à l'absorption atomique. Le domaines XANES correspond aux énergies proche du seuil (jusqu'à 50 eV) et le domaine EXAFS aux plus hautes énergies. La région juste avant le seuil est appelée *pré-seuil*.

Un exemple de spectre d'absorption au voisinage du seuil est représenté Figure 3. On distingue deux régions sur ce spectre :

- la région XANES (pour X-ray absorption near edge structure) proche du seuil et

- la région EXAFS (pour *extended X-ray absorption fine structure*) aux plus hautes énergies. Dans la région XANES, le spectre dépend de la structure électronique de l'atome absorbeur ainsi que de la structure locale autour de celui-ci alors que la région EXAFS permet surtout d'obtenir des information sur la structure locale (types de ligands, distances...). Dans cette thèse, nous nous intéressons seulement à la région XANES.

Un autre type de spectroscopie est la diffusion. Dans le cas de la diffusion, le faisceau lumineux est dévié de sa direction initiale (voir le schéma Figure 4).

**Fig. 4:** Principe d'une expérience de diffusion.  $\mathbf{k}_i$  et  $\mathbf{k}_f$  sont les vecteurs d'ondes incident et diffusé et  $\hbar\omega_i$  et  $\hbar\omega_f$  sont les énergies correspondantes.

L'énergie des photons diffusés peut être la même que celle des photons incidents (diffusion élastique) ou elle peut être différente (diffusion inélastique). On parle de diffusion résonante quand l'expérience est effectuée à une énergie proche de l'énergie d'un seuil d'absorption.

De même que pour l'absorption, la section efficace correspond au nombre de photons diffusés par centre absorbeur divisé par le nombre de photons incidents par unité de surface perpendiculaire à la direction de propagation. Comme la lumière émergente est recueillie par le détecteur dans une région donnée de l'espace et pour une énergie donnée, on considère en générale la section efficace différentielle  $\frac{d^2\sigma}{d\Omega d\omega}$  qui est la section efficace par unité d'énergie et par unité d'angle solide. Les spectroscopies de diffusion sont abordées de manière marginale dans cette thèse.

Le mot **dichroïsme** est, lui, dérivé du grec  $\delta$ í $\chi$ poo $\varsigma$  (de deux couleurs). Ce mot a plusieurs acceptions en physique, plus ou moins proches de son sens littéral. Le dichroïsme de rayons X qui nous intéresse ici désigne la propriété d'un matériau à absorber différemment des rayons X de polarisations orthogonales :<sup>2</sup> verticale et horizontale pour le dichroïsme linéaire ou droite et gauche pour le dichroïsme circulaire qui est le sujet principal de cette thèse.

Quand la lumière polarisée circulairement est absorbée différemment par un matériau selon qu'elle est droite ou gauche, cela signifie qu'une symétrie est brisée. On distingue deux types de dichroïsme circulaire selon la nature de la symétrie brisée :

- Le dichroïsme circulaire naturel (XNCD pour X-ray natural circular dichroism) est lié à la brisure de la symétrie d'inversion dans les échantillons non centrosymétriques.

- Le dichroïsme circulaire magnétique (XMCD pour X-ray magnetic circular dichroism) est lié à la brisure de la symétrie par renversement du temps dans les échantillons magnétiques.

Les mesures de XNCD sont encore assez rares mais le développement du magnétisme moléculaire et l'apparition de composés à la fois chiraux et magnétiques ont suscité un intérêt nouveau pour cette technique qui peut être facilement combinée avec les mesures de dichroïsme magnétique.

Le XMCD, en revanche, est très utilisé depuis les années 80 pour étudier le magnétisme de matériaux complexes. Le principe de l'expérience est illustré Figure 5. Son grand intérêt, par rapport aux mesures de magnétométries, réside dans le fait que l'on peut choisir le type d'atome que l'on veut sonder dans la structure. Pour cela, il suffit de se placer à une énergie

<sup>&</sup>lt;sup>2</sup>A. Rogalev, J. Goulon, F. Wilhelm and A. Bosak (2010) X-ray dectected optical activity. In E. Beaurepaire, H. Bulou, F. Scheurer and J.-P. Kappler, editors, *Magnetism and Synchrotron Radiation*, number 133 in Springer Proceedings in Physics, pages 169-190. Springer Berlin Heidelberg

correspondant au seuil d'absorption de l'élément choisi (on dit que la mesure est sélective en élément). Le succès du XMCD est lié à l'existence de règles de sommes magnéto-optiques. Aux seuils correspondant à des niveaux où il y a une levée de dégénérescence due au couplage spin orbite (comme les seuils  $L_{2,3}$  qui correspondent à des trous  $2p_{1/2}$  et  $2p_{3/2}$ ), ces règles de sommes lient des combinaisons d'intégrales des spectre XAS et XMCD aux moments d'orbite et de spin dus à certains électrons.<sup>3</sup> Pour les autres seuils, comme le seuil K (excitation d'un électron 1s), une seule règle de sommes magnétique existe (basée sur des hypothèses dont la validité est difficile à vérifier). Celle-ci lie l'intégrale du spectre au moment orbital dû aux électrons p alors que, dans les éléments de transition 3d, le magnétisme est principalement porté par les électrons d.



**Fig. 5:** Principe des mesures de dichroïsme circulaire magnétique de rayons X (XMCD): l'absorption des rayons X polarisés circulairement à droite (schématisés par la flèche rouge) ou à gauche (schématisés par la flèche verte) est mesurée. La différence de ces deux absorptions est le signal de dichroïsme circulaire. L'échantillon est positionné dans un champ magnétique dirigé le long de la direction de propagation des rayons X. Les spectres représentés ici sont les spectres obtenus au seuil K du Fe dans une feuille de fer métallique.

Le XMCD au seuil K est donc extrêmement complexe à interpréter. Une autre difficulté est que le signal est faible (la différence entre les absorptions au seuil K des rayons X polarisés circulairement à droite et à gauche atteint seulement quelques millièmes du saut au seuil). Pourtant, de nombreuses expériences de XMCD sont menées au seuil K car, dans les éléments de transition 3d, le seuil K correspond à des rayons X durs (> 5 keV). Ces rayons X très énergétiques peuvent pénétrer plus profondément la matière que les rayons X mous, ce qui permet de sonder des échantillons dans des environnements contraignants comme des cellules pour liquides ou encore des cellules à enclumes de diamant. Ces dernières sont utilisées pour appliquer une pression (pouvant atteindre des centaines de GPa) sur l'échantillon. Les études sous pression sont très importantes pour les sciences de la terre car, combinées aux hautes températures, elles permettent de reproduire des conditions proches de celles de l'intérieur terrestre. La pression est aussi beaucoup utilisée en science des matériaux pour son intérêt fondamental : en changeant les distances interatomiques, la pression modifie les interactions entre les atomes ce qui donne lieu à de nouvelles propriétés de la matière.

<sup>&</sup>lt;sup>3</sup>F. de Groot, A. Kotani (2008). *Core level spectroscopy of solids*, Advances in Condensed Matter Science. CRC Press

L'interprétation des spectres de XMCD au seuil K et la compréhension des spectres de XNCD, qui sont encore semi-quantitatives, nécessitent l'apport de nouveaux éléments théoriques. Ce que nous proposons est de calculer les spectres *ab initio*, c'est à dire à partir des premiers principes de la physique. Pour cela, il faut calculer les spectres d'absorption de la lumière polarisée circulairement droite et gauche avec une grande précision afin de pouvoir obtenir, en calculant leur différence, le dichroïsme circulaire.

Pour calculer la section efficace avec une précision suffisante, il est nécessaire de connaitre son expression formelle à un ordre suffisamment élevé. C'est l'objet du **Chapitre 1**. Dans ce chapitre, on exprime la section efficace d'absorption et de diffusion en incluant les effets relativistes. Un point fondamental est aussi discuté: l'invariance de jauge. On montre que la démarche usuelle ne garantit pas cette invariance et on propose une nouvelle manière de mener les calculs pour aboutir à un résultat correct. De cette manière, nous obtenons un terme supplémentaire par rapport aux approches proposées dans la littérature. Nous avons nommé ce terme "spin-position". Cette expression de la section efficace d'absorption semi-relativiste (Section 1.7) a été utilisée pour l'implémentation numérique présentée dans la suite de la thèse.

Le Chapitre 2 est consacré à la présentation de la méthode utilisée pour le calcul numérique des spectres. Une modélisation des matériaux utilisant la théorie de la fonctionnelle de la densité (DFT), qui a déjà été utilisée avec succès pour modéliser les spectres XAS, est appliquée. L'implémentation a été réalisée dans le code XSPECTRA de la suite QUANTUM ESPRESSO. L'idée de la méthode numérique est d'effectuer un calcul auto-cohérent de la structure électronique en présence d'un trou de cœur et de calculer la section efficace à partir des états obtenus. Dans ce chapitre, la manière dont la méthode est implémentée est discutée en détail ainsi que l'effet des différents paramètres de convergence.

Les résultats obtenus pour des systèmes modèles (XMCD du fer, cobalt et nickel et XNCD de l'iodate de lithium) sont présentés et discutés dans le **Chapitre 3**. Ces calculs sont l'occasion de discuter le contenu physique des spectres de dichroïsme circulaire et de clarifier les approximations de notre méthode numérique. On voit dans ce chapitre que le terme relativiste spin-position est loin d'être négligeable pour le XMCD. La dernière partie du chapitre est consacrée aux règles de sommes qui permettent de mieux comprendre l'importance des différentes contributions au XMCD au seuil K.

Enfin, le **Chapitre 4** est consacré à la présentation d'études du magnétisme sous pression par XAS et XMCD. On présente le dispositif expérimental utilisé mais aussi la manière d'inclure la pression dans les calculs. Deux matériaux d'intérêt technologique et fondamental sont étudiés: l'hydrure de fer FeH et le dioxyde de chrome. FeH est obtenu par compression d'une feuille de fer dans une atmosphère d'hydrogène, son étude concerne donc les hautes pressions de manière inhérente. Ici, les calculs appuient et enrichissent l'interprétation des expériences. Pour  $CrO_2$ , malgré le fait que ses propriétés ferromagnétiques ont été largement utilisées pour l'enregistrement magnétique sur les cassettes, sa structure électronique qui en fait un demimétal ferromagnétique est encore largement incomprise. Dans cette étude, les hautes pressions sont un moyen de changer l'environnement des atomes de chrome qui, d'après les modèles proposés, pourraient jouer un rôle très important pour expliquer le ferromagnétisme de  $CrO_2$ .

Quoique n'étant pas totalement indépendants, ces quatre chapitres peuvent être lus isolément. En particulier, la lecture du premier chapitre, qui est très formel, n'est pas nécessaire à la compréhension des trois autres chapitres. Dans ce manuscrit sont présentés des résultats théoriques, numériques et expérimentaux centrés sur le dichroïsme de rayons X. Pendant cette thèse, j'ai été affiliée au synchrotron SOLEIL et à l'IMPMC (Institut de Minéralogie de Physique des matériaux et de cosmochimie). En tant que membre de la ligne de lumière ODE, j'ai eu l'opportunité de mener et de participer à des expériences de XAS et XMCD sous pressions (en cellule à enclumes de diamant) qui se sont avérées très enrichissantes. J'ai aussi eu la chance de participer à des expériences de XAS et XMCD sur la ligne ID12 de l'ESRF et de RIXS-MCD sur les lignes ID26 de l'ESRF et la ligne GALAXIES à SOLEIL. A l'IMPMC, j'ai effectué aussi bien des calculs théoriques sur papier ou sur tableau blanc que des calculs numériques parallélisés qui nécessitaient l'utilisation de machines de calculs de hautes performances.

Cette diversité d'approches a rendu ce travail de thèse extrêmement intéressant. J'espère que, au-delà des nouveautés théoriques et méthodologiques qu'il présente, ce manuscrit saura clarifier les questions liées aux calculs de dichroïsme de rayons X.

NB: Ce manuscrit est rédigé en anglais. Il contient de nombreuses abréviations et des notations mathématiques diverses qui sont listées à la fin (p. 174). Comme les références bibliographiques se recoupent peu entre les chapitres, une bibliographie est présente à la fin de chacun d'entre eux.

## Introduction

The word **spectroscopy** is composed of a first Latin-based part, *spectrum* (appearance, vision) and a second part derived from the Greek word  $\sigma \times \sigma \pi \delta \omega$  (skopéō, to see). As the etymology indicates, spectroscopy techniques consist in studying ghosts.

It is, for example, the case of X-ray Absorption Spectroscopy (XAS) that we will study and discuss at length in this thesis. The principle of this experiment is far from being sophisticated: one simply illuminates a sample with an x-ray beam of chosen energy and measures the intensity of the transmitted light which, by comparison with the incident luminous intensity, allows to determine the quantity of absorbed light by the material. Therefore, we measure what has been withdrawn from the light by the material: it is the spectrum of the sample. More precisely if, as shown in Fig. 6, we denote  $I_0$  the incident intensity and I the transmitted one, absorption measurements consist in assessing a quantity proportional to  $\ln\left(\frac{I_0}{I}\right)$ . This absorption is generally plotted against the energy E of the incident X-ray beam (it would be equivalent to plot it against the wave length or the frequency of the X-rays as these quantities are linked to each other by very simple relations). When the intensity is weak enough, it decreases exponentially when the light propagates through the sample according to the relation  $I(z) = I_0 e^{-\mu z}$  with  $\mu$  being called the linear X-ray absorption coefficient.<sup>4</sup>



Fig. 6: Principle of X-ray Absorption Spectroscopy (XAS). The incident intensity  $I_0$  and the transmitted intensity I are measured.

At the microscopic scale, the absorption of X-rays by the sample is described by the cross section that is defined as the ratio of the number of photons absorbed per absorbing center divided by the number of incident photons per surface unit perpendicular to the propagation direction. It corresponds to the opaque surface that would provide the same effect as an absorbing center. The intensity absorbed in the sample slice dz is  $I(z)\mu dz$ , the number of absorbing centres is  $\rho_a S dz$  with  $\rho_a$  being the density of absorbing centers (expressed in atoms/volume) and the intensity per surface unit is I(z)/S. Therefore, the cross section is expressed in term of the linear absorption coefficient as  $\sigma = \frac{\mu}{\rho_a}$ .

<sup>&</sup>lt;sup>4</sup>S. Ravy (2013) Lecture notes "Structure de la matière condensée" (https://www.lps.u-psud.fr/spip.php?article531)

The cross section, usually expressed in barns (1 barn =  $10^{-28}$  m<sup>2</sup>), depends both on the nature of the absorbing atoms and on their environment. In the energy range of X-rays (100 eV to 100 keV), the process responsible for the absorption is the photoelectric effect, which principle is illustrated in Fig. 7.



Fig. 7: Illustration of the photoelectric effect. The photon is absorbed by the atom if its energy is higher than the binding energy of an electron. The excited electron is called a *photoelectron* and the empty space in its initial energy level is called a *core hole*.

When the energy of the incident photon matches the binding energy of an electron, there is a sudden increase in absorption. This resonance is called an *edge*. For example, the K-edge corresponds to the excitation of a 1s-electron.



Fig. 8: Absorption curve at the K-edge of iron in  $Fe_2O_3$ . The non-oscillating part corresponds to the atomic absorption. The XANES region corresponds to the energies close to the edge (up to 50 eV) whereas the highest energies are within the EXAFS region. The region right before the edge is qualified as *pre-edge*.

An example of absorption spectrum near the edge is shown on Fig. 8. Two regions can be seen on this spectrum:

- The XANES region (for X-ray absorption near edge structure) close to the edge and

- The EXAFS region (for extended X-ray absorption fine structure) corresponding to higher energies.

In the XANES region, the spectrum depends on the electronic structure of the absorbing atom and on its local environment and the EXAFS region provide mainly information on the local structure (ligands type, distances...). In this thesis, we will focus exclusively on XANES region.

Another type of spectroscopy is scattering. In the case of scattering, the light beam is deviated from its original course (see the schematic drawing Fig. 9).



Fig. 9: Principle of a scattering experiment.  $\mathbf{k}_i$  and  $\mathbf{k}_f$  are respectively the incident and emerging wave vectors,  $\hbar\omega_i$  and  $\hbar\omega_f$  are the corresponding energies.

Energy of the scattered photons can be the same as the one of incident photons (elastic scattering) or can be different (non-elastic scattering). The term resonant scattering is used when the experiment is done with an energy corresponding to an absorption edge.

As in the case of absorption, the scattering cross section corresponds to the ratio of the number of scattered photons per absorbing center divided by the number of incident photons per surface unit perpendicular to the propagation direction. As the outgoing light is collected by the detector for a spatially delimited region and a given energy, one generally considers the differential cross section  $\frac{d^2\sigma}{d\Omega d\omega}$  that is the cross section per unit of energy and unit of solid angle. Scattering spectroscopies are marginally addressed within this thesis.

The word **dichroism** is, for its part, derived from the Greek  $\delta(\chi\rho oo\zeta)$  (two-colored). This word has several meanings in physics, more or less in accordance with its literal signification. X-ray dichroism, that we are interested in here, designates the capacity for a material to absorb differently X-rays of orthogonal polarization.<sup>5</sup> vertical and horizontal for linear dichroism or right and left for circular dichroism which is the main subject of this thesis.

When circularly polarized light is differently absorbed by a material depending on whether it is left or right, it means that a symmetry is broken:

- X Ray Natural circular dichroism (XNCD) is linked to the breaking of inversion symmetry in non-centrosymmetric samples.

- X ray Magnetic circular dichroism (XMCD) is linked to the breaking of time inversion symmetry in magnetic samples.

XNCD experiments are still quite rare but the development of molecular magnetism and the appearance of both chiral and magnetic compounds have triggered a renewed interest for this technique that can easily be combined with magnetic dichroism measurements.

XMCD, on the other hand, is widely used since the eighties to study the magnetism of complex materials. The experiment principle is illustrated in Fig. 10. Its greatest strength, compared to magnetometry measurements, lies in the possibility to choose the type of atom to probe within the structure. In order to do so, it is sufficient to choose an energy corresponding to the absorption edge of the chosen element (the measurement is element selective). The success of XMCD comes from the existence of magneto-optic sum rules. At the edges corresponding to spin orbit split core levels (like the  $L_{2,3}$  edges which correspond to  $2p_{1/2}$  and  $2p_{3/2}$  core hole), these sum rules relate combinations of integrals of XAS and XMCD spectra to the orbital and

<sup>&</sup>lt;sup>5</sup>A. Rogalev, J. Goulon, F. Wilhelm and A. Bosak (2010) X-ray dectected optical activity. In E. Beaurepaire, H. Bulou, F. Scheurer and J.-P. Kappler, editors, *Magnetism and Synchrotron Radiation*, number 133 in Springer Proceedings in Physics, pages 169-190. Springer Berlin Heidelberg

spin moments due to some electrons.<sup>6</sup> For other edges, as K-edge (excitation of a 1s electron), there is only one magnetic sum rule (based on hypothesis which validity are hard to verify) that relates the integral of the spectrum to the orbital moment due to p-electrons whereas, in 3d transition elements, magnetism is mainly carried by d-electrons.



Fig. 10: Principle of X-ray magnetic circular dichroism (XMCD) measurements: the absorption of right- (illustrated by the red arrow) or left-(illustrated by the green arrow) circularly polarized X-rays are measured. The difference is circular dichroism. The sample is placed in a magnetic field along the direction of propagation of the X-rays. Spectra represented here are the ones obtained at the K-edge of Fe, within a metal iron foil.

XMCD at the *K*-edge is therefore extremely complex to interpret. Another difficulty is that the signal is weak (the difference between the absorption at the *K*-edge of right and left circularly polarized X-rays only reaches few thousandths). Nevertheless, many XMCD experiments are conducted at the *K*-edge because in 3d transition elements, the *K*-edge corresponds to hard X-rays (> 5 keV). Those highly energetic X-rays can penetrate matter deeper than soft X-rays and thus enable to probe materials in constraining environments as liquid cells or diamond anvil cells. The latter are used in order to apply pressure (that reaches hundreds of GPa) on the sample. Studies under pressure provide valuable information for Earth sciences because, combined with high temperatures, they offer the ability to reproduce conditions fairly similar to the ones existing in the inside of the Earth. Pressure is also much used in material science for its fundamental interest: by changing the interatomic distances, pressure modifies the interactions between atoms, which lead to new properties of matter.

<sup>&</sup>lt;sup>6</sup>F. de Groot, A. Kotani (2008). *Core level spectroscopy of solids*, Advances in Condensed Matter Science. CRC Press

The interpretation of K-edge XMCD spectra and the understanding of XNCD spectra, that are still semi-quantitative, require new theory tools. Our proposal is to calculate spectra *ab initio*, that is to say from first principles. In order to do so, we need to calculate the absorption spectrum of right- and left-circularly polarized light with a great accuracy in order to be able to obtain circular dichroism, by calculating their difference.

In order to calculate the cross section with sufficient accuracy, it is necessary to know its formal expression at a high enough order. This is the aim of **Chapter 1**. In this chapter, we express the absorption and scattering cross sections by taking into account relativistic effects. A fundamental point is discussed: gauge invariance. We prove that the usual approach does not ensure this invariance and we propose a new method to obtain a reliable result. With this method, we obtain an additional term compared to the usual calculations. We have named this term "spin-position". The expression of the semi-relativistic cross section (Section 1.7) has been used for the numerical implementation presented in the remainder of the thesis.

**Chapter 2** is dedicated to the presentation of the method used for numerical calculation of spectra. A modeling of materials using the density functional theory (DFT), which has been successfully employed to model XAS spectra, is applied. The implementation has been done in the code XSPECTRA of the QUANTUM ESPRESSO suite. The idea of the numerical method is to perform a self consistent calculation of the electronic structure in the presence of a core hole and to compute the cross section from the obtained wave functions. In this chapter, the way the method is implemented is presented and so are the effects of different convergence parameters.

The obtained results for model systems (XMCD of iron, cobalt, nickel and XNCD of lithium iodate) are presented and discussed in **Chapter 3**. These calculations serve as basis to discuss the physical content of circular dichroism and to clarify the approximations made in the numerical method. We find in this chapter that the relativistic spin-position term is far from being negligible for XMCD. The other part of this chapter is dedicated to sum rules which shed light on the importance of the different contributions to XMCD.

Finally, **Chapter 4** presents studies of magnetism under pressure by XAS and XMCD. The experimental set-up is presented and so is the way to include pressure in the calculations. Two materials of technological and fundamental interests are studied: iron hydride FeH and chromium dioxide. FeH is obtained by application of pressure on a Fe foil in a hydrogen atmosphere. A study of FeH therefore automatically includes high-pressure. Here, calculations confirm and enrich the interpretation of experiments. For  $CrO_2$ , despite the fact that its ferromagnetic properties have been widely used for magnetic recording in tapes, the electronic structure that makes it a ferromagnetic half metal remains largely misunderstood. In this study, high pressure is a way to change the environment of the chromium atoms and this environment, according to the proposed models, could play an important role to explain ferromagnetism in  $CrO_2$ .

While not being completely independent, these four chapters can be read separately. In particular, the first chapter, which is very formal, is not essential for the understanding of the other three chapters.

In this thesis, theoretical, numeric and experimental results about X-ray dichroism are presented. During my PhD, I was affiliated to both SOLEIL synchrotron and IMPMC (Institut de Minéralogie de Physique des matériaux et de cosmochimie). As a member of ODE beamline, I had the opportunity to both lead and be a member of experiments of XAS and XMCD under pressure (involving diamond anvil cells) which happened to be extremely enriching. I also had the chance to participate in XAS and XMCD experiments on the ID12 beamline at ESRF and in RIXS-MCD experiments on the beamline ID26 at ESRF and on the line GALAXIES at SOLEIL. At IMPMC, I made theoretical calculations on papers or on whiteboard as well as parallelized numerical calculations that required the use of high-performance computing clusters.

This diversity of approaches has made this PhD work extremely interesting. I hope that, beyond the theoretical and methodological innovations it introduces, this manuscript will be able to clarify questions related to X-ray dichroism calculations.

NB: This manuscript is written in English. It contains a large number of abbreviations and mathematical notations that are listed at the end (p.174). As the bibliographic references almost do not overlap between chapters, a bibliography is present at the end of each of them.

# Chapter 1

# Gauge invariance and relativistic effects in photon absorption and scattering by matter

#### Contents

| 1.1 | $\mathbf{Intr}$      | oduction  | 1 |
|-----|----------------------|---|---|
| 1.2 | $\operatorname{Tim}$ | e-dependent perturbation theory   | 1 |
| 1.3 | One                  | -body Foldy-Wouthuysen Hamiltonian and transition rate  | 1 |
|     | 1.3.1                | One-body Foldy-Wouthuysen transformation and Hamiltonian  |   |
|     |                      | Time-independent case   |   |
|     |                      | Time-dependent case   |   |
|     | 1.3.2                | Naive application to absorption cross section calculation $\ldots$ $\ldots$ $\ldots$                                |   |
|     |                      | Calculation from $H^{\rm FW}$   |   |
|     |                      | Calculation from $H^{\text{TFW}}$   |   |
|     | 1.3.3                | Subtleties in the application of time-dependent Foldy-Wouthuysen trans-<br>formation to transition rate calculation |   |
| 1.4 | Gau                  | ge Invariance   | 4 |
|     | 1.4.1                | Principle of gauge invariance in classical electrodynamics  |   |
|     |                      | Particle in a field   |   |
|     | 1.4.2                | Gauge invariance in quantum mechanics   |   |
|     | 1.4.3                | Gauge invariance of transition probabilities  |   |
|     | 1.4.4                | Gauge invariance in QED   |   |
| 1.5 | Rela                 | ativistic absorption and scattering cross sections  |   |
|     | 1.5.1                | Photon matter interaction and $S$ -matrix $\ldots \ldots \ldots \ldots \ldots \ldots$                               |   |
|     | 1.5.2                | Matrix-elements of $H_{e\gamma}$ , multipole expansion  |   |
|     | 1.5.3                | Fully relativistic absorption and scattering cross section $\ldots$ .   |   |
|     |                      | Absorption cross section  |   |
|     |                      | Scattering cross section  |   |
| 1.6 | Maı                  | ny-body Foldy-Wouthuysen transformation   | 4 |
|     | 1.6.1                | Time-independent many-body Foldy-Wouthuysen transformation  | 4 |

|     | 1.6.2          | Properties of this transformation  | 32        |
|-----|----------------|--|-----------|
| 1.7 | $\mathbf{Sem}$ | i-relativistic absorption cross sections   | <b>32</b> |
|     | 1.7.1          | Derivation of the transform of $T$ in the one-body case $\ldots \ldots \ldots$     | 32        |
|     | 1.7.2          | Many-body semi-relativistic absorption cross section $\ldots \ldots \ldots \ldots$ | 33        |
|     | 1.7.3          | Rewriting the cross section with large components of Dirac wave func-              |           |
|     |                | tions for the core state   | 34        |
| 1.8 | $\mathbf{Sem}$ | i-relativistic scattering cross section  | <b>35</b> |
|     | 1.8.1          | Many-body semi-relativistic scattering cross section                               | 36        |
|     |                | Positive energy intermediate states  | 36        |
|     |                | Negative energy intermediate states  | 36        |
|     | 1.8.2          | Comparison with the results from one-body semi-relativistic Hamilto-               |           |
|     |                | nians  | 37        |
|     |                | Comments on the spin-position operator   | 37        |
|     |                | Comments on $S_{<}$  | 37        |
| 1.9 | Con            | clusion  | 38        |

#### 1.1 Introduction

The description of relativistic effects in X-ray spectroscopy has become necessary because the recent experimental developments permit measurements with increased precision. The use of a semi-relativistic description is interesting because it allows to distinguish the effects related to the relativistic description of matter from the usual non-relativistic effects. Also, from a practical point of view, the use of two components wave functions opens the possibility to adapt a wide range of existing solid-state methods to spectroscopic calculation.

However, we show here that taking existing semi-relativistic Hamiltonian as a starting point to determine the absorption or scattering cross sections might be problematic. We propose an alternative approach in which we start from the fully-relativistic cross sections and we apply a Foldy-Wouthuysen transformation to the wave functions. Here, we stop the development at order  $c^{-2}$  but the orders beyond could be obtained with the same method. We also address the very problematic question of gauge invariance of transition probabilities and we propose a framework where gauge invariance is verified.

The outline of the chapter is the following:

- Section 1.2 introduces time-dependent perturbation theory in order to fix notation.
- Section 1.3 presents the Foldy-Wouthuysen (FW) transformation in the one-body case. In this section, we highlight why the straightforward way to obtain semi-relativistic cross sections (which consists in considering that the system is described by a FW Hamiltonian and computing the transition probabilities between eigenstates of this Hamiltonian) is, in fact, not valid. We nevertheless carry this naive calculation of the absorption cross section in order to have a basis for comparison.
- Section 1.4 reviews the question of gauge invariance of transition probabilities. We conclude that, to obtain a gauge invariant result, we must work in a framework where the states describe both the electronic system and the photons. We show that Quantum Electrodynamics (QED) provides such a framework in which gauge invariance of transition probabilities has been established for large classes of gauges.

- Section 1.5 is dedicated to the calculation of the fully relativistic absorption and scattering cross sections in QED.
- In Section 1.6, we derive a time-independent many-body Foldy-Wouthuysen calculation that is used in the next sections for the semi-relativistic expansion of the cross sections.
- Sections 1.7 and 1.8 present the semi-relativistic expansion of the absorption and scattering cross section. In both cases, we compare the results with the expressions that can be found in the literature.

This chapter contains numerous equations. Therefore, the proofs that are not essential for the understanding of the whole are written in footnotes.

#### **1.2** Introduction to time-dependent perturbation theory

Assume that a time-independent Hamiltonian  $H_0$  is subject to a perturbation described by a time-dependent term W(t) starting at  $t_0$ . To develop time-dependent perturbation theory for  $H(t) = H_0 + W(t)$ , we use the interaction representation:

$$|\psi(t)\rangle = e^{iH_0t/\hbar}|\psi_s(t)\rangle$$
 and  $|\psi(t_0)\rangle = |\phi_g\rangle \quad (|\psi_s(t_0)\rangle = e^{-iE_gt_0/\hbar}|\phi_g\rangle)$ 

where  $|\psi_s(t)\rangle$  is the Schrödinger wave function which verifies  $i\hbar\partial_t |\psi_s(t)\rangle = H |\psi_s(t)\rangle$ . In this representation the wave function obeys:

$$i\hbar\partial_t |\psi(t)\rangle = W_I |\psi(t)\rangle$$
 where  $W_I(t) = e^{iH_0 t/\hbar} W(t) e^{-iH_0 t/\hbar}$ .

As in the Schrödinger representation, the time evolution of  $|\psi(t)\rangle$  is governed by an *evolution* operator:

$$|\psi(t)\rangle = V_I(t,t_0)|\psi(t_0)\rangle$$

This operator is solution of  $i\hbar\partial_t V_I(t,t_0) = W_I(t)V_I(t,t_0)$ . Successive substitutions of the type  $V_I(t,t_0) = \tilde{I} - \frac{i}{\hbar} \int_{t_0}^t d\tau W_I(\tau)V_I(\tau,t_0)$  lead to (note that the n = 0 term is linked to the presence of the identity matrix  $\tilde{I}$ ):

$$V_{I}(t,t_{0}) = \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^{n} \int_{t_{0}}^{t} d\tau_{1} \dots \int_{t_{0}}^{\tau_{n-1}} d\tau_{n} W_{I}(\tau_{1}) \dots W_{I}(\tau_{n}) = \mathscr{T}\left[e^{\frac{i}{\hbar}\int_{t_{0}}^{t} d\tau W_{I}(\tau)}\right]$$
(1.1)

where  $\mathscr{T}$  is the time-ordering operator (because  $t_0 \leq \tau_n \leq ... \tau_1$ ).

If the system is at  $t = t_0$  in state  $|\phi_g\rangle$ , the transition probability to the eigenstate  $|\phi_n\rangle$  of  $H_0$  at time t is:

$$P_{ng}(t) = |\langle \phi_n | \psi(t) \rangle|^2 = |\langle \phi_n | V_I(t, t_0) | \phi_g \rangle|^2.$$
(1.2)

From Eq. (1.1),

$$\begin{aligned} \langle \phi_n | V_I(t,t_0) | \phi_g \rangle &= \delta_{ng} - \frac{i}{\hbar} \int_{t_0}^t d\tau \langle \phi_n | W_I(\tau) | \phi_g \rangle - \frac{1}{\hbar^2} \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \langle \phi_n | W_I(\tau_1) W_I(\tau_2) | \phi_g \rangle + \dots \\ &= \delta_{ng} - \frac{i}{\hbar} \int_{t_0}^t d\tau \langle \phi_n | W_I(\tau) | \phi_g \rangle - \frac{1}{\hbar^2} \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \sum_m \langle \phi_n | W_I(\tau_1) | \phi_m \rangle \langle \phi_m | W_I(\tau_2) | \phi_g \rangle + \dots \end{aligned}$$

Now,  $\langle \phi_n | W_I(\tau) | \phi_m \rangle = \langle \phi_n | e^{iH_0\tau/\hbar} W(\tau) e^{-iH_0\tau/\hbar} | \phi_m \rangle = \langle \phi_n | W(\tau) | \phi_m \rangle e^{i(E_n - E_m)\tau/\hbar}$  therefore the probability of transition to the state  $n \neq g$  to second order in perturbation theory is:

$$P_{ng}(t) = \frac{1}{\hbar^2} \Big| \int_{t_0}^t d\tau \langle \phi_n | W(\tau) | \phi_g \rangle e^{i\Delta E_{ng}\tau/\hbar} - \frac{i}{\hbar} \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \sum_m \langle \phi_n | W(\tau_1) | \phi_m \rangle \langle \phi_n | W(\tau_2) | \phi_g \rangle e^{i\Delta E_{nm}\tau_1/\hbar + i\Delta E_{mg}\tau_2/\hbar} \Big|^2.$$

where  $\Delta E_{ij} = E_i - E_j$ .

If one considers a monochromatic perturbation:  $W(t) = We^{-i\omega t}$ ,

$$\int_{t_0}^t d\tau \langle \phi_n | W(\tau) | \phi_g \rangle e^{i\Delta E_{ng}\tau/\hbar} = \langle \phi_n | W | \phi_g \rangle \int_{t_0}^t d\tau e^{i(\Delta E_{ng} - \hbar\omega)\tau/\hbar} = \langle \phi_n | W | \phi_g \rangle e^{i(\Delta E_{ng} - \hbar\omega)(t+t_0)/2\hbar} \frac{2i\sin((\Delta E_{ng} - \hbar\omega)(t-t_0)/2\hbar)}{i(\Delta E_{ng} - \hbar\omega)/\hbar}.$$

At first order,  $P_{ng}(t) = \frac{1}{\hbar^2} |\langle \phi_n | W | \phi_g \rangle|^2 \left( \frac{\sin(a(t-t_0))}{a} \right)^2$ , with  $a = (\Delta E_{ng} - \hbar \omega)/2\hbar$ . The transition rate, that is the probability of transition per unit of time:

$$w_{ng}(t) \equiv \lim_{t \to \infty} \frac{P_{ng}(t)}{t} = \frac{2\pi}{\hbar} |\langle \phi_n | W | \phi_g \rangle|^2 \delta(\Delta E_{ng} - \hbar\omega).$$
(1.3)

This formula due to Dirac was called the *Golden Rule* by Fermi. It is the formula that is usually used to derive absorption cross sections [Brouder, 1990].

### 1.3 One-body Foldy-Wouthuysen Hamiltonian and transition rate

Including relativistic effects in a semi-relativistic way has two advantages. The first one is that it permits to have a good physical insight on the origin of the relativistic phenomenon. The second one is that, in most solid state calculations, two-component wave functions are used and thus a semi-relativistic description can be more widely implemented. The Foldy-Wouthuysen (FW) transformation is widely used in particle physics but also in molecular and condensed matter physics [Ceresoli et al., 2010, Pickard and Mauri, 2001] to describe relativistic effects. After introducing the principle of this transformation, we discuss in this section the use of the FW Hamiltonian in perturbation theory.

#### 1.3.1 One-body Foldy-Wouthuysen transformation and Hamiltonian

#### Time-independent case

In the Dirac theory, the state of the particles is described by four-component wave functions  $\Psi_D = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}$ . The two spinors that compose these wavefunctions,  $\chi_1$  and  $\chi_2$ , are called upper and lower components.

For positive energy states, the upper component is called the large component and the lower component is called the small component. In the *non relativistic limit*  $(c \to \infty)$ , the small (i.e. lower) component vanishes. For negative energy states, on the other hand, it is the upper component that vanishes.

The Dirac Hamiltonian  $H^D$  has the form  $H^D = mc^2\beta + \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$ , where each  $H_{ij}$  is a 2×2 matrix. The idea of the Foldy-Wouthuysen (FW) transformation is to apply a unitary operator  $U_0$  such that, even for finite velocity of light, the upper and lower components of  $\psi_{\rm FW} = U_0 \Psi_D$  are decoupled. In other words, the transformed Hamiltonian  $H^{\rm FW} = U_0 H^D U_0^{\dagger}$ must be block diagonal.

For a one-body system, the Dirac matrix  $\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  allows defining odd and even vectors in the space of spinors: a vector  $|\Psi\rangle$  is said to be *even* if  $\beta|\Psi\rangle = |\Psi\rangle$  and it is said to be *odd* if  $\beta|\Psi\rangle = -|\Psi\rangle$ . *Odd and even operators* are defined by the way they transform vectors: an operator is said to be even if it transforms an even state into an even state and an odd state into an odd state (in other words it does not change parity). An odd operator, on the other hand, transforms an even state into an odd state into an even state into an even state. An odd operator  $\mathcal{O}$  verifies  $\beta \mathcal{O}\beta = -\mathcal{O}$  and an even operator  $\mathcal{E}$  verifies  $\beta \mathcal{E}\beta = \mathcal{E}$ . In a one-body system, even/odd components correspond to upper/lower components so a block diagonal Hamiltonian is even.

 $H^D$  can be written  $H^D = mc^2\beta + \mathcal{E} + \mathcal{O}$ , where its even part is  $mc^2\beta + \mathcal{E}$  with  $\mathcal{E} = \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix}$  and its odd part is  $\mathcal{O} = \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix}$ . A block diagonal representation of  $H^D$  was obtained with a step-by-step method in 1950 [Foldy and Wouthuysen, 1950] and corrected in 1952 [Foldy, 1952] by L.L. Foldy and S. A. Wouthuysen by applying successive transformations of the form  $e^{iS_p}$ . The successive steps are detailed in [Greiner, 2000, p285].

In 1958, E. Eriksen proposed a criteria to define a FW operator [Eriksen, 1958]:

$$U_0 = \beta U_0^{\dagger} \beta. \tag{1.4}$$

The transformation  $e^{iS_1}e^{iS_2}...e^{iS_n}$  used by Foldy and Wouthuysen does not satisfy Eriksen's condition [Silenko, 2016].<sup>1</sup>

Eriksen proposed another operator [Eriksen, 1958] that verifies condition (1.4) and that is linked to the operator sign of  $H^D$  through the relation:

$$U_0^{\dagger} \beta U_0 = \lambda = \operatorname{sign} H^D. \tag{1.5}$$

Thanks to this relation, the FW transformed wave functions of positive (negative) energy Dirac wavefunctions only have upper (lower) components. Indeed, it implies that  $\beta U_0 = U_0 \lambda$ . Therefore,  $|\psi_{\rm FW}\rangle$ , defined as  $|\psi_{\rm FW}\rangle = U_0|\Psi_D\rangle$ , is even  $(\beta|\psi_{\rm FW}\rangle = |\psi_{\rm FW}\rangle)$  if  $|\Psi_D\rangle$  is a positive energy state  $(\lambda|\Psi_D\rangle = |\Psi_D\rangle)$ . Therefore, only the upper components of the FW transform of a positive energy Dirac state are non-zero. Likewise, the FW transform of a negative energy state is odd, *i.e.* only its lower components are non-zero. In practice, these statements are exact up to the order to which the FW transformation is performed.

Eriksen, as Foldy and Wouthuysen, used  $m^{-1}$  as expansion parameter and stopped at order  $m^{-2}$ . At this order, both approaches yield the same transformed Hamiltonian  $H^{\text{FW}}$ . Indeed, the

$$S_p = \frac{-i}{2mc^2}\beta \mathcal{O}_p.$$

The operator  $U_p = e^{iS_p}$  verifies  $U_p = \beta U_p^{\dagger}\beta$ . Indeed,  $U_p^{\dagger} = e^{-iS_p} = \sum_k \frac{(-1)^k}{k!} \left(\frac{\beta \mathcal{O}_p}{2mc^2}\right)^k$ , so as  $\beta(\beta \mathcal{O}_p)^k \beta = (\beta \mathcal{O}_p)^k = (-1)^k (\beta \mathcal{O}_p)^k$ ,  $\beta e^{-iS_p} \beta = e^{iS_p}$ . However,  $\beta (U_1 U_2 ... U_n)^{\dagger} \beta = U_n ... U_2 U_1 \neq U_1 U_2 ... U_n$ .

<sup>&</sup>lt;sup>1</sup> At each step,  $S_p$  writes as a function of the odd part of the Hamiltonian:

correction to the operator proposed by Foldy and Wouthuysen for it to verify Eriksen's condition (Eq. (1.4)) is of order  $m^{-3}c^{-4}$  [Silenko, 2016, Eq.(31)]. The corresponding Hamiltonian is:

$$H^{\rm FW} = \beta mc^2 + \mathcal{E} + \frac{\beta}{2mc^2}\mathcal{O}^2 - \frac{1}{8m^2c^4}[\mathcal{O}, [\mathcal{O}, \mathcal{E}]].$$

The odd and even operators for the Dirac Hamiltonian are:

$$\mathcal{O} = c\boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{A}) \text{ and } \mathcal{E} = e\phi_0 \hat{I}.$$
 (1.6)

It leads, for positive energy states, to [Strange, 1998]:

$$H^{\rm FW} = mc^2 + \frac{\boldsymbol{\pi}^2}{2m} + e\phi_0 - \frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B} - \frac{e\hbar}{8m^2c^2}(\boldsymbol{\sigma}\cdot(\mathbf{E}\times\boldsymbol{\pi}) - \boldsymbol{\sigma}\cdot(\boldsymbol{\pi}\times\mathbf{E})) - \frac{e\hbar^2}{8m^2c^2}\nabla\cdot\mathbf{E} \quad (1.7)$$

where  $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$ . It is the sum of six terms:

(i) the rest energy of positive-energy eigenstates,

(*ii*) the kinetic energy of the electron,

(*iii*) the Coulomb interaction of the electron with the nuclei and the other electrons,

(iv) the Zeeman interaction with magnetic field,

(v) the spin-orbit interaction (for a spherical potential  $\phi_0$  and a static vector potential A,

$$\boldsymbol{\sigma} \cdot (\mathbf{E} \times \mathbf{p}) = rac{-1}{r} rac{d\phi_0}{dr} \boldsymbol{\sigma} \cdot (\mathbf{r} \times \mathbf{p}) = rac{-1}{r} rac{d\phi_0}{dr} \boldsymbol{\sigma} \cdot \mathbf{L}),$$

$$(vi)$$
 the Darwin term.

In the time-independent case, which is the object of this section,  $\mathbf{E} \times \mathbf{p} + \mathbf{p} \times \mathbf{E} = -i\hbar\nabla \times E = 0$ (Maxwell-Faraday equation) so the writing of the spin-orbit part of  $H^{\text{FW}}$  could be simplified by replacing the term in parenthesis by  $2\boldsymbol{\sigma} \cdot (\mathbf{E} \times \boldsymbol{\pi})$ .

Textbooks [Bjorken and Drell, 1965, Itzykson and Zuber, 1980] often derive a FW Hamiltonian  $H^{\text{TFW}}$  with expansion parameter  $c^{-1}$  up to order  $c^{-2}$  which is the same as  $H^{\text{FW}}$  except that  $\boldsymbol{\pi}$  is replaced by  $\mathbf{p}$  in the spin-orbit term. Because of the presence of  $\boldsymbol{\pi}$ , the Hamiltonian  $H^{\text{FW}}$  is often considered to be a *gauge invariant* Hamiltonian contrary to  $H^{\text{TFW}}$ . However, we will see in Section 1.4 that the concept of gauge invariance of transition probability is not straightforward. A mass-velocity term  $-(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is also often added in textbooks [Itzykson and Zuber, 1980].

#### Time-dependent case

If  $H^D$  is explicitly time dependent, its FW transformation writes [Foldy and Wouthuysen, 1950]:

$$H^{\rm FW} = U H^D U^{\dagger} - i\hbar U \frac{\partial U^{\dagger}}{\partial t}$$
(1.8)

in order for the time-dependent Schrödinger equation to remain valid. Note that the expression of  $H^{\rm FW}$  remains formally the same as in Eq. (1.7) but the part  $-i\hbar U \frac{\partial U^{\dagger}}{\partial t}$  is responsible for the  $\frac{\partial \mathbf{A}}{\partial t}$  term that appears in  $\mathbf{E} = -\nabla \phi_0 - \frac{\partial \mathbf{A}}{\partial t}$  [Goldman, 1977].

#### 1.3.2 Naive application to absorption cross section calculation

#### Calculation from $H^{\rm FW}$

In this section, we consider that the material is described by the FW Hamiltonian and we are going to use perturbation theory at first order to compute the absorption cross section. We

split the FW Hamiltonian between a static part,  $H_0$ , that describes the material in the absence of the electromagnetic wave and a time dependent part, W, that describes the interaction between the electron and the incident wave, which can be treated as a perturbation. The total vector potential **A** and fields **E** and **B** are decomposed as sums  $\mathbf{A} = \mathbf{A}_0 + \mathbf{a}$ ,  $\mathbf{E} = \mathbf{E}_0 + \mathbf{e}$ ,  $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}$  of their static part (letters with index 0) plus their dynamical part representing the incident electromagnetic wave (lowercase letters). We use Coulomb gauge ( $\nabla \cdot \mathbf{a} = 0$ ) and we choose the potential associated to the wave to be zero. The static Hamiltonian writes:

$$H_0^{\rm FW} = mc^2 + \frac{1}{2m}\boldsymbol{\pi}_0^2 + e\phi_0(\mathbf{r}) - \frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B}_0 - \frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{E}_0\times\boldsymbol{\pi}_0) + \frac{e\hbar^2}{8m^2c^2}(\Delta\phi_0)$$

where  $\boldsymbol{\pi}_0 = \mathbf{p} - e\mathbf{A}_0$ , and the perturbation writes:

$$W(t) = H^{\rm FW} - H_0^{\rm FW}$$
  
=  $\frac{-e}{m} \mathbf{a} \cdot \boldsymbol{\pi}_0 - \frac{e\hbar}{2m} \boldsymbol{\sigma} \cdot \mathbf{b} - \frac{e\hbar}{8m^2c^2} \boldsymbol{\sigma} \cdot (\mathbf{e} \times \boldsymbol{\pi}_0 - \boldsymbol{\pi}_0 \times \mathbf{e} - 2e\mathbf{E}_0 \times \mathbf{a}).$ 

We have used Maxwell's equation in vacuum  $\nabla \cdot \mathbf{e} = 0$ , we neglected the term proportional to  $\mathbf{a}^2$  because it gives a negligible contribution for available x-rays sources as was evaluated in [Brouder, 1990] and we have used the fact that in the Coulomb gauge  $\mathbf{a} \times \mathbf{p} = -\mathbf{p} \times \mathbf{a}$ .

We consider a monochromatic plane wave:  $\mathbf{a} = A\boldsymbol{\epsilon} \mathbf{e}^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ ,  $\mathbf{b} = \nabla \times \mathbf{a} = i\mathbf{k} \times \mathbf{a}$  and  $\mathbf{e} = -\frac{\partial \mathbf{a}}{\partial t} = i\omega \mathbf{a}$ . To develop further, it is useful to notice that the commutator of the static Hamiltonian with  $\mathbf{r}$  is:<sup>2</sup>

$$[H_0^{\rm FW}, \mathbf{r}] = \frac{-i\hbar}{m} \boldsymbol{\pi}_0 + \frac{e\hbar}{4m^2c^2} (i\hbar) (\boldsymbol{\sigma} \times \mathbf{E}_0).$$

Thus  $W(t) = W e^{-i\omega t}$  where,

$$W = \frac{e}{i\hbar} \mathbf{a} \cdot [\mathcal{H}_0^{\text{FW}}, \mathbf{r}] - \frac{e^2\hbar}{4m^2c^2} \mathbf{a} \cdot (\boldsymbol{\sigma} \times \mathbf{E}_0) - \frac{e\hbar}{2m} i\boldsymbol{\sigma} \cdot (\mathbf{k} \times \mathbf{a}) - \frac{e\hbar i\omega}{8m^2c^2} \boldsymbol{\sigma} \cdot (\mathbf{a} \times \boldsymbol{\pi}_0 - \boldsymbol{\pi}_0 \times \mathbf{a}) + \frac{e^2\hbar}{4m^2c^2} \boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \mathbf{a}). \quad (1.9)$$

In Eq. (1.9), the second term, which arises from the commutator of  $H_0^{\text{FW}}$  with **r**, cancels out the last term that arises from the spin-orbit term in  $H^{\text{FW}}$ . This fact is not simply anecdotal because, would one of these terms remain, it would be identical or opposite to the last  $c^{-2}$  term in the absorption case (which as we will see in the following, is responsible for a significant term that is called spin-position).<sup>3</sup>

<sup>2</sup> Using the derivation chain rule,  $\nabla^2 \mathbf{r} |\Psi\rangle = \nabla(\mathbf{r} \nabla |\Psi\rangle + \nabla(\mathbf{r}) |\Psi\rangle) = \nabla |\Psi\rangle + \nabla |\Psi\rangle + \mathbf{r} \nabla^2 |\Psi\rangle$ , that is to say  $[\nabla^2, \mathbf{r}] = 2\nabla$ . Therefore  $[\frac{\pi_0^2}{2m}, \mathbf{r}] = \frac{-i\hbar}{m} \pi_0$ . Now,

$$[\boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \boldsymbol{\pi}_0), \mathbf{r}] = (\boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \boldsymbol{\pi}_0))\mathbf{r} - \mathbf{r}(\boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \boldsymbol{\pi}_0)) = ((\boldsymbol{\sigma} \times \mathbf{E}_0) \cdot \boldsymbol{\pi}_0)\mathbf{r} - \mathbf{r}((\boldsymbol{\sigma} \times \mathbf{E}_0) \cdot \boldsymbol{\pi}_0)$$

The projection of this quantity on any axis i:

$$[\boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \boldsymbol{\pi}_0), \mathbf{r}]_i = \sum_j (\boldsymbol{\sigma} \times \mathbf{E}_0)_j (\boldsymbol{\pi}_{0j} \mathbf{r}_i - \mathbf{r}_i \boldsymbol{\pi}_{0j}) = \sum_j (\boldsymbol{\sigma} \times \mathbf{E}_0)_j (-i\hbar) \delta_{ij} = (-i\hbar) (\boldsymbol{\sigma} \times \mathbf{E}_0)_i$$

<sup>3</sup> Using the relation  $[\mathbf{p}, \phi_0] = i\hbar\nabla\phi_0$ , the electric field in matter writes at zeroth order in  $c^{-2}$  as a function of the commutator of  $\pi_0$  with  $H_0$ :

$$\mathbf{E}_0 = -\nabla E_0 = \frac{-i}{e\hbar} [\boldsymbol{\pi}_0, H_0].$$

In the case of absorption, the commutator transforms into a factor  $-\Delta E_{fg} = -\hbar\omega$  in the cross section so that

This expression of W can be used to determine the transition rate from the Golden Rule (Eq (1.3)). The absorption cross section is the ratio between the rate at which energy is removed from the photon beam ( $\hbar\omega \sum_f w_{fg}$  where the sum runs over unoccupied states f) and the rate at which energy in the photon beam crosses a unit area perpendicular to its propagation direction  $(I(\omega))$ :

$$\sigma(\hbar\omega) = \frac{\hbar\omega \sum_f w_{fg}}{I(\omega)} \text{ where } I(\omega) = 2\epsilon_0 c\omega^2 |A|^2.$$

As the core wavefunction is very localized (the evaluation of kr for several edges will be given in the next chapter), we use the quadrupole approximation  $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1 + i\mathbf{k}\cdot\mathbf{r}$  except in the  $c^{-2}$  term for which we use the dipole approximation  $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$ . In this term the commutator at zeroth order  $[\mathcal{H}_0^{\text{FW}}, \mathbf{r}] = \frac{-i\hbar}{m} \pi_0$  is sufficient to preserve the  $c^{-2}$  approximation.

Finaly, the semi-relativistic absorption cross section can be written with four operators:<sup>4</sup>

$$\sigma(\hbar\omega) = \frac{4\pi^2 \alpha_0 (\Delta E_{fg})^2}{\hbar\omega} \sum_f |\langle \phi_f| \underbrace{\boldsymbol{\epsilon} \cdot \mathbf{r}}_{\text{electric dipole}} + \underbrace{\frac{i}{2} (\mathbf{k} \cdot \mathbf{r}) (\boldsymbol{\epsilon} \cdot \mathbf{r})}_{\text{electric quadrupole}} + \underbrace{\frac{\hbar}{2m\Delta E_{fg}} (\mathbf{k} \times \boldsymbol{\epsilon}) \cdot (\mathbf{\Lambda} + \hbar\boldsymbol{\sigma})}_{\text{magnetic dipole}} + \underbrace{\frac{i\hbar\omega}{4mc^2} \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r})}_{\text{spin-position}} |\phi_g\rangle|^2 \delta(\Delta E_{fg} - \hbar\omega) \quad (1.10)$$

where  $\mathbf{\Lambda} = \mathbf{r} \times \mathbf{\pi}_0 = \mathbf{L} - e\mathbf{r} \times \mathbf{a}_0$  is sometimes named the moment of mechanical momentum [Cohen-Tannoudji et al., 1987].  $\mathbf{\Lambda}$  can be understood as a gauge invariant angular momentum. The first two operators are the usual electric dipole and electric quadrupole operators. The third one is the magnetic-dipole operator [Brouder, 1990] but with  $\mathbf{\Lambda}$  instead of the  $\mathbf{L}$  that usually enters it through the total angular momentum of the electron  $(\hbar \boldsymbol{\sigma} + \mathbf{L})$ . We name the  $c^{-2}$  term spin-position because it is proportional to  $\boldsymbol{\sigma} \cdot (\mathbf{r} \times \boldsymbol{\epsilon}) = \boldsymbol{\epsilon} \cdot (\boldsymbol{\sigma} \times \mathbf{r})$ .

A similar calculation was carried out in 2009 by Christos Gougoussis [Gougoussis, 2009, Chap. 3] in his PhD thesis. However, he obtained an additional contribution that he named SO1 and that wrote  $\frac{e\omega}{4m^2c^2}\boldsymbol{\sigma} \cdot [\nabla\phi_0 \times \boldsymbol{\epsilon}]$ . In our calculation, it cancels out with the  $c^{-2}$  term when  $\boldsymbol{\pi}_0$  is written as a function of the commutator of  $H_0^{\text{FW}}$  with  $\mathbf{r}$ . This term was not developed in C. Gougoussis' thesis but it yields the same contribution as the spin-position term

the term that cancels out would lead to the same contribution to the matrix element as the remaining  $c^{-2}$  term:

$$\frac{-ie\hbar\omega}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{a} imes\boldsymbol{\pi}_0).$$

<sup>4</sup> The rewriting of  $\mathbf{a} \cdot [H_0^{\text{FW}}, \mathbf{r}]$  within the quadrupole approximation requires a little trick:

$$\begin{split} \boldsymbol{\epsilon} \cdot (\mathbf{k} \cdot \mathbf{r})[H_0^{\text{FW}}, \mathbf{r}] &= (\mathbf{k} \cdot \mathbf{r}) \mathcal{H}_0^{\text{FW}}(\boldsymbol{\epsilon} \cdot \mathbf{r}) - (\mathbf{k} \cdot \mathbf{r})(\boldsymbol{\epsilon} \cdot \mathbf{r}) H_0^{\text{FW}} \\ &= \frac{1}{2} [H_0^{\text{FW}}, (\boldsymbol{\epsilon} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{r})] - \frac{1}{2} H_0^{\text{FW}}(\boldsymbol{\epsilon} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{r}) + \frac{1}{2} (\mathbf{k} \cdot \mathbf{r}) H_0^{\text{FW}}(\boldsymbol{\epsilon} \cdot \mathbf{r}) \\ &\quad - \frac{1}{2} (\boldsymbol{\epsilon} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{r}) H_0^{\text{FW}} + \frac{1}{2} (\mathbf{k} \cdot \mathbf{r}) H_0^{\text{FW}}(\boldsymbol{\epsilon} \cdot \mathbf{r}) \\ &= \frac{1}{2} \Big( [H_0^{\text{FW}}, (\boldsymbol{\epsilon} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{r})] + (\mathbf{k} \cdot \mathbf{r})(\boldsymbol{\epsilon} \cdot [H_0^{\text{FW}}, \mathbf{r}]) - (\mathbf{k} \cdot [H_0^{\text{FW}}, \mathbf{r}])(\boldsymbol{\epsilon} \cdot \mathbf{r}) \Big) \\ &= \frac{1}{2} \Big( [H_0^{\text{FW}}, (\boldsymbol{\epsilon} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{r})] + (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot ([H_0^{\text{FW}}, \mathbf{r}] \times \mathbf{r}) \Big). \end{split}$$

In this term, as it is proportional to  $\mathbf{k}$ , the commutator at zeroth order in  $c^{-2}$ :  $[H_0^{\text{FW}}, \mathbf{r}] = \frac{-i\hbar}{m} \pi_0$ , is also sufficient. Therefore,  $(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot ([H_0^{\text{FW}}, \mathbf{r}] \times \mathbf{r}) = \frac{i\hbar}{m} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \pi_0)$ .

(see footnote 3), called SO2 in his thesis. This illustrates the importance of a derivation of the commutator that is consistent in order of  $c^{-2}$  with the rest of the development. Here, it is particularly treacherous because, working in the Heisenberg picture, we could think that  $[H_0^{\text{FW}}, \mathbf{r}] = -i\hbar \frac{d}{dt}\mathbf{r} = \frac{-i\hbar}{m}\mathbf{p}$ , as it is used for example in [Takahashi and Hiraoka, 2015, Joly et al., 2012]. However, the relation  $m\frac{d}{dt}\mathbf{r} = \mathbf{p}$  is not valid. Indeed,  $m\frac{d}{dt}\mathbf{r} = \frac{-m}{i\hbar}[H_0, \mathbf{r}]$  is equal to  $\mathbf{p}$  only if  $H_0 = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r})$ . If  $H_0 = \frac{\pi^2}{2m} + V(\mathbf{r})$ ,  $m\frac{d}{dt}\mathbf{r} = \boldsymbol{\pi}$  and if the Hamiltonian includes the spin-orbit interaction proportional to  $\mathbf{E} \times \mathbf{p} + \mathbf{p} \times \mathbf{E}$ ,  $m\frac{d}{dt}\mathbf{r} = \pi - \frac{e\hbar}{4mc^2}(\boldsymbol{\sigma} \times \mathbf{E})$ .

#### Calculation from $H^{\text{TFW}}$

Surprisingly, if the same calculation is carried out from the FW Hamiltonian  $H^{\text{TFW}}$  defined p. 18, the result is not the same.

The fact that  $\boldsymbol{\pi}$  is replaced by  $\mathbf{p}$  in the spin-orbit term is quite important. Indeed, it leads to the absence of the term  $\boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \mathbf{a})$  in W in Eq. (1.9). The consequence is that there is no spin-position contribution to absorption cross section because it is cancelled out by the  $c^{-2}$ term arising from the commutator  $[H_0^{\text{TFW}}, \mathbf{r}]$ .

The additional mass-velocity term  $-(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  does not contribute to W but it is present in  $H_0^{\text{TFW}}$ . It leads to an additional contribution to the commutator  $[H_0^{\text{TFW}}, \mathbf{r}]$ :  $\frac{i\hbar(\mathbf{p} \cdot \mathbf{p})\mathbf{p}}{2m^3c^2}$ . It is small compared to  $\frac{i\hbar}{m}\mathbf{p}$  if the order of magnitude of the kinetic energy of the core state satisfies  $E_k \ll mc^2$  so we could neglect it here.

As the main difference is due to the replacement of  $\pi$  by **p**, it seems to be related to a problem of *gauge invariance*. Indeed, it is often said that  $H^{\text{TFW}}$  is the gauge invariant version of FW Hamilonian. It is not true that using  $H^{\text{TFW}}$  provides gauge invariance. The difficulties related to gauge invariance in transition probabilities calculations will be detailed in Section 1.4. Before that, we explain why the derivations presented in this section are, in any cases, invalid because there is a conflict between FW and time-dependent perturbation theory.

#### 1.3.3 Subtleties in the application of time-dependent Foldy-Wouthuysen transformation to transition rate calculation

It was noticed by Nieto as early as 1977 [Nieto, 1977, Goldman, 1977] that the FW Hamiltonian must not be used to calculate the Hamiltonian matrix elements. Indeed, as  $H^D = H^D(\mathbf{A}, V)$  is explicitly time dependent, its FW transformation [Foldy and Wouthuysen, 1950] writes:

$$H^{\rm FW} = U H^D U^{\dagger} - i\hbar U \frac{\partial U^{\dagger}}{\partial t}$$
(1.11)

in order for the time-dependent Schrödinger equation to remain valid for  $|\psi_{\rm FW}\rangle = U|\Psi_D\rangle$ . Therefore,  $\langle \Phi_D | H^D | \Psi_D \rangle = \langle \phi_{\rm FW} | U H^D U^{\dagger} | \psi_{\rm FW} \rangle \neq \langle \phi_{\rm FW} | H^{\rm FW} | \psi_{\rm FW} \rangle$ . The case of the transition probabilities yields another subtlety: the FW operator U corresponding to  $H^D$  is different from the FW operator  $U_0$  corresponding to  $H_0^D = H^D(\mathbf{A}_0, \phi_0)$ . The transition probability,

$$|\langle \Phi_D^n | \Psi_D(t) \rangle|^2 = |\langle \phi_{\rm FW}^n | U_0 U^{\dagger} | \psi_{\rm FW}(t) \rangle|^2 \neq |\langle \phi_{\rm FW}^n | \psi_{\rm FW}(t) \rangle|^2.$$
(1.12)

Therefore, the usual formula of perturbation theory cannot be used with FW transformed quantities. Tricks like transforming  $\langle \Phi_D^n |$  with U instead of  $U_0$  are not an option because it would make  $\langle \phi_{\rm FW}^n |$  time dependent so that it would not be an eigenstate of  $H_0^{\rm FW}$ .

This illustrates that the straightforward way to obtain semi-relativistic cross sections (that consists in considering that the system is described by a FW Hamiltonian and computing the transition probabilities between eigenstates of this Hamiltonian) is, in fact, not valid.

We propose an alternative approach in which we start from the fully-relativistic cross sections and we apply a FW transformation to the wave functions. In order for this approach to be correct, it is important to check its gauge invariance.

#### 1.4 Gauge Invariance

The principle of gauge invariance has become a cornerstone of particle physics. Since general relativity can also be considered as a gauge theory [Blagojević and Hehl, 2013], it may be safely said that gauge invariance was the guiding principle of most of the fundamental physics of the twentieth century. Therefore, we need to check that the cross section formulas are gauge invariant to ensure their true physical nature.

#### **1.4.1** Principle of gauge invariance in classical electrodynamics

Maxwell's equation  $\nabla \cdot \mathbf{B} = 0$  has for a consequence that **B** locally admits a vector potential **A** such that  $\mathbf{B} = \nabla \times \mathbf{A}$ . The magnetic field is unchanged if one adds the gradient of any continuously differentiable scalar function  $\Lambda$  to  $\mathbf{A}$ :  $\mathbf{A}' = \mathbf{A} + \nabla \Lambda(\mathbf{r}, t)$ . Similarly Maxwell's equation  $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$  that rewrites  $\nabla \times (\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}) = 0$  leads to the local existence of a scalar potential V such that  $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla V$ . The electric field remains unchanged if  $\mathbf{A}$  is changed to  $\mathbf{A}'$  and the time derivative of  $\Lambda$  is subtracted to V:  $V' = V - \frac{\partial \Lambda}{\partial t}$ . In classical electrodynamics gauge invariance means that the physics must remain the same when a gauge transformation is applied to the potentials.

#### Particle in a field

The force  $\mathbf{F}$  exerted on a particle of charge q and mass m in an electromagnetic field is the Lorentz force:  $\mathbf{F} = q(\mathbf{E}(\mathbf{r}, t) + \mathbf{v}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t))$ . Newton's law  $m \frac{d^2}{dt^2} \mathbf{r}(t) = \mathbf{F}$  governs the movement of this particle: in this formalism, the problem of gauge invariance does not arise.

However, the motion of this particle can also be described [Cohen-Tannoudji et al., 1987] by the Lagrangian:

$$\mathcal{L}(\mathbf{v}, \mathbf{r}, t) = \frac{1}{2}m\mathbf{v}^2 - q\left(V(\mathbf{r}, t) - \mathbf{v} \cdot \mathbf{A}(\mathbf{r}, t)\right)$$

The canonical momentum  $\mathbf{p} = \nabla_v \mathcal{L}(\mathbf{v}, \mathbf{r}, t) = m\mathbf{v} + q\mathbf{A}(\mathbf{r}, t)$ . The position  $\mathbf{r}$  and the velocity  $\mathbf{v}$ , that are driven by Newton's law, are independent of the gauge. Therefore  $\mathbf{p} - q\mathbf{A}(\mathbf{r}, t) = m\mathbf{v}$  is independent of the gauge but  $\mathbf{p}$  is not. The corresponding Hamiltonian is given by,

$$\mathcal{H}(\mathbf{p},\mathbf{r},t) = \sum p_i \dot{q}_i - \mathcal{L}(q_i,\dot{q}_i) = \frac{1}{2m} (\mathbf{p} - q\mathbf{A}(\mathbf{r},t))^2 + qV(\mathbf{r},t)$$

If the gauge is changed,  $\mathcal{H}$  is changed to  $\mathcal{H}' = \mathcal{H} - q \frac{\partial \Lambda}{\partial t}$  due to the presence of the scalar potential V in its expression. It can however be checked that the equations of motion associated to this Hamiltonian lead to Newton's law. Similarly, the Lagrangian is transformed under a gauge transformation to

$$\mathbf{L}' = \mathbf{L} + q \left( \frac{\partial \Lambda}{\partial t} + \mathbf{v} \cdot \nabla \Lambda \right) \tag{1.13}$$

so it differs only by a total time derivative  $\frac{d\Lambda}{dt}$  and the equations of motion remain the same. In conclusion, the quantities used to describe the physics are not necessarily physical quantities in the sense that physical quantities must be gauge invariant.

#### 1.4.2 Gauge invariance in quantum mechanics

In quantum mechanics, gauge change consists in both a change of the potentials and a change of the phase of the wavefunctions:

$$\mathbf{A} \to \mathbf{A}', V \to V' \text{ and } \psi(\mathbf{r}, t) \to M_{\Lambda} \psi(\mathbf{r}, t)$$
 (1.14)

where,  $M_{\Lambda} = e^{\frac{ie\Lambda(\mathbf{r},t)}{\hbar}}$  is a unitary operator.

By definition [Cohen-Tannoudji et al., 1973] an observable O describes a true physical quantity if it verifies  $M_{\Lambda}^{\dagger}O_{A'}M_{\Lambda} = O_A$  for every function  $\Lambda(\mathbf{r}, t)$ . A Hamiltonian  $H(\mathbf{A}, V) = f(\mathbf{p} - e\mathbf{A}) + eV$  where f is some function, is not such a gauge invariant operator:

$$H(\mathbf{A}', V') = M_{\Lambda} H(\mathbf{A}, V) M_{\Lambda}^{\dagger} - e \frac{\partial \Lambda}{\partial t}$$

(the  $\frac{\partial \Lambda}{\partial t}$  arises because of the potential). The time-dependent Schrödinger and Dirac equations are, however, invariant under gauge transformation:

$$M_{\Lambda}^{\dagger}(i\hbar\partial_t - H(\mathbf{A}', V'))M_{\Lambda} = i\hbar\partial_t - H(\mathbf{A}, V).$$

So if  $\psi$  is solution of  $i\hbar\partial_t\psi = H(\mathbf{A}, V)\psi$  then  $\psi' = M_\Lambda\psi$  is solution of  $i\hbar\partial_t\psi' = H(\mathbf{A}', V')\psi'$ .

#### **1.4.3** Gauge invariance of transition probabilities

Let us first consider the question of gauge invariance of the matrix element entering the Golden Rule Eq. (1.3). In the calculation presented in section 1.3.2, we separated the static field in the absence of perturbation  $(\mathbf{A}_0, V_0)$  from the incident electromagnetic wave field  $(\mathbf{a}, v)$ . When a gauge change transforms the static potentials  $\mathbf{A}_0$  and  $V_0$ , the interaction W is changed to  $M_{\Lambda}WM_{\Lambda}^{\dagger}$  and the static eigenstates are transformed to  $M_{\Lambda}|\Psi_n\rangle$ . As a consequence, the matrix element  $\langle \Psi^m|W|\Psi^n\rangle$  is conserved under a gauge transformation for the static field. The gauge function  $\Lambda$  must however be time independent in order for  $H_0$  to remain a static Hamiltonian.

Now, if the gauge for the incident electromagnetic wave is changed, W is transformed whereas the eigenstates are unchanged. In general,

$$\langle \Psi^m | H' - H_0 | \Psi^n \rangle \neq \langle \Psi^m | H - H_0 | \Psi^n \rangle.$$

In order to determine when gauge invariance can still be achieved we compute the difference between these two quantities:

$$\langle \Psi^m | H' - H | \Psi^n \rangle = \langle \Psi^m | M_\Lambda H M^{\dagger}_\Lambda - H - e \frac{\partial \Lambda}{\partial t} | \Psi^n \rangle.$$

Using Baker-Campbell-Hausdorff formula,

$$M_{\Lambda}HM_{\Lambda}^{\dagger} = H + \left[\frac{ie\Lambda}{\hbar}, H\right] + \frac{1}{2}\left[\frac{ie\Lambda}{\hbar}, \left[\frac{ie\Lambda}{\hbar}, H\right]\right] + \dots$$

If one considers W and  $\Lambda$  as perturbations, at first order

$$M_{\Lambda}HM_{\Lambda}^{\dagger} \approx H + \left[\frac{ie\Lambda}{\hbar}, H_0\right]$$
so that [Kazes et al., 1982, Feuchtwang et al., 1986]:

$$\langle \Psi^m | H' - H | \Psi^n \rangle \approx i e \frac{E_i - E_f}{\hbar} \langle \Psi^m | \Lambda | \Psi^i \rangle - e \langle \Psi^f | \frac{\partial \Lambda}{\partial t} | \Psi^n \rangle.$$
(1.15)

Note that the equality is exact in the case of the Dirac Hamiltonian, with no need to make any approximation on the order of magnitude of the perturbation or the gauge because  $H'^D - H^D = c\boldsymbol{\alpha} \cdot \nabla \Lambda + \frac{\partial \Lambda}{\partial t}$  and  $c\boldsymbol{\alpha} \cdot \nabla \Lambda = (i/\hbar)[H_0^D, \Lambda]$ .

During the absorption process, energy conservation implies that  $E_f = E_i + \hbar \omega$ , thus if  $\Lambda$  verifies  $\frac{\partial \Lambda}{\partial t} = -i\omega\Lambda$  then  $\langle \Psi^m | H' - H | \Psi^n \rangle = 0$ . This explains the fact that derivations that have been performed with several different gauges (for example in [Yang, 1988]) lead to consistent results when  $E_f - E_i = \hbar \omega$ . This condition is however not met in scattering experiments [Yang, 1988, Stokes, 2013] and, even for absorption, the restriction to gauges that verify  $\frac{\partial \Lambda}{\partial t} = -i\omega\Lambda$  is not enough to provide gauge invariance.

The right term in Eq. (1.15) is also zero in the elastic case  $(E_i = E_f)$  with time-independent gauge transformations  $(\frac{\partial \Lambda}{\partial t} = 0)$  which is also a very narrow restriction.

In conclusion, in the semi-classical approach, where the photon is represented by an external field, the matrix elements of W are not gauge-invariant. Given the remark of the first paragraph, that these matrix elements are invariant under a gauge transformation of  $\mathbf{A}_0$  and  $V_0$ , it could be tempting to apply a method that would consist in subtracting  $\partial \Lambda$  from  $\mathbf{A}_0$  and  $V_0$  instead of  $\mathbf{a}$  and v. This would, indeed, have the same effect on the total field. This method, that was proposed in [Forney et al., 1977, Epstein, 1979] is questionable for a time-dependent gauge change because it requires to define 'time-dependent eigenvectors'  $|\phi'_n\rangle = M_{\Lambda}|\phi_n\rangle$  which is not consistent with perturbation theory as presented in Section 1.2. The problem for time-independent gauge changes is subtle and was noticed by Yang [Yang, 1982]. If one decides to apply all the gauge changes to  $\mathbf{A}_0$  and  $V_0$ , it means that the gauge is fixed for  $\mathbf{a}$  and v. In other words, it requires to arbitrary define the "good gauge" for  $\mathbf{a}$  and v which is in conflict with gauge invariance.

More generally, to check if this problem of gauge invariance is not related to the order of approximation in perturbation theory, we can check whether the transition probability  $P_{ng}$  is gauge invariant. For a matter of simplicity, we use Schrödinger representation instead of the interaction representation. In the Schrödinger representation, the time evolution operator obeys  $i\hbar\partial_t V(t,t_0) = H(t)V(t,t_0)$  and under a gauge transformation of **a**, it becomes [Kobe and Yang, 1985]:

$$V'(t,t_0) = e^{ie\Lambda(t)/\hbar} V(t,t_0) e^{-ie\Lambda(t_0)/\hbar}$$

(the proof is given in [Kobe and Yang, 1985] paragraph V). The transition probability  $P_{ng}(t) = |\langle \phi_n | V(t, t_0) | \phi_g \rangle|^2$  becomes

$$P_{ng}'(t) = |\langle \phi_n | \mathrm{e}^{ie\Lambda(t)/\hbar} V(t, t_0) \mathrm{e}^{-ie\Lambda(t_0)/\hbar} |\phi_g \rangle|^2 = |\langle \phi_n | \mathrm{e}^{ie/\hbar \int_{t_0}^t d\tau \dot{\Lambda}(\tau) + ie\Lambda(t_0)/\hbar} V(t, t_0) \mathrm{e}^{-ie\Lambda(t_0)/\hbar} |\phi_g \rangle|^2$$

The two probabilities are equal if  $\dot{\Lambda}(t) = 0$  and  $[\Lambda(t_0), V] = 0$  but in general  $P'_{ng}(t) \neq P_{ng}(t)$  which is alarming. Many papers describe the discrepancy between the probability calculated in two different gauges, which can be large [W. E. Lamb, 1952, Stokes, 2013, for example].

The fact that the right term of Eq. (1.15) is zero in the elastic case with time-independent gauge transformations, suggests that gauge invariance could be achieved in a framework where the states described both the electronic system and the photons so that the final and initial energy are always equal. Such a framework is provided by quantum electrodynamics (QED) which is a quantized theory of light interacting with matter.

### 1.4.4 Gauge invariance in QED

In QED, the electromagnetic field is quantized and the photons are excitations of the quantized field. A scattering experiment is described by the transition from an initial state involving both the electronic system in its ground state and the incident photon, to a final state involving both the electronic system in its (possibly) excited state and the scattered photon. The transition probabilities are now described through the so-called S-matrix. In the Schrödinger picture, the gauge transformation can be expressed in terms of time-independent operators so both conditions that are sufficient for  $\langle \Psi^f | H' - H | \Psi^i \rangle$  (Eq. (1.15)) to be zero seem to be met. This is a hand-waving argument that suggests that gauge invariance could be satisfied. However, a review of the literature on quantum electrodynamics is rather confusing. For standard textbooks "the S-matrix is gauge invariant by construction" [Peskin and Schroeder, 1995]. For mathematicallyminded authors, "an even approximately complete solution [of the gauge invariance problem] does not exist" [Steinmann, 2000]. The difficulty comes from the fact that the state spaces have different natures depending on the gauge (Coulomb and Lorenz gauge are a very common illustration of this problem: in the Lorenz gauge there are four polarization vectors whereas in the Coulomb gauge there are two polarization vectors [Cohen-Tannoudji et al., 1997]).

The gauge invariance of the renormalized S-matrix is however established for infinitesimal gauge transformations [Weinberg, 1995, Hollands, 2008] and for reasonably large classes of gauges [Haller and Sohn, 1979, Matsuda and Kubo, 1980, Voronov et al., 1982, Manoukian, 1988, Haller and Lim-Lombridas, 1994, Lenz et al., 1994, Kashiwa and Tanimura, 1997, Grigore, 2001, Das et al., 2013]. In other words, it is proved at a reasonable level of rigour for a physicist.

The most studied gauges are the Lorenz and Coulomb gauges. In the Coulomb gauge, the physical degrees of freedom are manifest: the photon states form a Fock space built by acting on the vacuum with creation operators of left and right polarized photons. Coulomb gauge is used in most of the low-energy many-body calculations and, in this gauge, the Coulomb interaction is exactly accounted for. For these reasons, we will use the Coulomb gauge in the following derivation of the cross section.

# **1.5** Relativistic absorption and scattering cross sections

Cross sections are expressed in terms of the S-matrix and  $\mathcal{T}$ -matrix elements. We present their calculation in a fully relativistic QED framework in the Coulomb gauge.

## **1.5.1** Photon matter interaction and S-matrix

The quantum field Hamiltonian describing the interaction of light with matter in the Coulomb gauge is [Greiner and Reinhardt, 1996, Cohen-Tannoudji et al., 1997]:

$$H = H_e + H_\gamma + H_{e\gamma}.$$

The electron Hamiltonian is

$$H_e = \int d\mathbf{r} \psi^{\dagger}(\mathbf{r}) \left( c \boldsymbol{\alpha} \cdot (-i\hbar \nabla - e\mathbf{A}_0) + \beta mc^2 + e\phi \right) \psi(\mathbf{r}) + \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{8\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}$$

where  $\phi$  is a time-independent scalar external potential (for instance the nuclear potential),  $\mathbf{A}_0$  is a time-independent vector potential (describing an external magnetic field) and  $\psi$  are fermion field operators. The photon Hamiltonian is

$$H_{\gamma} = \frac{\epsilon_0}{2} \int d\mathbf{r} |\mathbf{e}^{\perp}|^2 + c^2 |\mathbf{e}|^2 = \sum_{\mathbf{k},l} \hbar \omega_{\mathbf{k},l} a_{\mathbf{k},l}^{\dagger} a_{\mathbf{k},l},$$

where l stands for the polarization of a mode (there are two independent directions for a given wavevector  $\mathbf{k}$ ). Finally, the photon-matter interaction is described by:

$$H_{e\gamma} = -ec \int d\mathbf{r} \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \cdot \mathbf{a}(\mathbf{r}) \psi(\mathbf{r}).$$

The S-matrix is

$$S = \lim_{\epsilon \to 0} \mathscr{T}(e^{-\frac{i}{\hbar} \int_{-\infty}^{\infty} H_{\epsilon}(t)dt}), \qquad (1.16)$$

where  $H_{\epsilon}(t) = e^{-\epsilon|t|} e^{iH_0 t} H_{e\gamma} e^{-iH_0 t}$ . The adiabatic switching factor  $e^{-\epsilon|t|}$  enables us to describe physical processes as matrix elements of S between eigenstates of  $H_0 = H_e + H_{\gamma}$ .

In the non-covariant approach [Heitler, 1984], using matrix elements of  $H_{\epsilon}(t)$  between eigenstates of  $H_0$ , cross sections are expressed in terms of the S-matrix and  $\mathcal{T}$ -matrix elements related by:

$$\langle m|S|n\rangle = \delta_{mn} - 2i\pi\delta(e_m - e_n)\langle m|\mathcal{T}|n\rangle.$$

Up to second order,

$$\langle m|\mathcal{T}|n\rangle = \langle m|H_{e\gamma}|n\rangle + \sum_{p} \frac{\langle m|H_{e\gamma}|p\rangle\langle p|H_{e\gamma}|n\rangle}{e_{p} - e_{n} + i\gamma},$$
(1.17)

where  $|m\rangle$ ,  $|p\rangle$  and  $|n\rangle$  are eigenstates of  $H_0$  with energy  $e_m$ ,  $e_p$  and  $e_n$ , respectively. The term  $i\gamma$  was added as a heuristic way to avoid divergence at resonance (*i.e.* when  $e_n = e_p$ ). In this approach, the operators are independent of time (Schrödinger picture).

## 1.5.2 Matrix-elements of $H_{e\gamma}$ , multipole expansion

In the Schrödinger picture, the expression for the photon field is [Strange, 1998]:

$$\mathbf{a}(\mathbf{r}) = \sum_{\mathbf{k},l} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega_{\mathbf{k}}}} \left( \boldsymbol{\epsilon}_{\mathbf{k},l} a_{\mathbf{k},l} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} + \boldsymbol{\epsilon}_{\mathbf{k},l}^{\star} a_{\mathbf{k},l}^{\dagger} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{r}} \right).$$

We denote  $|n\rangle = a_{\mathbf{k},l}^{\dagger}|0\rangle|\Psi_n\rangle$  an eigenstate of  $H_0$  where one photon is present in mode  $\mathbf{k}, l$ and the electrons are in state  $|\Psi_n\rangle$  with energy  $E_n$ . The energy of  $|n\rangle$  is  $e_n = \hbar\omega_{\mathbf{k},l} + E_n$ . The interaction Hamiltonian  $H_{e\gamma}$  is linear in  $\mathbf{A}$  which is linear in photon creation and annihilation operators so that only one-photon transitions are possible. From now on, we denote  $\omega = \omega_{\mathbf{k},l}$ ,  $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_{\mathbf{k},l}, \, \omega' = \omega_{\mathbf{k}',l'}$  and  $\boldsymbol{\epsilon}' = \boldsymbol{\epsilon}_{\mathbf{k}',l'}$ .

The state  $|n\rangle$  can make transitions to  $|a\rangle = |0\rangle |\Psi_m\rangle$  by absorption and to  $|e\rangle = a^{\dagger}_{\mathbf{k},l}a^{\dagger}_{\mathbf{k}',l'}|0\rangle |\Psi_m\rangle$  by emission. The corresponding matrix elements are:

$$\langle a|H_{e\gamma}|n\rangle = -ec\sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\boldsymbol{\epsilon} \cdot \langle \Psi_m| \int \psi^{\dagger}(\mathbf{r})\boldsymbol{\alpha}\psi(\mathbf{r})\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r}|\Psi_n\rangle,$$

and

$$\langle e|H_{e\gamma}|n\rangle = -ec\sqrt{\frac{\hbar}{2\epsilon_0 V\omega'}} \boldsymbol{\epsilon}'^{\star} \cdot \langle \Psi_m| \int \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) e^{-i\mathbf{k}'\cdot\mathbf{r}} d\mathbf{r} |\Psi_n\rangle.$$

To carry the multipole expansion we use the following result:

let  $F = \int \psi^{\dagger}(\mathbf{r}) f(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r}$ , where f is some function of **r** then,<sup>5</sup>

$$[H_0, F] = -i\hbar c \int \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) \cdot \nabla f(\mathbf{r}) d\mathbf{r}.$$
(1.18)

When expanding to first order  $e^{i\mathbf{k}\cdot\mathbf{r}} \simeq 1 + i\mathbf{k}\cdot\mathbf{r}$ , we apply Eq. (1.18) with  $f(\mathbf{r}) = \boldsymbol{\epsilon}\cdot\mathbf{r}$  and  $f(\mathbf{r}) = \boldsymbol{\epsilon}\cdot\mathbf{r} \mathbf{k}\cdot\mathbf{r}$ :

$$\begin{array}{lll} [H_e, \psi^{\dagger} \boldsymbol{\epsilon} \cdot \mathbf{r} \psi] &=& -i\hbar c \psi^{\dagger} \boldsymbol{\alpha} \psi \cdot \boldsymbol{\epsilon} \\ [H_e, \psi^{\dagger} \boldsymbol{\epsilon} \cdot \mathbf{r} \ \mathbf{k} \cdot \mathbf{r} \psi] &=& -i\hbar c \psi^{\dagger} \boldsymbol{\alpha} \psi \cdot (\boldsymbol{\epsilon} \ \mathbf{k} \cdot \mathbf{r} + \mathbf{k} \ \boldsymbol{\epsilon} \cdot \mathbf{r}) \end{array}$$

where we removed the integral signs for notational convenience. The second relation leads to  $2\psi^{\dagger}\boldsymbol{\epsilon}\cdot\boldsymbol{\alpha} \mathbf{k}\cdot\mathbf{r}\psi = \frac{i}{\hbar c}[H_{e},\psi^{\dagger}\boldsymbol{\epsilon}\cdot\mathbf{r} \mathbf{k}\cdot\mathbf{r}\psi] - \psi^{\dagger}(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}\times\boldsymbol{\alpha})\psi$ . This leads to:

$$\langle a|H_{e\gamma}|n\rangle = \frac{e\Delta E_{an}}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V\omega}} \langle \Psi_m| \int \psi^{\dagger}(\mathbf{r}) T(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} |\Psi_n\rangle, \qquad (1.19)$$

where

$$T(\mathbf{r}) = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \mathbf{k} \cdot \mathbf{r} - \frac{\hbar c}{2\Delta E_{an}} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \boldsymbol{\alpha}).$$

The first term of T is the usual electric-dipole operator, the second one is the electric-quadrupole operator and the third one will turn to be the magnetic-dipole operator (see Section 1.6). Similarly for emission,

$$\langle e|H_{e\gamma}|n\rangle = \frac{e\Delta E_{en}}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega'}} \langle \Psi_m| \int \psi^{\dagger}(\mathbf{r}) T'(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} |\Psi_n\rangle, \qquad (1.20)$$

where

$$T'(\mathbf{r}) = \boldsymbol{\epsilon}^{\prime \star} \cdot \mathbf{r} - \frac{i}{2} \boldsymbol{\epsilon}^{\prime \star} \cdot \mathbf{r} \mathbf{k}^{\prime} \cdot \mathbf{r} + \frac{\hbar c}{2\Delta E_{en}} (\boldsymbol{\epsilon}^{\prime \star} \times \mathbf{k}^{\prime}) \cdot (\mathbf{r} \times \boldsymbol{\alpha}).$$

### **1.5.3** Fully relativistic absorption and scattering cross section

The transition probability per unit time from state m to state n is related to the  $\mathcal{T}$ -matrix elements by [Walecka, 2010]:

$$w = \frac{2}{\hbar} \delta_{mn} \operatorname{Im} \langle m | \mathcal{T} | m \rangle + \frac{2\pi}{\hbar} \delta(e_m - e_n) | \langle m | \mathcal{T} | n \rangle |^2.$$
(1.21)

<sup>5</sup> To obtain this, we go to the interaction picture and define  $F_I(t) = e^{iH_0t/\hbar}Fe^{-iH_0t/\hbar}$ . Then, the timederivative  $\dot{F}_I$  of  $F_I$  is given by  $-i\hbar\dot{F}_I(t) = [H_0, F_I(t)]$ . Now, we notice that F is related to the density operator  $\rho(\mathbf{r}) = \psi^{\dagger}(\mathbf{r})\psi(\mathbf{r})$  by  $F = \int \rho(\mathbf{r})f(\mathbf{r})d\mathbf{r}$ . Thus,  $-i\hbar\dot{F}_I(t) = -i\hbar\int\dot{\rho}(\mathbf{r},t)f(\mathbf{r})d\mathbf{r} = [H_0, F_I(t)]$ . If  $H_0$  conserves the electric charge, the continuity equation  $e\dot{\rho}(\mathbf{r}) = -\nabla \cdot \mathbf{j}$  holds, where  $\mathbf{j}$  is the electric current operator. By taking t = 0 to recover the operators in the Schrödinger picture, we obtain:

$$[H_0, F] = \frac{i\hbar}{e} \int \nabla \cdot \mathbf{j}(\mathbf{r}) f(\mathbf{r}) d\mathbf{r} = -\frac{i\hbar}{e} \int \mathbf{j}(\mathbf{r}) \cdot \nabla f(\mathbf{r}) d\mathbf{r}$$

Finally,  $\mathbf{j}(\mathbf{r}) = ec\psi^{\dagger}(\mathbf{r})\boldsymbol{\alpha}\psi(\mathbf{r})$  leads to Eq. (1.18).

### Absorption cross section

The absorption cross section is derived by assuming that initially the electrons are in state  $|I\rangle$  with energy  $E_i$  and that a photon  $\mathbf{k}, \boldsymbol{\epsilon}, \boldsymbol{\omega}$  is present. In the final state, there is no photon and the electrons are in state  $|F\rangle$ . The transition rate must be divided by c/V (rate at which the photon crosses a unit of surface) to obtain the cross section. Since we consider real transitions (*i.e.*  $m \neq n$ ), only the second term in Eq. (1.21) is present. From Eq. (1.19) the absorption cross section is:

$$\sigma = 4\pi^2 \alpha_0 \frac{\Delta E_{fi}^2}{\hbar \omega} \sum_f |\langle F|T|I \rangle|^2 \delta(E_f - E_i - \hbar \omega)$$
  
=  $4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle F|T|I \rangle|^2 \delta(E_f - E_i - \hbar \omega).$  (1.22)

In a many-body formalism, T writes as a sum over the electrons labeled n:

$$T = \sum_{n} \boldsymbol{\epsilon} \cdot \mathbf{r}_{n} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r}_{n} \mathbf{k} \cdot \mathbf{r}_{n} - \frac{\hbar c}{2\Delta E_{fi}} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r}_{n} \times \boldsymbol{\alpha}_{n}).$$

#### Scattering cross section

The scattering cross section is derived by assuming that initially the electrons are in state  $|I\rangle$  with energy  $E_i$  and that a photon  $\mathbf{k}_i, \boldsymbol{\epsilon}_i, \omega_i$  is present. In the final state, a photon  $\mathbf{k}_f, \boldsymbol{\epsilon}_f, \omega_f$  is present and the electrons are in state  $|F\rangle$ . We do not consider the special case when  $\mathbf{k}_i, \boldsymbol{\epsilon}_i = \mathbf{k}_f, \boldsymbol{\epsilon}_f$ .

The scattering cross section is related to w by [Als-Nielsen and McMorrow, 2000, p.265]:

$$\frac{d^2\sigma}{d\Omega d\omega_f} = \frac{V^2}{(2\pi)^3}\omega_f^2 \frac{1}{\hbar c^4}w.$$

As only one-photon processes are possible, there can not be a contribution from  $\langle m|H_{e\gamma}|n\rangle$  in Eq. (1.17) and the transition rate writes as a sum over intermediate states that contain either no photon ( $\mathbf{k}_i, \boldsymbol{\epsilon}_i$  absorbed) or two photons ( $\mathbf{k}_f, \boldsymbol{\epsilon}_f$  emitted).

Without carrying the multipole expansion:

$$\frac{d^{2}\sigma}{d\Omega d\omega_{f}} = (r_{e}mc^{2})^{2} \frac{\omega_{f}}{\omega_{i}} \sum_{F} \delta(E_{f} + \hbar\omega_{f} - E_{i} - \hbar\omega_{i})$$

$$\left| \sum_{L} \frac{\langle F|\sum_{n} e^{-i\mathbf{k}_{f} \cdot \mathbf{r}_{n}} \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\epsilon}_{f}^{\star}|L\rangle \langle L|\sum_{n} e^{i\mathbf{k}_{i} \cdot \mathbf{r}_{n}} \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\epsilon}_{i}\psi|I\rangle}{E_{i} - E_{l} + \hbar\omega_{i} + i\gamma} + \frac{\langle F|\sum_{n} e^{i\mathbf{k}_{i} \cdot \mathbf{r}_{n}} \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\epsilon}_{i}\psi|L\rangle \langle L|\sum_{n} e^{-i\mathbf{k}_{f} \cdot \mathbf{r}_{n}} \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\epsilon}_{f}^{\star}|I\rangle}{E_{i} - E_{l} - \hbar\omega_{f}} \right|^{2} (1.23)$$

where  $r_e$  is the classical electron radius.

In this expression, the sum over  $|L\rangle$  involves a complete set of states, with positive and negative energies. If  $|L\rangle$  is a positive energy state, only the first sum can be resonant. Indeed as  $|I\rangle$  is the ground state  $E_i - E_l < 0$  can not be equal to  $\hbar\omega_f$ . If  $|L\rangle$  is a negative energy state, none of the two sums can be resonant. Indeed,  $E_i - E_l \approx 2mc^2$  (the transition from L to I corresponds to the emission of an electron positron pair) can not be equal to  $\hbar\omega$  in standard experimental conditions.

# 1.6 Many-body Foldy-Wouthuysen transformation

For the semi-relativistic expansion, we have seen in Section 1.3 that it was not correct to start from FW Hamiltonians to determine the cross sections. However, now that we have the expression for the cross sections in a fully-relativistic framework, we can apply a time-independent FW transformation to the wave functions:

$$\langle M|T|N\rangle = \langle m|U_0TU_0^{\dagger}|n\rangle \tag{1.24}$$

where m and n are the FW transform of M and N and  $U_0$  is the time-independent FW operator.

The static FW operator  $U_0$  is known for one-body systems but not for many-body ones. Therefore, we derive a many-body version of FW transformation here.

# 1.6.1 Time-independent many-body Foldy-Wouthuysen transformation

In the literature, the Foldy-Wouthuysen transformation was studied for two-body Hamiltonians [Chraplyvy, 1953, Eriksen, 1958], but the results were rather complicated and not easy to extend to many bodies.

Here, our aim is to obtain wave functions that are even or odd. The generalization of  $\beta$  for a many-body systems is  $\eta = \beta_1 \otimes \cdots \otimes \beta_n$  that verifies  $\eta^{\dagger} = \eta$  and  $\eta^2 = 1$  which is enough to define parity.<sup>6</sup>

For many-body systems, even or odd wave functions do not correspond to wave functions with only lower or upper components.<sup>7</sup>

The many-body Dirac Hamiltonian:

...

$$H_N^D = \sum_{n=1}^N \left[ \beta_n m c^2 + e\phi_0(\mathbf{r}_n) + c \boldsymbol{\alpha}_n \cdot (\mathbf{p}_n - e\mathbf{A}_0(\mathbf{r}_n)) + \sum_{m \neq n} eV(\mathbf{r}_m - \mathbf{r}_n) \right]$$

where  $V(\mathbf{r}) = \frac{e}{8\pi\epsilon_0|\mathbf{r}|}$  is the Coulomb potential and  $\phi_0(\mathbf{r})$  describes the Coulomb interaction of the electron with the nuclei. The even and odd part of  $H_N^D$  are therefore  $\sum_{n=1}^N \beta_n mc^2 + \mathcal{E}$  and

<sup>7</sup> For example, for a two body system, if  $|\psi_D\rangle = \begin{pmatrix} \phi_1 \\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} \phi_2 \\ \psi_2 \end{pmatrix}$ . Then, the even part of  $|\psi_D\rangle$  is

$$\left(\begin{array}{c}\phi_1\\0\end{array}\right)\otimes\left(\begin{array}{c}\phi_2\\0\end{array}\right)+\left(\begin{array}{c}0\\\psi_1\end{array}\right)\otimes\left(\begin{array}{c}0\\\psi_2\end{array}\right),$$

while its odd part is

$$\left(\begin{array}{c}\phi_1\\0\end{array}\right)\otimes\left(\begin{array}{c}0\\\psi_2\end{array}\right)+\left(\begin{array}{c}0\\\psi_1\end{array}\right)\otimes\left(\begin{array}{c}\phi_2\\0\end{array}\right).$$

<sup>&</sup>lt;sup>6</sup> Let's consider a self-adjoint operator  $\eta$  such that  $\eta^2 = 1$ . It can be used to define projectors  $B_{\pm} = (1 \pm \eta)/2$ . It is clear that  $B_+ + B_- = 1$ ,  $B_{\pm}^2 = B_{\pm}$ ,  $B_{\pm}^{\dagger} = B_{\pm}$  and  $B_+B_- = B_-B_+ = 0$ . A vector  $|\psi\rangle$  is said to be even (odd) if  $\eta|\psi\rangle = |\psi\rangle \ (\eta|\psi\rangle = -|\psi\rangle)$ . Then, any vector  $|\psi\rangle$  can be written as the sum of its even part  $B_+|\psi\rangle$ and its odd part  $B_-|\psi\rangle$ . An operator H is said to be even (odd) if it transforms an even state into an even (odd) state and an odd state into an odd (even) state. An operator O is even (odd) if and only if  $\eta O \eta = O$  $(\eta O \eta = -O)$ . Any operator H can be written as the sum of its even part  $B_+OB_+ + B_-OB_-$  and its odd part  $B_+OB_- + B_-OB_+$ .

 $\mathcal{O}$  with:

$$\mathcal{E} = e \sum_{n=1}^{N} \phi_0(\mathbf{r}_n) + e \sum_{m \neq n} V(\mathbf{r}_m - \mathbf{r}_n) \text{ and } \mathcal{O} = \sum_{n=1}^{N} c \boldsymbol{\alpha}_n \cdot \boldsymbol{\pi}_n = \sum_{n=1}^{N} \mathcal{O}_n$$

where  $\boldsymbol{\pi}_n = \mathbf{p}_n - e\mathbf{A}_0(\mathbf{r}_n)$ .

In analogy with the one-body case, at the first order in  $c^{-1}$ , the operator is  $U_0 = e^{iS^{(1)}}$  where

$$S^{(1)} = -\frac{i}{2mc^2} \sum_n \beta_n \mathcal{O}_n$$

is chosen as the Foldy-Wouthuysen operator. Indeed, when expanding  $H^{FW(1)} = e^{iS^{(1)}}H_N^D e^{-iS^{(1)}}$  using Baker-Campbell-Hausdorf formula:

$$H^{FW(1)} = H_N^D + i[S^{(1)}, H_N^D] + \frac{i^2}{2}[S^{(1)}, [S^{(1)}, H_N^D]] + \dots$$

the term  $i[S^{(1)}, \sum_n \beta_n mc^2]$  compensates exactly for  $\mathcal{O}$ . Moreover, the conditions  $S^{(1)\dagger} = S^{(1)}$ and  $\eta S^{(1)}\eta = -S^{(1)}$  are satisfied.<sup>8</sup>

From this point, the FW transformation can be carried as in [Greiner, 2000, p285] and we find:

$$H_0^{\rm FW} = \sum_n \beta_n mc^2 + \mathcal{E} + \frac{1}{2mc^2} \sum_n \beta_n \mathcal{O}_n^2 + \frac{1}{8m^2c^4} \left[\sum_m \beta_m \mathcal{O}_m, \left[\sum_n \beta_n \mathcal{O}_n, \mathcal{E}\right]\right] - \frac{1}{8m^3c^6} \sum_n \beta_n \mathcal{O}_n^4$$

In the double commutator, the two cases m = n and  $m \neq n$  can be separated which leads to:

$$H_0^{\rm FW} = \sum_n \beta_n m c^2 + \mathcal{E} + \frac{1}{2mc^2} \sum_{n=1}^N \beta_n \mathcal{O}_n^2 - \frac{1}{8m^2c^4} \sum_n \left[ \mathcal{O}_n[\mathcal{O}_n, \mathcal{E}] \right] + \frac{1}{8m^2c^4} \sum_{m \neq n} \beta_m \beta_n [\mathcal{O}_m[\mathcal{O}_n, \mathcal{E}]] - \frac{1}{8m^3c^6} \sum_n \beta_n \mathcal{O}_n^4$$

Using the formula for  $\mathcal{O}$  and  $\mathcal{E}$ :<sup>9</sup>

$$H_0^{\rm FW} = \sum_{n=1}^N H_n^{\rm FW} + H_{MB}^{\rm FW}$$
(1.26)

 $^{8}$  The commutator,

$$[S^{(1)}, \sum_{n} \beta_{n} mc^{2}] = \frac{-i}{2} \sum_{n,n'} \beta_{n} \mathcal{O}_{n} \beta_{n'} - \beta_{n'} \beta_{n} \mathcal{O}_{n} = \frac{-i}{2} \Big( \sum_{n \neq n'} \beta_{n} \beta_{n'} \mathcal{O}_{n} - \beta_{n'} \beta_{n} \mathcal{O}_{n} + \sum_{n} -2\mathcal{O}_{n} \Big)$$
$$= i\mathcal{O}$$

because  $[\beta_{n'}, \beta_n] = 0.$ 

<sup>9</sup> The double commutator  $[\mathcal{O}_n[\mathcal{O}_n, \mathcal{E}]]$  rewrites:

$$[\mathcal{O}_n[\mathcal{O}_n, \mathcal{E}]] = e[\mathcal{O}_n[\mathcal{O}_n, \sum_p \phi_0(\mathbf{r}_p)]] + e[\mathcal{O}_n[\mathcal{O}_n, \sum_{p,q \neq p} V(\mathbf{r}_q - \mathbf{r}_p)]]$$

In this expression,

$$\begin{aligned} [\mathcal{O}_n, \phi_0(\mathbf{r}_p)] &= -i\hbar c [\boldsymbol{\alpha}_n \cdot \nabla_n, \phi_0(\mathbf{r}_p)] = -i\hbar c \delta_{p,n} \boldsymbol{\alpha}_p \cdot [\nabla_p, \phi_0(\mathbf{r}_p)] = -i\hbar c \delta_{p,n} \boldsymbol{\alpha}_p \cdot \nabla \phi_0(\mathbf{r}_p), \\ [\mathcal{O}_n, V(\mathbf{r}_q - \mathbf{r}_p)] &= -i\hbar c \delta_{p,n} \boldsymbol{\alpha}_p \cdot \nabla V(\mathbf{r}_p - \mathbf{r}_q) + i\hbar c \delta_{q,n} \boldsymbol{\alpha}_q \cdot \nabla V(\mathbf{r}_p - \mathbf{r}_q). \end{aligned}$$

where  $H_n^{\rm FW}$  are the usual one-body FW Hamiltonians:

$$H_n^{\rm FW} = \beta_n mc^2 + e\phi_0(\mathbf{r}_n) + e \sum_{m \neq n} V(\mathbf{r}_m - \mathbf{r}_n) + \beta_n \frac{\boldsymbol{\pi}_n^2}{2m} - \frac{e\hbar}{2mc} \boldsymbol{\Sigma}_n \cdot \mathbf{B}_0(\mathbf{r}_n) - \frac{e\hbar^2}{8m^2c^2} \nabla \mathbf{E}_0(\mathbf{r}_n) + \frac{e\hbar}{8m^2c^2} \boldsymbol{\Sigma}_n \cdot (\boldsymbol{\pi}_n \times \mathbf{E}_0(\mathbf{r}_n) - \mathbf{E}_0(\mathbf{r}_n) \times \boldsymbol{\pi}_n) - \beta_n \frac{(\boldsymbol{\pi}_n \cdot \boldsymbol{\pi}_n)^2}{8m^3c^2}$$

with  $\mathbf{B}_0(\mathbf{r}_n) = \nabla \times \mathbf{A}_0(\mathbf{r}_n), \ \mathbf{E}_0(\mathbf{r}_n) = -\nabla \phi_0(\mathbf{r}_n) - \sum_{q \neq n} \nabla V(\mathbf{r}_n - \mathbf{r}_q) \text{ and } \mathbf{\Sigma}_n = \begin{pmatrix} \boldsymbol{\sigma}_n & 0 \\ 0 & \boldsymbol{\sigma}_n \end{pmatrix}.$ The other term  $H_{MB}^{\text{FW}}$  arises because  $V(\mathbf{r}_p - \mathbf{r}_q)$  is a two-body operator:

$$\begin{split} H_{MB}^{\mathrm{FW}} &= \sum_{n,q \neq n} \frac{e\hbar^2}{8m^2c^2} \Delta V(\mathbf{r}_n - \mathbf{r}_q) - \frac{e\hbar}{8m^2c^2} \boldsymbol{\Sigma}_n \cdot (\boldsymbol{\pi}_n \times \nabla V(\mathbf{r}_n - \mathbf{r}_q) - \nabla V(\mathbf{r}_n - \mathbf{r}_q) \times \boldsymbol{\pi}_n) \\ &+ \frac{e\hbar^2}{4m^2c^2} \sum_{n,m \neq n} \beta_n \beta_m (\boldsymbol{\alpha}_m \cdot \nabla) (\boldsymbol{\alpha}_n \cdot \nabla) V(\mathbf{r}_n - \mathbf{r}_m). \end{split}$$

The first line in this expression is due to the factor of 2 in front of  $\nabla V$  in  $[\mathcal{O}_n, \mathcal{E}]$  (Eq. (1.25)) and the second line is due to the sum over  $m \neq n$  in the expression of  $H_0^{\text{FW}}$ .

As 
$$\sum_{q} \nabla V(\mathbf{r}_{n} - \mathbf{r}_{q}) = -\sum_{p} \nabla V(\mathbf{r}_{p} - \mathbf{r}_{n})$$
 due to the fact that  $\nabla V$  is a odd function,  

$$[\mathcal{O}_{n}, \mathcal{E}] = -ie\hbar c \big( \boldsymbol{\alpha}_{n} \cdot \nabla \phi_{0}(\mathbf{r}_{n}) + 2\sum_{q \neq n} \boldsymbol{\alpha}_{n} \cdot \nabla V(\mathbf{r}_{n} - \mathbf{r}_{q}) \big).$$
(1.25)

We must therefore calculate two commutators for which we use the formula in Appendix B.4:

• 
$$[\mathcal{O}_n, \boldsymbol{\alpha}_n \cdot \nabla \phi_0(\mathbf{r}_n)] = c[\boldsymbol{\alpha}_n \cdot \boldsymbol{\pi}_n, \boldsymbol{\alpha}_n \cdot \nabla \phi_0(\mathbf{r}_n)]$$
  
 $= c \sum_i [\boldsymbol{\pi}_{n,i}, \nabla_i \phi_0(\mathbf{r}_n)] + ic \boldsymbol{\Sigma} \cdot (\boldsymbol{\pi}_n \times \nabla \phi_0(\mathbf{r}_n) - \nabla \phi_0(\mathbf{r}_n) \times \boldsymbol{\pi}_n)$   
 $= -i\hbar c \Delta \phi_0(\mathbf{r}_n) + ic \boldsymbol{\Sigma} \cdot (\boldsymbol{\pi}_n \times \nabla \phi_0(\mathbf{r}_n) - \nabla \phi_0(\mathbf{r}_n) \times \boldsymbol{\pi}_n)$   
•  $[\mathcal{O}_n, \boldsymbol{\Sigma}, \boldsymbol{\alpha}_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_n)] = c \sum_i [\boldsymbol{\alpha}_n \cdot \boldsymbol{\pi}_n, \boldsymbol{\alpha}_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_n)]$ 

• 
$$[\mathcal{O}_n, \sum_{q \neq n} \alpha_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_q)] = c \sum_{q \neq n} [\alpha_n \cdot \pi_n, \alpha_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_q)]$$
  
=  $\sum_{q \neq n} -i\hbar c \Delta V(\mathbf{r}_n - \mathbf{r}_q) + ic \mathbf{\Sigma} \cdot (\pi_n \times \nabla V(\mathbf{r}_n - \mathbf{r}_q) - \nabla V(\mathbf{r}_n - \mathbf{r}_q) \times \pi_n)$ 

The second double commutator with  $m \neq n$ :

$$\beta_m \beta_n [\mathcal{O}_m, [\mathcal{O}_n, \mathcal{E}]] = \beta_m \beta_n [\mathcal{O}_m, -ie\hbar c \big( \boldsymbol{\alpha}_n \cdot \nabla \phi_0(\mathbf{r}_n) + 2 \sum_{q \neq n} \boldsymbol{\alpha}_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_q) \big)]$$

As  $\alpha_n \cdot \nabla \phi_0(\mathbf{r}_n)$  commutes with  $\mathcal{O}_m$ , only the second term in  $[\mathcal{O}_n, \mathcal{E}]$  must be considered:

• 
$$[\mathcal{O}_m, \sum_{q \neq n} \alpha_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_q)] = c \sum_{q \neq n} [\alpha_m \cdot \pi_m, \alpha_n \cdot \nabla V(\mathbf{r}_n - \mathbf{r}_q)]$$
  
 $= c \sum_{q \neq n} \sum_{ij} \alpha_m^i \alpha_n^j [\pi_{m,i}, \nabla_j V(\mathbf{r}_n - \mathbf{r}_q)]$   
 $= -i\hbar c \sum_{q \neq n} \sum_{ij} \alpha_m^i \alpha_n^j \delta_{mq} (-\nabla_i \nabla_j V(\mathbf{r}_n - \mathbf{r}_m))$   
 $= i\hbar c (\alpha_m \cdot \nabla) (\alpha_n \cdot \nabla) V(\mathbf{r}_n - \mathbf{r}_m)$ 

The expression for  $\mathcal{O}_n^2$  and  $\mathcal{O}_n^4$  can be found in [Strange, 1998, p.220].

### **1.6.2** Properties of this transformation

At the first step, the FW operator,  $U_0 = e^{iS^{(1)}}$  is a tensor product of one-body FW operators:  $U_0 = e^{iS_1^{(1)}} \otimes ... \otimes e^{iS_N^{(1)}}$ . At the next step, however,  $U_0$  can not be written as a tensor product because  $S^{(2)}$  contains two-body operators. With expansion parameter  $m^{-2}$ , in analogy with the formula in [Eriksen, 1958],  $U_0$  writes:

$$U_0 = 1 + \frac{1}{2mc^2} \sum_n \beta_n \mathcal{O}_n - \frac{1}{8m^2c^4} \left(\sum_n \beta_n \mathcal{O}_n\right)^2 + \frac{1}{4m^2c^4} \sum_n \beta_n \left[\sum_m \beta_m \mathcal{O}_m, \mathcal{E}\right]$$
(1.27)

and it obeys  $U_0 = \eta U_0^{\dagger} \eta$ . Also, if  $H_N^D$  is changed to  $-H_N^D$ <sup>10</sup>, then  $U_0^{\dagger} \eta U_0 = \eta U_0^2$  is changed to  $-U_0^{\dagger} \eta U_0$ . This property is required to prove that

$$U_0^{\dagger} \eta U_0 = \operatorname{sign} H_N^D = \lambda \tag{1.28}$$

which is the many-body version of Eq. (1.5) (see the paper in Appendix F).

As in the one-body case, this relation implies that the transform of a wave function corresponding to a positive energy state is even and the transform of a wave function corresponding to a negative energy state is odd.

# **1.7** Semi-relativistic absorption cross sections

As the transition operator T in Eq. (1.22) is of order 0 in c, we can use  $U_0 = e^{iS^{(1)}}$  to compute  $U_0TU_0^{\dagger}$  at order  $c^{-2}$ . Indeed,  $S^{(2)} \propto \frac{1}{c^3}$ .

As  $T = \sum_{n} T^{n}$  where  $T^{n}$  are one-body operators and  $U_{0} = e^{iS_{1}^{(1)}} \otimes ... \otimes e^{iS_{N}^{(1)}}$ ,

$$U_0 T U_0^{\dagger} = \sum_n e^{iS_n^{(1)}} T^n e^{-iS_n^{(1)}}$$
(1.29)

In other words, the action of  $U_0$  can be calculated for each variable independently. We present the derivation in the one-body case omitting the index n that will be restored at the end of the calculation.

## 1.7.1 Derivation of the transform of T in the one-body case

Using Baker-Campbell-Hausdorff formula,

$$U_0 T U_0^{\dagger} = e^{iS^{(1)}} T e^{-iS^{(1)}} = T + \frac{1}{2mc^2} [\beta \mathcal{O}, T] + \frac{1}{8m^2c^4} [\mathcal{O}, [T, \mathcal{O}]]$$
(1.30)

For the electric dipole term, we stop the development at order  $c^{-2}$  and for the term proportional to **k**, at order  $c^{-1}$ . Moreover, as FW wave functions are even  $(|F\rangle)$  and  $|I\rangle$  in Eq. (1.22) are positive energy states), all the odd operators in Eq. (1.30) have a zero contribution to the matrix elements. This is the case for the operators  $[\beta \mathcal{O}, \boldsymbol{\epsilon} \cdot \mathbf{r}], [\beta \mathcal{O}, (\boldsymbol{\epsilon} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{r})]$  and  $(\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})$ . We do not consider them so that,

$$U_0 T U_0^{\dagger} = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{1}{8m^2c^4} [\mathcal{O}, [\boldsymbol{\epsilon} \cdot \mathbf{r}, \mathcal{O}]] + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \, \mathbf{k} \cdot \mathbf{r} - \frac{\hbar c}{2\Delta E} \frac{1}{2mc^2} [\beta \mathcal{O}, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})].$$

<sup>10</sup> Formally, to mimic a change  $H_N^D \to -H_N^D$ , we change all  $\beta_n$  to  $-\beta_n$ ,  $\mathcal{O}$  to  $-\mathcal{O}$  and  $\mathcal{E}$  to  $-\mathcal{E}$  in the formula for  $U_0$  Eq. (1.27).

Using the canonical commutation relation and the properties of Dirac alpha matrix, we obtain:<sup>11</sup>

$$[\mathcal{O}, [\boldsymbol{\epsilon} \cdot \mathbf{r}, \mathcal{O}]] = -2\hbar c^2 (\mathbf{p} - e\mathbf{A}_0) \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}).$$

As  $\{\beta, \alpha\} = 0$ , the other commutator writes as a sum of anticommutators:

$$[\beta \mathcal{O}, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})] = c\beta(\{\boldsymbol{\alpha} \cdot \mathbf{p}, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})\} - e\{\boldsymbol{\alpha} \cdot \mathbf{A}_0, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})\})$$

The two anticommutators rewrite:<sup>12</sup>

$$\{\boldsymbol{\alpha} \cdot \mathbf{p}, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})\} = 2(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma} + \mathbf{r} \times \mathbf{p})$$
  
and  $\{\boldsymbol{\alpha} \cdot \mathbf{A}_0, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})\} = 2(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \mathbf{A}_0).$ 

Finally,

$$U_0 T U_0^{\dagger} = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} (\boldsymbol{\epsilon} \cdot \mathbf{r}) (\mathbf{k} \cdot \mathbf{r}) - \frac{\hbar \beta}{2m\Delta E} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma} + \boldsymbol{\Lambda}) - \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi} \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}).$$

where  $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}_0$  and  $\boldsymbol{\Lambda} = \mathbf{r} \times (\mathbf{p} - e\mathbf{A}_0) = \mathbf{L} - e\mathbf{r} \times \mathbf{A}_0$  is the moment of mechanical momentum [Cohen-Tannoudji et al., 1973].

# 1.7.2 Many-body semi-relativistic absorption cross section

Using Eq. (1.29) and Eq. (1.22) and the formula for  $e^{iS_n^{(1)}}T^n e^{-iS_n^{(1)}}$  the absorption cross section can be expressed in terms of the many-body FW wave-functions  $|i\rangle$  and  $|f\rangle$ :

$$\sigma(\hbar\omega) = \frac{4\pi^2 \alpha_0 (\Delta E_{fi})^2}{\hbar\omega} \sum_f |\langle f|T_{\rm FW}|i\rangle|^2 \delta(\Delta E_{fi} - \hbar\omega)$$
(1.31)

<sup>11</sup> The commutator  $[\boldsymbol{\epsilon} \cdot \mathbf{r}, \mathcal{O}] = c[\boldsymbol{\epsilon} \cdot \mathbf{r}, \boldsymbol{\alpha} \cdot \mathbf{p}] - ec[\boldsymbol{\epsilon} \cdot \mathbf{r}, \boldsymbol{\alpha} \cdot \mathbf{A}_0] = c \sum_{i,j} (\epsilon_j r_j \alpha_i p_i - \alpha_i p_i \epsilon_j r_j) = c \sum_{i,j} \alpha_i \epsilon_j [r_j, p_i] = i\hbar c \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}$  so that using the formula in Appendix B.4:

$$\begin{aligned} [\mathcal{O}, [\boldsymbol{\epsilon} \cdot \mathbf{r}, \mathcal{O}]] &= i\hbar c^2 [\boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{A}_0), \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}] = i\hbar c^2 (\sum_i [(\mathbf{p} - e\mathbf{A}_0)_i, \boldsymbol{\epsilon}_i] + i\boldsymbol{\Sigma} \cdot ((\mathbf{p} - e\mathbf{A}_0) \times \boldsymbol{\epsilon}) \\ &= -2\hbar c^2 \boldsymbol{\Sigma} \cdot ((\mathbf{p} - e\mathbf{A}_0) \times \boldsymbol{\epsilon}) \end{aligned}$$

<sup>12</sup> The first anticommutator:

$$\begin{aligned} \{\boldsymbol{\alpha} \cdot \mathbf{p}, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})\} &= \sum_{ijkl} \alpha_i p_i (\boldsymbol{\epsilon} \times \mathbf{k})_j \epsilon_{jkl} r_k \alpha_l + (\boldsymbol{\epsilon} \times \mathbf{k})_j \epsilon_{jkl} r_k \alpha_l \alpha_i p_i = \sum_{ijkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j (\alpha_i \alpha_l p_i r_k + r_k p_i \alpha_l \alpha_i) \\ &= \sum_{ijkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j (\alpha_i \alpha_l p_i r_k + (i\hbar \delta_{ik} + p_i r_k) \alpha_l \alpha_i) \\ &= i\hbar \sum_{jkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j \alpha_l \alpha_k + \sum_{ijkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j p_i r_k (\alpha_i \alpha_l + \alpha_l \alpha_i) \\ &= i\hbar \sum_{jkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j (\delta_{lk} + i \sum_m \epsilon_{lkm} \Sigma_m) + \sum_{ijkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j p_i r_k 2\delta_{il} \\ &= i\hbar \sum_j -2i(\boldsymbol{\epsilon} \times \mathbf{k})_j \Sigma_j + (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \mathbf{p}) = 2(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \Sigma + \mathbf{L}). \end{aligned}$$

To obtain the last line, we used two properties of Levi-Civita symbols:  $\epsilon_{jll} = 0$  and  $\sum_{kl} \epsilon_{klj} \epsilon_{klm} = 2\delta_{jm}$ . The second anticommutator requires less steps:

$$\{\boldsymbol{\alpha} \cdot \mathbf{A}_{0}, (\boldsymbol{\epsilon} \times \mathbf{k})(\mathbf{r} \times \boldsymbol{\alpha})\} = \sum_{ijkl} \epsilon_{ikl} (\boldsymbol{\epsilon} \times \mathbf{k})_{i} A_{0j} r_{k} \{\alpha_{j}, \alpha_{l}\} = 2 \sum_{ikl} \epsilon_{ikl} (\boldsymbol{\epsilon} \times \mathbf{k})_{i} A_{0l} r_{k} = 2(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \mathbf{A}_{0})$$

where

$$T_{\rm FW} = \sum_{n} \boldsymbol{\epsilon} \cdot \mathbf{r}_n + \frac{i}{2} (\boldsymbol{\epsilon} \cdot \mathbf{r}_n) (\mathbf{k} \cdot \mathbf{r}_n) - \frac{\hbar \beta_n}{2m \Delta E_{fi}} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n) - \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}_n).$$

The first two operators are the usual electric-dipole and electric-quadrupole operators. The third one is the magnetic-dipole operator but with  $\Lambda_n = \mathbf{L}_n - e\mathbf{r}_n \times \mathbf{A}_0(\mathbf{r}_n)$  instead of the  $\mathbf{L}_n$  that usually enters it through the total angular momentum of the electron  $(\hbar \Sigma_n + \mathbf{L}_n)$ . The amplitude of the  $-e\mathbf{A}_0$  part of this term depends on the choice of the space origin in the Coulomb gauge for  $\mathbf{A}_0$  but it does not make the cross section gauge dependent. Indeed, when choosing the origin of the gauge, the states are changed accordingly (see Section 1.4). If the origin of the gauge is chosen at the atom position (which is the usual choice), fields way beyond laboratory accessible values (larger than  $10^6$  T) are required for this term to be significant. We rewrite the last term using  $\pi_n = (m/i\hbar)[\mathbf{r}_n, H_0^{\text{FW}}] + O(c^{-2})$ :

$$\langle f| - \frac{\hbar}{4m^2c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}_n) |i\rangle = \langle f| \frac{i\Delta E_{fi}}{4mc^2} \boldsymbol{\Sigma}_n \cdot (\boldsymbol{\epsilon} \times \mathbf{r}_n) |i\rangle.$$

We call spin-position operator the operator  $\Sigma_n \cdot (\boldsymbol{\epsilon} \times \mathbf{r}_n)$ .

In the one-body case, we find the same cross section as the one obtained from the naive calculation in Subsection 1.3.2 (Eq. (1.10)) by projection of  $T_{\rm FW}$  on large components:

$$T_{\mathrm{FW},p} = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} (\boldsymbol{\epsilon} \cdot \mathbf{r}) (\mathbf{k} \cdot \mathbf{r}) - \frac{\hbar}{2m\Delta E_{fi}} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\sigma} + \boldsymbol{\Lambda}) - \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi} \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\sigma})$$

With our method, we eliminated the uncertainty related to the choice of the starting Hamiltonian. We intentionally decided not to use the fact that  $\Delta E_{fi} = \hbar \omega$  in the expressions for the cross section to highlight the fact that the spin-position operator obtained from a fullyrelativistic calculation differs from the spin-position operator obtained from a naive calculation by a factor  $\frac{\Delta E_{fi}}{\hbar \omega}$  which is equal to 1 in the case of absorption but not in other kind of experiments like scattering.

# 1.7.3 Rewriting the cross section with large components of Dirac wave functions for the core state

As this subsection concerns a problem related to numerical implementation, we use the one-body cross section. For practical implementation, it can be useful to consider a different situation: the use of FW wave function for the final states and large components of the Dirac wave function for the initial (core) state. In practice, this difference in treatment is linked to the fact that the core wave function can be determined from a relativistic atomic code whereas the unoccupied states are often calculated with a semi-relativistic condensed-matter code.

We note  $\phi_i$  and  $\chi_i$  the large and small components of  $|I\rangle$ . The order of magnitude of the ratio between small and large components is v/c where v is the velocity of the particle.[Strange, 1998] Up to order  $c^{-1}$ , the small component writes [Strange, 1998, van Lenthe et al., 1996]:

$$\chi_i = \frac{1}{2mc} \boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{a}_0)\phi_i. \tag{1.32}$$

We remind, that at order  $c^{-2}$ , the FW operator is:

$$U_0 = 1 + \frac{\beta}{2mc^2}\mathcal{O} - \frac{1}{8m^2c^4}\mathcal{O}^2.$$
 (1.33)

Only the second term in  $U_0$  couples the small and the large components. From Eqs (1.32) and (1.33), the large component of the FW transformed wave function can be expressed as a function of the large components of Dirac wave functions up to order  $c^{-2}$ ,

$$\phi_i^{\text{FW}} = (1 - \frac{1}{8m^2c^4} [\mathcal{O}^2]_p)\phi_i + \frac{1}{4mc^3} \mathcal{O}_p \boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{a}_0)\phi_i.$$

 $[\mathcal{O}^2]_p$  is the projection of  $\mathcal{O}^2$  on large components[Strange, 1998]:  $[\mathcal{O}^2]_p = c^2(\mathbf{p} - e\mathbf{a}_0)^2 - c^2 e\hbar \boldsymbol{\sigma}.\mathbf{B}_0$  and  $\mathcal{O}_p = c\boldsymbol{\sigma}.(\mathbf{p} - e\mathbf{a}_0)$  is the projection of  $\beta \mathcal{O}$  on the upper right components.

The fact that  $c\mathcal{O}_p\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{a}_0)=[\mathcal{O}^2]_p$  leads to:

$$\phi_i^{\text{FW}} = (1 + \frac{1}{8m^2c^4}[\mathcal{O}^2]_p)\phi_i.$$

From this relation, the cross section obtained in the previous subsection can be adapted to the case that we consider here:

$$\sigma(\hbar\omega) = 4\pi^2 \alpha_0 \hbar\omega \sum_f |\langle \phi_f^{\rm FW} | T'_{\rm FW,p} | \phi_i \rangle|^2 \delta(E_f - E_i - \hbar\omega)$$
(1.34)

where:  $T'_{\rm FW} = T_{\rm FW,p} (1 + \frac{1}{8m^2c^4} [\mathcal{O}^2]_p) = T_{\rm FW,p} (1 + \frac{1}{8m^2c^2} ((\mathbf{p} - e\mathbf{a}_0)^2 - e\hbar\boldsymbol{\sigma}.\mathbf{B}_0)).$ The expansion for  $T_{\rm FW,p}$  was made to order  $c^{-2}$  for the dipole contribution and to order kr

The expansion for  $T_{FW,p}$  was made to order  $c^{-2}$  for the dipole contribution and to order krand  $c^0$  for multipole contributions. At the same order,

$$T'_{\mathrm{FW},p} = T_{\mathrm{FW},p} + \frac{1}{8m^2c^2}((\mathbf{p} - e\mathbf{a}_0)^2 - e\hbar\boldsymbol{\sigma}.\mathbf{B}_0).$$

This formula is the one used in our numerical implementation presented in the next chapter. In this chapter, we will also study the relative order of magnitude of the different terms in  $T'_{\text{FW},p}$ .

# **1.8** Semi-relativistic scattering cross section

The fully relativistic scattering cross section Eq. (1.23) can be written:

$$\frac{d^2\sigma}{d\Omega d\omega_f} = (r_e m c^2)^2 \frac{\omega_f}{\omega_i} \sum_F \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i) \Big| S_{<} + S_{>} \Big|^2,$$

where

$$S_{} = \sum_{L,E_l0} \frac{\langle F|\sum_n e^{-i\mathbf{k}_f \cdot \mathbf{r}_n} \boldsymbol{\alpha}_n \cdot \boldsymbol{\epsilon}_f^{\star}|L\rangle \langle L|\sum_n e^{i\mathbf{k}_i \cdot \mathbf{r}_n} \boldsymbol{\alpha}_n \cdot \boldsymbol{\epsilon}_i \psi |I\rangle}{E_i - E_l + \hbar\omega_i + i\gamma} + \frac{\langle F|\sum_n e^{i\mathbf{k}_i \cdot \mathbf{r}_n} \boldsymbol{\alpha}_n \cdot \boldsymbol{\epsilon}_i \psi |L\rangle \langle L|\sum_n e^{-i\mathbf{k}_f \cdot \mathbf{r}_n} \boldsymbol{\alpha}_n \cdot \boldsymbol{\epsilon}_f^{\star}|I\rangle}{E_i - E_l - \hbar\omega_f}.$$

We have separated the sum over intermediate state of positive energy from the sum over over intermediate state of negative energy and we are going to treat them separately.

## 1.8.1 Many-body semi-relativistic scattering cross section

### Positive energy intermediate states

When  $E_l > 0$ , some of the matrix elements are the same as in the absorption case. The others correspond to emission and, as illustrated by Eq. (1.20), they are formally very close to the absorption ones but with  $\epsilon \to \epsilon'^*$  and  $\mathbf{k} \to -\mathbf{k}'$ . Therefore,

$$S_{>} = \sum_{l,E_{l}>0} \frac{-\Delta E_{li}\Delta E_{fl}}{\hbar^{2}c^{2}} \left( \frac{\langle f|T_{\rm FW}^{\prime fl}(\boldsymbol{\epsilon}_{f})|l\rangle \langle l|T_{\rm FW}^{li}(\boldsymbol{\epsilon}_{i})|i\rangle}{E_{i} - E_{l} + \hbar\omega_{i} + i\gamma} + \frac{\langle f|T_{\rm FW}^{fl}(\boldsymbol{\epsilon}_{i})|l\rangle \langle l|T_{\rm FW}^{\prime li}(\boldsymbol{\epsilon}_{f})|i\rangle}{E_{i} - E_{l} - \hbar\omega_{f}} \right)$$
(1.35)

with

$$T_{\rm FW}^{ji}(\boldsymbol{\epsilon}_i) = \sum_n \boldsymbol{\epsilon}_i \cdot \mathbf{r}_n + \frac{i}{2} \boldsymbol{\epsilon}_i \cdot \mathbf{r}_n \, \mathbf{k}_i \cdot \mathbf{r}_n - \frac{\hbar \beta_n}{2m\Delta E_{ji}} (\boldsymbol{\epsilon}_i \times \mathbf{k}_i) \cdot (\hbar \boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n) - \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}_i \times \boldsymbol{\Sigma}_n)$$

and

$$T_{\rm FW}^{\prime ij}(\boldsymbol{\epsilon}_f) = \sum_n \boldsymbol{\epsilon}_f^{\star} \cdot \mathbf{r}_n - \frac{i}{2} \boldsymbol{\epsilon}_f^{\star} \cdot \mathbf{r}_n \, \mathbf{k}_f \cdot \mathbf{r}_n + \frac{\hbar\beta}{2m\Delta E^{ij}} (\boldsymbol{\epsilon}_f^{\star} \times \mathbf{k}_f) \cdot (\hbar\boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n) - \frac{\hbar}{4m^2c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}_f^{\star} \times \boldsymbol{\Sigma}_n).$$

### Negative energy intermediate states

If  $|L\rangle$  is a negative energy state,  $E_i - E_l \pm \hbar \omega = 2mc^2 + E'_i + E'_l \pm \hbar \omega$  with  $E'_i + E'_l \ll 2mc^2$ . None of the two terms in  $S_{<}$  can be resonant. Their denominators write:

$$\frac{1}{E_i - E_l \pm \hbar\omega} \approx \frac{1}{2mc^2} (1 \mp \frac{\hbar\omega}{mc^2}). \tag{1.36}$$

We can therefore apply the completeness relation on the sum over states of negative energy that involves the projector onto the space of negative energy  $\Gamma_{-}$ :

$$\sum_{L,E_l < 0} |L\rangle \langle L| = \Gamma_{-} = \frac{1-\lambda}{2} = \frac{1-U_0^{\mathsf{T}} \eta U_0}{2},$$

the last equality being a consequence of Eq. (1.28). If P and Q are odd operators,

$$P\Gamma_{-}Q = \frac{1}{2}(PQ - PU_{0}^{\dagger}\eta U_{0}Q) = \frac{1}{2}(PQ + \eta PU_{0}^{2}Q)$$

If one performs the semi-relativistic expansion with the principle of Eq. (1.24):

$$U_0 P \Gamma_- Q U_0^{\dagger} = \frac{1}{2} (U_0 P Q U_0^{\dagger} + \eta U_0^{\dagger} P U_0^2 Q U_0^{\dagger}).$$

This expansion is made easier by the fact that we stop the development at order  $c^{-2}$  and the energy denominator is already of this order. It is therefore enough to consider  $U_0 = 1$  so that

$$\sum_{L,E_l<0} \langle I|P|L\rangle \langle L|Q|F\rangle = \langle i|\frac{1+\eta}{2}PQ|f\rangle = \langle i|PQ|f\rangle$$
(1.37)

because  $|i\rangle$  is even  $(|I\rangle)$  is a positive energy state), that is to say  $\eta|i\rangle = |i\rangle$ .

With  $P = \sum_{n} e^{i\mathbf{k}\cdot\mathbf{r}_{n}} \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\epsilon}$  and  $Q = \sum_{m} e^{-i\mathbf{k}'\cdot\mathbf{r}_{m}} \boldsymbol{\alpha}_{m} \cdot \boldsymbol{\epsilon}'^{\star}$ , the product PQ is:

$$PQ = \sum_{n} \epsilon^{\prime \star} \cdot \epsilon + i \Sigma_{n} \cdot (\epsilon \times \epsilon^{\prime \star}) e^{i(\mathbf{k} - \mathbf{k}^{\prime}) \cdot \mathbf{r}_{n}} + \sum_{n, m \neq n} \alpha_{n} \cdot \epsilon \, \alpha_{m} \cdot \epsilon^{\prime \star} e^{i\mathbf{k} \cdot \mathbf{r}_{n} - i\mathbf{k}^{\prime} \cdot \mathbf{r}_{m}}.$$

So that, using Eq. (1.36) and Eq. (1.37) in both parts of the sum constituting  $S_{<}$ , it rewrites:

$$S_{<} = \frac{1}{mc^{2}} \langle f | \sum_{n} e^{i\mathbf{K}\cdot\mathbf{r}_{n}} \left( \boldsymbol{\epsilon}_{i} \cdot \boldsymbol{\epsilon}_{f}^{\star} + \frac{i\hbar}{2mc^{2}} (\omega_{i} + \omega_{f}) \boldsymbol{\Sigma}_{n} \cdot (\boldsymbol{\epsilon}_{i} \times \boldsymbol{\epsilon}_{f}^{\star}) \right) \\ + \sum_{n, m \neq n} \left( 1 + \frac{\hbar(\omega_{f} - \omega_{i})}{2mc^{2}} \right) \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\epsilon}_{i} \; \boldsymbol{\alpha}_{m} \cdot \boldsymbol{\epsilon}_{f}^{\star} e^{i\mathbf{k}_{i} \cdot \mathbf{r}_{n} - i\mathbf{k}_{f} \cdot \mathbf{r}_{m}} |i\rangle \quad (1.38)$$

where  $\mathbf{K} = \mathbf{k}_i - \mathbf{k}_f$ .

# 1.8.2 Comparison with the results from one-body semi-relativistic Hamiltonians

There are two important differences between the usual scattering cross section [Blume, 1985] and the one obtained here : the spin-position operators in  $S_{>}$  and the two-body term in  $S_{<}$ .

### Comments on the spin-position operator

The additional spin-position operator  $\pi_n \cdot (\boldsymbol{\epsilon}_i \times \boldsymbol{\Sigma}_n)$  in  $S_>$  does not appear in [Blume, 1985, Eq.(13)]. However, if one starts the calculation from the Hamiltonian that Blume proposed [Blume, 1985, Eq.(1)] and if one does the same kind of calculation as in Subsection 1.3.2 but for the scattering case, terms similar to the spin-position terms appear in the matrix elements. Indeed, Blume's Hamiltonian  $H^{\rm B}$  present only two differences with  $H^{\rm FW}$ :

(i) Blume omitted the term  $\nabla \cdot \mathbf{E}$  that does not contribute to the perturbation W in the Coulomb gauge,

(ii) he assumed that **E** commuted with  $\pi_0$  which is generally not true in the time-dependent case. However, in the dipole approximation **E** is uniform so that, within the approximation made in the calculation, it has no impact on the result.

However, as was mentioned in subsection 1.7.2, there is a factor  $\Delta E/\hbar\omega$  between the spinposition term obtained from  $H^{\rm FW}$  or  $H^{\rm B}$  and the one obtained from our semi-relativistic expansion of the fully-relativistic cross section. The spin-position terms obtained from  $H^{\rm B}$  are, therefore, not the correct ones. Anyway, these terms were neglected by Blume because he considered them to be small. Indeed, Blume neglected the terms proportional to  $c^{-2}$  in the second order part of the Kramers-Heisenberg formula.

### Comments on $S_{<}$

The terms on the first line of Eq. (1.38) describe usual Thomson scattering and non-resonant magnetic scattering. In usual calculations, they arise from terms proportional to  $\mathbf{a}^2$  and  $\partial_t \mathbf{a} \times \mathbf{a}$ in the perturbation [Joly et al., 2012, Eq. (3) and (4)]. Here, the physical picture is completely changed because  $H_{e\gamma}$  is linear in  $\mathbf{a}$  and these terms arise from the sum over negative energy intermediate states. This was already noticed by Strange [Strange, 1998, p.525]. However, Strange did not obtain the two-body term that appears in our calculation because the indexes of the  $\boldsymbol{\alpha}$  matrices (that label the electrons) are omitted in the book which prevents from getting it.

This two-body term must be computed in order to evaluate its order of magnitude and if there exists conditions where it could be measured experimentally. In any case, we can already stress its conceptual importance. Such a two-body term could not arise from an  $\mathbf{a}^2$  term in the perturbation because  $\mathbf{a}^2$  becomes a one-body operator. Therefore it can not be present in a calculation starting from a semi-relativistic Hamiltonian. It is a relativistic many-body effect the physical interpretation of which is not straightforward but that could lead to new kinds of phenomena.

# 1.9 Conclusion

After presenting the two problems that arise from semi-relativistic calculations of transition probabilities, we proposed a method which consists in performing the semi-relativistic expansion of the wave-functions in the fully-relativistic matrix elements. We took great care of the formalism that we used in order to insure gauge invariance and, for this, we worked in a quantum electrodynamic (QED) framework. The results that we obtained for the cross sections present several differences with the usual ones. The most noticeable ones are a relativistic term in the transition operator for absorption that we called *spin-position*, that is evaluated in the next chapter, and a two-body term for scattering that needs to be evaluated in the future.

# Bibliography

- [Als-Nielsen and McMorrow, 2000] J. Als-Nielsen and D. McMorrow (2000). Elements of Modern X-Ray Physics. Wiley, New York.
- [Bialynicki-Birula, 1984] I. Bialynicki-Birula (1984). The Hamiltonian of quantum electrodynamics. In A. O. Barut, editor, *Quantum Electrodynamics and Quantum Optics*, volume 110 of NATO ASI Series B, pages 41–61, New York. Plenum Press.
- [Bjorken and Drell, 1965] J. D. Bjorken and S. D. Drell (1965). Relativistic Quantum Fields. McGraw-Hill, New York.
- [Blagojević and Hehl, 2013] M. Blagojević and F. W. Hehl (2013). *Gauge Theories of Gravi*tation. Imperial College Press, London.
- [Blume, 1985] M. Blume (1985). Magnetic scattering of x-rays. J. Appl. Phys., 51:3615-8.
- [Brouder, 1990] C. Brouder (1990). Angular dependence of x-ray absorption spectra. J. Phys.: Condens. Matter, 2:701–38.
- [Ceresoli et al., 2010] D. Ceresoli, U. Gerstmann, A. P. Seitsonen, and F. Mauri (2010). Firstprinciples theory of orbital magnetization. *Phys. Rev. B*, 81:060409(R).
- [Chraplyvy, 1953] Z. V. Chraplyvy (1953). Reduction of relativistic two-body wave equations to approximate forms. I. Phys. Rev., 91:388–91.
- [Cohen-Tannoudji et al., 1973] C. Cohen-Tannoudji, B. Diu, and F. Laloë (1973). *Mécanique Quantique*. Hermann, Paris.
- [Cohen-Tannoudji et al., 1987] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg (1987). *Photons et atomes. Introduction à l'électrodynamique quantique.* InterEditions/Editions du CNRS, Paris.
- [Cohen-Tannoudji et al., 1997] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg (1997). Photons and Atoms. Introduction to Quantum Electrodynamics. Wiley-VCH, New York.
- [Das et al., 2013] A. K. Das, R. R. Francisco, and J. Frenkel (2013). Gauge independence of the fermion pole mass. *Phys. Rev. D*, 88:085012.
- [Epstein, 1979] S. T. Epstein (1979). On the gauge invariance of transition probabilities. Chem. Phys. Lett., 65:417–8.
- [Eriksen, 1958] E. Eriksen (1958). Foldy-Wouthuysen transformation. Exact solution with generalization to the two-particle problem. *Phys. Rev.*, 111:1011–6.

- [Feuchtwang et al., 1986] T. E. Feuchtwang, E. Kazes, and P. H. Cutler (1986). Generalized gauge independence and the physical limitations on the von Neumann measurement postulate. *Found. Phys.*, 16:1263–84.
- [Foldy, 1952] L. L. Foldy (1952). The Electromagnetic Properties of Dirac Particles. Phys. Rev., 87:688-693.
- [Foldy and Wouthuysen, 1950] L. L. Foldy and S. A. Wouthuysen (1950). On the Dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys. Rev.*, 78:29–36.
- [Forney et al., 1977] J. J. Forney, A. Quattropani, and F. Bassani (1977). Choice of gauge in optical transitions. *Nuovo Cimento*, 37B:78-88.
- [Goldman, 1977] T. Goldman (1977). Gauge invariance, time-dependent Foldy-Wouthuysen transformation, and the Pauli Hamiltonian. *Phys. Rev. D*, 15:1063–7.
- [Gougoussis, 2009] C. Gougoussis (2009). Excitations électroniques et magnétisme des matériaux: calcul ab initio de l'absorption X et du dichroïsme circulaire magnétique au seuil K. Université Pierre et Marie Curie-Paris 6. Ph.D. thesis.
- [Greiner, 2000] W. Greiner (2000). *Relativistic Quantum Mechanics*. Springer, Berlin, third edition.
- [Greiner and Reinhardt, 1996] W. Greiner and J. Reinhardt (1996). *Field Quantization*. Springer, Berlin.
- [Grigore, 2001] D. Grigore (2001). Gauge invariance of quantum electrodynamics in the causal approach to renormalization theory. Ann. Phys. (Leipzig), 10:439–71.
- [Haller and Lim-Lombridas, 1994] K. Haller and E. Lim-Lombridas (1994). Quantum gauge equivalence in QED. Found. Phys., 24:217–47.
- [Haller and Sohn, 1979] K. Haller and R. B. Sohn (1979). Interaction of Schrödinger electrons and photons. *Phys. Rev. A*, 20:1541–52.
- [Heitler, 1984] W. Heitler (1984). The Quantum Theory of Radiation. Dover, New York, third edition.
- [Hollands, 2008] S. Hollands (2008). Renormalized quantum Yang-Mills fields in curved spacetime. Rev. Math. Phys., 20:1033–172.
- [Itzykson and Zuber, 1980] C. Itzykson and J.-B. Zuber (1980). Quantum Field Theory. McGraw-Hill, New York.
- [Joly et al., 2012] Y. Joly, S. DiMatteo, and O. Bunau (2012). Resonant x-ray diffraction: Basic theoretical principles. *Eur. Phys. J. Special Topics*, 208:21–38.
- [Kashiwa and Tanimura, 1997] T. Kashiwa and N. Tanimura (1997). Proof for gauge independence of the energy-momentum tensor in quantum electrodynamics. *Fortschr. Phys.*, 45:381–409.
- [Kazes et al., 1982] E. Kazes, T. E. Feuchtwang, P. H. Cutler, and H. Grotch (1982). Gauge invariance and gauge independence of the S-matrix in nonrelativistic quantum mechanics and relativistic quantum field theories. *Annals Phys.*, 142:80–94.

- [Kobe and Yang, 1985] D. H. Kobe and K.-H. Yang (1985). Gauge transformation of the timeevolution operator. *Phys. Rev. A*, 32:952.
- [Lenz et al., 1994] F. Lenz, H. W. L. Naus, K. Ohta, and M. Thies (1994). Quantum mechanics of gauge fixing. Annals Phys., 233:17–50.
- [Manoukian, 1988] E. B. Manoukian (1988). Gauge invariance properties of transition amplitudes in gauge theories. I. Int. J. Theor. Phys., 27:787–800.
- [Matsuda and Kubo, 1980] H. Matsuda and R. Kubo (1980). Gauge independence of transition amplitudes in quantum electrodynamics. *Prog. Theor. Phys.*, 63:275–86.
- [Nieto, 1977] M. M. Nieto (1977). Hamiltonian expectation value for time-dependent Foldy-Wouthuysen transformations: Implications for electrodynamics and resolution of the external-field  $\pi N$  ambiguity. *Phys. Rev. Lett.*, 38:1042–5.
- [Peskin and Schroeder, 1995] M. E. Peskin and D. V. Schroeder (1995). An Introduction to Quantum Field Theory. Addison-Wesley, Reading.
- [Pickard and Mauri, 2001] C. Pickard and F. Mauri (2001). All-electron magnetic response with pseudopotentials: NMR chemical shifts. *Phys. Rev. B*, 63:245101.
- [Silenko, 2016] A. J. Silenko (2016). General properties of the Foldy-Wouthuysen transformation and applicability of the corrected original Foldy-Wouthuysen method. *Phys. Rev. A*, 93:022108.
- [Steinmann, 2000] O. Steinmann (2000). Perturbative Quantum Electrodynamics and Axiomatic Field Theory. Springer, Berlin.
- [Stokes, 2013] A. Stokes (2013). On the gauge of the natural lineshape. J. Phys. B: At. Mol. Opt. Phys., 46:145505.
- [Strange, 1998] P. Strange (1998). *Relativistic Quantum Mechanics*. Cambridge University Press, Cambridge.
- [Takahashi and Hiraoka, 2015] M. Takahashi and N. Hiraoka (2015). Theory of magnetic circular dichroism of non-resonant x-ray Raman scattering. *Phys. Rev. B*, 92:094441.
- [van Lenthe et al., 1996] E. Lenthe, E. J. Baerends, and J. G. Snijders (1996). Construction of the Foldy-Wouthuysen transformation and solution of the Dirac equation using large components only. J. Chem. Phys., 105:2373-7.
- [Voronov et al., 1982] B. L. Voronov, P. M. Lavrov, and I. V. Tyutin (1982). Canonical transformations and the gauge dependence in general gauge theories. Sov. J. Nucl. Phys., 36:292–7.
- [W. E. Lamb, 1952] J. W. E. Lamb (1952). Fine structure of the hydrogen atom. III. Phys. Rev., 85:259–76.
- [Walecka, 2010] J. D. Walecka (2010). Advanced Modern Physics. Theoretical Foundations. World Scientific, Singapore.
- [Weinberg, 1995] S. Weinberg (1995). The Quantum Theory of Fields II. Modern Applications. Cambridge University Press, Cambridge.

- [Yang, 1982] K.-H. Yang (1982). Gauge-invariant time-dependent perturbation theory: I. Nondegenerate case. J. Phys. A: Math. Gen., 15:1201–16.
- [Yang, 1988] K.-H. Yang (1988). Comparison between the gauge invariant and the multipolar gauge formulations of quantum mechanics. Annals Phys., 186:209–29.

# Chapter 2

# Method for *ab initio* calculation of XAS, XMCD and XNCD spectra at K-edge

## Contents

| 2.1 Inti | roduction  | 44        |
|----------|--|-----------|
| 2.2 X-r  | ay Circular Dichroism cross section  | 46        |
| 2.2.1    | Order of magnitude of the terms in the absorption cross-section $% \mathcal{A} = \mathcal{A}$        | 46        |
| 2.2.2    | Terms implemented  | 48        |
| 2.3 Me   | $\operatorname{thod}$  | 49        |
| 2.3.1    | DFT self-consistent field calculation  | 50        |
|          | pwscf code   | 50        |
|          | Spin-orbit coupling within PAW formalism   | 52        |
| 2.3.2    | Cross-section calculation  | 54        |
|          | Principle of the XSpectra code   | 54        |
|          | Details related to the cross terms   | 56        |
|          | (i) SP operator in the case of collinear spins $\ldots \ldots \ldots \ldots \ldots \ldots$           | 56        |
|          | (ii) Calculation of the cross terms $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$ | 57        |
|          | (iii) Alternative way to compute the term D-SP   | 57        |
| 2.3.3    | Discussion on the necessity to include spin orbit coupling in the scf calculation                    | 58        |
| 2.4 Cor  | nvergence with calculation parameters  | <b>59</b> |
| 2.4.1    | Parameters of the scf calculation  | 59        |
|          | k-points sampling  | 60        |
|          | Energy cut-off   | 62        |
|          | Smearing   | 63        |
|          | Convergence threshold  | 65        |
| 2.4.2    | Parameters of XSpectra   | 67        |
|          | k-point grid   | 67        |
|          | Convergence error  | 68        |
| 2.5 Incl | lusion of the core hole  | 69        |

|            | 2.5.1 | Principle | 69 |
|------------|-------|-----------|----|
|            | 2.5.2 | Supercell | 69 |
| <b>2.6</b> | Conc  | elusion   | 71 |

# 2.1 Introduction

X-ray spectroscopy is a tangible example of a field where the developments of theory and experiment are performed in conjunction with a constant mutual enrichment. Several well-established codes [Haverkort et al., 2012, Ebert et al., 2011, Laskowski and Blaha, 2010a, Rehr et al., 2010, Bunău and Joly, 2009, Taillefumier et al., 2002, for selected examples] aim at reproducing X-ray absorption spectra in order to analyze in detail the experimental data. Indeed, in calculations effects can be turned on or off selectively (for example, calculations can be performed with or without the electric quadrupole transition or with or without spin-orbit coupling) and calculations also permit a comparison of the spectra with quantities that are computed otherwise (projected densities of states for example) in order to propose a physical interpretation to the observed features. The experimental developments (such as increase in brilliance and energy resolution or progress in beam stability) that lead to an increasing precision and also to new possibilities in the way that the sample is probed, challenge the theoreticians to improve their tools in order to include finer effects.

X-ray magnetic circular dichroism (XMCD) at K-edge is a good example of an effect that requires the crucial support of theoretical calculations for its analysis. XMCD is a powerful tool for the study of the magnetic structure of complex systems as it provides element- and orbital-specific information. The well-established magneto-optical sum-rules allow to obtain the orbital and spin (up to another term called  $T_z$ ) contribution to the magnetic moment directly from the integral of the spectra [Thole et al., 1992, Carra et al., 1993, Altarelli, 1993], see also [de Groot and Kotani, 2008, Section 7.3]. These sum-rules are widely and successfully applied at spin-orbit split  $L_{2,3}$ -edges of transition metals [Vogel et al., 1997, Stöhr, 1999, Edmonds et al., 2005, Prado et al., 2013] and  $M_{4,5}$ -edges of rare earths [Schillé et al., 1994] and actinides [Wilhelm et al., 2013]. In the absence of spin-orbit splitting of the core state (as for K- and  $L_1$ -edges), only the orbital magnetization sum-rule [Thole et al., 1992, Altarelli, 1993] can apply. This sum-rule relates the integral of the XMCD spectrum at the K-edge to the orbital magnetic moment of occupied p states. The practical application of this sum-rule is however difficult (see section 3.4.1) and it provides only indirect information on the magnetic properties of the sample that are mainly governed by d or f electrons. In this context, calculations of the spectra could be very useful for interpreting experimental data beyond the fingerprint approach which consists in comparing spectra recorded in different samples or under several conditions and deducing similarities or differences in the nature of the sample.

There exist, at the present time, at least two very different codes available to the public that include calculation of K-edge XMCD spectra in the near-edge region.

• The first code is **SPR-KKR** [SPRKKR, site] developed by the group of Hubert Ebert in München. It is based on the fully relativistic Korringa-Kohn-Rostoker (KKR) band structure method. In this method, the electronic structure is represented in terms of the single particle Green's function of the Kohn-Sham equation. This Green's function is expressed within the multiple scattering formalism in terms of scattering path. The expression for the absorption cross section in this formalism is given in [Ebert, 1996, Eq. (77)]. Two approaches are implemented: (i) a real space approach in which the calculations are performed for a finite cluster centred on the absorbing atom and (ii) a reciprocal space approach in which a system with periodic boundary conditions is considered. The published results of calculated XMCD spectra based on this method are obtained without core hole on the absorbing atom (for example, [Šipr and Ebert, 2005]).



Fig. 2.1: Illustration of the muffin-tin approximation in the case of  $SrTiO_3$ . Within this approximation, the material is separated in different regions. In spheres around the atoms, the potential is spherical and in the interstitial region it is constant. Figure is taken from [Schwarz, 2015].

The main limitation of this approach is the use of muffin-tin potentials (see Fig. 2.1). This is a significant approximation on the shape of the potential and its influence on the spectra cannot really be controlled. The fact that it deals with Green's functions makes it, however, a method of choice to study disordered materials [Ebert et al., 2011]. As it is a fully relativistic method, there is no semi-relativistic approximation in the treatment of the Dirac equation contrary to the other codes.

• The second code is **FDMNES** [FDMNES, site] developed by the team of Yves Joly (Institut Néel - Grenoble) and it is a free open source code. The conceptual basis of this code is the following: the real-space basis wave functions  $\phi(\mathbf{r}) = \begin{pmatrix} \phi_{\uparrow}(\mathbf{r}) \\ \phi_{\downarrow}(\mathbf{r}) \end{pmatrix}$  are computed on a 3D grid of points. The unknowns are, therefore, a set of values of the wave function on each point:  $\phi_l = \phi(\mathbf{r}_l)$  and these values are used directly to compute the cross-section. The volume of the calculation, called cluster, is a sphere around the absorbing atom. The shape of the potential is free. To determine  $\phi_l$ , a semi-relativistic Schrödinger-like equation is solved using the finite difference method. This equation is the one derived by Wood and Boring [Wood and Boring, 1978] which is a two-component relativistic equation that improves Pauli equation by adding  $1/c^2$  terms:

$$\left[-\nabla^2 + V(\mathbf{r}) - E - \frac{\alpha_0^2}{4}(V(\mathbf{r}) - E)^2 - \frac{\alpha_0^2}{4}B_{wb}\Big(\nabla V(\mathbf{r}) \cdot \nabla + i\boldsymbol{\sigma}(\nabla V \times \nabla)\Big)\right]\boldsymbol{\phi}(\mathbf{r}) = 0$$

with  $B_{wb} = \left(1 - \frac{\alpha_0^2}{4}(V(\mathbf{r}) - E)\right)^{-1}$ . This is an equation for the large components of Dirac wave-functions.

The shape of the potential  $V(\mathbf{r}) = \begin{pmatrix} V_{\uparrow}(\mathbf{r}) \\ V_{\downarrow}(\mathbf{r}) \end{pmatrix}$  that enters this equation can be determined *a priori* (in a first step) self-consistently with a method based on multiple scattering

[Bunău and Joly, 2009]. The main difficulty in the practical use of this method is that the calculated spectra depend on the size of the clusters used to perform the self-consistent calculation (first step) and the finite difference method calculation (second step).

Here, we present a new method to compute XMCD in which the calculations are performed in reciprocal space using plane waves and pseudopotentials. The method has been implemented within a highly efficient solid-state code so that it can be applied to a large range of systems.

# 2.2 X-ray Circular Dichroism cross section

X-ray circular dichroism describes the dependence of the absorption cross section on the state of circularly-polarized light (left/right polarization). In the case of XMCD at the K-edge of 3d transition elements, the asymmetry in absorption is, at most, of the order  $10^{-3}$  compared to the edge jump. To compute XMCD spectra, it is therefore mandatory to compute the absorption cross sections for right and left circular polarization very accurately. This is the reason why we start from the semi-relativistic absorption cross section obtained in the previous Chapter in Section 1.7.

Our numerical calculations are based on a monoelectronic framework with 2-components wave functions. More precisely, for the core state (initial state) we compute the large components of Dirac wave function whereas for the unoccupied states (final states) we compute the Foldy-Wouthuysen wave functions. In other words, we are in the situation described in Subsection 1.7.3.

### 2.2.1 Order of magnitude of the terms in the absorption cross-section

The formula obtained in Subsection 1.7.3 Eq. (1.34) for the absorption cross section is:

$$\sigma(\hbar\omega) = 4\pi^2 \alpha_0 \hbar\omega \sum_f |\langle \phi_f^{\rm FW} | T_{\rm D} + T_{\rm Q} + T_{\rm MD} + T_{A_0} + T_{\rm SP} + T^e |\phi_i\rangle|^2 \delta(E_f - E_i - \hbar\omega).$$
(2.1)

In this expression we simply named the operators constituting  $T'_{FW,p}$ . These operators are detailed hereafter.

As the core wave function is very localized, we can have an idea of the relative order of magnitude of the operators in this cross-section by evaluating them at the radius corresponding to the core state. We give these evaluations in Table 2.1 by taking as a reference the order of magnitude of the dominant operator that is the *electric dipole* operator (usually called E1):

$$T_{\rm D} = \boldsymbol{\epsilon} \cdot \mathbf{r}$$

The second column of the table gives the formula used to determine these orders of magnitude.

The *electric quadrupole* operator (usually called E2) writes:

$$T_{\rm Q} = \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \, \mathbf{k} \cdot \mathbf{r}.$$

Its order of magnitude compared to  $T_{\rm D}$  is given by  $kr_c/2$  where k is the norm of the incident wave vector (the incident electromagnetic wave is described by a plane wave  $e^{i\mathbf{k}\cdot\mathbf{r}}$ ) and  $r_c$  the mean radius of the core orbital. It is related to the spatial inhomogeneity of the x-ray electric field at the scale of the core state radius. Working in the electric-dipole approximation

**Table 2.1:** Orders of magnitude of the operators in Eq. (2.1) evaluated at the core state radius  $r_c$  with respect to the electric dipole operator. The mean radius of core orbitals is deduced from the effective nuclear charge:  $r_c = \frac{3}{2} \frac{a_0}{Z_{eff}}$  [Clementi and Raimondi, 1963, Clementi et al., 1967]. In the table,  $B_0$  has been fixed to  $2 \times 10^4$  T (1.2eV) which is two orders of magnitude larger than the exchange splitting calculated for Fe K-edge. The Coulomb potential is  $V = \frac{-Z_{eff}e}{4\pi\epsilon_0 r_c}$  and the core state energy  $E_i$  is evaluated in a planetary model  $E_i = \frac{-Z_{eff}e^2}{8\pi\epsilon_0 r_c}$ .

| <br>Fdgo               |                             | $L_1$                |                      | L_2                  |                      | K                      |                      |
|------------------------|-----------------------------|----------------------|----------------------|----------------------|----------------------|------------------------|----------------------|
| Edge                   |                             | Ι                    | Fe                   | Gd                   | Bi                   | 0                      | Fe                   |
| Energy                 | $(\mathrm{keV})$            | 5.19                 | 0.72                 | 7.898                | 15.71                | 0.53                   | 7.11                 |
| $Z_{eff}$              |                             | 39.067               | 22.089               | 29.8527              | 39.2335              | 7.6579                 | 25.381               |
| $r_c$                  | $(a_0)$                     | 0.038                | 0.068                | 0.050                | 0.038                | 0.20                   | 0.059                |
| $T_{\rm Q}/T_{\rm D}$  | $\frac{kr_c}{2}$            | $2.7{\times}10^{-2}$ | $6.6 	imes 10^{-3}$  | $5.3 \times 10^{-2}$ | $8.1 \times 10^{-2}$ | $1.4 \times 10^{-2}$   | $5.7{\times}10^{-2}$ |
| $T_{A_0}/T_{\rm D}$    | $\frac{ekr_cB_0}{4m\omega}$ | $6.0 \times 10^{-6}$ | $1.1 \times 10^{-5}$ | $7.8 \times 10^{-6}$ | $6.0 \times 10^{-6}$ | $3.0 \times 10^{-5}$   | $9.2 \times 10^{-6}$ |
| $T_{\rm SP}/T_{\rm D}$ | $\frac{\hbar\omega}{4mc^2}$ | $2.6 \times 10^{-3}$ | $3.5 \times 10^{-4}$ | $3.9 \times 10^{-3}$ | $7.7 \times 10^{-3}$ | $2.6\!\times\!10^{-4}$ | $3.5 \times 10^{-3}$ |
| $T^e/T_{\rm D}$        | $\frac{E_i - eV}{4mc^2}$    | $6.7 \times 10^{-3}$ | $2.2 \times 10^{-3}$ | $3.9 \times 10^{-3}$ | $6.8 \times 10^{-3}$ | $2.6{\times}10^{-4}$   | $2.8 \times 10^{-3}$ |

corresponds to considering that the x-ray electric field is spatially constant at the scale of the wave function of the core electron. In the quadrupole approximation  $(e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1 + i\mathbf{k}\cdot\mathbf{r})$ , one partially takes into account the spatial variation of the x-ray electric field. Given the order of magnitude of  $kr_c$  (see Table 2.1), it is a good approximation in XAS.

The magnetic dipole operator  $T_{\rm MD}$  (usually called M1) is:

$$T_{\rm MD} = \frac{1}{2m\omega} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\sigma} + \mathbf{L}).$$

It is proportional to the total magnetic moment operator  $(\hbar \boldsymbol{\sigma} + \mathbf{L}) = (2\mathbf{S} + \mathbf{L})$  where  $\mathbf{S}$  is the spin operator.  $T_{\rm MD}$  is also present in common non-relativistic derivations [Brouder, 1990, Di Matteo et al., 2005]. This operator does not incorporate the radial variable r and it conserves the orbital angular momentum. Its selection rules are therefore [Brouder, 1990]:<sup>1</sup>  $l_i = l_f$ and  $n_i = n_f$ . It vanishes in the X-ray energy range because the states involved in the transitions have different principal quantum numbers. Note that the argument to justify the selection rules of  $T_{\rm MD}$  is not rigorous. In a monoelectronic framework where we consider that  $|f\rangle$  is a stationary state in the presence of the core-hole (so it is not exactly an eigenstate of the same Hamiltonian as  $|i\rangle$ ), the proof does not apply. In a many-body framework, the matrix elements of the magnetic dipole operator write as sums of one-body matrix elements to which the selection rules apply. However, it is not sure that a transition that obeys  $l_i = l_f$  and  $n_i = n_f$  can not occur. Finally, the picture is further complicated in a relativistic framework in which l is not a good quantum number.<sup>2</sup> In fact, the electric dipole-magnetic dipole contribution is very small

<sup>&</sup>lt;sup>1</sup>The initial state writes as the product of a radial function by a sum of the product of spherical harmonics and of spinors:  $|i\rangle = \mathcal{R}_{n_i l_i}(r) \sum_{m,s} c_{m,s} |l_i m\rangle |s\rangle$ . As the MD operator does not act on the radial part and does not change  $l_i$ ,  $T_{\rm MD}|i\rangle = \mathcal{R}_{n_i l_i}(r) \sum_{m',s'} c_{m',s'} |l_i m'\rangle |s'\rangle$ . The final state writes as a linear combination of  $|\phi_f\rangle = \mathcal{R}_{n_f l_f}(r) \sum_{m,s} c_{m,s} |l_f m\rangle |s\rangle$ ,  $T_{\rm MD}|i\rangle$  and  $|\phi_f\rangle$  are orthogonal except if  $l_i = l_f$  (otherwise the angular parts are orthogonal) and  $n_i = n_f$  (otherwise the radial parts are orthogonal).

 $<sup>^{2}</sup>$ For example, in a relativistic framework there is some p character in the core wave function (see

but non zero. Such a contribution to the XNCD has, for example, been measured at the Ni K-edge in  $\alpha$ -NiSO<sub>4</sub> × 6H<sub>2</sub>O [Rogalev et al., 2010] and it reaches  $3 \times 10^{-5}$  of the absorption edge jump. As it is small, we neglect this term in all this thesis. It is absent in Table 2.1 because no evaluation *a priori* of the operator can be given.

The correction to magnetic dipole term due to the external static potential  $A_0$  is:

$$T_{A_0} = \frac{e}{2m\omega} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \mathbf{A}_0)$$

This operator is evaluated using  $\mathbf{A}_0 = \frac{1}{2}\mathbf{r} \times \mathbf{B}_0$  and considering a magnetic field  $B_0 = 2 \times 10^4 \text{ T}$  so that  $\mu_B B_0 = 1.16 \text{ eV}$ . The exchange splitting calculated at the 1s state of Fe is 0.015 eV [Y. Joly private communication, 2016]. The chosen value is, therefore, larger than the exchange magnetic field in matter and it is also way larger than accessible values in laboratory for a magnetic field (which reaches several 100 T for pulsed magnetic field).

The *spin-position* transition operator writes:

$$T_{\rm SP} = rac{i\hbar\omega}{4mc^2} \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} imes \mathbf{r}).$$

It is proportional to the ratio between the energy of the incident photons and the rest energy of the electron:  $\frac{\hbar\omega}{mc^2}$ . As  $mc^2 = 511$  keV,  $\frac{\hbar\omega}{4mc^2}$  can reach  $10^{-3}$  only if the incident energy is larger than few keV.

Finally, the extra operator that is related to the use of large components of the Dirac operator for the core state writes:

$$T^{e} = \frac{1}{4mc^{2}} \left( \boldsymbol{\epsilon} \cdot \mathbf{r} \ H_{0}^{0} - e\boldsymbol{\epsilon} \cdot \mathbf{r} \ V(\mathbf{r}) \right)$$

where  $H_0^0 = \frac{(\mathbf{p}-e\mathbf{A}_0)^2}{2m} + eV(\mathbf{r}) - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}_0$ . At zero order in  $c^{-2}$ , the states  $|\phi_f^{\text{FW}}\rangle$  and  $|\phi_i\rangle$  are both eigenstates of  $H_0^0$ . Moreover the potential felt by the core electron is spherical (the core states are almost unaffected by the chemical environment). The matrix element of  $T^e$  therefore rewrites:

$$\langle \phi_f^{\rm FW} | T^e | \phi_i \rangle = \frac{1}{4mc^2} \langle \phi_f^{\rm FW} | \boldsymbol{\epsilon} \cdot \mathbf{r} (E_i - V(r)) | \phi_i \rangle.$$

This operator concerns transitions to the same orbitals as the electric dipole operator, it has the same angular dependency and it does not incorporate a spin variable. For this reason, even in XMCD, it will only yield a very small correction (see the order of magnitude in Table 2.1) to the electric dipole term.

When expanding the square modulus of the matrix elements in Eq. (2.1), we keep the terms with contributions higher than  $10^{-3}$  compared to the dominant electric dipole term (except  $T^e$  for the reason mentioned above). Altogether, four terms remain that are listed in the next subsection (D-D, Q-Q, D-Q and D-SP).

### 2.2.2 Terms implemented

The electric dipole-electric dipole (D-D) term is:

$$\sigma_{\rm D-D}(\hbar\omega) = 4\pi^2 \alpha_0 \hbar\omega \sum_f |\langle f | \boldsymbol{\epsilon} \cdot \mathbf{r} | i \rangle|^2 \delta(E_f - E_i - \hbar\omega).$$
(2.2)

[Thaller, 1992, p.126]).

It is usually the only term that is taken into account in the calculation of XAS and XMCD spectra at the  $L_{2,3}$ -edges and it is also sometimes the only term that is taken into account in the calculations at the K-edge [Natoli et al., 1980, Brouder et al., 1996, Rehr and Albers, 2000, Fujikawa and Nagamatsu, 2003].

The electric quadrupole-electric quadrupole (Q-Q) term is:

$$\sigma_{\rm Q-Q}(\hbar\omega) = \pi^2 \alpha_0 \hbar\omega \sum_f |\langle f|(\mathbf{k} \cdot \mathbf{r})(\boldsymbol{\epsilon} \cdot \mathbf{r})|i\rangle|^2 \delta(E_f - E_i - \hbar\omega).$$
(2.3)

At the K-edge, it can reach a few percent of  $\sigma_{D-D}$  in amplitude (square of the value for  $T_Q$  in Table 2.1). Its contribution is mainly significant in the pre-edge region of the absorption spectra where the D-D contribution is small or nonexistent. It is sometimes included in X-ray absorption calculations [Taillefumier et al., 2002, Bunău and Joly, 2009].

If the orbital parts of the wave functions can be chosen real (no spin-orbit coupling or applied magnetic field)<sup>3</sup> in the terms D-D and Q-Q,  $\sigma(\epsilon) = \sigma(\epsilon^*)$ , which leads to a zero contribution to circular dichroism. For this reason it is crucial to take into account the spin orbit coupling effects in the wave functions calculation in order to compute XMCD.

On the other hand, the two following terms can give a non vanishing contribution to the circular dichroism cross section even when the orbital parts of the wave functions are real.

The electric dipole-electric quadrupole cross term (D-Q) is:

$$\sigma_{\rm D-Q}(\hbar\omega) = -4\pi^2 \alpha_0 \hbar\omega \sum_f \operatorname{Im}[\langle f | (\mathbf{k} \cdot \mathbf{r}) (\boldsymbol{\epsilon} \cdot \mathbf{r}) | i \rangle \langle i | \boldsymbol{\epsilon}^* \cdot \mathbf{r} | f \rangle] \delta(E_f - E_i - \hbar\omega).$$
(2.4)

If  $|i\rangle$  and  $|f\rangle$  are parity invariant (*i.e.* if inversion  $\mathbf{r} \to -\mathbf{r}$  is a symmetry of the system) then  $\sigma_{D-Q} = 0$ . It is however this term that is responsible for XNCD [Natoli et al., 1998] because the electric dipole-magnetic dipole term (that is responsible for optical activity in the optical range) is very small in the X-ray range.

The electric dipole-spin-position (D-SP) cross term is:

$$\sigma_{\rm D-SP}(\hbar\omega) = -\frac{2\pi^2 \alpha_0 \hbar^2 \omega^2}{mc^2} \sum_f {\rm Im}[\langle f | \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}) | i \rangle \langle i | \boldsymbol{\epsilon}^* \cdot \mathbf{r} | f \rangle] \delta(E_f - E_i - \hbar\omega).$$
(2.5)

It can exist only in materials that exhibit a spin polarization. Like the spin-orbit coupling term in the FW Hamiltonian, it arises from the coupling of the small components of the Dirac wave functions. To our knowledge, it has never been evaluated before our work. We will show in section 3.2 of the next chapter that, despite the small prefactor of this term, its contribution to XMCD at the K-edge of 3d metals can account for up to one third of the XMCD intensity near the edge.

# 2.3 Method

In order to compute the contribution of each of the previously listed terms to the total cross section, we have modified the code XSpectra [Gougoussis et al., 2009] of the suite QUANTUM ESPRESSO [Giannozzi et al., 2009].

<sup>&</sup>lt;sup>3</sup>The orbital parts of the wave-functions can be chosen real if the states  $\pm m_l$  are equivalent. We will see in Section 3.2.1 of the next chapter that the separation of the state  $\pm m_l$  occurs in the presence of spin-orbit coupling and magnetic ordering but also if an external magnetic field is applied (Zeeman effect). This second phenomenon is however too weak to contribute significantly to a XMCD of the order  $10^{-3}$  the edge jump.

The idea of the method is to proceed in two steps:

(1) Determine the electron density of the material using an approximation of Density Functional Theory (DFT) with a self-consistent field (scf) method. For this, we use a semi-relativistic version of the code pwscf.

(2) Compute all the contributions to the absorption cross section for given  $\boldsymbol{\epsilon}$  and  $\mathbf{k}$ .

## 2.3.1 DFT self-consistent field calculation

### pwscf code

The code pwscf [Giannozzi et al., 2009] is based on density-functional theory, plane waves, and pseudopotentials and allows performing self-consistent field calculations. In this thesis, we give only an overview of the principle of the method which is otherwise developed in several textbooks. For example in the book by R.M. Martin [Martin, 2004], chapters 6 and 7 present the foundations of DFT and of the Kohn-Sham ansatz. In chapter 9, a description of the self-consistent solving of the Kohn-Sham equation can be found and chapter 11 is dedicated to the topic of pseudopotentials.

The principle behind density functional theory is that the energy of a system of electrons (many-body system) can be viewed as a functional of the ground state density  $n_0(\mathbf{r})$ [Hohenberg and Kohn, 1964]. In principle, a functional E[n] for the total energy could be defined and  $n_0(\mathbf{r})$  could be found by minimizing this total energy. Nevertheless, in practice the direct minimization of E[n] is impracticable because the functional E[n] is not known explicitly. The success of DFT is due to the Kohn-Sham (KS) approach [Kohn and Sham, 1965] that provides a powerful method for electronic structure calculation. In this approach, the many-body system of interacting electrons is replaced by an auxiliary system of independent electrons. The ground state energy functional writes in the form [Martin, 2004, p.137]:

$$E_{KS} = T_s[n] + \int d\mathbf{r} V_{ext}(\mathbf{r}) n(\mathbf{r}) + E_{Hartree}[n] + E_{II} + E_{xc}[n]$$

where  $T_s$  is the independent-particle kinetic energy,  $V_{ext}$  is the external potential due to the nuclei and external fields,  $E_{Hartree}$  is the classical Coulomb interaction of the electron density  $n(\mathbf{r})$  with itself,  $E_{II}$  is the interaction between the nuclei and  $E_{xc}$  is called the *exchange-correlation* energy.  $E_{xc}$  is the difference of the sum of the kinetic energy plus the internal energy of the interacting system and  $T_s[n] + E_{Hartree}[n]$  (sum of the kinetic energy and of the Hartree energy for the fictitious non-interacting system). In other words, all many body effects are grouped into  $E_{xc}[n]$ . If  $E_{xc}[n]$  was known exactly, the exact ground state energy and density of the interacting system could be obtained from the study of the Kohn-Sham auxiliary system. In practice it must be approximated and this is what is called *choice of the functional*. Then, the energy can be determined numerically by iterations with the scf method [Martin, 2004, p.173].

In the Kohn-Sham approach, the kinetic energy of the auxiliary system is treated in terms of orbitals:

$$T_s = -\frac{\hbar^2}{2} \sum_{\sigma} \sum_{i=1}^{N} \langle \psi_i^{\sigma} | \nabla^2 | \psi_i^{\sigma} \rangle.$$

where  $\psi_i^{\sigma}$  are the eigenvectors with lowest eigenvalues of the Hamiltonian of the system of independent particles. The density of the auxiliary system in the ground state is given by:

$$n(\mathbf{r}) = \sum_{\sigma} \sum_{i=1}^{N} |\psi_i^{\sigma}(\mathbf{r})|^2.$$

In order to solve the KS equation, the wavefunctions must be expanded in a basis set. *Plane waves* are particularly appropriate for periodic crystals that can be represented by a cell with periodic boundary conditions.



**Fig. 2.2:** Example of norm-conserving pseudopotentials and pseudofunctions for the element Mo. Figure is taken from [Martin, 2004].

Top: All-electron valence radial functions (dashed lines) and norm-conserving pseudofunctions (solid lines). These wavefunctions coincide for radius larger than the cut-off radius:  $r > R_c$ .

Bottom: Pseudopotential  $V_l^{ps}(r)$  in Rydbergs for angular momentum l = 0, 1, 2 compared to  $\frac{-Z_{ion}}{r}$ (dashed). Note that the pseudopotential is a nonlocal operator that can be written in a semilocal form (non local in angular variables and local in radial variable):  $V^{ps} = \sum_{lm} V_l^{ps}(r)\delta(r - r')Y_{lm}(\mathbf{r})Y_{lm}(\mathbf{r}')$ . The components  $V_l^{ps}(r)$  represent how it acts on *l*-symmetry valence electrons. Each potential  $V_l^{ps}(r)$  behaves as  $\frac{-Z_{ion}}{r}$  for  $r \to \infty$ . Traditionally pseudopotentials are split into a *l*independent (local) long range part behaving as  $\frac{-Z_{ion}}{r}$  for  $r \to \infty$  and a short-ranged semilocal part (see [Martin, 2004, p.235]).

In *pseudopotential*-based methods the strong Coulomb potential near the nuclei and the core-electrons is replaced by a fictitious smooth potential (see Figure 2.2). The underlying idea is that the core electrons do not participate in bonding. Therefore, an atomic calculation is sufficient to determine their wavefunctions and, to study the materials, they can be replaced by an effective potential. The valence electrons wave functions are replaced by pseudo-wavefunctions that are exempt from the rapid oscillations near the core which are otherwise required to provide orthogonality. The size of the plane-waves basis set needed to describe the system is therefore considerably lowered which leads to a better computational efficiency compared to an all-electron approach.

### Spin-orbit coupling within PAW formalism

Inclusion of the spin orbit coupling (SOC) in the wave functions calculation plays a crucial role for the evaluation of XMCD spectra. For this, we use an approach [Gerstmann et al., 2014] where the spin orbit interaction term of  $H^{\text{FW}}$  (Eq. (1.7)) is implemented through Projector Augmented Wave (PAW) formalism.

In the PAW formalism, as described by Blöchl [Blöchl, 1994], the physical valence wave functions  $|\Psi\rangle$  can be reconstructed from the pseudo-wave-functions  $|\tilde{\Psi}\rangle$  as they are related through a linear operator  $\mathcal{T}$ :

$$|\Psi\rangle = \mathcal{T}|\Psi\rangle \tag{2.6}$$

This linear transformation is chosen to differ from identity only by a sum of local, atom-centered operators  $\hat{\mathcal{T}}_{\mathbf{R}}$  (**R** : atomic positions):

$$\mathcal{T} = 1 + \sum_{\mathbf{R}} \widehat{\mathcal{T}}_{\mathbf{R}}.$$

Each local operator  $\widehat{\mathcal{T}}_{\mathbf{R}}$  acts only in an augmentation region surrounding the atom. Outside the augmentation regions, all-electron and pseudo-wave-functions coincide.

Around each atom located at  $\mathbf{R}$ , a partial wave basis  $\{|\phi_{\mathbf{R},n}\rangle\}$  is built<sup>4</sup> (in our case,  $|\phi_{\mathbf{R},n}\rangle$  are solutions of the Dirac equation for the isolated atom within a scalar relativistic approximation [Koelling and Harmon, 1977]). For each partial wave, a pseudo partial wave  $|\tilde{\phi}_{\mathbf{R},n}\rangle$ , that is identical to  $|\phi_{\mathbf{R},n}\rangle$  outside the augmentation region, is chosen. These pseudo-partial-waves form a complete set of functions within the augmentation region. Then, the local reconstruction operators  $\widehat{\mathcal{T}}_{\mathbf{R}}$  are defined by the fact that  $|\phi_{\mathbf{R},n}\rangle$  is the target function of  $|\tilde{\phi}_{\mathbf{R},n}\rangle$  *i.e.*  $|\phi_{\mathbf{R},n}\rangle =$  $(1 + \widehat{\mathcal{T}}_{\mathbf{R}})|\tilde{\phi}_{\mathbf{R},n}\rangle$  The transformation operator consequently writes:<sup>5</sup>

$$\mathcal{T} = 1 + \sum_{\mathbf{R},n} (|\phi_{\mathbf{R},n}\rangle - |\tilde{\phi}_{\mathbf{R},n}\rangle) \langle \tilde{p}_{\mathbf{R},n}|.$$
(2.7)

where the  $\langle \tilde{p}_{\mathbf{R},n} |$  form a complete set of projector functions.

The pseudo-Hamiltonian is given by [Ceresoli et al., 2010]:

$$\tilde{\mathcal{H}}_0 = \frac{\mathbf{p}^2}{2m} + eV_{\rm ps}^{\rm loc}(\mathbf{r}) + \sum_{\mathbf{R}} eV_{\mathbf{R}}^{\rm nl} + \tilde{\mathcal{H}}_{\rm SO}$$

 $^{5}$  Within the augmentation region, one can expand any pseudo wave-function into pseudo partial waves :

$$|\tilde{\Psi}\rangle = \sum_{n} c_{n} |\tilde{\phi}_{\mathbf{R},n}\rangle$$

Since  $|\tilde{\phi}_{\mathbf{R},n}\rangle = \mathcal{T}|\phi_{\mathbf{R},n}\rangle$ ,  $|\Psi\rangle = \mathcal{T}|\tilde{\Psi}\rangle = \sum_{n} c_{n}|\phi_{\mathbf{R},n}\rangle = \sum_{n} c_{n}|\phi_{\mathbf{R},n}\rangle + |\tilde{\Psi}\rangle - \sum_{n} c_{n}|\tilde{\phi}_{\mathbf{R},n}\rangle$ . For the transformation  $\mathcal{T}$  to be linear, the coefficients  $c_{n}$  must be linear functionals of the pseudo-wave-function. We call projector functions the functions  $\langle \tilde{p}_{\mathbf{R},n}|$  such that  $c_{n} = \langle \tilde{p}_{\mathbf{R},n}|\tilde{\Psi}\rangle$ .

For the relationship  $|\tilde{\Psi}\rangle = \sum_i \langle \tilde{p}_i^{\mathbf{R}} | \tilde{\Psi} \rangle | \tilde{\phi}_i^{\mathbf{R}} \rangle$  to be true within the augmentation region, the projector functions must fulfill the condition  $\mathbb{1} = \sum_i | \tilde{\phi}_i^{\mathbf{R}} \rangle \langle \tilde{p}_i^{\mathbf{R}} |$  within the augmentation region. This implies that:

$$\langle \tilde{p}_i^{\mathbf{R}} | \tilde{\phi}_j^{\mathbf{R}'} \rangle = \delta_{i,j} \delta_{\mathbf{RR}'}.$$

In practice, the projectors are determined iteratively from an initial guess to fulfill this condition (See [Blöchl, 1994, Sec. VI C.]). As there is no restriction outside the augmentation spheres, the projector functions are chosen to be zero.

<sup>&</sup>lt;sup>4</sup> The index *n* refers to the angular momentum quantum number (l,m) and to an additional number,  $\varsigma$ , used to label different partial waves for the same site and angular momentum (if there is more than one projector per channel).



Fig. 2.3: Schematic drawing of the augmentation region defined as sphere around the atoms in the PAW framework. Outside the augmentation regions, pseudo wave functions coincide with allelectron wave functions.

where  $V_{\rm ps}^{\rm loc}$  and  $V_{\rm R}^{\rm nl}$  are the local part and the nonlocal part in separable form of the pseudopotential.  $\tilde{\mathcal{H}}_{\rm SO}$  is the pseudo-Hamiltonian corresponding to the time independent spin-orbit term in the Foldy-Wouthuysen transformed Hamiltonian:

$$\tilde{\mathcal{H}}_{\rm SO} = \mathcal{T}^{\dagger} \left( \frac{e\hbar}{4m^2c^2} \boldsymbol{\sigma} \cdot (\nabla V(\mathbf{r}) \times \mathbf{p}) \right) \mathcal{T} = \frac{e\hbar}{4m^2c^2} \left( \boldsymbol{\sigma} \cdot (\nabla V_{\rm ps}^{\rm loc}(\mathbf{r}) \times \mathbf{p}) + \sum_{\mathbf{R}} F_{\mathbf{R}}^{\rm nl} \right).$$
(2.8)

The expression for  $F_{\mathbf{R}}^{\text{nl}}$  is given in Ref. [Pickard and Mauri, 2002, Eq. (11)]:

$$F_{\mathbf{R}}^{\mathrm{nl}} = \sum_{n,m} |\tilde{p}_{n}^{\mathbf{R}}\rangle \boldsymbol{\sigma} \cdot (\langle \phi_{\mathbf{R},n} | \nabla v_{\mathrm{ae}}(\mathbf{r}) \times \mathbf{p} | \phi_{\mathbf{R},n} \rangle - \langle \tilde{\phi}_{\mathbf{R},n} | \nabla v_{\mathrm{ps}}^{\mathrm{loc}}(\mathbf{r}) \times \mathbf{p} | \tilde{\phi}_{\mathbf{R},n} \rangle) \langle \tilde{p}_{m}^{\mathbf{R}} | \boldsymbol{\phi}_{\mathbf{R},n} \rangle$$

where  $v_{ae}$  and  $v_{ps}^{loc}$  are the atomic all-electron and local pseudopotentials respectively. As these potentials are spherical,  $F_{\mathbf{R}}^{nl}$  rewrites:

$$F_{\mathbf{R}}^{\mathrm{nl}} = \sum_{n,m} |\tilde{p}_{n}^{\mathbf{R}}\rangle \boldsymbol{\sigma} \cdot (\langle \phi_{\mathbf{R},n} | \frac{1}{r} \frac{\partial v_{\mathrm{ae}}}{\partial r} \mathbf{L} | \phi_{\mathbf{R},n} \rangle - \langle \tilde{\phi}_{\mathbf{R},n} | \frac{1}{r} \frac{\partial v_{\mathrm{ps}}^{\mathrm{loc}}}{\partial r} \mathbf{L} | \tilde{\phi}_{\mathbf{R},n} \rangle) \langle \tilde{p}_{m}^{\mathbf{R}} | \mathbf{L} | \tilde{\phi}_{\mathbf{R},n} \rangle = \langle \tilde{\phi}_{\mathbf{R},n} | \tilde{\sigma}_{\mathbf{R},n} \rangle \langle \tilde{p}_{m}^{\mathbf{R}} | \mathbf{L} | \tilde{\phi}_{\mathbf{R},n} \rangle \langle \tilde{p}_{m}^{\mathbf{R}} | \mathbf{L} | \tilde{\phi}_{\mathbf{R},n} \rangle = \langle \tilde{\phi}_{\mathbf{R},n} | \tilde{\sigma}_{\mathbf{R},n} \rangle \langle \tilde{p}_{m}^{\mathbf{R}} | \mathbf{L} | \tilde{\phi}_{\mathbf{R},n} \rangle \langle \tilde{p}_{m}^{$$

In Eq. (2.8), the local potential is the sum of the atomic local potentials:

$$V_{\mathrm{ps}}^{\mathrm{loc}}(\mathbf{r}) = \sum_{\mathbf{R}} v_{\mathrm{ps}}^{\mathrm{loc}}(\mathbf{r}).$$

As  $\frac{1}{r} \frac{\partial v_{\rm ps}^{\rm loc}}{\partial r}$  decreases as  $1/r^3$ , the action of the operator  $\nabla V_{\rm ps}^{\rm loc}(\mathbf{r}) \times \mathbf{p}$  in the augmentation region is, at first order, the same as the action of  $\nabla v_{\rm ps}^{\rm loc}(\mathbf{r}) \times \mathbf{p}$  (in other words, in a given augmentation region, the contributions of the neighboring atoms to the sum in  $V_{\rm ps}^{\rm loc}(\mathbf{r})$  is small). In the augmentation region, the pseudo-wave-function can be expanded according to:

$$|\tilde{\Psi}\rangle = \sum_{n} |\tilde{\phi}_{n,\mathbf{R}}\rangle \langle \tilde{p}_{n}^{\mathbf{R}} |\tilde{\Psi}\rangle$$

Therefore, the term proportional to  $v_{\rm ps}^{\rm loc}$  and the term proportional to  $V_{\rm ps}^{\rm loc}$  in Eq. (2.8) partially compensate each other so that the dominant contribution to  $\tilde{\mathcal{H}}_{\rm SO}$  arises from the following term (Term 1 on Fig. (2.4)):

$$\frac{e\hbar}{4m^2c^2}\sum_{n\mathbf{R}m}\boldsymbol{\sigma}\cdot|\tilde{p}_n^{\mathbf{R}}\rangle\langle\phi_{\mathbf{R},n}|\frac{1}{r}\frac{\partial v_{\mathrm{ae}}}{\partial r}\mathbf{L}|\phi_{\mathbf{R},m}\rangle\langle\tilde{p}_m^{\mathbf{R}}|.$$



Fig. 2.4: Computed XMCD spectrum (test calculation for a *bcc*-Fe structure) by considering either the total  $\tilde{\mathcal{H}}_{SO}$  or each term constituting it separatly.

In Rydberg atomic units, with e included in the potentials (*i.e.* the way it is in the code) the terms are:

1) 
$$\frac{\alpha_{0}^{2}}{4} \sum_{n\mathbf{R}m} \boldsymbol{\sigma} \cdot |\tilde{p}_{n}^{\mathbf{R}}\rangle \langle \phi_{\mathbf{R},n}| \frac{1}{r} \frac{\partial v_{ae}}{\partial r} \mathbf{L} |\phi_{\mathbf{R},m}\rangle \langle \tilde{p}_{m}^{\mathbf{R}}|$$
2) 
$$\frac{\alpha_{0}^{2}}{4} \sum_{n\mathbf{R}m} \boldsymbol{\sigma} \cdot |\tilde{p}_{n}^{\mathbf{R}}\rangle \langle \phi_{\mathbf{R},n}| \frac{1}{r} \frac{\partial v_{ps}^{\text{loc}}}{\partial r} \mathbf{L} |\phi_{\mathbf{R},m}\rangle \langle \tilde{p}_{m}^{\mathbf{R}}|$$
3) 
$$\frac{\alpha_{0}^{2}}{4} \boldsymbol{\sigma} \cdot (\nabla V_{ps}^{\text{loc}}(\mathbf{r}) \times \mathbf{p})$$

We see that, as expected, Term 2 and Term 3 almost compensate each other.

Fig. 2.4 illustrates this point but it also shows that it would be a rather rough approximation to consider only this term in the calculation. Therefore, even if it would represent a substantial gain in computing time, we do not use this approach and we keep all three terms in  $\tilde{\mathcal{H}}_{SO}$ .

Our calculations are, however, performed with collinear spins along one direction (named z) and within the *diagonal spin-orbit coupling* approximation: only  $\sigma_z \mathbf{e}_z$  is considered in  $\boldsymbol{\sigma}$  which allows to conserve the decoupling into two equations (one for spin up and one for spin down) of the Kohn-Sham equation.

### 2.3.2 Cross-section calculation

We have adapted the XSPECTRA code of QUANTUM ESPRESSO [Gougoussis et al., 2009] for the calculation of XMCD and XNCD spectra.

### Principle of the XSpectra code

The main features of the code XSpectra are presented in [Taillefumier et al., 2002]. We briefly recall them here.

The aim is to compute the contribution of a given operator O to the absorption cross-section:

$$\sigma(\omega) = 4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle f|O|i\rangle|^2 \delta(E_f - E_i - \hbar \omega).$$

The core state  $|i\rangle$  (for example, at K-edge, the 1s state) is determined from an all-electron isolated atom calculation because the core states are almost unaffected by the chemical bond (this is called the *frozen core* approximation).

The operators O are known. We give the expression of the spin-position operator - in the collinear spin and diagonal spin-orbit coupling case- in the next subsection and the expression of all operators in terms of spherical harmonics in Appendix D.2.

The scf calculation described in the previous subsection aims at determining the wave functions of the empty states  $|f\rangle$  but there are two problems:

(i) the scf calculation is pseudopotential-based so that pseudo-wave-functions  $|\hat{f}\rangle$  are determined instead of  $|f\rangle$ . They are not suitable to compute the matrix elements.

(ii) the calculation of empty states by DFT is computationally expensive and it would require

huge resources to compute the empty states in a supercell (which is needed when including the core-hole, see Section 2.5).

The first problem is addressed by using PAW reconstruction as presented in the previous subsection. In the PAW formalism, the matrix elements can be rewritten:<sup>6</sup>

$$\langle f|\mathcal{O}|i\rangle = \langle \tilde{f}|\tilde{\varphi}_{\mathbf{R}_0}\rangle \text{ with } |\tilde{\varphi}_{\mathbf{R}_0}\rangle = \sum_n |\tilde{p}_n^{\mathbf{R}_0}\rangle \langle \phi_n^{\mathbf{R}_0}|O|i\rangle.$$

In addition to the initial state wave function and to the operator, the vector  $\tilde{\varphi}_{\mathbf{R}_0}$  contains quantities related to the PAW reconstruction,  $|\tilde{p}_n^{\mathbf{R}_0}\rangle$  and  $|\phi_n^{\mathbf{R}_0}\rangle$ , which can be read in the pseudopotential file. At this point, the cross section can be rewritten as a sum over  $|\tilde{f}\rangle$ :

$$\sigma(\omega) = 4\pi^2 \alpha_0 \hbar \omega \langle \tilde{\varphi}_{\mathbf{R}_0} | \left( \sum_{\tilde{f}} |\tilde{f}\rangle \delta(E_f - E_i - \hbar \omega) \langle \tilde{f} | \right) | \tilde{\varphi}_{\mathbf{R}_0} \rangle,$$

which solves problem (i).

The second problem is addressed by a recursion method so that the cross section can be determined without having to compute any empty state. Using the Sokhotski–Plemelj theorem,

$$\lim_{\gamma \to 0^+} \frac{1}{z \pm i\gamma} = \mathcal{P}\left(\frac{1}{z}\right) \mp i\pi\delta(z)$$

(where P is the Cauchy principal value). Therefore,

$$\sum_{\tilde{f}} |\tilde{f}\rangle \delta(E_f - E_i - \hbar\omega) \langle \tilde{f}| = -\frac{1}{\pi} \sum_{f} |\tilde{f}\rangle \operatorname{Im} \frac{1}{E_i + \hbar\omega - E_f + i\gamma} \langle \tilde{f}| = -\frac{1}{\pi} \operatorname{Im}[\tilde{G}(E)]$$

where  $\tilde{G}(E)$  is the Green's operator associated with the pseudo-Hamiltonian  $\tilde{\mathcal{H}}_0$ :  $\tilde{G}(E)=(E-\tilde{\mathcal{H}}_0+i\gamma)^{-1}$  and the energy  $E=E_i+\hbar\omega$ . Note that, here, the Green's function is used only as a convenient way to write the invert of a matrix element. With this rewriting, we transformed the problem of a sum over empty states into a problem of large matrix inversion. In order to proceed to this inversion in a reasonable amount of time, Lanczos algorithm [Lanczos, 1952] is used to make  $\tilde{\mathcal{H}}_0$  tridiagonal with  $a_i$  on the main diagonal and  $b_i$  on the first

$$\langle f|O|i\rangle = \langle \tilde{f}|O|i\rangle + \sum_{\mathbf{R},j} \langle \tilde{f}|\tilde{p}_{j}^{\mathbf{R}}\rangle \langle \phi_{j}^{\mathbf{R}}|O|i\rangle - \sum_{\mathbf{R},j} \langle \tilde{f}|\tilde{p}_{j}^{\mathbf{R}}\rangle \langle \tilde{\phi}_{j}^{\mathbf{R}}|O|i\rangle$$

As the initial wave-function  $\langle \mathbf{r}|i\rangle$  is localized around the absorbing atom  $\mathbf{R}_0$ , we consider that it is zero outside the augmentation region  $\Omega_{\mathbf{R}_0}$  so that only  $\mathbf{R} = \mathbf{R}_0$  have to be considered in the previous expression. Moreover,  $\langle \mathbf{r}|O|i\rangle$  can be expanded on the  $|\tilde{\phi}_{\mathbf{R},n}\rangle$  basis:

$$O|i\rangle = \sum_{n} |\tilde{\phi}_{n,\mathbf{R}_{0}}\rangle \langle \tilde{p}_{n}^{\mathbf{R}_{0}}O|i\rangle$$

The first and the last terms in  $\langle f|O|i\rangle$  cancel each other, which leads to:

$$\langle f|\mathcal{O}|i\rangle = \langle \tilde{f}|\sum_{n} |\tilde{p}_{n}^{\mathbf{R}_{0}}\rangle \langle \phi_{n}^{\mathbf{R}_{0}}|O|i\rangle.$$

 $<sup>^{6}</sup>$  Using Eq. (2.6) and Eq. (2.7),

diagonals:7

The inversion leads to a form in continued fraction of the matrix elements:

$$\langle \tilde{\varphi}_{\mathbf{R}_{0}} | (\tilde{\mathcal{H}}_{0} - E - i\gamma)^{-1} | \tilde{\varphi}_{\mathbf{R}_{0}} \rangle = \frac{\langle \tilde{\varphi}_{\mathbf{R}_{0}} | \tilde{\varphi}_{\mathbf{R}_{0}} \rangle}{a_{0} - E - i\gamma - \frac{b_{1}^{2}}{a_{1} - E - i\gamma - \frac{b_{2}^{2}}{\cdots}}} \equiv \langle \tilde{\varphi}_{\mathbf{R}_{0}} | \tilde{\varphi}_{\mathbf{R}_{0}} \rangle \operatorname{cont}(E, \mathbf{a}, \mathbf{b})$$

We define this way  $cont(E, \mathbf{a}, \mathbf{b})$  where **a** and **b** are vectors of components  $a_i$  and  $b_i$  respectively.

Finally, the contribution of O to the absorption cross section is computed as:

$$\sigma(E) = 4\pi \alpha_0 \hbar \omega \langle \tilde{\varphi}_{\mathbf{R}_0} | \tilde{\varphi}_{\mathbf{R}_0} \rangle \operatorname{cont}(E, \mathbf{a}, \mathbf{b}).$$

In the code, for each k-point,  $|\tilde{\varphi}_{\mathbf{R}_0}\rangle$  is determined for the chosen operator with the formula given in Appendix D.3. Then, the Lanczos procedure is applied to determine the vectors **a** and **b**, and a spectrum is calculated. The total spectrum is obtained as a sum over the k-points of these spectra. If the calculation is spin-polarized the number of k-points is multiplied by two so that each k-point is associated to a given spin. This allows to obtain spin-up and spin-down spectra by summing over the corresponding k-points.

### Details related to the cross terms

### (i) SP operator in the case of collinear spins

In order to compute the D-SP cross term, the calculation must be spin-polarized. The only case implemented, for now, is that of collinear spins and diagonal SOC ( $\boldsymbol{\sigma}$  is replaced by  $\sigma_z \mathbf{e}_z$ ). In the code pwscf spin channels remain independent because spin-up and spin-down electronic states are solutions of separate Kohn-Sham equations. We consider that the collinear spins are

along z. The spin part of the wave functions  $|s\rangle$  can either be the spin up spinor  $|\uparrow\rangle = \begin{pmatrix} 1\\ 0 \end{pmatrix}$ ,

or the spin down spinor  $|\downarrow\rangle = \begin{pmatrix} 0\\ 1 \end{pmatrix}$ .

$$\begin{cases} |u_{-1}\rangle = 0\\ |u_{0}\rangle = \frac{|\tilde{\varphi}_{\mathbf{R}_{0}}\rangle}{\sqrt{\langle \tilde{\varphi}_{\mathbf{R}_{0}} | \tilde{\varphi}_{\mathbf{R}_{0}} \rangle}}\\ \tilde{\mathcal{H}}_{0}|u_{i}\rangle = a_{i}|u_{i}\rangle + b_{i+1}|u_{i+1}\rangle + b_{i}|u_{i-1}\rangle \end{cases}$$

In practice, at each iteration,  $|t\rangle = \tilde{\mathcal{H}}_0 |u_i\rangle - b_i |u_{i-1}\rangle$ , that is equal to  $b_{i+1} |u_{i+1}\rangle + a_i |u_i\rangle$ , is computed. Then,  $a_i = \langle t | u_i \rangle$  can be determined and  $|t'\rangle = |t\rangle - a_i |u_i\rangle$  is computed. Then  $b_{i+1}$  is equal to the norm of  $|t'\rangle$ :  $b_{i+1} = \sqrt{\langle t' | t' \rangle}$  and  $|u_{i+1}\rangle = \frac{|t'\rangle}{b_{i+1}}$ . Note that in the code, the indexes for  $a_i$  start from 1 whereas, in our notation they start from 0. This only implies a shift of 1 in the index of  $a_i$  in each formula.

 $<sup>^7</sup>$  The Lanczos basis can be determined recursively as it obeys :

The D-SP term is a cross term between the electric dipole and the spin-position operator. Spin does not appear in the electric dipole operator, so it is diagonal in spin

$$\langle \phi_i s | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | \phi_f s' \rangle = \langle \phi_i | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | \phi_f \rangle \delta_{ss'}.$$

This imposes s' = s. On the other hand, the vector of Pauli matrices  $\boldsymbol{\sigma}$  appears explicitly in the Spin-Position operator:

$$\langle \phi_i s | \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}) | \phi_f s \rangle = \langle \phi_i | (\boldsymbol{\epsilon} \times \mathbf{r}) | \phi_f \rangle \cdot \langle s | \boldsymbol{\sigma} | s \rangle.$$

As  $\langle s|\sigma_x|s\rangle = \langle s|\sigma_y|s\rangle = 0$ , we can exclude a priori the terms that are proportional to  $\sigma_x$  and  $\sigma_y$  in the Spin-Position operator. In that case the spin position operator rewrites:

$$SP_{col}(\boldsymbol{\epsilon}) = \boldsymbol{\sigma}.(\boldsymbol{\epsilon} \times \mathbf{r}) = \sigma_z(\epsilon_x y - \epsilon_y x).$$

Its expression in spherical harmonics, as it is implemented in the code, is presented in Appendix D.2.

### (ii) Calculation of the cross terms

To compute the D-Q and D-SP cross terms, we cannot simply apply the procedure presented above (p. 54) because they do not write as square modulus. Let us consider the electric dipole operator D and another operator B, which is either the electric quadrupole Q or the spinposition  $SP_{col}$  operator. We want to compute:

$$I_{\rm D-B} = \pi \sum_{f} {\rm Im}[\langle f|B|i\rangle\langle i|D^*|f\rangle]\delta(E_f - E_i - \hbar\omega).$$

Two calculations are performed in order to obtain:  $I_1 = \pi \sum_f |\langle f|D + iB|i\rangle|^2 \delta(E_f - E_i - \hbar\omega)$ and  $I_2 = \pi \sum_f |\langle f|D - iB|i\rangle|^2 \delta(E_f - E_i - \hbar\omega)$ . The difference of  $I_1$  and  $I_2$  is proportional to  $I_{\text{D-B}}$ :

$$I_1 - I_2 = -4I_{\rm D-B}.$$

The prefactors implemented in the code in order to obtain from  $I_1$  and  $I_2$  the contribution of the cross terms to the absorption cross section are given in Appendix D.4. The fact that two Lanczos procedures are needed to determine the cross terms means that they require twice as much computing time as the other terms.

### (iii) Alternative way to compute the term D-SP

If one considers full circular polarization such that  $\mathbf{k}$  is along the quantisation axis z, there is an alternative way to compute the D-SP term. This case corresponds to the calculation of XMCD. Indeed, in XMCD experiments a magnetic field is usually applied parallel to the beam, which justifies to consider  $\mathbf{k}$  parallel to the quantisation axis.

The calculation must be performed with left-handed and right-handed polarization:

$$\boldsymbol{\epsilon}_{L} = \boldsymbol{\epsilon}_{2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\\ i\\ 0 \end{pmatrix} \text{ and } \boldsymbol{\epsilon}_{R} = \boldsymbol{\epsilon}_{L}^{\star} = \boldsymbol{\epsilon}_{1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\\ -i\\ 0 \end{pmatrix}$$
 (2.9)

Using Eq. (D.2), as  $Y_1^{-1}(\boldsymbol{\epsilon}_1) = 0$ ,  $Y_1^{-1}(\boldsymbol{\epsilon}_2) = \sqrt{3/4\pi}$ ,  $Y_1^0(\boldsymbol{\epsilon}_1) = Y_1^0(\boldsymbol{\epsilon}_2) = 0$ ,  $Y_1^1(\boldsymbol{\epsilon}_1) = -\sqrt{3/4\pi}$  and  $Y_1^1(\boldsymbol{\epsilon}_2) = 0$ :

$$SP_{col}(\boldsymbol{\epsilon}_1) = i\sqrt{\frac{4\pi}{3}}rY_1^{-1}(\mathbf{u_r})\sigma_z = \sigma_z i\boldsymbol{\epsilon}_1.\mathbf{r}$$

$$SP_{col}(\boldsymbol{\epsilon}_2) = i\sqrt{\frac{4\pi}{3}}rY_1^1(\mathbf{u_r})\sigma_z = -\sigma_z i\boldsymbol{\epsilon}_2.\mathbf{r}.$$

Hence,

$$\sigma_{\rm D-SP}(\boldsymbol{\epsilon}_1) = -\frac{\hbar\omega}{2mc^2} (\sigma_{\rm D-D}^{\uparrow}(\boldsymbol{\epsilon}_1) - \sigma_{\rm D-D}^{\downarrow}(\boldsymbol{\epsilon}_1))$$
  
$$\sigma_{\rm D-SP}(\boldsymbol{\epsilon}_2) = \frac{\hbar\omega}{2mc^2} (\sigma_{\rm D-D}^{\uparrow}(\boldsymbol{\epsilon}_2) - \sigma_{\rm D-D}^{\downarrow}(\boldsymbol{\epsilon}_2))$$
(2.10)

with

$$\sigma_{\rm D-D}^{s}(\boldsymbol{\epsilon}) = 4\pi^{2}\alpha_{0}\hbar\omega\sum_{f}|\langle f^{s}|\boldsymbol{\epsilon}\cdot\mathbf{r}|i^{s}\rangle|^{2}\delta(E_{f}-E_{i}-\hbar\omega)$$

where  $s = \uparrow$  or  $\downarrow$ . Therefore, in the collinear case, the D-SP term can be computed from the D-D cross section for the up and down spin channels. In other words, the calculation of the spin-dependent D-D term is sufficient to obtain the D-SP term without requiring more computing time. We have checked that this approach gives the same result as the double Lanczos procedure described in paragraph (ii).

# 2.3.3 Discussion on the necessity to include spin orbit coupling in the scf calculation



Fig. 2.5: Comparison of the D-D contribution to the spectra (test calculation for a bcc Fe structure) obtained by including (red) or not (green) spin-orbit coupling in the first step of the method (*i.e.* the self-consistent field calculation). On the left, spectra are convoluted by a Lorentzian of width at half maximum 1.6 eV and on the right, by a Lorentzian of width at half maximum 0.2 eV. In both cases, red and green curves are the same.

Due to the rewriting of the sum over  $|\tilde{f}\rangle$  with the delta function in terms of  $(E - \tilde{\mathcal{H}}_0 + i\gamma)^{-1}$ , in our method, the aim of the scf calculation is only to determine the electron wave functions to be able to compute  $\tilde{\mathcal{H}}_0$  in the second step of our calculation. In order to better understand the effect of spin orbit coupling (SOC) in the XMCD effect, we tried to perform this scf calculation, that is to say the electron relaxation, without SOC. We included it, however, in the  $\tilde{\mathcal{H}}_0$  used by XSpectra (otherwise no D-D or Q-Q XMCD could be computed).

The result of this test is shown in Figure 2.5 and it is quite surprising at first sight: the result is exactly the same if SOC is included or not in the scf relaxation.



**Fig. 2.6:** *p*-projected densities of states (DOS) for *bcc* Fe obtained with the relativistic version of **pwscf**. In purple for  $j = \frac{1}{2}$  and in green for  $j = \frac{3}{2}$ . The DOS for  $j = \frac{1}{2}$  multiplied by two has been added in transparent purple for easy comparison between  $j = \frac{1}{2}$  and  $j = \frac{3}{2}$  DOS.

This is in fact understandable by considering the *p*-projected density of states in *bcc* Fe (Fig. 2.6). From this plot, we can evaluate the order of magnitude of the value of the spin orbit splitting of the 4*p* states in Fe: it is at most a few meV which is smaller than the energy resolution needed for the calculation of the spectra (for the examples of Figure 2.5, the spacing between points is  $\Delta E = 0.06$  eV which is enough even with a 0.2eV broadening).

In other words, it seems that, as long as the spin orbit splitting of the probed states is small with respect to the energy resolution, the electron relaxation can be performed without including spin orbit coupling. This represents a significant gain in computing time.

# 2.4 Convergence with calculation parameters: example of Fe

When running the two steps of a cross-section calculation, several parameters must be set in inputs. Computation time and required memory depend on these parameters which limit their accessible values. For this reason, it is important to find the values that allow convergence of the properties we are interested in. This convergence must be checked for any system under study. Here, to have an idea of the dependence of the spectra on these parameters, we have checked in the case of *bcc*-Fe the convergence of the D-D XAS spectrum and of the XMCD corresponding to the terms D-D, D-SP and Q-Q.

The correspondence between the quantities defined in Subsection 2.2.2 and the quantities plotted below (in this chapter only - in other chapters XAS spectrum will be the sum of several terms) are:

"XAS" 
$$\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{1})$$
  
"D-D XMCD"  $\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{2}) \cdot \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{1})$   
"Q-Q XMCD"  $\sigma_{\mathrm{Q-Q}}(\boldsymbol{\epsilon}_{2}) \cdot \sigma_{\mathrm{Q-Q}}(\boldsymbol{\epsilon}_{1})$  with  $\boldsymbol{\epsilon}_{2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\\i\\0 \end{pmatrix}$  and  $\boldsymbol{\epsilon}_{1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\\-i\\0 \end{pmatrix}$ .  
"D-SP XMCD"  $\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}_{2}) \cdot \sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}_{1})$ 

# 2.4.1 Parameters of the scf calculation

In this section, calculations for a unitary *bcc* Fe cell (without core hole) with the experimental lattice parameter  $a_{\text{lat}} = 2.87$  Å are presented as a test example. All parameters are varied independently while others are kept fixed.
When relevant, the convergence of the total energy, the absolute magnetization and the total magnetization at the end of the self-consistent field calculation is plotted. The total magnetization is the integral of the magnetization in the cell and the absolute magnetization is the integral of the magnetization in the cell.

A norm-conserving pseudopotential with GGA-PBE functional is used (see Appendix E). The values chosen as references for the parameters are: a Methfessel-Paxton smearing of 0.01 Ry, a  $10 \times 10 \times 10$  uniform k-point grid, an energy cut-off  $E_{cutoff} = 180$  Ry and a convergence threshold set to  $10^{-9}$  Ry. For the spectra, the reference calculation is performed with a  $20 \times 20 \times 20$  uniform k-point grid and a convergence error set to  $5.10^{-4}$ .

Spectra are convolved with a Lorentzian broadening function with full width at half maximum set to 0.8 eV.

#### k-points sampling



Fig. 2.7: Absolute magnetization, total magnetization (left Y-axis) and total energy (right Y-axis, red) for a unitary *bcc*-Fe cell as a function of the size of the *k*-point grid in each direction. The smearing is of Methfessel-Paxton type with a spreading  $\sigma = 0.01$  Ry.

We consider a crystal whose translations write  $\mathbf{T} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$  (where  $n_1, n_2, n_3$  are integers) and we name  $\Omega$  the volume of the cell defined by  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  and  $\mathbf{a}_3$ . We note  $\mathbf{G}$  the reciprocal lattice vectors ( $\mathbf{G} = n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2 + n_3 \mathbf{b}_3$  with  $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij}$ ). In the presence of a periodic potential,  $V(\mathbf{r} + \mathbf{T}) = V(\mathbf{r})$ , Bloch's theorem states that the eigenfunctions can be written in the form:

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{i,\mathbf{k}}(\mathbf{r})$$

where  $u_{i,\mathbf{k}}$  has the periodicity of the crystal and writes:

$$u_{i,\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{i\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}.$$

The Hamiltonian is block diagonal in  $\mathbf{k}$  such that a Schrödinger equation can be written and must be solved for any  $\mathbf{k}$ . As  $\psi_{i,\mathbf{k}+\mathbf{G}}(\mathbf{r})$  is the same as  $\psi_{i,\mathbf{k}}(\mathbf{r})$ , a restriction to the first Brillouin zone (BZ - primitive cell in reciprocal space that contain the points closer to the origin) is possible. In practice a finite number of k-points in the first Brillouin zone is used. To choose these points, the most widely used scheme is the method proposed by Monkhorst and Pack [Monkhorst and Pack, 1976] which consists in building a uniform set of points in each direction of the reciprocal lattice:

$$\mathbf{k}_{n_1, n_2, n_3} = \sum_{i} \frac{2n_i - N_i - 1}{2N_i} \mathbf{b}_i$$
(2.11)



Fig. 2.8: Calculated XAS and XMCD spectra for a unitary *bcc*-Fe cell with different *k*-point grids, all other parameters being equal to the reference values. The smearing is of type Methfessel-Paxton with a spreading  $\sigma = 0.01$  Ry. In inset: zoom near the extrema of the spectra.

with  $n_i = 1, 2, ..., N_i$  where  $N_1, N_2, N_3$  are chosen in input. It is called a  $N_1 \times N_2 \times N_3$  uniform grid. A sum over these points exactly integrates a periodic function that has Fourier components that extend to  $N_i \mathbf{b}_i$  in each direction [Martin, 2004, p.93].

Symmetry can be used to reduce the calculation to the irreducible Brillouin zone (IBZ). The weight  $w_{\mathbf{k}}$  of the remaining k-points is defined by the ratio of the total number of k-points related by symmetry to the given k-point divided by the total number of k-points  $N_k$ . The average value per cell for a general function  $f(\mathbf{k})$  is:

$$\overline{f} = \frac{1}{N_k} \sum_{\mathbf{k} \in \mathrm{BZ}} f(\mathbf{k}) = \sum_{\mathbf{k} \in \mathrm{IBZ}} w_{\mathbf{k}} f(\mathbf{k}).$$

Note that in this thesis the k-point grid dimensions are given for the full Brillouin zone, as it is customary.

Fig 2.7, the total energy, total magnetization and absolute magnetization at the end of the scf calculation are plotted as a function of the size of the k-point grid. The total magnetization is defined as  $\mathbf{M}_{\text{tot}} = \int_{\text{cell}} (\rho^{\uparrow}(\mathbf{r}) - \rho^{\downarrow}(\mathbf{r})) dr$  and the absolute magnetization as  $\mathbf{M}_{\text{abs}} = \int_{\text{cell}} |\rho^{\uparrow}(\mathbf{r}) - \rho^{\downarrow}(\mathbf{r})| dr$ . The fact that they are almost equal indicates that *bcc*-Fe is ferromagnetic.

The convergence is not monotonous and the relative variation of the magnetization is still rather large ( $\approx 4\%$ ) with dense k-point grids for which the total energy is converged with a precision of the order of  $10^{-3}$  Ry. For a metal, convergence with the k-point mesh must be checked together with the convergence with the spreading corresponding to the chosen smearing (see Subsection 2.4.1). A combined test can be done by increasing the number of k-points with several chosen values of smearing and then determine the smearing and k-point grid that allow to obtain the desired precision on the total energy and magnetization.

Fig. 2.8 shows that the calculated XAS spectrum is not sensitive to the size of the k-point grid in the chosen interval. The three contributions to XMCD depend more strongly on this parameter. For the D-D contribution it results only in a small variation of the intensity of the signal but for the D-SP and the Q-Q contributions, the shape of the spectra very close to the Fermi level depends on the number of k-points  $N_k$  used in the calculation. The convergence of the XMCD spectra, like the convergence of magnetization, is not monotonous when increasing  $N_k$ .

# Energy cut-off



Fig. 2.9: Absolute magnetization, total magnetization (left Y-axis) and total energy (right Y-axis, red) for a unitary *bcc*-Fe cell as a function of the energy cut-off.

The number of plane waves used to describe the wave functions is limited by the available computing resources. Therefore the sum over **G** is restrained to the reciprocal lattice vectors such that  $|\mathbf{G}| < G_{\max}$  (sphere in the reciprocal space) where  $G_{\max}$  is indirectly chosen in input through the choice of the cut-off energy.

The kinetic energy associated with a plane wave  $\phi_{\mathbf{G}} = \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$  is:

$$E_{\mathrm{K}} = \frac{-1}{2} \nabla^2 \phi_{\mathbf{k},\mathbf{G}}(\mathbf{r}) = \frac{1}{2} |\mathbf{k} + \mathbf{G}|^2 \phi_{\mathbf{k},\mathbf{G}}(\mathbf{r}) \approx \frac{1}{2} |\mathbf{G}|^2 \phi_{\mathbf{k},\mathbf{G}}(\mathbf{r}) \text{ for } |\mathbf{G}| \text{ large enough.}$$

Therefore, the restriction  $|\mathbf{G}| < G_{\text{max}}$  leads to  $E_{\text{K}} < \frac{G_{\text{max}}^2}{2}$ . The cut-off energy is defined as:

$$E_{cutoff} = \frac{G_{\max}^2}{2}.$$

For a given cut-off energy, the number of plane waves is  $N_g \approx \frac{1}{2\pi^2} \Omega E_{cutoff}^{3/2}$ . In other words, the size of the basis set used to describe the wave functions depends on the cut-off energy and on the volume of the cell.



Fig. 2.10: Calculated XAS and XMCD spectra for a unitary *bcc*-Fe cell with different cut-off energies, all other parameters being equal to the reference values.

The kinetic energy cut-off for the charge density (that writes as sum of squared modulus of wave functions) is set to  $4E_{cutoff}$ .<sup>8</sup>

The total energy has a monotonous behaviour with  $E_{cutoff}$  (see Fig 2.9) which makes the convergence with respect to this parameter easier than for the k-point grid size. The combination of Fig. 2.10 and Fig 2.9 illustrates the fact that, in the case considered here, a precision of  $10^{-2}$  Ry, which corresponds to a relative error of  $4.10^{-5}$  on the total energy, is enough for all spectra to be well converged.

# Smearing

If the material under study is metallic, the quantities to be integrated over the first Brillouin zone are multiplied by a sharp function equal to 1 if  $E < E_F$  and equal to 0 if  $E > E_F$ . This discontinuity at the Fermi level is problematic because a very dense k-point grid would be required to get an acceptable precision on the calculated values. Such a problem does not arise in the case of insulators because the density of states cancels smoothly before the gap. The practical solution in the case of metals is to replace the step function by a smoother function

<sup>&</sup>lt;sup>8</sup>If ultrasoft pseudopotentials were used instead of norm-conserving ones it should be set to a higher value.

 $\beta$ :

$$\sum_{\mathbf{k}} w_{\mathbf{k}} f(\mathbf{k}) \Theta(E(\mathbf{k}) - E_F) \text{ with } \Theta(x) = \begin{cases} 0 & \text{if } x < 0\\ 1 & \text{if } x > 0 \end{cases}$$
$$\longrightarrow \sum_{\mathbf{k}} w_{\mathbf{k}} f(\mathbf{k}) \beta\left(\frac{E(\mathbf{k}) - E_F}{\sigma}\right).$$

A gaussian smearing corresponds to the choice  $\beta(x) = \frac{1}{2} (1 - \operatorname{erf}(x))$  where  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ .



Fig. 2.11: Absolute magnetization, total magnetization (left Y-axis) and total energy (right Y-axis) for a unitary *bcc*-Fe cell as a function of the logarithm of the spreading  $\sigma$  in the Methfessel-Paxton scheme. The *k*-point grid is a  $10 \times 10 \times$ 10 uniform grid.

In the more sophisticated *Methfessel-Paxton scheme* [Methfessel and Paxton, 1989], the function  $\beta$  is obtained recursively as it writes as a function of Hermite-polynomials:

$$\beta_0(x) = \frac{1}{2} \left( 1 - \operatorname{erf}(x) \right)$$
  
$$\beta_N(x) = \beta_0(x) + \sum_{n=1}^N A_n H_{2n-1}(x) e^{-x^2} \text{ with } A_n = \frac{(-1)^n}{n! 4^n \sqrt{\pi}}$$
  
and  $H_0(x) = 1, H_1(x) = 2x, H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x)$ 

These  $\beta_n$  functions have the interesting property that:

$$\int_{E} dEf(E)\beta_{N}\left(\frac{E-E_{F}}{\sigma}\right) = \int_{E} dEf(E)\Theta(E-E_{F})$$

if f is a polynomial of 2N - th or smaller degree. In the code pwscf, this scheme is used at first order.

A spreading  $\sigma$  that is too large can result in a wrong total energy and a wrong magnetization but the smaller the spreading is, the larger the k-point grid must be.

To be rigorous, the convergence of the calculation when decreasing the spreading must be checked together with the choice of the k-point grid (the same plot as in Fig. 2.11 should be plotted with other grids). Our point here is to show the impact of each parameter on the spectra, hence we only make  $\sigma$  vary.

As illustrated in Fig. 2.12, the calculated XAS spectrum does not depend much on  $\sigma$  whereas the three contributions to XMCD depend a little more appreciably on this parameter. For the D-D and D-SP contribution it results mostly in a small variation of the intensity of the signal but for the Q-Q contribution, the shape of the spectra very close to the Fermi level changes when  $\sigma$  is changed.



Fig. 2.12: Calculated XAS and XMCD spectra for a unitary *bcc*-Fe cell with different spreading  $\sigma$ , all other parameters being equal to the reference values.

#### Convergence threshold

During the scf calculation, a numerical iterative procedure is used, which changes successively the effective potential  $V_{eff}$  and the density n. The flow chart shown in Fig 2.13 illustrates the idea of this procedure.

The convergence threshold for self-consistency set in input of pwscf corresponds to the limit on the error on the total energy at the end of the cycle. It is the criterion that puts an end to the cycle (box "Self-consistent ?" in Fig 2.13).

The calculated spectra are insensitive to a change in convergence threshold if it is smaller than  $10^{-5}$  Ry as illustrated in Fig. 2.14. This corresponds to a relative error on the total energy that is smaller than  $4.10^{-8}$ . Note that, as the total energy is extensive, to obtain the same relative error on a system that would contain more Fe atoms, a smaller convergence threshold would be required.



Fig. 2.13: Schematic representation of the self-consistent loop for solving the Kohn-Sham equation. In the spinpolarized case, one must iterate two of such loops simultaneously for the two spin channels. Figure is redrawn from [Martin, 2004, Fig.9.1].



Fig. 2.14: Calculated XAS and XMCD spectra for a unitary *bcc*-Fe cell with different convergence threshold, all other parameters being equal to the reference values.

# 2.4.2 Parameters of XSpectra

Here, we study the effect of two parameters that can be set in the input of the code XSpectra: the number of k-points and the convergence error parameter.



# k-point grid

Fig. 2.15: Calculated XAS and XMCD spectra for a unitary bcc-Fe cell with different k-point grids for the cross section calculation (input of XSpectra).

In XSpectra, the k-point mesh is of type Monkhorst-Pack (*i.e.* the k-points are distributed homogeneously in the Brillouin zone) and the number of k-points is not reduced by symmetry. In other words, the calculation is performed for all k-points of the mesh and not for the non-symmetry-equivalent k-points only.

The number of k-points used to compute the absorption cross section has a strong impact on the result as visible in Fig. 2.15. If the k-point grid is not dense enough, there is a lot of numerical noise in the XAS spectra, which results in lots of spurious peaks in XMCD. As a convolution of the spectra with a larger broadening at high energy would smooth them, the fact that the calculation is not converged might go unnoticed if the calculation is directly performed with a too large broadening.

# **Convergence** error



Fig. 2.16: Calculated XAS and XMCD spectra for a unitary *bcc*-Fe cell with different values for the convergence error parameter set in input of XSpectra.

For each k-point, when  $\mathcal{H}_0$  is made tridiagonal with Lanczos algorithm, a spectrum is periodically computed and compared to the spectrum obtained at the previous iteration in order to determine if the calculation is converged or not. The number of iterations that separates two checks of this kind can be set in input (keyword: xcheck\_conv). The error (keyword: xerror) corresponds to the criterion for the spectra to be considered as converged: for a given k-point, at step l, the spectrum is considered as converged if the relation

$$\frac{\sum_{E} |\operatorname{cont}(E, \mathbf{a}_{l}, \mathbf{b}_{l}) - \operatorname{cont}(E, \mathbf{a}_{l-\text{xcheck}\_\text{conv}}, \mathbf{b}_{l-\text{xcheck}\_\text{conv}})|}{\sum_{E} |\operatorname{cont}(E, \mathbf{a}_{l}, \mathbf{b}_{l})|} \leq \operatorname{xerror}$$

is verified.

As visible in Fig. 2.16, XMCD spectra are more sensitive to this parameter than XAS spectra. If **xerror** is set to a too large value, peaks at high energy appear, especially in the D-D contribution to XMCD.



Fig. 2.17: Example of a  $3 \times 3 \times 3$  supercell of the *bcc* Fe cubic cell (conventional cell). With this cell, periodically reproduced (due to periodic boundary conditions) atoms with a core hole are 8.61 Å away from each other.

# 2.5 Inclusion of the core hole

In X-ray absorption experiments, a core-hole is created by excitation of a core electron. The spectra can be strongly affected by core-hole effects (see for example [Taillefumier et al., 2002], Fig.4 or [Cabaret et al., 2010]). It is therefore important to calculate the screened core-hole potential that is experienced by an excited electron.

# 2.5.1 Principle

In the framework of the final state rule [von Barth and Grossmann, 1982] the final states  $|f\rangle$  that enter the cross section formula Eq. (2.1) are stationary unoccupied states in the presence of a static core hole whereas  $|i\rangle$  is the core state without core hole.

We therefore run the scf calculation with one electron missing in the core state (and a background charge to insure electrical neutrality). In practice, a core electron is removed from the pseudopotential of the absorbing atom. For example for Fe, the pseudopotential for the absorbing Fe is generated with the configuration  $1s^1 2s^2 2p^6 3s^2 3p^6 3d^6$  (with the 3s, 3p, 3d electrons in valence).

The response of the electrons to the presence of the core is therefore considered at all orders. In other words, we do not use linear-response theory to determine the core-hole screening as it is done within the random-phase approximation (RPA) [Shirley et al., 2005] but we compute it self-consistently. This independent particle approximation has proven successful to compute K-edge XAS spectra but it fails to reproduce  $L_{2,3}$  edges for 3d elements [Laskowski and Blaha, 2010b] because the effects related to the electron-hole interactions are strong. A comparison of the calculation using the final-state rule with an approach using Bethe-Salpeter equation to describe the electron-hole interaction is carried out in Ref. [Rehr et al., 2005]. In this paper, a close connection between the two approaches is established.

# 2.5.2 Supercell

The core-hole lifetime  $\tau$  in XAS is of the order of several  $10^{-15}$  s = 1 fs. In XAS experiments, the created core-holes are too distant to interact with each other.<sup>9</sup>

<sup>&</sup>lt;sup>9</sup>For example, at the Fe K-edge the lifetime broadening is  $\Gamma = 1.25$  eV which corresponds to  $\tau \approx 20$  fs ( $\Gamma \tau = \hbar/2$ ). The absorption cross section above the edge (7113 eV) is  $\sigma \approx 38000$  barns/atom [database CXRO, site]. On ODE beamline at Soleil synchrotron, the photon flux on the sample at 7 keV is  $f = 4 \times 10^{22}$  photons.m<sup>-2</sup>.s<sup>-1</sup> =  $4 \times 10^{-6}$  photons.barns<sup>-1</sup>.s<sup>-1</sup>. The number of photons absorbed per second per atom,  $n_{abs} = \sigma f$ . For the mean number of absorbed photons during time  $\tau$  to be equal to 1, there must be a total of  $N = \frac{1}{n_{abs}\tau} \approx 3 \times 10^{14}$  atoms. It means, that, on average the core hole are  $N^{1/3} \approx 10^5$  atoms away



Fig. 2.18: XAS and XMCD spectra calculated with an absorbing atom containing a core-hole (ch) in *bcc*-Fe cells of different sizes. The distance between periodically reproduced core-holes ranges from 5.73 to 9.96 Å.

When including a core-hole in the cell used to describe the material, it is important, that the periodically repeated core holes (due to periodic boundary conditions) do not interact with each other. A supercell that contains a large number of atoms, among which one atom with a core-hole, must therefore be built. The k-point grid can be reduced accordingly ( $\times n$  in real space corresponds to  $\times 1/n$  in reciprocal space). To verify that core-holes do not interact with each other, one must check that the spectra do not change when the supercell size is increased. Previous works demonstrated that a distance between core-holes of 8 to 10 Å is usually enough to reach convergence of the XAS spectrum [Cabaret et al., 2010] and that this distance is material dependent. One sees, in the example of *bcc*-Fe (Fig. 2.18), that the XMCD spectra are more sensitive to the interaction between core-holes than the XAS spectrum. A distance between core-holes of approximatively 10 Å is, however, enough to obtain converged spectra.

one from each other. With an interatomic distance d = 2 Å, it corresponds to a distance between core holes of approximatively 10  $\mu$ m. Therefore, the core holes do not interact with each other.

# 2.6 Conclusion

In this chapter, we have presented the significant terms of the absorption cross section and the method that we use to compute them. To summarize, the cross section is computed in the presence of a static core-hole with a full relaxation of the valence electrons. For this, we work in a DFT framework with plane-waves and pseudopotentials. For XMCD, the calculation must be spin-polarized and include spin-orbit coupling. We consider collinear spins and the spinorbit coupling is accounted for within a diagonal spin-orbit coupling approximation. The cross section is computed using a recursion method that avoids the heavy workload of computing empty states by DFT.

The effect of several parameters that can be set in input of the calculation on the obtained spectra has been discussed.

The results for model systems are presented in the next chapter.

# Bibliography

- [Altarelli, 1993] M. Altarelli (1993). Orbital-magnetization sum rule for x-ray circular dichroism: A simple proof. *Phys. Rev. B*, 47(2):597–598.
- [Blöchl, 1994] P. E. Blöchl (1994). Projector augmented-wave method. *Phys. Rev. B*, 50(24):17953-17979.
- [Brouder, 1990] C. Brouder (1990). Angular dependence of X-ray absorption spectra. J. Phys.: Condens. Matter, 2(3):701.
- [Brouder et al., 1996] C. Brouder, M. Alouani, and K. H. Bennemann (1996). Multiplescattering theory of x-ray magnetic circular dichroism: Implementation and results for the iron K edge. Phys. Rev. B, 54(10):7334-7349.
- [Bunău and Joly, 2009] O. Bunău and Y. Joly (2009). Self-consistent aspects of x-ray absorption calculations. J. Phys.: Condens. Matter, 21(34):345501.
- [Cabaret et al., 2010] D. Cabaret, A. Bordage, A. Juhin, M. Arfaoui, and E. Gaudry (2010). First-principles calculations of X-ray absorption spectra at the K-edge of 3d transition metals: an electronic structure analysis of the pre-edge. *Phys. Chem. Chem. Phys.*, 12(21):5619.
- [Carra et al., 1993] P. Carra, B. T. Thole, M. Altarelli, and X. Wang (1993). X-ray circular dichroism and local magnetic fields. *Phys. Rev. Lett.*, 70(5):694–697.
- [Ceresoli et al., 2010] D. Ceresoli, U. Gerstmann, A. P. Seitsonen, and F. Mauri (2010). Firstprinciples theory of orbital magnetization. *Phys. Rev. B*, 81(6):060409.
- [Clementi and Raimondi, 1963] E. Clementi and D. L. Raimondi (1963). Atomic Screening Constants from SCF Functions. J. Chem. Phys., 38(11):2686-2689.
- [Clementi et al., 1967] E. Clementi, D. L. Raimondi, and W. P. Reinhardt (1967). Atomic Screening Constants from SCF Functions. II. Atoms with 37 to 86 Electrons. J. Chem. Phys., 47(4):1300-1307.
- [database CXRO, site] CXRO (website). CXRO database accessed February 02, 2017. http: //henke.lbl.gov/cgi-bin/pert\_cgi.pl.
- [de Groot and Kotani, 2008] F. Groot and A. Kotani (2008). Core Level Spectroscopy of Solids. Advances in Condensed Matter Science. CRC Press.
- [Di Matteo et al., 2005] S. Di Matteo, Y. Joly, and C. R. Natoli (2005). Detection of electromagnetic multipoles by x-ray spectroscopies. *Phys. Rev. B*, 72:144406.
- [Ebert, 1996] H. Ebert (1996). Magneto-optical effects in transition metal systems. Reports on Progress in Physics, 59(12):1665.

- [Ebert et al., 2011] H. Ebert, D. Ködderitzsch, and J. Minár (2011). Calculating condensed matter properties using the KKR-Green's function method-recent developments and applications. *Reports on Progress in Physics*, 74(9):096501.
- [Edmonds et al., 2005] K. W. Edmonds, N. R. S. Farley, T. K. Johal, G. Laan, R. P. Campion, B. L. Gallagher, and C. T. Foxon (2005). Ferromagnetic moment and antiferromagnetic coupling in (Ga,Mn)As thin films. *Phys. Rev. B*, 71:064418.
- [FDMNES, site] FDMNES (website). The FDMNES project accessed March 06, 2017. http: //neel.cnrs.fr/spip.php?rubrique1007&lang=en.
- [Fujikawa and Nagamatsu, 2003] T. Fujikawa and S. Nagamatsu (2003). Relativistic multiple scattering theory of K-edge X-ray magnetic circular dichroism. J. Electron. Spectrosc. Relat. Phenom., 129(1):55–69.
- [Gerstmann et al., 2014] U. Gerstmann, N. J. Vollmers, A. Lücke, M. Babilon, and W. G. Schmidt (2014). Rashba splitting and relativistic energy shifts in In/Si(111) nanowires. *Phys. Rev. B*, 89(16):165431.
- [Giannozzi et al., 2009] P. Giannozzi et al. (2009). QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. J. Phys.: Condens. Matter, 21(39):395502.
- [Gougoussis et al., 2009] C. Gougoussis, M. Calandra, A. P. Seitsonen, and F. Mauri (2009). First-principles calculations of x-ray absorption in a scheme based on ultrasoft pseudopotentials: From  $\alpha$ -quartz to high-Tc compounds. *Phys. Rev. B*, 80(7):075102.
- [Haverkort et al., 2012] M. W. Haverkort, M. Zwierzycki, and O. K. Andersen (2012). Multiplet ligand-field theory using Wannier orbitals. *Phys. Rev. B*, 85:165113.
- [Hohenberg and Kohn, 1964] P. Hohenberg and W. Kohn (1964). Inhomogeneous Electron Gas. *Phys. Rev.*, 136:B864–B871.
- [Koelling and Harmon, 1977] D. D. Koelling and B. N. Harmon (1977). A technique for relativistic spin-polarised calculations. J. Phys. C, 10(16):3107.
- [Kohn and Sham, 1965] W. Kohn and L. J. Sham (1965). Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133–A1138.
- [Lanczos, 1952] C. Lanczos (1952). Solution of systems of linear equations by minimized iterations. J. Res. Natl. Bur. Stand, 49:33-53.
- [Laskowski and Blaha, 2010a] R. Laskowski and P. Blaha (2010a). Understanding the  $L_{2,3}$  x-ray absorption spectra of early 3d transition elements. *Phys. Rev. B*, 82:205104.
- [Laskowski and Blaha, 2010b] R. Laskowski and P. Blaha (2010b). Understanding the  $L_{2,3}$  x-ray absorption spectra of early 3d transition elements. *Phys. Rev. B*, 82:205104.
- [Martin, 2004] R. Martin (2004). *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press.
- [Methfessel and Paxton, 1989] M. Methfessel and A. T. Paxton (1989). High-precision sampling for Brillouin-zone integration in metals. *Phys. Rev. B*, 40:3616–3621.

- [Monkhorst and Pack, 1976] H. J. Monkhorst and J. D. Pack (1976). Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192.
- [Natoli et al., 1998] C. R. Natoli, C. Brouder, P. Sainctavit, J. Goulon, C. Goulon-Ginet, and A. Rogalev (1998). Calculation of X-ray natural circular dichroism. *Eur. Phys. J. B*, 4(1):1– 11.
- [Natoli et al., 1980] C. R. Natoli, D. K. Misemer, S. Doniach, and F. W. Kutzler (1980). Firstprinciples calculation of x-ray absorption-edge structure in molecular clusters. *Physical Re*view A, 22(3):1104–1108.
- [Pickard and Mauri, 2002] C. J. Pickard and F. Mauri (2002). First-Principles Theory of the EPR g Tensor in Solids: Defects in Quartz. *Phys. Rev. Lett.*, 88(8):086403.
- [Prado et al., 2013] Y. Prado, M. Arrio, F. Volatron, E. Otero, C. Moulin, P. Sainctavit, L. Catala, and T. Mallah (2013). Magnetic Anisotropy of Cyanide-Bridged Core and Core-Shell Coordination Nanoparticles Probed by X-ray Magnetic Circular Dichroism. C Chem. Eur. J., 19(21):6685-6694.
- [Rehr and Albers, 2000] J. J. Rehr and R. C. Albers (2000). Theoretical approaches to x-ray absorption fine structure. *Rev. Mod. Phys.*, 72(3):621–654.
- [Rehr et al., 2010] J. J. Rehr, J. J. Kas, F. D. Vila, M. P. Prange, and K. Jorissen (2010). Parameter-free calculations of X-ray spectra with FEFF9. *Phys. Chem. Chem. Phys.*, 12:5503-5513.
- [Rehr et al., 2005] J. J. Rehr, J. A. Soininen, and E. L. Shirley (2005). Final-state rule vs the Bethe-Salpeter equation for deep-core x-ray absorption spectra. *Physica Scripta*, 2005(T115):207.
- [Rogalev et al., 2010] A. Rogalev, J. Goulon, F. Wilhelm, and A. Bosak (2010). X-Ray Detected Optical Activity. In E. Beaurepaire, H. Bulou, F. Scheurer, and J.-P. Kappler, editors, *Magnetism and Synchrotron Radiation*, number 133 in Springer Proceedings in Physics, pages 169–190. Springer Berlin Heidelberg.
- [Schillé et al., 1994] J. Ph. Schillé, F. Bertran, M. Finazzi, Ch. Brouder, J.P. Kappler, and G. Krill (1994). 4f orbital and spin magnetism in cerium intermetallic compounds studied by magnetic circular x-ray dichroism. *Phys. Rev. B*, 55(5):2985.
- [Schwarz, 2015] K. Schwarz (2015). Innovation and Intellectual Property Rights. In M. R. Pahlavani, editor, Selected Topics in Applications of Quantum Mechanics, chapter 10.
- [Shirley et al., 2005] E. L. Shirley, J. A. Soininen, and J. J. Rehr (2005). Modeling core-hole screening in core-excitation spectroscopies. *Physica Scripta*, 2005(T115):31.
- [SPRKKR, site] SPRKKR (website). Munich SPRKKR band structure program package accessed March 06, 2017. http://ebert.cup.uni-muenchen.de/index.php?option=com\_ content&view=article&id=8&catid=4&Itemid=7.
- [Stöhr, 1999] J. Stöhr (1999). Exploring the microscopic origin of magnetic anisotropies with X-ray magnetic circular dichroism (XMCD) spectroscopy. J. Magn. Magn. Mater., 200(1-3):470 - 497.

[Taillefumier et al., 2002] M. Taillefumier, D. Cabaret, A.-M. Flank, and F. Mauri (2002). Xray Absorption Near-Edge Structure calculations with pseudopotentials. Application to Kedge in diamond and alpha-quartz. *Phys. Rev. B*, 66(19).

[Thaller, 1992] B. Thaller (1992). The Dirac Equation. Springer-Verlag Berlin Heidelberg.

- [Thole et al., 1992] B. T. Thole, P. Carra, F. Sette, and G. Laan (1992). X-ray circular dichroism as a probe of orbital magnetization. *Phys. Rev. Lett.*, 68(12):1943-1946.
- [Vogel et al., 1997] J. Vogel, A. Fontaine, V. Cros, F. Petroff, J.-P. Kappler, G. Krill, A. Rogalev, and J. Goulon (1997). Structure and magnetism of Pd in Pd/Fe multilayers studied by x-ray magnetic circular dichroism at the Pd L<sub>2.3</sub>sedges. *Phys. Rev. B*, 55:3663–3669.
- [von Barth and Grossmann, 1982] U. Barth and G. Grossmann (1982). Dynamical effects in x-ray spectra and the final-state rule. *Phys. Rev. B*, 25(8):5150–5179.
- [Šipr and Ebert, 2005] O. Šipr and H. Ebert (2005). Theoretical Fe  $L_{2,3}$  and K-edge x-ray magnetic circular dichroism spectra of free iron clusters. *Phys. Rev. B*, 72(13):134406.
- [Wilhelm et al., 2013] F. Wilhelm, R. Eloirdi, J. Rusz, R. Springell, E. Colineau, J.-C. Griveau, P. M. Oppeneer, R. Caciuffo, A. Rogalev, and G. H. Lander (2013). X-ray magnetic circular dichroism experiments and theory of transuranium Laves phase compounds. *Phys. Rev. B*, 88:024424.
- [Wood and Boring, 1978] J. H. Wood and A. M. Boring (1978). Improved Pauli Hamiltonian for local-potential problems. *Phys. Rev. B*, 18:2701–2711.

# Chapter 3

# Calculation of XMCD and XNCD: application to the 3d ferromagnetic metals and to $\alpha$ -LiIO<sub>3</sub>

# Contents

| 3.2 | XM    | CD at K-edge in Fe, Co and Ni            |
|-----|-------|--|
|     | 3.2.1 | XMCD phenomenon                          |
|     | 3.2.2 | Calculations and experiments             |
|     |       | Structures                               |
|     |       | Details about the calculations           |
|     |       | Calculated absorption cross-section      |
|     |       | Calculated XMCD                          |
|     |       | Comparison with experimental spectra     |
|     | 3.2.3 | Discussion on the calculated spectra     |
|     |       | Effect of the core-hole                  |
|     |       | Comparison with FDMNES                   |
|     |       | Discussion on the amplitude of XMCD      |
|     |       | Comparison to a quasi-particle approach  |
|     |       | Exchange-correlation functional          |
| 3.3 | XNO   | CD in $\alpha$ -LiIO <sub>3</sub>        |
|     | 3.3.1 | XNCD effect                              |
|     | 3.3.2 | Calculations of the spectra              |
|     |       | Structure of $\alpha$ -LiIO <sub>3</sub> |
|     |       | Details about the calculation            |
|     |       | Results at I $L_1$ -edge                 |
|     |       | Angular dependence                       |
|     | 3.3.3 | Effect of the core-hole on XNCD          |

|                | 3.4.1 | Sum-rules for the D-D term $\ldots \ldots \ldots$ |  |
|----------------|-------|---|--|
|                |       | XAS sum-rule  |  |
|                |       | XMCD sum-rule   |  |
|                | 3.4.2 | Sum-rules for the Q-Q term $\ldots \ldots 109$                  |  |
|                |       | XAS sum-rule  |  |
|                |       | XMCD sum-rule   |  |
|                | 3.4.3 | Sum-rules for the D-Q term $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $111$   |  |
|                |       | XAS sum-rule  |  |
|                |       | XNCD sum-rule   |  |
|                | 3.4.4 | Sum-rules for the D-SP term $\ldots \ldots 113$                        |  |
|                |       | XAS sum-rule  |  |
|                |       | XMCD sum-rule   |  |
|                |       | Total dipole XMCD sum-rule $\ldots \ldots 114$                                |  |
|                | 3.4.5 | Partial density of states and spectra in the case of Fe and Ni $\ldots$ 115   |  |
| 3.5 Conclusion |       |   |  |

# 3.1 Introduction

The application of the method presented in the previous chapter to well-known systems was required to serve several purposes. The first objective was to calculate the terms listed in section 2.2 in order to evaluate their relative contribution to absorption and circular dichroism. Another goal was to assess the capability of the DFT-based method to reproduce the complex phenomena that are XMCD and XNCD and to try to get a good understanding of its limits. Finally, the studies presented in this chapter serve to clarify the physical content of the magnetic and natural circular dichroism spectra.

The outline of the chapter is the following:

- In section 3.2, after the origin of the XMCD phenomenon is recalled, the calculated contributions to XAS and XMCD spectra at the *K*-edge for Fe, Co and Ni are detailed. Then the calculated and experimental spectra are compared and several features of the method are discussed.
- Section 3.3 is dedicated to the calculation of XAS and XNCD at the  $L_1$ -edge of iodine in LiIO<sub>3</sub>. The effect of the core-hole on XNCD is specifically discussed.
- Finally, in section 3.4 the sum-rules for all the terms are derived in order to get an insight into their physical content. This allows, in particular, to understand the importance of the D-SP term that was presented in Chapter 2.

# 3.2 XMCD at *K*-edge in Fe, Co and Ni

Early spin-polarized DFT-based calculations found accurately that the ferromagnetic state of Fe, Co and Ni was more stable than the non-magnetic one (by comparison with non-spinpolarized calculation) and gave values for the spin magnetic moment in agreement with experiment (see for example [Moruzzi et al., 1978, p.161-185]). Here, we try to compute the XMCD spectra with the DFT-based method presented in the previous chapter. XMCD is a good way to test the validity of the modelling of magnetic materials because the calculation of a spectrum is more challenging than the calculation of a magnetic moment.

# 3.2.1 XMCD phenomenon

In XMCD experiments, the absorption by a magnetic sample of right- and left-circularly polarized X-rays is successively measured and circular dichroism corresponds to the difference between them. Generally, an external magnetic field is applied parallel to the direction of propagation of the incident beam in order to maximize the signal. In a centrosymetric system, the terms that, a priori, can contribute to XMCD are those that were named D-D, Q-Q and D-SP. The selection rules for the operators constituting these terms are given in Table 3.1. At the K-edge, the D-D term and the D-SP term probe the p (l = 1) final states whereas the Q-Q term probes the d (l = 2) final states.

**Table 3.1:** Selection rules in absorption of the operators electric dipole, electric quadrupole and spin-position.

| Operator            |              | Selection rule                             | K-edge final states |
|---------------------|--------------|--|---------------------|
| electric dipole     | D            | $\Delta l = \pm 1 \ \Delta m_s = 0$        | p states            |
| electric quadrupole | $\mathbf{Q}$ | $\Delta l = \pm 2 \ \Delta m_s = 0$        | d states            |
| spin-position       | SP           | $\Delta l = \pm 1 \ \Delta m_s = \pm 1, 0$ | p states            |

Among the three terms, D-SP has the characteristic that the spin is explicitly included in the transition operator. For this reason, the XMCD related to this term has not the same origin as the two other terms.

As will be mathematically found in section 3.4, when the core state is not spin-orbit split (s state), the right-handed photons (that carry a helicity  $-\hbar$ ) probe the final states with orbital magnetic quantum number m = -1 and the left-handed photons (that carry a helicity  $+\hbar$ ) probe the final states with m = 1. For the contributions D-D and Q-Q to exhibit circular dichroism there must be a difference in occupation between the states m = 1 and m = -1. Such a difference can only happen when time-reversal symmetry is broken.<sup>1</sup> There are two main phenomena that can lead to the splitting of the states m = 1 and m = -1: the Zeeman effect and spin-orbit coupling in the presence of a spin magnetic moment.

The Zeeman effect occurs in the presence of an external magnetic field and it is related to the presence of a term  $H_Z = \frac{\mu_B}{\hbar} \mathbf{B}_0 \cdot \mathbf{L}$  in the Hamiltonian (that arises from the term  $\frac{1}{2m} (\mathbf{p} - e\mathbf{A}_0)^2$ ).<sup>2</sup> The Bohr magneton  $\mu_B = \frac{-e\hbar}{2m} \approx 5.8 \times 10^{-5} \text{ eV}.\text{T}^{-1}$  so that for  $m = \pm 1$ , the order of magnitude of the Zeeman term with an external magnetic field  $B_0 = 2T$ ,  $\langle H_Z \rangle (B_0 = 2T) \approx 0.1$  meV.

To evaluate the order of magnitude of the XMCD effect that would be induced by the Zeeman effect, we consider that the transition towards the state m = +1 and the state m = -1 yield two identical XAS spectra only shifted by  $\Delta E_Z = 0.2$  meV. The difference between the

<sup>&</sup>lt;sup>1</sup>In quantum mechanics, time-reversal is related to complex conjugation: if  $\psi(x,t)$  is solution of the Schrödinger equation  $i\hbar \frac{\partial \psi}{\partial t} = (\frac{\mathbf{p}^2}{2m} + V)\psi$ , then  $\psi^*(x, -t)$  is also a solution. If a system obeys time-reversal symmetry, the wave-functions to describe it can be chosen real and the wavefunction  $R_{n,l}(r) \sum_m c_m Y l^m(\theta, \phi)$  is real if  $c_m^* = (-1)^m c_{-m}$ .

<sup>&</sup>lt;sup>2</sup>In the radiation gauge,  $\mathbf{A}_0 = \frac{\mathbf{B}_0 \times \mathbf{r}}{2m}$  leads to  $(\mathbf{p} - e\mathbf{A}_0)^2 = \mathbf{p}^2 - e(\mathbf{B}_0 \times \mathbf{r}) \cdot \mathbf{p} + \frac{e^2}{4}(\mathbf{B}_0 \times \mathbf{r})^2 = \mathbf{p}^2 - e\mathbf{B}_0 \cdot \mathbf{L} + \frac{e^2}{4}(\mathbf{B}_0 \times \mathbf{r})^2$ . The term proportional to  $\mathbf{B}_0 \cdot \mathbf{L}$  is called the Zeeman term or paramagnetic term and the last term is called the diamagnetic term. In general the diamagnetic term is substantially smaller than the paramagnetic term (but exceptions exist).



Fig. 3.1: Diagram illustrative of the degeneracy lifting due to Zeeman effect on a p state.

two spectra (that is XMCD) is given by their derivative multiplied by  $\Delta E_Z$ . In the case of Fe, for example, the maximum of derivative of the normalized XAS is approximately  $1.6 \times 10^{-1}$  eV<sup>-1</sup> so the Zeeman effect would induce a XMCD of magnitude of  $3 \times 10^{-5}$ . This corresponds approximately to the limit of detection of the most precise existing XMCD beamlines and it can not explain the large XMCD observed in Fe. More generally, in the examples considered in this chapter, for which the XMCD amplitude reaches several  $10^{-3}$  compared to the main edge jump, the Zeeman effect cannot be the origin of the  $m=\pm 1$  splitting that induces circular dichroism. However, on beamlines that combine a very high signal to noise ratio with a strong magnetic field the measurement of a XMCD signal entirely due to Zeeman effect is possible (for example, with a 7 T magnet,  $\Delta E_Z = 0.8$  meV which can lead to a XMCD amplitude of several  $10^{-4}$  which is measurable [Sessoli et al., 2015]).



Fig. 3.2: Diagram illustrative of the degeneracy lifting due to spin-orbit coupling (SOC) on a p state.

Spin-orbit coupling (SOC) is a relativistic effect that, in a one-electron description, is often described by a model term in the Hamiltonian:  $\lambda \mathbf{S} \cdot \mathbf{L}$  where  $\lambda$  is positive and increases with atomic number Z. In the presence of spin-orbit coupling, m and  $m_s$  are no longer good quantum numbers. In the absence of magnetic field,<sup>3</sup> the appropriate quantum numbers to describe the eigenstates of the Hamiltonian are  $j, m_j, l, s$  (j and  $m_j$  are the eigenvalues of  $\mathbf{J}^2$  and  $J_z$  where  $\mathbf{J}$  is the operator for total angular momentum  $\mathbf{J} = \mathbf{L} + \mathbf{S}$ ). Spin-orbit coupling induces a degeneracy lifting on the j (see Fig.3.2 for an illustration). The relation between the states  $|lsjm_j\rangle$  and the states  $|lmsm_s\rangle$  is known and tabulated under the name of Clebsh-Gordan coefficients. From the table in Fig. 3.3 we deduce, for example, that  $|l = 1, s = \frac{1}{2}, j = \frac{1}{2}, m_j = \frac{1}{2}\rangle$  writes  $\frac{1}{\sqrt{3}}(\sqrt{2}Y_1^1|\downarrow\rangle - Y_1^0|\uparrow\rangle)$ .

In the presence of spin-orbit coupling, a spin polarization (difference between spin-up and spin-down) induces an orbital polarization (difference between m = 1 and m = -1) which is responsible for the XMCD of the terms D-D and Q-Q.

Note that, for the D-SP term, due to the expression of the transition operator  $\boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r})$ , left-circularly polarized X-rays probe the difference of the densities of state for spin-up and spindown  $(\rho^{\uparrow} - \rho^{\downarrow})$  of the *p* states, whereas right-circularly polarized X-rays probe its opposite (see

<sup>&</sup>lt;sup>3</sup>In a N-electrons framework, the spin-orbit coupling term is proportional to the sum over electrons:  $\sum_{i} \mathbf{l}_{i} \cdot \mathbf{s}_{i}$ . As  $\mathbf{L} \cdot \mathbf{S} \neq \sum_{i} \mathbf{l}_{i} \cdot \mathbf{s}_{i}$ , the l, s states are coupled (see [Cowan, 1981, p.335]).



Fig. 3.3: Table of Clebsch-Gordan coefficients for l = 1 and  $s = \frac{1}{2}$  [Group, ].

section 3.4). For this reason, XMCD for this term is not related to a difference of population between m=1 and m=-1 states but directly to a difference between spins up and spins down.

Calculation of XMCD spectra in the near-edge region requires that the density of spin and orbital polarization for each energy above the Fermi level up to several 10 eV is well described. This is a task that goes way beyond the usual magnetic calculations performed with DFT which are often limited to the determination of the spin magnetization of the ground state.

# 3.2.2 Calculations and experiments

The K-edge XMCD in metallic Fe is the first XMCD signal that was experimentally recorded in 1987 [Schütz et al., 1987]. Ever since, a large number of calculations have been reported for the XMCD spectra at Fe K-edge in bcc Fe in the near-edge region, for example in [Gotsis and Strange, 1994, Igarashi and Hirai, 1994, Brouder et al., 1996, Guo, 1996, Ebert, 1996, Fujikawa and Nagamatsu, 2003, Šipr and Ebert, 2005, Dixit and Alouani, 2016]. Calculations of XMCD at the K-edge in fcc Ni and hcp Co are fewer [Igarashi and Hirai, 1994, Igarashi and Hirai, 1996, Guo, 1996, Torchio et al., 2011] and they are not really conclusive. These calculations have been performed with various methods, often within the electric dipole approximation and with muffin-tin potentials.

In this section, we will present and discuss the results of ab-initio calculations of K-edge XMCD for the three 3d ferromagnetic metals: Fe, Co and Ni. We present the calculation of the three terms (D-D, Q-Q and D-SP) that are likely to contribute to the XMCD cross-section using the method presented in the previous chapter, which is not based on the muffin-tin approximation.

# Structures

Fe, Co and Ni crystallize respectively in body-centered cubic (*bcc*), hexagonal close-packed (*hcp*) and face-centered cubic (*fcc*, which is also a close-packed arrangement) structures. The three metals exhibit a long-range ferromagnetic ordering and a significant XMCD (with amplitude larger than  $10^{-3}$  of the edge jump) at *K*-edge.

### Details about the calculations

For the calculation, the following experimental lattice parameters were used: a = 2.87 Å for *bcc*-Fe, a = 3.52 Å for *fcc*-Ni and a = 2.51 Å and c = 4.07 Å for *hcp*-Co. The number of atoms per supercell was 64 atoms for Fe and Ni and 96 atoms for Co, so the smallest distance between the periodically repeated core-holes was 9.84 Å in Fe, 9.97 Å in Ni and 10.03 Å in Co. A Methfessel-Paxton smearing of 0.14 eV (0.01 Ry) and a  $\Gamma$ -centered 2×2×2 *k*-point

grid were used for the self-consistent charge density calculation. The spectra calculation was performed with a  $6 \times 6 \times 6$  grid for Fe and Co and a  $8 \times 8 \times 8$  grid for Ni. PBE norm-conserving pseudopotentials (parameters detailed in Appendix E) were used with cutoff energies 180 Ry for Fe, 200 Ry for Co and 190 Ry for Ni.

These calculations were performed with collinear spins along the easy axis of the crystal, that is to say [001] for *bcc*-Fe and *hcp*-Co and [111] for *fcc*-Ni [O'Handley, 1999] and the wavevector  $\mathbf{k}$  was set along the same axis.

The spectra were convolved with a Lorentzian broadening function to simulate the effect of the finite lifetime of the core-hole (constant in energy) and of the inelastic scattering of the photoelectron (additional energy-dependent broadening). The exact energy dependence of this broadening is governed by the imaginary part of the self-energy (see subsection 3.2.3) which is a many-body effect that can not be modeled within our DFT approach. In this work, we chose the values for  $\Gamma$  published in [Müller et al., 1982] and depicted in Fig. 3.4. No additional Gaussian broadening was added to account for the experimental resolution, so that the calculations can be compared with experimental data measured on various beamlines.



Fig. 3.4: Energy-dependent half energy-width  $\Gamma(E) = \Gamma_c + \Gamma_{\rm MFP}(E)$  where  $\Gamma_c$  accounts for the core-hole lifetime [Fuggle and Inglesfield, 1992] and  $\Gamma_{\rm MFP}(E)$  is related to the mean-free path  $\lambda$  of the photoelectron [Müller et al., 1982]:  $\Gamma_{\rm MFP}(E) = 2\hbar\sqrt{2E/m\lambda^{-1}(E)}$ .

The calculated spectra are normalized such that the edge jump in absorption is equal to 1.

#### Calculated absorption cross-section

The different terms of the absorption cross-section listed in section 2.2.2, except the cross term electric dipole-electric quadrupole (D-Q), were computed with two different polarization vectors corresponding to left-handed ( $\epsilon_2$ ) and right-handed ( $\epsilon_1$ ) circularly polarized light:

$$\boldsymbol{\epsilon}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\\ i\\ 0 \end{pmatrix}$$
 and  $\boldsymbol{\epsilon}_1 = \boldsymbol{\epsilon}_2^{\star} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1\\ -i\\ 0 \end{pmatrix}$ 

where the easy axis of the considered crystal defines the z axis. The cross term D-Q was not computed because it is known to be zero in centrosymmetric systems (systems for which the inversion  $\mathbf{r} \rightarrow -\mathbf{r}$  is a symmetry) which is the case for all three considered structures.

The calculation of the electric quadrupole-electric quadrupole (Q-Q) terms requires the definition of  $\hat{\mathbf{k}}$  which is the direction of the wave vector. It was set along z.

In the calculation, the zero of energy is set to the Fermi level. The occupied states, that do not contribute to the absorption cross section, are cut according to the method described in paragraph III-B of [Brouder et al., 1996].



Fig. 3.5: Calculated contributions to the *K*-edge absorption for *bcc*-Fe and *hcp*-Co and *fcc*-Ni. For each term (D-D, Q-Q, D-SP), the spectrum calculated with the polarization vector  $\epsilon_1$  is represented in solid line and the one calculated with the polarization vector  $\epsilon_2$  is represented in dashed line. Inset: zoom in the vertical direction near zero to better see the small contributions.

The calculated contributions to XAS are plotted in Fig. 3.5. The fact that magnetic circular dichroism is a very small effect at K-edge results in an impossibility to visually distinguish the spectra for left-handed and right-handed circular light without zooming in. The electric dipole-electric dipole term (D-D) is clearly prevailing. For example in Co, the maximum of the Q-Q term reaches barely 3 % of the D-D term at the same energy. The D-SP term is even one order of magnitude smaller than the Q-Q term.

We observe that the D-SP term is almost entirely circular-dichroic in the sense that it is opposite for left- and right-handed polarization (as it does not write as a square modulus, the D-SP term can be negative). If the sum of the spectra for both polarizations would be computed  $(\frac{1}{2}(\sigma(\epsilon_1) + \sigma(\epsilon_2)))$  - which is roughly the isotropic spectrum in cubic symmetry), the contribution of the D-SP term would be almost exactly zero. The equality  $\sigma_{D-SP}(\epsilon_1) = -\sigma_{D-SP}(\epsilon_2)$  is exact if the orbital parts of the wave functions can be chosen real. Indeed, if the orbital parts of  $|i\rangle$ and  $|f\rangle$  are real, as

$$\sigma_{\rm D-SP}(\boldsymbol{\epsilon}) = -\frac{2\pi^2 \alpha_0 \hbar^2 \omega^2}{mc^2} \sum_f {\rm Im}[\langle f | \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}) | i \rangle \langle i | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | f \rangle] \delta(E_f - E_i - \hbar \omega),$$

then  $\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}^{\star}) = -\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}).$ 



**Fig. 3.6:** Calculated contributions to the *K*-edge XMCD for *bcc*-Fe and *hcp*-Co and *fcc*-Ni. For each term (D-D, Q-Q, D-SP), the spectrum is calculated as the difference  $\sigma(\epsilon_2) - \sigma(\epsilon_1)$ .

# Calculated XMCD

The contributions to XMCD (Fig. 3.6) are given by the difference of the XAS spectra obtained with left- and right-handed circular polarization.

The contribution of the Q-Q term to XMCD is small but not negligible. In Fe, it presents a maximum that has almost the same energy as the maximum of the D-D term. At this energy, the Q-Q contribution reaches 17% of the D-D XMCD contribution.

As the D-SP term is almost entirely circular-dichroic, its contribution to the XMCD spectra is significant: it reaches 40% of the D-D term in amplitude. This can also be understood considering the sum-rules (see Section 3.4): in the XMCD cross-section, the D-SP term probes the spin polarization of the *p* states whereas the D-D term probes their orbital polarization. In [Igarashi and Hirai, 1996] the 4*p* orbital magnetic moment in Co, Fe and Ni is evaluated to a few  $10^{-4}\mu_B$  (Fe:  $5 \times 10^{-4}\mu_B$ , Co:  $16 \times 10^{-4}\mu_B$ , Ni:  $6 \times 10^{-4}\mu_B$ ) and in [Chen et al., 1995] the 4*p* spin magnetic moment in Fe and Co is evaluated to several  $10^{-2}\mu_B$  (Fe:  $5 \times 10^{-2}\mu_B$ , Co:  $6 \times 10^{-2}\mu_B$ ) in the opposite direction. This difference in orders of magnitude compensates for the smallness of the prefactor ( $\hbar\omega/4mc^2$ ) of the D-SP term.

To compute XMCD at the K-edge, it is required to compute all three contributions D-D, Q-Q and D-SP as none of them is negligible. For comparing the result of the calculations with experimental spectra, sums of the contributions presented above are calculated.<sup>4</sup>

 $<sup>^{4}</sup>$  The D-SP contribution is negligible in XAS so, in practice, it can be omitted. Also, as XMCD is a very small

A rigid shift in energy is applied to all the calculated spectra to make the maxima of the calculated XAS correspond to the maxima of the experimental spectra.



# Comparison with experimental spectra

Fig. 3.7: Comparison between the experimental XAS and XMCD spectra for Fe, Co and Ni and the total calculated spectra. The experimental spectra were recorded on ODE beamline (SOLEIL). They are corrected for the rate of circular polarization  $P_c = 0.7$ . For the calculation, the wave vector and the magnetization axis were set to the easy axis of the crystals. The calculated spectra were shifted in energy by 7113 eV for Fe, 7712 eV for Co and 8333.5 eV for Ni.

The experimental XAS and XMCD spectra for Fe, Co and Ni polycrystalline metallic foils

effect the absorption of left- and right-circular polarization are almost equal so that  $\sigma_{D-D}(\epsilon_1) + \sigma_{Q-Q}(\epsilon_1) + \sigma_{D-SP}(\epsilon_1) \approx \frac{1}{2}(\sigma_{D-D}(\epsilon_1) + \sigma_{Q-Q}(\epsilon_1) + \sigma_{D-SP}(\epsilon_1) + \sigma_{D-D}(\epsilon_2) + \sigma_{Q-Q}(\epsilon_2) + \sigma_{D-SP}(\epsilon_2))$ . In other words, the average of the spectra for both polarizations is almost identical to the spectrum corresponding to one polarization.

were recorded on ODE beamline at SOLEIL (see Section 4.2) with a Si(311) crystal polychromator which corresponds to an energy resolution  $\Delta E/E \approx 4.10^{-5}$ . The tabulated edge energy for the elements Fe, Co and Ni are:

| Element       | K-edge              |
|---------------|---------------------|
| $\mathrm{Fe}$ | $7.1120~{\rm keV}$  |
| Co            | $7.7089~{\rm keV}$  |
| Ni            | $8.3328~{\rm keV}.$ |

The foils were tilted by  $45^{\circ}$  compared to the direction of propagation of the beam which was also the direction of the 2.1T applied magnetic field. The experimental spectra, as the calculated spectra, are normalized such that the absorption edge jump far from the edge is equal to 1.

The calculated and experimental spectra are depicted in Fig. 3.7. The agreement between the calculated and the experimental XAS spectra (top panels Fig. 3.7) is fair: all the features of the experimental spectra are reproduced by the calculations.

The shape of XMCD (bottom panels in Fig. 3.7) near the edge is also well reproduced: a positive peak followed by a negative peak for Fe and a main negative peak for Co and Ni. All the secondary peaks in Fe also exhibit a good agreement between experiment and calculation. For Co and Ni, however, some secondary peaks that appear in the calculation do not seem to correspond to experimental features. In all three compounds, a positive peak around 10 eV above the edge (at 7122 eV for Fe, 7722 eV for Co and 8340 eV for Ni) appears in the calculation that overestimates an experimental feature. Another discrepancy between the calculated and experimental spectra is the amplitude of the XMCD effect: the calculated XMCD amplitude is too large for Fe and too small for Ni.

In order to narrow down the possible causes of discrepancies between the calculated and the experimental spectra, we discuss several elements of the method in the next subsection.

# 3.2.3 Discussion on the calculated spectra

#### Effect of the core-hole

To model the absorption process, the calculation presented above was performed in the presence of a core-hole and the electronic response to this core-hole was determined self-consistently.

In order to have an idea of the effect of the core-hole on the calculated spectra, the comparison of the spectra with and without core-hole is shown in Fig. 3.8. For the case without core-hole, the self-consistent calculation is performed without a core-hole but all the parameters of the spectra calculation remain the same. In particular, the core-hole lifetime broadening is still included. For a given element, the same rigid shift in energy was applied to all the calculated spectra. The multiplication factor to obtain an edge-jump equal to 1, on the other hand, is different for each spectra (e.g. for Co: 820 without core hole and 950 with core-hole).

For Fe and Co K-edges, the presence of the core-hole has a weak effect on the XAS spectra which means that the core-hole is efficiently screened by the electrons. Yet, the XMCD is visibly impacted by the presence of the core-hole, in particular in the case of Fe. On the right panel of Fig. 3.8, we see that the Q-Q contribution to XMCD, that probes the orbital polarization of the empty d states is shifted to lower energy in the presence of the core-hole (by about 0.5 - 0.7 eV). On the other hand, the energy positions of the D-D contribution in the case of Co and of the D-SP contributions in both cases, are almost unmoved by the presence of the core-hole (for the D-D contribution of Fe, the situation is not so clear). As these contributions



Fig. 3.8: Comparison of the calculated XAS and XMCD spectra obtained in the presence of a 1s core-hole (solid line) and of the spectra obtained without core-hole (dashed line) at Fe and Co K-edge. For XMCD, the experimental spectra multiplied by 1600 for Fe and 1100 for Co are plotted in light gray whereas the calculated spectra are multiplied by 1000. On the right-hand side: detail of the different terms composing the XMCD spectra.

probe the properties of empty p states, this is in agreement with the idea that the core-hole is more screened for the p states than for the d states. We see also that the shape of the D-D contribution is quite different in the presence or in the absence of the core-hole contrary to the D-SP contribution that is almost unaffected. It means, that the orbital polarization is more sensitive than the spin polarization to the core-hole-induced perturbation. For this reason, it is crucial to include the core-hole in the calculations for XMCD even in cases where the XAS seems well described in the absence of core-hole.

#### **Comparison with FDMNES**

To check possible numerical problems, in particular for the newly discovered D-SP term, Yves Joly performed the same calculations using the FDMNES code [Bunău and Joly, 2009] in which he implemented this term. Preliminary results at Ni *K*-edge are depicted in Fig.3.9. In FDMNES, the calculation is performed in real space for a cluster of atoms and the potential is not pseudized. Despite a significant difference in method, the result is very similar to the one obtained with XSpectra, in particular for the shape and relative amplitude of the D-SP contribution.

This calculation allowed us to assess the possible impact of several approximations of our method, detailed below, that were otherwise difficult to check.

One questionable approximation in our method is the diagonal spin-orbit coupling approximation (see p. 54). The fact that with an implementation that includes full spin-orbit coupling, the result is consistent with ours implies that this approximation is reasonable in the cases considered here.



Fig. 3.9: Preliminary calculation for Ni *K*-edge with FDMNES. Energy is relative to Fermi level. Figure provided by Yves Joly.

Another question concerned the spin-polarization of the core state: in the 1s state there is a splitting between the spin up and the spin down band that has been evaluated by Yves Joly to approximatively 0.015 eV in Fe. Calculations with FDMNES showed that the influence of this splitting on the spectra is negligible at the K-edge in the considered 3d metals.

Finally, the terms up to the octupole approximation  $(e^{i\mathbf{k}\cdot\mathbf{r}} = 1 + i\mathbf{k}\cdot\mathbf{r} - \frac{(\mathbf{k}\cdot\mathbf{r})^2}{2})$  were implemented in FDMNES. It allowed Yves Joly to check that the electric octupole term was negligible in accordance with the assumption made in our derivation of the cross-section where we neglected all  $\mathbf{k}^2$  terms.

#### Discussion on the amplitude of XMCD

From the bottom parts of Fig. 3.7, we see that the amplitude of the XMCD effect is not well reproduced by the calculation. The amplitude of the calculated XMCD for Fe is approximately 170% the amplitude of the experimental spectrum, for Ni it is 74%. We show in the following that the fact that the experimental conditions are quite different from the case treated in the calculation cannot explain these differences.

In the calculations, the rate of circular polarization of the light is  $P_c = 1$ . Experimentally, the circular polarized light is obtained by selecting a portion of beam below the orbit plane of the radiation emitted by a bending magnet. The rate of circular polarization depends on the position and on the size of the slit used to select this portion.  $P_c$  is then modified by the optics following this selector and this modification is difficult to quantify. For this reason, the experimental  $P_c$  is not known exactly. It is however evaluated to approximately 0.7 in the considered energy range. The experimental XMCD spectra plotted Fig. 3.7 are already multiplied by 1/0.7 to account for  $P_c$ . Despite this, the calculated amplitude of XMCD for Fe



is largely overestimated whereas for Ni it is underestimated.

Fig. 3.10: Top: crystal structures showing easy and hard magnetization directions for (a) Fe, (b) Ni and (c) Co. Bottom: Respective magnetization curves. Figure extracted from [O'Handley, 1999]. For cobalt, the field required to reach saturation in the basal plane is more than one order of magnitude larger than along the easy axis (whereas anisotropy in iron and nickel is very weak).

Besides the rate of circular polarization, another difference between the experimental conditions and the calculated case is the fact that experiments are performed with polycrystalline metal foils under an applied magnetic field whereas calculations are performed for a single crystal with full 3*d* spin polarization. According to the magnetization curves shown in Fig. 3.10, saturation is fully reached with the 2.1 T (21 kOe) applied magnetic field in the three considered ferromagnetic metals. The fact that the measured samples are foils could result in a greater difficulty to fully polarize them perpendicularly (along the thin direction). That is one of the reasons why experimentally the foils are tilted at 45 ° with respect to the direction of the magnetic field that corresponds to the direction of the beam (the other reason is to increase the path length of the beam through the sample). Given the order of magnitude of the field required to reach saturation in iron and nickel (respectively 0.06 T and 0.03 T) it seems unlikely that the material is not fully polarized under a 2.1 T applied field. For cobalt, the saturating field is much higher (1T) than in Fe and Ni. We observed experimentally that the XMCD amplitude is 16 % larger with a 2.1 T field than with a 1.3 T field (see Fig. 3.11, left). It seems, however, reasonable to consider that 2.1 T is enough to reach saturation.

Another specificity of Co compared to Fe and Ni is that it does not crystallize in a cubic symmetry (see Table 3.2). For this reason, the absorption of linear polarized light by a crystal of Co depends on the angle between the polarization vector and the symmetry axis of the crystal. In cobalt, the angular dependence of the electric dipolar absorption of linear polarization is dichroic. The spectra calculated with  $\epsilon_1$  or  $\epsilon_2$  correspond (up to circular dichroism which is very small) to  $\sigma_{\perp}$  (see section 3.4 p.106). The isotropic spectrum (spectrum measured on a powder) is given by  $\sigma = \frac{1}{3}(\sigma_{\parallel} + 2\sigma_{\perp})$ . For this reason, we calculated the spectra for two directions of magnetization: along the axis [001] and along the axis [100],<sup>5</sup> with **k** parallel to

<sup>&</sup>lt;sup>5</sup>In terms of Bravais Miller indices, the Miller direction [001] corresponds to [0001] and [100] to  $[\overline{2}110]$ .



Fig. 3.11: XMCD spectra recorded on ODE beamline (SOLEIL) at Co K-edge in a Co foil. Left: spectra acquired with two different applied magnetic field at room temperature on a foil positionned at 45°. Right: spectra acquired at room temperature and T=4K on a foil positioned perpendicular to the beam with a 1.3T field.

**Table 3.2:** Space groups and point groups of the *bcc*-Fe, *hcp*-Co and *fcc*-Ni crystals. The point groups are given with their short Hermann-Mauguin symbols and the point groups with their Schoenflies symbols. Angular dependence of electric dipole absorption were determined in the case of a *linear* polarization in [Brouder, 1990]. For cubic symmetries, the absorption cross section is isotropic whereas in the hexagonal case it is dichroic ( $\sigma(\epsilon) = \sigma_{\parallel} \sin^2 \theta + \sigma_{\perp} \cos^2 \theta$  where  $\theta$  is the angle between  $\epsilon$  and the 6-fold rotation axis.).

|                | Space group                               | Point group of the crystal | Angular dependence of elect-           |
|----------------|---|----------------------------|--|
|                |   |                            | ric dipolar absorption [Brouder, 1990] |
| $bcc	ext{-}Fe$ | $\mathrm{Im}\overline{3}\mathrm{m}$ (229) | $O_h$                      | Isotropy                               |
| <i>hcp</i> -Co | $P6_3/mmc$ (194)                          | $D_{6h}$                   | $\operatorname{Dichroism}$             |
| fcc-Ni         | $Fm\overline{3}m$ (225)                   | $O_h$                      | Isotropy                               |

the magnetization axis. In practice, the crystal cell used in input of the calculation is rotated so that the axis [001] or [100] corresponds to the z axis. In both cases, the total magnetization obtained at the end of the calculation is 1.665  $\mu_B$  per atom. The results are depicted in Fig. 3.12. As expected, the XAS and XMCD spectra are not the same for both directions. This angular dependence affects the positions of the peaks but it impacts very weakly the XMCD amplitude.

Another difference is that the calculations are performed for the ground state (at T=0K) whereas the experimental spectra are acquired at room temperature. We observed experimentally that the effect of temperature on the amplitude of the experimental XMCD spectra at Co K-edge is very weak (7 % variation - see Fig. 3.11 Right).

Finally, none of the above mentioned differences between the conditions in which the calculation is performed and the experimental conditions seems to explain the differences of the XMCD amplitude between our calculations and experiments. Note that the broadening applied to the spectra, discussed below, can also influence the calculated amplitude.



Fig. 3.12: Calculated XAS and XMCD spectra at the Co K-edge without core-hole for two directions of magnetization: along the axis [001] and along the axis [100] in both cases with **k** parallel to the magnetization axis (in practice, for the [100] axis, crystal rotated by 90° is used and the calculation is again performed with z as quantization axis and  $\epsilon_1$  and  $\epsilon_2$  as polarization vectors).

### Comparison to a quasi-particle approach

Kohn-Sham density functional theory describes the ground-state properties of an interactingelectron system. The final states in X-ray absorption are excited states so, in principle, they would require the use of a many-body method to treat them. DFT is still widely used to compute XAS at K-edge because it is successful to reproduce the experimental spectra in a large number of cases [Cabaret et al., 2010, Bordage et al., 2010, Ducher et al., 2016, for example].

To be more accurate and still keep one electron equations, the *self-energy*  $\Sigma$  (complex quantity) can be introduced to include many-body effects and losses during the propagation of the photoelectron. Within the quasi-particle theory, the final states  $|f\rangle$  are eigenstates of the equation [Rehr and Ankudinov, 2005]:

$$\left[\frac{p^2}{2m} + V'_{coul} + \Sigma(E)\right]|f\rangle = E|f\rangle$$

where  $V'_{coul}$  is the Coulomb potential in the presence of the core-hole. This equation is formally similar to a DFT Schrödinger equation but contrary to the exchange-correlation potential of DFT, the self-energy is non local and energy dependent. The self-energy can be computed within the GW approximation as, for example, in [Kas et al., 2007] however the calculations are time consuming [Rehr and Albers, 2000] (in practice, it is more convenient to use an analytic representation of Hedin-Lunquvist self-energy [Mustre de Leon et al., 1991]).

Note that this approach does not include all many-body effects. For example, the treatment of the electron-hole interaction would require the use of two-body operators. Bethe-Salpeter equation, that links the interacting polarization propagator to the independent-particle propagator is an appropriate tool for that. Also the losses due to the excitations in the medium that arise from the sudden creation of the core-hole, which are called intrinsic losses in [Kas et al., 2007], are not taken into account.

As illustrated by the case of Cu that was studied in [Kas et al., 2007] the real and imaginary part of the self-energy  $\Sigma(E)$  vary by several eV over a XAS spectrum. The imaginary part of the self-energy is negative (it describes losses).

The fact that the real part of the self-energy depends on energy results for the XAS spectra



Fig. 3.13: Quasiparticle self-energy for Cu. Figure taken from [Kas et al., 2007].

in a deformation of the energy axis: the energy-dependent shift is given by  $\Delta(E) = \operatorname{Re}[\Sigma(E) - \Sigma(E_f)]$  [Kas et al., 2009]. From Fig. 3.13, we see that it is positive (which means that, in DFT calculations, the peaks are calculated at an energy which is too low, namely the energy axis is compressed) and that, overall, it increases with energy. In the XANES region, the variation of  $\operatorname{Re}[\Sigma(E)]$  is small and it is the reason why it is usually not taken into account. In the EXAFS region, however, it is important to rescale the energy axis in order to obtain accurate peak positions. As  $\operatorname{Re}[\Sigma(E)]$  is nearly linear in this region, a simplified solution to apply this correction is a phenomenological linear rescaling as proposed in [Materlik et al., 1983] and discussed in [Mustre de Leon et al., 1991].

The energy dependence of the imaginary part of the self-energy results in an energy dependent Lorentzian broadening of half width at half maximum  $\Gamma(E) = -\text{Im}[\Sigma(E)]$  [Kas et al., 2009]. In our method it is taken into account by applying the broadening depicted in Fig. 3.4. The convolution of the spectra with a Lorentzian function with an energy dependent width results in the modification of the amplitude for the calculated peaks, while their areas remain the same. The fact that  $\Gamma$  increases with energy and that, a few 10 eV above the edge, it reaches several eV, is a life-saver from a numerical point of view. Indeed, as illustrated in the previous chapter, if the calculation parameters are not completely converged, spurious peaks appear at high energies. These spurious peaks all disappear when applying the energy-dependent broadening.

As a conclusion, we included by hand the main effect of energy dependence of the selfenergy which is the energy-dependent broadening. Thus, our method gives very similar results compared to a quasi-particle approach. Contrary to some existing studies in which  $\Gamma(E)$  is calculated [Sainctavit et al., 1989, Mustre de Leon et al., 1991], here  $\Gamma(E)$  is a parameter of the calculation. We must either arbitrarily define its shape (in a reasonable manner, that is to say with a curve similar to the one in Fig. 3.13) or use existing calculations for the system under study. Such calculations are quite rare so the first solution is almost always used. Our method is, therefore, not entirely parameter-free.

#### **Exchange-correlation functional**

Density functional theory calculations would be exact for the ground state energy if the exchange-correlation functional  $E_{xc}[n]$  introduced in section 2.3.1 was known exactly. The interest of separating out the independent-particle kinetic energy and the Hartree terms from the energy of the interacting system is that the remaining  $E_{xc}[n]$  can reasonably be approximated as a local or nearly local functional of the density.

In the calculations presented above we use the form of the functional proposed by Perdew, Burken and Enzerhof [Perdew et al., 1996] called PBE. This functional belongs to the family of Generalized-Gradient Approximations (GGA). It means that it writes as an integral over space of a function of the local density and of its gradient:

$$E_{xc}^{\text{GGA}}[n_{\uparrow}, n_{\downarrow}] = \int \,\mathrm{d}\mathbf{r} \, n(\mathbf{r}) \epsilon_{xc}^{\text{PBE}}(n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r}), \nabla n^{\uparrow}(\mathbf{r}), \nabla n^{\downarrow}(\mathbf{r}))$$

where  $n = n^{\uparrow} + n^{\downarrow}$ .

This is an improved functional compared to LSDA (local spin density approximation) that was already introduced in the paper by Kohn and Sham [Kohn and Sham, 1965]:

$$E_{xc}^{\text{LSDA}}[n_{\uparrow}, n_{\downarrow}] = \int \,\mathrm{d}\mathbf{r} \, n(\mathbf{r}) \epsilon_{xc}^{\text{unif}}(n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r}))$$

where  $\epsilon_{xc}^{\text{unif}}$  is the exchange-correlation energy per electron of a uniform electron gas. The analytic form of the exchange part of this energy is known [Martin, 2004, p.106] and the correlation part has been calculated with Monte-Carlo methods so that  $\epsilon_{xc}^{\text{unif}}(n^{\uparrow}, n^{\downarrow})$  is considered a known function. In the seminal paper on PBE, the enhancement factor  $F_{xc}$  defined by  $\epsilon_{xc}^{\text{PBE}}(n^{\uparrow}, n^{\downarrow}, \nabla n^{\uparrow}, \nabla n^{\downarrow}) = F_{xc}(n^{\uparrow}, n^{\downarrow}, \nabla n^{\uparrow}, \nabla n^{\downarrow}) \epsilon_{xc}^{\text{unif}}(n^{\uparrow}, n^{\downarrow})$  is depicted and not directly  $\epsilon_{xc}^{\text{PBE}}$ .



Fig. 3.14: Total calculated *bcc*-Fe *K*-edge XMCD spectra without core hole using PBE and LDA functionals (all other technical parameters identical). Here, the broadening was taken constant (0.8 eV) along the whole energy range.

In practice, the form of the functional used in our calculation is chosen when generating the pseudopotential files (the same functional must be used for all atoms in the structure). Fig. 3.14 shows a comparison between the spectra obtained with PBE and LSDA exchange-correlation functional in a unit cell of *bcc*-Fe. All other parameters in the pseudopotential generation, in the scf calculation and in the spectra calculations are the same. We observe that the choice of the functional does not really affect the position or sign of the peaks that appear in the XMCD spectra but it affects their amplitude. In particular, the positive peak around 10 eV above the edge, which is overestimated in our calculation, compared to experiment (see Fig.3.7), is even

more enhanced with LSDA. We can therefore suggest that the overestimation of this peak is linked to approximations made in the exchange-correlation functional.

XMCD calculation requires a very good modeling of magnetism and this study illustrates that the usual functionals used in solid-state DFT are not sufficient to precisely describe such a fine and complex phenomenon.

Finally, all the cross-checks presented in this section allow to narrow down the origin of the failure of our method to reproduce detailed features of the XMCD spectra - even if their overall shape is correctly described - to *inherent limitations* of Kohn-Sham DFT. In the next section, we will see that XNCD, on the other hand, is well described. This illustrates that efforts should be pursued to improve the description of magnetism in matter.

# 3.3 XNCD in $\alpha$ -LiIO<sub>3</sub>

# 3.3.1 XNCD effect

In non-centrosymmetric crystals, the breaking of inversion symmetry permits the existence of X-ray Natural Circular Dichroism. In the domain of molecular magnetism, a renewed interest for this technique has recently grown [Train et al., 2008, Sessoli et al., 2015] with the emergence of new materials that are both chiral and magnetic.

In the X-ray range, where magnetic dipole transitions are negligible, a single mechanism largely dominates in the natural circular dichroism phenomenon: the interference of electric dipole and electric quadrupole transitions [Rogalev et al., 2010, Natoli et al., 1998]. At the  $L_1$ edge, XNCD is an element-specific measurement of the mixing between p and d states (see subsection 3.4.3). It is a parity-odd property in the sense that no such mixing can exist in centrosymmetric systems. There are 13 crystal classes that may exhibit XNCD associated with the term D-Q, they are listed in [Rogalev et al., 2010, Table 6.1] and are boxed among the other crystal classes in Appendix C of this thesis. The point group to be considered is the point group resulting from the space group of the crystal structure. Note that there exist noncentrosymetric point groups (no inversion center) that are not chiral (a chiral point group has no rotoinversion). Among these groups, several may exhibit XNCD (see Appendix C) so that XNCD can exist in some non-chiral materials.

The first XNCD spectra were recorded at the ESRF ten years after G. Schütz et al. recorded the first XMCD spectrum. In 1997, XNCD was measured in a stereogenic organometallic complex [Alagna et al., 1998] and in the inorganic crystal  $\alpha$ -LiIO<sub>3</sub> [Goulon et al., 1998].  $\alpha$ -LiIO<sub>3</sub> belongs to the crystal class "6" so that XNCD associated to the D-Q term is possible. The average of the term D-Q over all directions is zero, so XNCD must be measured on a single crystal and this is what has been done for the experiments reported in [Goulon et al., 1998].

# 3.3.2 Calculations of the spectra

Previous calculations [Natoli et al., 1998, Goulon et al., 1998, Ankudinov and Rehr, 2000] were able to reproduce the overall peak positions and intensities of XNCD (see Fig. 3.15 as an example) by computing the interference of electric dipole and electric quadrupole transitions. The agreement is however not entirely satisfactory for the absorption spectra. This discrepancy has been attributed to the use of muffin-tin potentials [Ankudinov and Rehr, 2000].

The method presented in the previous chapter, that does not rely on the muffin-tin approximation, was applied to compute the XAS and XNCD spectra for  $\alpha$ -LiIO<sub>3</sub>. Since  $\alpha$ -LiIO<sub>3</sub>



Fig. 3.15: Figure from [Natoli et al., 1998]. Experimental (solid line) and calculated within a multiple-scattering formalism (dotted line) absorption and XNCD spectra at the  $L_1$ -edge of iodine in LiIO<sub>3</sub>. The XNCD spectra were multiplied by 50.

is diamagnetic, the calculation is not spin-polarized and the Hamiltonian without spin-orbit coupling is used for the self-consistent field electron relaxation.

# Structure of $\alpha$ -LiIO<sub>3</sub>



Fig. 3.16: Top: Hexagonal  $\alpha$ -LiIO<sub>3</sub> unit cell for a)  $\Delta$  and b)  $\Lambda$  enantiomers [Goulon et al., 1998]. Bottom: top view of the cells (projection on (001)).

The  $\Delta$  enantiomer of  $\alpha$ -LiIO<sub>3</sub> crystallizes [Svensson et al., 1983] in the hexagonal space group  $P6_3$  (173). The lattice parameters are a = 5.48 Å and c = 5.17 Å. The atomic positions [Goulon et al., 1998] are Li 2(a) (0,0,0.076), I 2(b) (1/3,2/3,0) and O 6(c) (0.247,0.342,0.838). The  $\Lambda$  enantiomer is the mirror image of the  $\Delta$  one (see Fig. 3.16) and it belongs to the same space group. As was done in [Stadnicka et al., 1985], it is possible to draw a helix of alternating I and O atoms. In Fig. 3.16 the sense of the helices is denoted by arrows with the convention that the helices come toward the reader (to make the construction, we simply sort the four concerned atoms by order of z coordinates and draw an arrow from the smallest z to the largest z).

The absolute configuration of a  $\text{LiIO}_3$  crystal can be resolved using anomalous scattering effect [Stadnicka et al., 1985]. It was not done for the crystal studied in [Goulon et al., 1998] so the determination of which enantiomer was used must be done by calculating the XNCD spectra.
#### Details about the calculation

The experimental structure was used in the calculation. A  $2\times2\times2$  supercell (80 atoms) with one atom containing a core-hole (2s for  $L_1$ -edge or 1s for K-edge) was used so that the smallest distance between a core-hole and its periodic image is 10.344 Å. We used Gamma-centered k-point grids,  $3\times3\times3$  for the self-consistent charge density calculation and  $9\times9\times9$  for the spectra calculation. A constant Lorentzian broadening, with full width at half maximum set to the core-hole lifetime broadening 3.46 eV [Fuggle and Inglesfield, 1992], was applied.

Calculated spectra were normalized such that the edge jump in absorption is equal to 1 as it is the case for experimental spectra.

During the calculation of the spectra the origin of energy is set to the energy of the highest occupied level that we note  $E_{ho}$ . For the spectra to be compared with experiment, a rigid shift in energy was applied to all the calculated spectra to make the maxima of the calculated XAS correspond to the maxima of the experimental spectra. In the plots for I  $L_1$ -edge, the origin of energy  $E_0$  is therefore the one chosen in [Goulon et al., 1998].



#### Results at I $L_1$ -edge

Fig. 3.17: Calculated contributions to the XAS at I  $L_1$ -edge in  $\alpha$ -LiIO<sub>3</sub>. For the D-Q term, the spectrum calculated with the polarization vector  $\epsilon_1$  is in solid line and the one calculated with  $\epsilon_2$  in dashed line (for the terms D-D and Q-Q, there is only one spectrum because these terms do not exhibit any natural circular dichroism). Inset: zoom in the vertical direction near zero.

The calculated contributions of the D-D, Q-Q and D-Q terms (as the calculation is not spin-polarized, there is no contribution from the D-SP term) to the absorption at I  $L_1$ -edge in  $\alpha$ -LiIO<sub>3</sub> are shown in Fig. 3.17. The absorption is largely dominated by the D-D term.

For the terms D-D and Q-Q, the spectra obtained with right-circular polarization ( $\epsilon_1$ ) and the spectra obtained with left-circular polarization ( $\epsilon_2$ ) are exactly the same. In other words, in accordance with theory, these terms do not contribute to XNCD. The D-Q term, on the other hand is entirely circular-dichroic. This was also expected because  $|i\rangle$  and  $|f\rangle$  can be chosen real so that:

$$\sigma_{\rm D-Q} = -4\pi^2 \alpha_0 \hbar \omega \sum_f {\rm Im}[\langle f | (\mathbf{k} \cdot \mathbf{r}) (\boldsymbol{\epsilon} \cdot \mathbf{r}) | i \rangle \langle i | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | f \rangle] \delta(E_f - E_i - \hbar \omega)$$

obeys  $\sigma_{\mathrm{D-Q}}(\boldsymbol{\epsilon}^{\star}) = -\sigma_{\mathrm{D-Q}}(\boldsymbol{\epsilon}).$ 

The calculated XAS is mostly composed of the D-D contribution and the XNCD is entirely due to the D-Q term. In [Goulon et al., 1998]  $\sigma^{XNCD} = \sigma^R - \sigma^L$  where  $\sigma^R$  and  $\sigma^L$  are the

spectra obtained with right- and left-circularly polarized X-rays so the spectra that we compare with experiment are:



Fig. 3.18: Comparison of experimental [Goulon et al., 1998] and calculated XAS and XNCD spectra at the  $L_1$ -edge of iodine in LiIO<sub>3</sub>. The calculation was performed for both enantiomers with **k** parallel to **c**. Dashed black line: calculation for the  $\Lambda$  enantiomer. Solid black line: calculation for  $\Delta$  enantiomer. The calculated XAS spectrum is the same for both enantiomers.

As illustrated in Fig. 3.18, both the calculated XAS and XNCD spectra at the  $L_1$ -edge of iodine are in good agreement with experiment. However, the amplitude of the calculated XNCD is  $4 \times 10^{-2}$  of the edge jump (set to 1) while the amplitude of the experimental spectra is  $6 \times 10^{-2}$ . Such an underestimation of the XNCD spectrum by the calculation was observed in [Natoli et al., 1998] within a multiple-scattering approach (though to a lesser extent). The sign of the calculated XNCD for  $\Delta$  enantiomer is also in agreement with the result of [Natoli et al., 1998]. Finally, the main difference between the result obtained with our pseudopotentials-based DFT method and those obtained with a multiple-scattering muffin-tin calculation in [Natoli et al., 1998] is that the XAS spectrum is better described in our calculation.

In Fig. 3.18, we see that the XNCD spectra for both enantiomers are opposite. Indeed, it has the same effect for XNCD to change an enantiomer for the other ( $\Delta \leftrightarrow \Lambda$ ) as for XMCD to change the sign of the magnetic field ( $\mathbf{B} \leftrightarrow -\mathbf{B}$ ).

<sup>&</sup>lt;sup>6</sup>Note that this is exactly equal to the average of the XAS spectra for left- and right-circular polarization:  $\frac{1}{2}(\sigma_{D-D}(\boldsymbol{\epsilon}_1) + \sigma_{Q-Q}(\boldsymbol{\epsilon}_1) + \sigma_{D-Q}(\boldsymbol{\epsilon}_1) + \sigma_{D-D}(\boldsymbol{\epsilon}_2) + \sigma_{Q-Q}(\boldsymbol{\epsilon}_2) + \sigma_{D-Q}(\boldsymbol{\epsilon}_2)) = \sigma_{D-D}(\boldsymbol{\epsilon}_1) + \sigma_{Q-Q}(\boldsymbol{\epsilon}_1)$ because  $\sigma_{D-D}(\boldsymbol{\epsilon}_1) = \sigma_{D-D}(\boldsymbol{\epsilon}_2), \ \sigma_{Q-Q}(\boldsymbol{\epsilon}_1) = \sigma_{Q-Q}(\boldsymbol{\epsilon}_2)$  and  $\sigma_{D-Q}(\boldsymbol{\epsilon}_1) = -\sigma_{D-Q}(\boldsymbol{\epsilon}_2)$ .

#### Angular dependence



Fig. 3.19: Angular dependence of the calculated XAS (on the left-hand side) and XNCD (on the right-hand side) at the I  $L_1$ -edge in LiIO<sub>3</sub>. In inset: XNCD amplitude as a function of the angle following the law  $3\cos^2\theta - 1$  where  $\theta$  is the angle between **c** and the incident wave-vector **k**.

The angular dependence of the calculated XAS and XNCD spectra is shown in Fig. 3.19 and the amplitude of the XNCD is plotted in inset as a function of  $\theta$ , the angle between **k** and the **c**-axis of the crystal. This amplitude varies as  $3\cos^2\theta - 1$ . This dependence is consistent with the formula derived in [Natoli et al., 1998] for point group  $C_6$  (point group of the space group of the crystal). The amplitude of XNCD is maximal in the case **k**  $\parallel$  **c** ( $\theta = 0$ ) and decreases to reach zero when  $\theta = \cos^{-1}(1/\sqrt{3}) = 54.74^{\circ}$ . This angle is sometimes named the magic angle (note that for XNCD it corresponds to the angle between **k** and **c** whereas for XNLD the magic angle corresponds to an angle between  $\epsilon$  and **c**). If the angle between **k** and **c** is further increased, the XNCD amplitude rises again but the sign of the signal is opposite compared to the case  $\theta < \cos^{-1}(1/\sqrt{3})$ .

Note that, as  $\epsilon_1$  and  $\epsilon_2$  are kept perpendicular to **k** and  $C_6$  is a dichroic point group for the electric dipole transition [Brouder, 1990], the XAS spectra also present an angular dependence. Nevertheless, it does not prevent a direct comparison of the amplitude of the XNCD spectra because the edge jump of the isotropic spectra remains unchanged.

#### 3.3.3 Effect of the core-hole on XNCD

It is known that the presence of the core-hole affects differently the p and d states. As XNCD probes the mixed p-d states, the presence of the core-hole is expected to lead to important changes in the XNCD spectra. In Fig. 3.20 the contributions to the XAS at the  $L_1$ -edge of iodine with and without core-hole are shown. Note that, as  $\sigma_{D-Q}$  is entirely circular dichroic,  $\sigma^{XNCD} = \sigma_{D-Q}(\epsilon_1) - \sigma_{D-Q}(\epsilon_2) = 2\sigma_{D-Q}(\epsilon_1)$ .

The D-D contribution probes the empty p states, the Q-Q contribution probes the empty d states and the D-Q contribution probes the mixed p - d states. Yet, Fig. 3.20 illustrates that the D-Q contribution is not some product of the D-D and of the Q-Q contribution. The maximum of the XNCD effect does not correspond in energy to the main peak of the XAS (white line) but it lies approximately 15 eV higher in energy.

A comparison of the two plots from Fig. 3.20 shows that, as usual, in the calculation without core-hole the white line is less intense than with a core-hole. We see also that the Q-Q contribution is larger in the absence of a core-hole. Finally, the XNCD is totally different: with core-hole, the main peak is positive whereas it is negative without core-hole.



Fig. 3.20: Contributions to the absorption of right-handed polarized light ( $\epsilon_1$ ) at the  $L_1$ -edge of iodine in the  $\Delta$  enantiomer of  $\alpha$ -LiIO<sub>3</sub>. On the left-hand side: in the presence of a 2s core-hole. On the right-hand side: without core-hole.

When the pseudopotential for the absorbing atom in the presence of a core-hole is generated, it is possible to include fractions of core-hole instead of a full core-hole: one uses  $2s^x$  in the configuration for the atom with, for example, x=1.5 for 0.5 core-hole, x=1.25 for 0.75 core-hole and x=1.1 for 0.9 core-hole. Fractions of core-hole allows to some extent to check the influence of the strength of the core-hole screening but the exact physical meaning is not so clear.



Fig. 3.21: Calculated XAS and XNCD at the I  $L_1$ -edge in the  $\Delta$  enantiomer of  $\alpha$ -LiIO<sub>3</sub> with different fractions of core-hole.

The XNCD spectra obtained with different fractions of core-hole are presented in Fig. 3.21. The XAS spectra for fractions of core-hole larger than 0.5 are very similar one to each other and they are quite different from the spectrum without core-hole. On the other hand, the XNCD spectra with 0.5 core-hole is almost the same as the spectrum without core-hole. When the fraction of core-hole is increased the amplitude of XNCD decreases and then increases back with an opposite sign for the main peak. This illustrates the fact that the sign and the amplitude of the XNCD effect is highly dependent on slight changes in the electronic structure.

In order to better understand the core-hole effect, we computed the spectra at the I K-edge. It corresponds to an energy  $\hbar\omega = 33.2$  keV (for the  $L_1$ -edge  $\hbar\omega = 5.2$  keV) and no experimental



Fig. 3.22: Comparison of the calculated spectra with (solid line) or without core-hole (dashed line) at the I  $L_{I^-}$  (top) and  $K^-$  (bottom) edge in the  $\Delta$  enantiomer of  $\alpha$ -LiIO<sub>3</sub>.

data are available for this edge. The motivation for this calculation is to see if the XNCD spectra would be less sensitive to a deeper core-hole. As the aim is to compare the calculated spectra at the  $L_1$ - and K-edges, the same broadening is applied to the spectra at the K-edge as to the spectra at the  $L_1$ -edge (3.46 eV) even if the 1s core-hole lifetime broadening is 10.6 eV (which would result in a more broadened spectra). In Fig. 3.22, for both edges, the same rigid 2.8 eV horizontal shift to the left was applied to the XAS and XNCD spectra without core-hole in order to make the XAS maxima coincide.



**Fig. 3.23:** Radial part  $R_{nl}(r)$  of the 1s, 2s, 5d and 5p wave functions for an hydrogen-like atom. The functions  $R_{5,2}(r)$  and  $R_{5,1}(r)$  are multiplied by 20.

The first notable point is that, in the absence of core-hole (that is to say with the same

electronic structure), XNCD spectra at K- and  $L_1$ -edges are opposite. The only thing that differs between these two calculations is the core wavefunction  $|i\rangle$  used in the cross-section calculation that is a 1s atomic wavefunction in one case and a 2s atomic wavefunction in the other case. It results in an opposite sign of the quadrupole radial integral (Q - see Eq. (3.12)) whereas the dipole radial integral (D - see Eq. (3.3)) has the same sign in both cases. Now, the D-Q term is proportional to the product DQ of these two values (see Table 3.5). The difference in the sign of Q is due to the fact that the 2s radial part is negative in the region where the dradial part is significant whereas the 1s radial wave function is always positive (see Fig. 3.23). This explains why the XNCD is opposite at the K-edge compared to the  $L_1$ -edge despite the fact that they probe the same valence properties.

Otherwise, the effect of a 1s core-hole is as spectacular as the effect of a 2s core-hole. As visible in Fig. 3.23, the 1s and 2s states are spatially very localized near the nucleus:  $\frac{Zr}{a_0} = 5$  with  $Z = 43 \Rightarrow r \approx 0.05$  Å which is very small compared to the interatomic distances (the distance I-O  $d_{I-O} = 1.8$  Å). From the point of view of the valence electrons, removing a 1s or a 2s electron is therefore roughly equivalent to adding a proton on the nucleus (Z+1 approximation). We observe however in Fig. 3.22 that the XNCD amplitude in the presence of the core hole is significantly smaller at the K-edge than at the  $L_1$ -edge, which highlights the limitation of this Z+1 approximation that would result in identical spectra. By comparison with Fig. 3.21, we see that the fraction of 2s core hole that would correspond to the effect of the 1s core hole lies between 0.9 and 1. This might suggest that a 1s core hole is more efficiently screened than a 2s core hole.

The sensitivity of the XNCD to the core hole relaxation creates a possibility for it to be used as a benchmark to test relaxation processes.

# 3.4 Sum-rules at K- and $L_1$ -edge

A sum-rule is a formula in which the integral of the circular dichroism spectrum due to a given term of the cross-section is expressed as a function of the ground state expectation value of some operator. The circular dichroism sum-rules for the dipole-dipole term at  $L_{2,3}$ -edges are well established [Thole et al., 1992, Carra et al., 1993, Wu et al., 1993, Chen et al., 1995] and are widely used to extract quantitative magnetic ground state properties. At spin-orbit split edges, two magneto optical sum-rules exist: an orbital sum-rule that links the integral of the spectra to the ground-state expectation value of  $\langle L_z \rangle$  and a spin sum-rule that links the difference of the integrals corresponding to the contribution from each spin-orbit split core levels, to the ground-state expectation value of  $\langle S_z \rangle$  up to a  $\langle T_z \rangle$  value that can be significant [Wu and Freeman, 1994].

At K-edge or  $L_1$ -edge, only the orbital magnetization sum-rule can apply. This sum rule was originally derived by B. T. Thole [Thole et al., 1992] using a graphical method but Altarelli [Altarelli, 1993] proposed later a "physically transparent" proof by expressing the electronic states in a basis of wave-functions that are eigenstates of  $L_z$ . These derivations are based on several approximations among which the fact that the radial integrals are spin independent and energy independent [Altarelli, 1998].

Here, we use a second quantization formalism to retrieve this sum-rule at the K- and  $L_1$ -edge and we do not try to express the result in terms of spherical tensors. For a matter of simplicity, the circular dichroism sum-rules for the terms Q-Q, D-Q and D-SP are presented in the case we considered in the calculation, that is to say with **k** parallel to the axis z and full circular polarization and, for D-SP, diagonal spin-orbit coupling. In this section, we use many-body wave functions:  $|G\rangle$  for the ground state and  $|F\rangle$  for the final state. The operators in the cross section are replaced by many-body operators that write as sums of one-body operators:

$$D = \sum_{i} \boldsymbol{\epsilon} \cdot \mathbf{r}_{i}$$

$$Q = \frac{1}{2} \sum_{i} (\boldsymbol{\epsilon} \cdot \mathbf{r}_{i}) (\mathbf{k} \cdot \mathbf{r}_{i})$$

$$SP = \sum_{i} \boldsymbol{\sigma}_{i} (\boldsymbol{\epsilon} \times \mathbf{r}_{i})$$

As the derivation is very similar for the four terms (D-D, Q-Q, D-Q and D-SP terms), the details of the calculations and the main ideas are given for the D-D term whereas, for the other terms, only the most important results are given.

#### 3.4.1 Sum-rules for the D-D term

As the spin does not appear explicitly in the electric dipole operator, it will be omitted in the notations for the derivation of the sum-rule for the D-D term and it will be reintroduced at the end of the derivation.

#### XAS sum-rule

Using the expression of the electric dipole operator of Appendix D.2, the contribution to the cross section of the D-D term writes:

$$\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}) = 4\pi^2 \alpha_0 \hbar \omega \sum_F |\langle G| \sum_i \frac{4\pi}{3} r_i \sum_{\lambda=-1}^{\lambda=1} (-1)^{\lambda} Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star}) Y_1^{\lambda}(\mathbf{u}_{\mathbf{r}_i}) |F\rangle|^2 \delta(E_f - E_g - \hbar \omega)$$
$$= 4\pi^2 \alpha_0 \hbar \omega \sum_{\lambda, \lambda'=-1}^1 Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star}) Y_1^{\lambda'}(\boldsymbol{\epsilon}) (-1)^{\lambda'} \boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda, \lambda'}$$

where

$$\boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'} = \frac{(4\pi)^2}{9} (-1)^{\lambda+\lambda'} \sum_{F} \langle G| \sum_{i} r_i Y_1^{\lambda}(\mathbf{u}_{\mathbf{r}_i}) |F\rangle \langle F| \sum_{i} (r_i Y_1^{\lambda'}(\mathbf{u}_{\mathbf{r}_i}))^* |G\rangle \delta(E_f - E_g - \hbar\omega). \quad (3.1)$$

Note that in this expression the final state is in the ket and the ground state is in the bra which is not the same convention as the one used in Eq. 2.2. To be consistent the operator  $\boldsymbol{\epsilon} \cdot \mathbf{r}$  is transformed into  $\boldsymbol{\epsilon}^* \cdot \mathbf{r}$ .

The operators  $\sum_{i} r_i Y_1^{\lambda}(\mathbf{u}_{\mathbf{r}_i})$  are one-body operators in the sense that they have the form  $\hat{O} = \sum_{i} \hat{o}[i]$ . Such operators, in the language of second quantization, take the form:

$$\hat{O} = \sum_{\alpha,\beta} \langle \alpha | \hat{o} | \beta \rangle a_{\alpha}^{\dagger} a_{\beta}$$

where  $\langle \alpha | \hat{o} | \beta \rangle$  are the matrix elements of the single-particle operator in a complete singleparticle basis set. We choose a basis of eigenstates of  $L_z$  and  $\mathbf{L}^2$  to make this expansion:<sup>7</sup>

$$\sum_{i} r_i Y_1^{\lambda}(\mathbf{u}_{\mathbf{r}_i}) = \sum_{lml_0m_0} \left( \int d\mathbf{r} \, Y_{l_0}^{m_0 \star}(\mathbf{u}_{\mathbf{r}}) R_{l_0}^{\star}(r) r Y_1^{\lambda}(\mathbf{u}_{\mathbf{r}}) Y_l^m(\mathbf{u}_{\mathbf{r}}) R_l(r) \right) a_{l_0,m_0}^{\dagger} a_{l,m}$$

<sup>&</sup>lt;sup>7</sup>Note that this is not a complete basis set because the quantum number n is missing. We could justify it by the fact that, at a specific edge, the transition occurs from a given  $n_0$  to a given n. At *K*-edge, it is however not true that the transition occurs to a given n. This is related to the problem on the upper limit of the integral that we describe below.

where m and l are the usual quantum numbers. Using the formula for the integral of the product of three spherical harmonics (Appendix D.1):

$$\langle G|\sum_{i} r_{i} Y_{1}^{\lambda}(\mathbf{u}_{\mathbf{r}_{i}})|F\rangle = \sum_{lml_{0}m_{0}} \sqrt{\frac{3(2l+1)}{4\pi(2l_{0}+1)}} (10l0|l_{0}0)(1\lambda lm|l_{0}m_{0})\langle G|a_{l_{0},m_{0}}^{\dagger}a_{l,m}|F\rangle \mathcal{D}_{l_{0},k_{0}}$$

where  $\mathcal{D}_{l_0,l} = \int \mathrm{d}r \, r^3 R_{l_0}^{\star}(r) R_l(r)$ . The experimental procedure (choice of the edge) enables to obtain the signal corresponding to a specific  $l_0$  so the sum over  $l_0$  can be removed. At K- and  $L_1$ -edges,  $l_0 = 0$  and  $m_0 = 0$ .

The first Clebsch–Gordan coefficient,  $(10l0|00) = \delta_{l,1}(-1)/\sqrt{2l+1} = \delta_{l,1}(-1)/\sqrt{3}$ , therefore, consistently with the selection rules of the dipole operator, only l = 1 will give non-zero contributions to the sum. The second Clebsch–Gordan coefficient with l = 1,  $(1\lambda 1m|00) = \delta_{\lambda,-m}(-1)^{1-\lambda}/\sqrt{3}$ . So that,

$$\langle G|\sum_{i} r_{i} Y_{1}^{\lambda}(\mathbf{u}_{\mathbf{r}_{i}})|F\rangle = (-1)^{\lambda} \sqrt{\frac{1}{4\pi}} \langle G|a_{0,0}^{\dagger}a_{1,-\lambda}|F\rangle \mathcal{D}$$
(3.2)

where

$$\mathcal{D} = \mathcal{D}_{l_0,l} = \int r^3 R_0^\star(r) R_1(r) \,\mathrm{d}r.$$
(3.3)

This leads to:

$$\boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'} = \frac{4\pi}{9} \sum_{F} \langle G | a_{0,0}^{\dagger} a_{1,-\lambda} | F \rangle \langle F | a_{1,-\lambda'}^{\dagger} a_{0,0} | G \rangle \delta(E_f - E_g - \hbar\omega) | \mathcal{D} |^2.$$
(3.4)

The next step towards the derivation of a sum rule is to integrate  $\zeta_{D-D}^{\lambda,\lambda'}$  over  $\hbar\omega$ . The quantity  $\langle G|a_{0,0}^{\dagger}a_{1,-\lambda}|F\rangle$  is zero if the state  $\{l=1, m=-\lambda\}$  is fully occupied in  $|G\rangle$  but also if it is fully empty in  $|F\rangle$ . Therefore, it is equivalent to integrate the spectra from 0 or from an energy  $E_0$  that represents the minimal energy for the level  $\{1, -\lambda\}$  to be populated:

$$\int_{E_0}^{\infty} \boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'} \,\mathrm{d}\hbar\omega = \int_{0^+}^{\infty} \boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'} \,\mathrm{d}\hbar\omega$$

This is due to the fact that a given  $l_0$  was chosen for the core state. It is equivalent to what is done experimentally: the experimental spectra for a given edge exhibit interesting features starting from a given energy that corresponds more or less to the edge energy. The baseline below this energy is set to zero.

To derive the integral over the energy, we use the completeness relation:

$$\int_{0^+}^{\infty} \sum_{F} |F\rangle \langle F|\delta(E_f - E_g - E) \,\mathrm{d}E = 1 - |G\rangle \langle G|.$$
(3.5)

So that the integral of  $\boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'}$  writes:

$$\int_{E_0}^{\infty} \boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'} \,\mathrm{d}E = \frac{4\pi}{9} \langle G | a_{0,0}^{\dagger} a_{1,-\lambda} | \left( \mathbbm{1} - |G\rangle\langle G| \right) | a_{1,-\lambda'}^{\dagger} a_{0,0} | G \rangle | \mathcal{D} |^2$$
$$= \frac{4\pi}{9} \left( \langle G | a_{00}^{\dagger} a_{1-\lambda} a_{1,-\lambda'}^{\dagger} a_{0,0} | G \rangle - \langle G | a_{0,0}^{\dagger} a_{1,-\lambda} | G \rangle \langle G | a_{1,-\lambda'}^{\dagger} a_{0,0} | G \rangle \right) | \mathcal{D} |^2$$

The second term is zero because, using fermionic anticommutation relations,  $\langle G|a_{0,0}^{\dagger}a_{1,-\lambda}|G\rangle = -\langle G|a_{1,-\lambda}a_{0,0}^{\dagger}|G\rangle$  and the core shell is full in the ground state so  $a_{00}^{\dagger}|G\rangle = 0$ . The use of the anticommutation relations in the first term leads to:

$$\int_{E_0}^{\infty} \boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'} \mathrm{d}E = \frac{4\pi}{9} \langle G|a_{1-\lambda}a_{1-\lambda'}^{\dagger}a_{00}^{\dagger}a_{00}|G\rangle |\mathcal{D}|^2$$
$$= \frac{4\pi}{9} \langle G|a_{1,-\lambda}a_{1,-\lambda'}^{\dagger}(1-a_{0,0}a_{0,0}^{\dagger})|G\rangle |\mathcal{D}|^2$$
$$= \frac{4\pi}{9} \langle G|a_{1,-\lambda}a_{1,-\lambda'}^{\dagger}|G\rangle |\mathcal{D}|^2.$$

Finally, the absorption sum-rule for the D-D term is:

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon})}{\hbar\omega} \,\mathrm{d}\hbar\omega = \frac{16\pi^3}{9} \alpha_0 \sum_{s=\uparrow,\downarrow} \sum_{\lambda,\lambda'=-1}^{1} |\mathcal{D}^s|^2 (-1)^{\lambda'} Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star}) Y_1^{\lambda'}(\boldsymbol{\epsilon}) \langle G|a_{1,-\lambda,s} a_{1,-\lambda',s}^{\dagger}|G\rangle.$$
(3.6)

where we reintroduced the spin in the notations.

In Table 3.3 the results for several remarkable polarization vectors are presented. It shows that the right-handed circular light probes the p states with m=-1 and the left-handed circular light probes the p states with m=1. This is in accordance with the fact that the angular momentum along z of the photons is either  $-\hbar$  in the case of a right-handed circular polarization or  $+\hbar$  for a left-handed circular polarization [Stöhr and Siegmann, 2006, p.151]<sup>8</sup> and with the fact that the total angular momentum is conserved during the transition.

**Table 3.3:** Electric dipole XAS sum-rule at the K and  $L_1$ -edges for five remarkable polarization vectors.

| Type of polarization | Polarization vector   | $\int_{E_0}^{\infty} \frac{\sigma_{\rm D-D}(\boldsymbol{\epsilon})}{\hbar\omega} \mathrm{d}\hbar\omega$  |
|----------------------|---|--|
| Right-Handed         | $oldsymbol{\epsilon}_1 = rac{1}{\sqrt{2}} egin{pmatrix} 1 \ -i \ 0 \end{pmatrix}$      | $\frac{4\pi^2}{3}\alpha_0 \mathcal{D} ^2\langle G a_{1,-1}a_{1,-1}^{\dagger} G\rangle$   |
| Left-Handed          | $\boldsymbol{\epsilon}_2 = rac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$ | $\frac{4\pi^2}{3}\alpha_0 \mathcal{D} ^2\langle G a_{1,1}a_{1,1}^{\dagger} G\rangle$   |
| Linear along z       | $\boldsymbol{\epsilon}_0 = \begin{pmatrix} 0\\ 0\\ 1 \end{pmatrix}$                     | $\frac{4\pi^2}{3}\alpha_0 \mathcal{D} ^2\langle G a_{1,0}a_{1,0}^{\dagger} G\rangle$   |
| Linear along x       | $\boldsymbol{\epsilon}_x = egin{pmatrix} 1 \ 0 \ 0 \end{pmatrix}$                       | $\frac{2\pi^2}{3}\alpha_0  \mathcal{D} ^2 (\langle G a_{1,-1}a_{1,-1}^{\dagger} G\rangle + \langle G a_{1,1}a_{1,1}^{\dagger} G\rangle - \langle G a_{1,-1}a_{1,1}^{\dagger} G\rangle - \langle G a_{1,1}a_{1,-1}^{\dagger} G\rangle)$     |
| Linear along y       | $oldsymbol{\epsilon}_y = egin{pmatrix} 0 \ 1 \ 0 \end{pmatrix}$                         | $\frac{2\pi^{2}}{3}\alpha_{0} \mathcal{D} ^{2}(\langle G a_{1,-1}a_{1,-1}^{\dagger} G\rangle + \langle G a_{1,1}a_{1,1}^{\dagger} G\rangle + \langle G a_{1,-1}a_{1,1}^{\dagger} G\rangle + \langle G a_{1,1}a_{1,-1}^{\dagger} G\rangle)$ |

<sup>&</sup>lt;sup>8</sup>We use the convention for right and left polarization that is used in the optics community and that consists in defining the sense of circular polarization "from the sample point of view". In [Stöhr and Siegmann, 2006] they use the other convention (see p.149).

Link with the number of holes The ground state writes:  $|G\rangle = \sum_{k} c_k \Pi_k$  where  $\Pi_k = \prod_{l,m} a_{l,m}^{\dagger} |0\rangle$  corresponds to a Slater determinant. The quantity  $\langle G|a_{1,-\lambda}a_{1,-\lambda}^{\dagger}|G\rangle$  is zero if the state  $\{1, -\lambda\}$  is fully occupied in  $|G\rangle$  that is to say if all  $\Pi_k$  include the operator  $a_{1,-\lambda}^{\dagger}$ . It is one if no  $\Pi_k$  include this operator, that is to say if the state  $\{1, -\lambda\}$  is fully empty in  $|G\rangle$ . In the general case:

$$\langle G|a_{1,-\lambda}a_{1,-\lambda}^{\dagger}|G\rangle = \sum_{\substack{k \\ \text{with } k \text{ such that} \\ \Pi_k \text{ does not contain } a_{1,-}^{\dagger}}} |c_k|^2$$

This quantity represents the number of holes in the level  $\{1, -\lambda\}$  in the ground state. For example, the integral of the XAS spectrum for  $\epsilon_1$  is proportional to the number of holes in states l=1, m=-1. Here, we immediately see a problem related to the upper limit of the integral: if we integrate up to infinity, the number of holes would be infinite so the integral must be restricted to a given n. For 3d transition elements, it would be interesting to restrict it to the 4p states in order to "count" the number of  $\{4p, m=-1\}$  holes but it is impossible in practice because the 4p states are in the continuum.

Link with the partial densities of states If we denote  $E = E_g + \hbar \omega$ , Eq. 3.4 writes:

$$\boldsymbol{\zeta}_{\mathrm{D-D}}^{\lambda,\lambda'}(E) = \frac{4\pi}{9} |\mathcal{D}|^2 \sum_F |\langle G|a_{0,0}^{\dagger}a_{1,-\lambda}|F\rangle|^2 \delta(E_f - E).$$

We define  $Dos_{|G|}^{1,\lambda}(E) = \sum_{F} |\langle G|a_{0,0}^{\dagger}a_{1,\lambda}|F\rangle|^2 \delta(E_f - E)$ . It corresponds to a G-conditioned partial density of states (in states/eV) in the presence of a core-hole because it counts in an energy interval the number of states in which the level  $\{1, -\lambda\}$  is occupied and the core level  $\{0, 0\}$  is unoccupied and that are compatible with  $|G\rangle$ .<sup>9</sup>

The contribution to the spectra from the D-D term now writes:

$$\frac{\sigma_{\mathrm{D-D}}(E)}{\hbar\omega} = \frac{16\pi^3}{9}\alpha_0 |\mathcal{D}|^2 \sum_{\lambda=-1}^1 |Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star})|^2 Dos^{1,-\lambda}_{|G}(E).$$
(3.7)

It means that the electric dipole-electric dipole XAS probes the density of empty states with l=1.

$$\begin{split} \langle F|a_{1,\lambda}^{\dagger}a_{0,0}|G\rangle\langle G|a_{0,0}^{\dagger}a_{1,\lambda}|F\rangle &= |f_k|^2 \text{ where } k \text{ is such that } \Pi_k = a_{0,0}a_{1,\lambda}^{\dagger}\Pi \\ \langle F|a_{1,\lambda}^{\dagger}a_{0,0}a_{0,0}^{\dagger}a_{1,\lambda}|F\rangle &= \sum_{\substack{k \\ n_k \text{ does not contain } a_{0,0}^{\dagger}} |f_k|^2. \end{split}$$

Hence, the difference between a partial density of states in the presence of a core-hole and  $Dos_{|G|}^{1,\lambda}(E)$  is that the final states must be related to the ground state by a one-body transition. In the following, the mention "conditioned to  $|G\rangle$ " will be left implicit when mentioning density of states.

<sup>&</sup>lt;sup>9</sup>A natural definition of the partial density of states in the presence of a core-hole would be:  $\sum_{F} \langle F | a_{1,\lambda}^{\dagger} a_{0,0} a_{0,0}^{\dagger} a_{1,\lambda} | F \rangle \delta(E_{f} - E) \text{ whereas here we have } \sum_{F} \langle F | a_{1,\lambda}^{\dagger} a_{0,0} | G \rangle \langle G | a_{0,0}^{\dagger} a_{1,\lambda} | F \rangle \delta(E_{f} - E). \text{ If } | G \rangle$ writes as  $\Pi$  where  $\Pi = \prod_{l,m} a_{l,m}^{\dagger} | 0 \rangle$  and  $| F \rangle = \sum_{k} f_{k} \Pi_{k}$  where  $\Pi_{k} = \prod_{l,m} a_{l,m}^{\dagger} | 0 \rangle$  then

**Isotropic spectrum** To get the expression for the isotropic spectrum, we consider an arbitrary linear polarization vector:

$$\boldsymbol{\epsilon}_{l} = \begin{pmatrix} \sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta \end{pmatrix}.$$

If this polarization vector is used in the sum-rule of Eq.(3.6) and the integration is performed over  $\theta \in [0, \pi]$  and  $\phi \in [0, 2\pi]$ , the result is:<sup>10</sup>

$$A^{0} = \int_{\theta,\phi} \left( \int_{E_{0}}^{\infty} \frac{\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{l})}{\hbar\omega} \,\mathrm{d}\hbar\omega \right) \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\phi = \frac{16\pi^{3}}{9} \alpha_{0} |\mathcal{D}|^{2} \langle G|a_{1,0}a_{1,0}^{\dagger} + a_{1,1}a_{1,1}^{\dagger} + a_{1,-1}a_{1,-1}^{\dagger}|G\rangle.$$

This yields the sum-rule for the isotropic spectrum:

$$\int_{E_0}^{\infty} \frac{\sigma_{\rm D-D}^0}{\hbar\omega} \mathrm{d}\hbar\omega = \frac{A^0}{\int \sin\theta \mathrm{d}\theta \mathrm{d}\phi} = \frac{4\pi^2}{9} \alpha_0 |\mathcal{D}|^2 \langle G|a_{1,0}a_{1,0}^{\dagger} + a_{1,1}a_{1,1}^{\dagger} + a_{1,-1}a_{1,-1}^{\dagger}|G\rangle = \frac{4\pi^2}{9} \alpha_0 |\mathcal{D}|^2 h_p$$

The integral from  $E_0$  of the isotropic spectrum is proportional to the total number of holes in the occupied p states  $h_p$ . In practice, it is difficult to apply this sum-rule because of the uncertainty on the upper limit of the integral already mentioned above. However, a similar sum-rule at the  $L_{2,3}$ -edges has been successfully used to compare the occupation of d states in a material compared to a reference (Au<sub>4</sub>Mn compared to Au in [Rogalev et al., 2006, p.93] or Mo<sup>IV</sup> compared to Mo<sup>V</sup> in [Arrio et al., 2010]).

The spectrum itself is a representation of the *total density of empty* p states (see previous paragraph).

From the equality of the sum-rules we can deduce that:

$$\sigma_{\mathrm{D-D}}^{0} = rac{1}{3} \left( \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{0}) + \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{2}) + \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_{1}) 
ight)$$

In other words, in the general case, the isotropic spectrum (that corresponds to the spectrum measured on a powder sample) is obtained by computing three different spectra and taking the average.

**Linear dichroism** If one consider a non magnetic crystal with a rotation axis of order strictly higher than two, it is known that the angular dependence of the spectra is [Brouder, 1990]:

$$\sigma_{\rm D-D}(\boldsymbol{\epsilon}) = \cos^2\theta\sigma_{\parallel} + \sin^2\theta\sigma_{\perp}$$

with  $\sigma_{\parallel}$  the spectra obtained with a linear polarization vector along the direction of the rotation axis and  $\sigma_{\perp}$  the spectra obtained with a polarization vector perpendicular to the rotation axis. Here we find the same result by considering that  $\langle G|a_{1-\lambda}a_{1-\lambda'}^{\dagger}|G\rangle = \delta_{\lambda,\lambda'}\langle G|a_{1-\lambda}a_{1-\lambda}^{\dagger}|G\rangle$ : this is the condition to remove all azimuthal angular dependence. In this case if one defines

$$\sigma_{\perp} = \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_x) = \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_y) = \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_1) = \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_2) \quad \text{and} \quad \sigma_{\parallel} = \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}_0),$$

the isotropic spectrum writes:  $\sigma_{D-D}^0 = \frac{1}{3}(2\sigma_{\perp} + \sigma_{\parallel})$ . Linear dichroism is defined as the difference between  $\sigma_{\parallel}$  and  $\sigma_{\perp}$ .

Note that if the symmetry of the crystal is cubic, the  $\sigma_{D-D}$  spectrum presents no angular dependence (so no linear dichroism) and the isotropic spectrum is  $\sigma_{D-D}^0 = \sigma_{D-D}(\epsilon_0) = \sigma_{D-D}(\epsilon_1) = \sigma_{D-D}(\epsilon_2)$ .

<sup>&</sup>lt;sup>10</sup>The calculation was done with Mathematica. Note that, with an arbitrary linear polarization, the sum-rule presents both polar- and azimuthal-angle dependence and it contains 9 terms (all the combinations of  $\lambda$  and  $\lambda'$ ). This complexity can be reduced by considering the symmetry of the system as done in [Brouder, 1990].

#### XMCD sum-rule

The sum-rule for circular dichroism can be deduced from the XAS sum-rule (Eq. (3.6)):

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}) - \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}^{\star})}{\hbar\omega} \,\mathrm{d}\hbar\omega = \frac{16\pi^3}{9} \alpha_0 \sum_{\lambda,\lambda'=-1}^{1} (-1)^{\lambda'} \left( Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star}) Y_1^{\lambda'}(\boldsymbol{\epsilon}) - Y_1^{-\lambda}(\boldsymbol{\epsilon}) Y_1^{\lambda'}(\boldsymbol{\epsilon}^{\star}) \right) \langle G|a_{1,-\lambda} a_{1,-\lambda'}^{\dagger}|G\rangle.$$

The quantity  $(Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star})Y_1^{\lambda'}(\boldsymbol{\epsilon}) - Y_1^{-\lambda}(\boldsymbol{\epsilon})Y_1^{\lambda'}(\boldsymbol{\epsilon}^{\star}))$  depends on the rate of circular polarization defined by:

$$\boldsymbol{\epsilon} \times \boldsymbol{\epsilon}^{\star} = i P_c \hat{\mathbf{k}} \tag{3.8}$$

where  $\hat{\mathbf{k}}$  is the direction of the wave vector (it is a vector of norm 1).

The formula [Natoli et al., 1998]

$$Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star})Y_1^{\lambda'}(\boldsymbol{\epsilon}) - Y_1^{-\lambda}(\boldsymbol{\epsilon})Y_1^{\lambda'}(\boldsymbol{\epsilon}^{\star}) = -\sqrt{\frac{3}{2\pi}}P_c\sum_{\mu}(1\lambda'1-\lambda|1\mu)Y_1^{\mu}(\hat{\mathbf{k}})$$

allows to compute  $(-1)^{\lambda'} \frac{4\pi}{3} (Y_1^{-\lambda}(\boldsymbol{\epsilon}^*) Y_1^{\lambda'}(\boldsymbol{\epsilon}) - Y_1^{-\lambda}(\boldsymbol{\epsilon}) Y_1^{\lambda'}(\boldsymbol{\epsilon}^*))$  for each  $\lambda$  and  $\lambda'$ . The Clebsh Gordan coefficient is non zero when  $\mu = \lambda' - \lambda$  and  $\lambda \neq -\lambda'$ :

Therefore,

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}) - \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}^{\star})}{\hbar\omega} \, \mathrm{d}\hbar\omega = \frac{4\pi^2}{3} \alpha_0 |\mathcal{D}|^2 P_c \Big( \hat{k}_z \langle G | a_{1,-1} a_{1,-1}^{\dagger} - a_{1,1} a_{1,1}^{\dagger} | G \rangle \\ + \hat{k}_y \frac{i}{\sqrt{2}} \langle G | a_{1,0} a_{1,1}^{\dagger} - a_{1,1} a_{1,0}^{\dagger} + a_{10} a_{1,-1}^{\dagger} - a_{1,-1} a_{1,0}^{\dagger} | G \rangle \\ + \hat{k}_x \frac{-1}{\sqrt{2}} \langle G | a_{1,0} a_{1,1}^{\dagger} + a_{1,1} a_{1,0}^{\dagger} + a_{1,0} a_{1,-1}^{\dagger} + a_{1,-1} a_{1,0}^{\dagger} | G \rangle \Big).$$

We can apply the anti commutation relation to the pairs of fermion operators which only add a minus sign, even in the term proportional to  $\hat{k}_z$  because:

$$\langle G|a_{1,\lambda}^{\dagger}a_{1,\lambda} - a_{1,\mu}^{\dagger}a_{1,\mu}|G\rangle = \langle G|1 - a_{1,\lambda}a_{1,\lambda}^{\dagger} - 1 + a_{1,\mu}a_{1,\mu}^{\dagger}|G\rangle = -\langle G|a_{1,\lambda}a_{1,\lambda}^{\dagger} - a_{1,\mu}a_{1,\mu}^{\dagger}|G\rangle.$$

Using the matrix elements for the angular momentum given in Appendix B.2, we define the components  $Lz_{l=1}^s, Ly_{l=1}^s$  and  $Lx_{l=1}^s$  of the operator contribution to the orbital polarization of the p states as:

$$Lz_{l=1}^{s} = \sum_{m=-1}^{1} \hbar m a_{1,m,s}^{\dagger} a_{1,m,s}$$

$$Ly_{l=1}^{s} = \frac{\hbar}{i\sqrt{2}} (a_{1,1,s}^{\dagger} a_{1,0,s} - a_{1,0,s}^{\dagger} a_{1,1,s} + a_{1,-1,s}^{\dagger} a_{1,0,s} - a_{1,0,s}^{\dagger} a_{1,-1,s})$$

$$Lx_{l=1}^{s} = \frac{\hbar}{\sqrt{2}} (a_{1,1,s}^{\dagger} a_{1,0,s} + a_{1,0,s}^{\dagger} a_{1,1,s} + a_{1,-1,s}^{\dagger} a_{1,0,s} + a_{1,0,s}^{\dagger} a_{1,-1,s}).$$

The XMCD sum-rule for the electric dipole term writes as a function of the ground state expectation value of the angular momentum along the direction of propagation of the incident beam:

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}) - \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}^{\star})}{\hbar\omega} \,\mathrm{d}\hbar\omega = \frac{4\pi^2}{3\hbar} \alpha_0 \sum_{s=\uparrow,\downarrow} |\mathcal{D}^s|^2 P_c \langle G | \hat{\mathbf{k}} \cdot \mathbf{L}_{l=1}^s | G \rangle$$

If the radial matrix element is assumed to be spin-independent  $(\mathcal{D}^{\uparrow} = \mathcal{D}^{\downarrow})$  and with  $\mathbf{L} = \mathbf{L}^{\uparrow} + \mathbf{L}^{\downarrow}$ ,

$$\int_{E_0}^{\infty} \frac{\sigma_{\rm D-D}(\boldsymbol{\epsilon}) - \sigma_{\rm D-D}(\boldsymbol{\epsilon}^{\star})}{\hbar\omega} \,\mathrm{d}\hbar\omega = \frac{4\pi^2}{3\hbar} \alpha_0 |\mathcal{D}|^2 P_c \langle G|\hat{\mathbf{k}} \cdot \mathbf{L}_{l=1}|G\rangle$$
(3.9)

This corresponds to the sum-rule that has been derived by several authors [Thole et al., 1992, Ankudinov and Rehr, 1995, Carra et al., 1993, Altarelli, 1993] and that is called the *orbital sum-rule*. It can be normalized with the sum-rule obtained for the isotropic spectrum:

$$\frac{\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}) - \sigma_{\mathrm{D-D}}(\boldsymbol{\epsilon}^{\star})}{\hbar\omega} \,\mathrm{d}\hbar\omega}{\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}^0}{\hbar\omega} \,\mathrm{d}\hbar\omega} = \frac{3P_c \langle G | \hat{\mathbf{k}} \cdot \mathbf{L}_{l=1} | G \rangle}{\hbar h_p}.$$

Full circular polarization case In the case of a fully circularly polarized light with **k** along z, the sum-rule for  $\sigma_{D-D}^{\text{XMCD}} = \sigma_{D-D}(\epsilon_2) - \sigma_{D-D}(\epsilon_1)$  is  $(P_c = -1)$ :

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-D}}^{\mathrm{XMCD}}}{\hbar\omega} \,\mathrm{d}\hbar\omega = -\frac{4\pi^2}{3\hbar} \alpha_0 \sum_{s=\uparrow,\downarrow} |\mathcal{D}^s|^2 \langle G|Lz_{l=1}^s|G\rangle$$
$$= -\frac{4\pi^2}{3\hbar} \alpha_0 \sum_{s=\uparrow,\downarrow} |\mathcal{D}^s|^2 \langle G|a_{1,1,s}^{\dagger}a_{1,1,s} - a_{1,-1,s}^{\dagger}a_{1,-1,s}|G\rangle$$

This result can also be deduced directly from the XAS sum-rule for  $\epsilon_1$  and  $\epsilon_2$  displayed in Table 3.3.

The spectrum is a representation of the orbital angular momentum density of the empty p states, that is to say

$$\sigma_{\rm D-D}^{\rm XMCD}(E) = \frac{4\pi^2}{3} \alpha_0 \omega(|\mathcal{D}^{\uparrow}|^2 \mathcal{O}_p^{\uparrow}(E) + |\mathcal{D}^{\downarrow}|^2 \mathcal{O}_p^{\downarrow}(E))$$
(3.10)

with  $\mathcal{O}_p^{\sigma}(E) = Dos_{|G|}^{1,1}(E) - Dos_{|G|}^{1,-1}(E)$  where the  $Dos_{|G|}^{1,\lambda}(E)$  are defined p. 105. In other words it probes the difference between the densities of the levels  $\{l=1, m=1\}$  and  $\{l=1, m=-1\}$ .

**Difficulties to apply this sum-rule** This sum rule is impossible to apply in practice at the K-edge of 3d transition elements. The first difficulty, as in the XAS case, is that the upper limit of the integral is not well defined. Moreover, the 4p states are almost unoccupied so that  $\langle Lz_{l=1} \rangle = \langle G | Lz_{l=1} | G \rangle$  is very small and it does not contribute to the total magnetic moment of the material. In other words, this sum-rule relates the integral of the XMCD spectra to a quantity that is considered zero in most of the models for magnetism. Yet, even if the quantitative measurement of  $\langle Lz_{l=1} \rangle$  is of moderate interest in itself, the orbital polarization of the p states carries information on the magnetic structure of the material in its whole which justifies the great interest for XMCD at K-edge beyond the applicability of the sum-rules.

### 3.4.2 Sum-rules for the Q-Q term

#### XAS sum-rule

Using the expression of the electric quadrupole operator of Appendix D.2, the contribution to the cross section of the Q-Q term writes:

$$\begin{split} \sigma_{\mathbf{Q}-\mathbf{Q}}(\boldsymbol{\epsilon},\mathbf{k}) &= \pi^2 \alpha_0 \hbar \omega \left(\frac{\omega}{c}\right)_{\boldsymbol{\lambda},\boldsymbol{\lambda}',\boldsymbol{\mu},\boldsymbol{\mu}'=-1}^{2} (-1)^{\boldsymbol{\lambda}'+\boldsymbol{\mu}'} (1\boldsymbol{\lambda} 1\boldsymbol{\mu} | 2\boldsymbol{\lambda} + \boldsymbol{\mu}) (1\boldsymbol{\lambda}' 1\boldsymbol{\mu}' | 2\boldsymbol{\lambda}' + \boldsymbol{\mu}') \\ & Y_1^{-\boldsymbol{\lambda}}(\boldsymbol{\epsilon}^{\star}) Y_1^{-\boldsymbol{\mu}}(\hat{\mathbf{k}}) Y_1^{\boldsymbol{\lambda}'}(\boldsymbol{\epsilon}) Y_1^{\boldsymbol{\mu}'}(\hat{\mathbf{k}}) \boldsymbol{\zeta}_{\mathbf{Q}-\mathbf{Q}}^{\boldsymbol{\lambda}+\boldsymbol{\mu},\boldsymbol{\lambda}'+\boldsymbol{\mu}'} \\ & \text{with } \boldsymbol{\zeta}_{\mathbf{Q}-\mathbf{Q}}^{\boldsymbol{\lambda}+\boldsymbol{\mu},\boldsymbol{\lambda}'+\boldsymbol{\mu}'} = \left(\frac{4\pi}{3}\right)^4 \frac{3}{10\pi} (-1)^{\boldsymbol{\lambda}+\boldsymbol{\mu}+\boldsymbol{\lambda}'+\boldsymbol{\mu}'} \sum_F \langle G| \sum_i r_i^2 Y_2^{\boldsymbol{\lambda}+\boldsymbol{\mu}}(\mathbf{u}_{r_i}) | F \rangle \\ & \langle F| \sum_i (r_i^2 Y_2^{\boldsymbol{\lambda}'+\boldsymbol{\mu}'}(\mathbf{u}_{r_i}))^{\star} | G \rangle \delta(E_f - E_g - \hbar \omega). \end{split}$$

As in the previous section we can rewrite the operator within the second quantization formalism:

$$\langle G|\sum_{i} r_{i}^{2} Y_{2}^{\lambda+\mu}(\mathbf{u}_{\mathbf{r}_{i}})|F\rangle = \sum_{lml_{0}m_{0}} \sqrt{\frac{5(2l+1)}{4\pi(2l_{0}+1)}} (20l0|l_{0}0)(2\lambda+\mu lm|l_{0}m_{0})\langle G|a_{l_{0},m_{0}}^{\dagger}a_{l,m}|F\rangle \mathcal{Q}_{l_{0},l_{0}}$$

with  $\mathcal{Q}_{l_0,l} = \int \mathrm{d}r \, r^4 R_{l_0}^{\star}(r) R_l(r)$ . Again, we only consider K- and L<sub>1</sub>-edges where  $l_0 = 0$  and  $m_0 = 0$ . As  $(20l0|00) = \delta_{l,2} \frac{(-1)^2}{\sqrt{2l+1}} = \delta_{l,2} \frac{1}{\sqrt{5}}$  and  $(2\lambda + \mu lm|00) = \delta_{\lambda+\mu,-m} \frac{(-1)^{2-(\lambda+\mu)}}{\sqrt{5}}$ ,

$$\langle G|\sum_{i} r_{i}^{2} Y_{2}^{\lambda+\mu}(\mathbf{u}_{\mathbf{r}_{i}})|f\rangle = (-1)^{2-(\lambda+\mu)} \frac{1}{\sqrt{4\pi}} \langle G|a_{0,0}^{\dagger}a_{2,-(\lambda+\mu)}|f\rangle \mathcal{Q}$$
(3.11)

with

$$Q = Q_{0,2} = \int r^4 R_0^*(r) R_2(r) \,\mathrm{d}r.$$
 (3.12)

Making the spin explicit:

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathbf{Q}-\mathbf{Q}}(\boldsymbol{\epsilon}, \mathbf{k})}{(\hbar\omega)^3} d\hbar\omega = \left(\frac{4\pi}{3}\right)^3 \frac{\pi\alpha_0}{10} \left(\frac{1}{\hbar c} \sum_{\boldsymbol{\lambda}, \boldsymbol{\lambda}', \boldsymbol{\mu}, \boldsymbol{\mu}'=-1}^{2} (-1)^{\boldsymbol{\lambda}'+\boldsymbol{\mu}'} (1\lambda 1\mu|2\boldsymbol{\lambda}+\boldsymbol{\mu}) (1\boldsymbol{\lambda}' 1\mu'|2\boldsymbol{\lambda}'+\boldsymbol{\mu}')$$
$$Y_1^{-\boldsymbol{\lambda}}(\boldsymbol{\epsilon}^{\star}) Y_1^{-\boldsymbol{\mu}}(\hat{\mathbf{k}}) Y_1^{\boldsymbol{\lambda}'}(\boldsymbol{\epsilon}) Y_1^{\boldsymbol{\mu}'}(\hat{\mathbf{k}}) \sum_{s=\uparrow,\downarrow} \langle G|a_{2,-(\boldsymbol{\lambda}+\boldsymbol{\mu}),s} a_{2,-(\boldsymbol{\lambda}'+\boldsymbol{\mu}'),s}^{\dagger}|G\rangle |\mathcal{Q}^s|^2 \quad (3.13)$$

In Table 3.4 the results for right-handed and left-handed circularly polarized light are presented. It shows that the right-handed circular light probes the d states with m = -1 and the left-handed circular light probes the d states with m = 1. This is, as in the electric dipole case, in accordance with what is expected from the fact that the angular momentum along z of the photons is either  $-\hbar$  in the case of a right-handed circular polarization or  $\hbar$  for a left-handed circular polarization.

**Table 3.4:** Electric quadrupole XAS sum-rule at the K- and  $L_1$ -edges for left and right circularly polarized X-rays with  $\mathbf{k}$  along z.

| Wave vector  | Polarization vector  | $\int_{E_0}^{\infty} \frac{\sigma_{\rm Q-Q}(\epsilon)}{(\hbar\omega)^3} \mathrm{d}\hbar\omega$                    |
|--|--|---|
| $\hat{\mathbf{k}} = \begin{pmatrix} 0\\0\\1 \end{pmatrix}$ | $\boldsymbol{\epsilon}_1 = rac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}$ | $\left(\frac{\pi^2 \alpha_0}{15(\hbar c)^2}\right)  \mathcal{Q} ^2 \langle G a_{2,-1}a_{2,-1}^{\dagger} G\rangle$ |
|  | $\boldsymbol{\epsilon}_2 = rac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$  | $\left(\frac{\pi^2 \alpha_0}{15(\hbar c)^2}\right)  \mathcal{Q} ^2 \langle G a_{2,1} a_{2,1}^{\dagger} G\rangle$  |

**Isotropic spectrum** To get the expression for the isotropic spectrum, we consider, as in [Brouder, 1990], an arbitrary linear polarization vector (both angles  $\theta$  and  $\phi$  are needed) and a wave vector that is perpendicular to the polarization vector (one extra angle  $\psi$  is needed):

$$\boldsymbol{\epsilon}_{l} = \begin{pmatrix} \sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta \end{pmatrix} \text{ and } \hat{\mathbf{k}} = \begin{pmatrix} \cos\theta\cos\phi\cos\psi - \sin\phi\sin\psi\\ \cos\theta\sin\phi\cos\psi + \cos\phi\sin\psi\\ -\sin\theta\cos\psi \end{pmatrix}.$$

The integral over  $\theta$  in  $[0, \pi]$  and  $\phi$  and  $\psi$  in  $[0, 2\pi]$  of the XAS sum-rule (Eq. (3.13)) yields:

$$\iiint \sin(\theta) \mathrm{d}\theta \mathrm{d}\phi \mathrm{d}\psi \int_{E_0}^{\infty} \frac{\sigma_{\mathrm{Q-Q}}(\boldsymbol{\epsilon}_l, \mathbf{k})}{(\hbar\omega)^3} \mathrm{d}\hbar\omega$$
$$= \frac{\pi\alpha_0}{(\hbar c)^2} \left(\frac{4\pi}{3}\right)^3 \frac{1}{10} \frac{9}{20} \langle G|a_{2,0}a_{2,0}^{\dagger} + a_{2,1}a_{2,1}^{\dagger} + a_{2,-1}a_{2,-1}^{\dagger} + a_{2,2}a_{2,2}^{\dagger} + a_{2,-2}a_{2,-2}^{\dagger}|G\rangle |\mathcal{Q}|^2$$

As  $\iiint \sin(\theta) d\theta d\phi d\psi = 8\pi^2$ , this means that the isotropic spectrum  $\sigma_{Q-Q}^0$  is such that:

$$\int_{E_0}^{\infty} \frac{\sigma_{\rm Q-Q}^0}{(\hbar\omega)^3} \,\mathrm{d}\hbar\omega = \frac{\pi^2}{75} \frac{\alpha_0}{(\hbar c)^2} |\mathcal{Q}|^2 \langle G| \sum_{m=-2}^2 a_{2,m} a_{2,m}^{\dagger} |G\rangle = \frac{\pi^2}{75} \frac{\alpha_0}{(\hbar c)^2} |\mathcal{Q}|^2 h_d \tag{3.14}$$

The integral of the isotropic spectrum is proportional to the total number of holes in the occupied d states  $h_d$ . The spectrum itself,  $\sigma^0_{Q-Q}(E)$ , probes the total d density of empty states in the presence of the core-hole.

Interpretation of XAS pre-edge The electric quadrupole contribution can be significant in the pre-edge region of the spectrum. In this case, it can be useful to compare the spectra with the p and d partial density of unoccupied states in the presence of a core-hole in order to assign the peaks either to electric dipole or to electric quadrupole transition. Indeed, the spectra recorded on powders (isotropic spectra) are proportional to this density of states up to the radial matrix element. The relative order of magnitude of the two contributions can be roughly evaluated by multiplying the d density of state by  $\frac{9}{4\pi^2} \frac{\pi^2}{75} \mathbf{k}^2 \frac{|\mathcal{Q}|^2}{|\mathcal{D}|^2} = 0.03 \mathbf{k}^2 \frac{|\mathcal{Q}|^2}{|\mathcal{D}|^2}$ . The order of magnitude of  $\frac{|\mathcal{Q}|^2}{|\mathcal{D}|^2}$  is  $r_c^2$  where  $r_c$  is the core state radius.

#### XMCD sum-rule

We explore the XMCD sum-rule for the Q-Q term. In order to get an understandable sumrule, the result presented here is limited to the case considered in our numerical calculation: **k** along z and a full circular polarization. The general XMCD sum-rule would be obtained by taking the difference of the sum-rule Eq. (3.13) with  $\boldsymbol{\epsilon}$  and  $\boldsymbol{\epsilon}^{\star}$ . This sum-rule can be found in [Brouder, 1994].

The sum-rule for  $\sigma_{Q-Q}^{XMCD} = \sigma_{Q-Q}(\epsilon_2) - \sigma_{Q-Q}(\epsilon_1)$  can be directly deduced from the results presented in Table 3.4:

$$\int_{E_0}^{\infty} \frac{\sigma_{\rm Q-Q}^{\rm XMCD}}{(\hbar\omega)^3} \,\mathrm{d}\hbar\omega = \left(\frac{\pi^2 \alpha_0}{15(\hbar c)^2}\right) |\mathcal{Q}|^2 \langle G|a_{2,1}a_{2,1}^{\dagger} - a_{2,-1}a_{2,-1}^{\dagger}|G\rangle \tag{3.15}$$

Now, we define the operator 1-partial contribution to the orbital polarization along z of the d states as:

$$Lz_{l=2}^{1} = \sum_{m=-1}^{1} m a_{2,m}^{\dagger} a_{2,m}.$$
(3.16)

After restoration of the spin indices, the XMCD sum-rule for the Q-Q term writes:

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathbf{Q}-\mathbf{Q}}^{\mathrm{XMCD}}}{(\hbar\omega)^3} \,\mathrm{d}\hbar\omega = -\left(\frac{\pi^2 \alpha_0}{15(\hbar c)^2}\right) \sum_{s=\uparrow,\downarrow} |\mathcal{Q}^s|^2 \langle G|Lz_{l=2}^{1\,s}|G\rangle.$$
(3.17)

Since the Q-Q term is almost always smaller than an accompanying D-D term, this sum-rule cannot be easily applied to experimental spectra. When applied to calculated spectra, it can however yield valuable physical information.

#### 3.4.3 Sum-rules for the D-Q term

#### XAS sum-rule

The expression of the contribution to the cross section of the cross term D-Q is:

$$\sigma_{\mathrm{D-Q}}(\boldsymbol{\epsilon}, \mathbf{k}) = -4\pi^2 \alpha_0 \hbar \omega \sum_F \mathrm{Im} \left[ \langle F | \sum_i ((\mathbf{k} \cdot \mathbf{r}_i)(\boldsymbol{\epsilon}^* \cdot \mathbf{r}_i))^* | G \rangle \langle G | \sum_i \boldsymbol{\epsilon}^* \cdot \mathbf{r}_i | F \rangle \right] \delta(E_f - E_g - \hbar \omega).$$

Using the expression for the electric dipole and the electric quadrupole operators in terms of spherical harmonics (Appendix D.2),  $\sigma_{D-Q}$  rewrites:

$$\sigma_{\mathrm{D-Q}} = -4\pi^2 \hbar \alpha_0 \omega k \sum_F \mathrm{Im} \left[ \sum_{\lambda,\nu,\mu=-1}^{1} (1\lambda 1\mu | 2\lambda + \mu) (Y_1^{-\lambda}(\boldsymbol{\epsilon}^{\star}) Y_1^{-\mu}(\hat{\mathbf{k}}))^{\star} Y_1^{-\nu}(\boldsymbol{\epsilon}^{\star}) \boldsymbol{\zeta}_{\mathrm{D-Q}}^{\lambda+\mu,\nu} \right] \delta(E_f - E_g - \hbar \omega)$$

where

$$\boldsymbol{\zeta}_{\mathrm{D-Q}}^{\lambda+\mu,\nu} = \left(\frac{4\pi}{3}\right)^3 (-1)^{\lambda+\mu+\nu} \sqrt{\frac{3}{10\pi}} \langle F| \sum_i \left(r_i^2 Y_2^{\lambda+\mu}(\mathbf{u_{r_i}})\right)^* |G\rangle \langle G| \sum_i r_i Y_1^{\nu}(\mathbf{u_{r_i}}) |F\rangle \delta(E_f - E_g - \hbar\omega).$$

Using Eq. (3.2) and (3.11), we obtain:

$$\int_{E_0}^{\infty} \boldsymbol{\zeta}_{\mathrm{D-Q}}^{\lambda+\mu,\nu} \,\mathrm{d}E = \left(\frac{4\pi}{3}\right)^2 \sqrt{\frac{1}{30\pi}} \langle G|a_{1-\nu}a_{2-(\lambda+\mu)}^{\dagger}|G\rangle \mathcal{Q}^{\star}\mathcal{D}$$

So that,

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-Q}}(\boldsymbol{\epsilon})}{(\hbar\omega)^2} \,\mathrm{d}\hbar\omega = -\frac{(4\pi)^3}{9\sqrt{30\pi}\hbar c} \pi\alpha_0 \sum_{\lambda,\nu,\mu=-1}^{1} (1\lambda 1\mu | 2\lambda + \mu)(-1)^{\lambda+\mu} \\ \mathrm{Im}\left[Y_1^{\lambda}(\boldsymbol{\epsilon})Y_1^{\mu}(\hat{\mathbf{k}})Y_1^{-\nu}(\boldsymbol{\epsilon}^{\star})\langle G|a_{1-\nu}a_{2-(\lambda+\mu)}^{\dagger}|G\rangle \mathcal{Q}^{\star}\mathcal{D}\right]$$

What is noticeable about this sum-rule is that the right term is non zero only if there exist in  $|G\rangle$  mixed states l = 1 and l = 2, that is to say mixed p - d states. These kind of states can exist only in non-centrosymmetric systems. The result for circular polarization and **k** along z is given in table 3.5.

**Table 3.5:** Cross term electric dipole-electric quadrupole XAS sum-rule at the K- and  $L_1$ -edges for left and right circularly polarized X-rays with  $\mathbf{k}$  along z.

| Wave vector  | Polarization vector  | $\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D}-\mathrm{Q}}(\boldsymbol{\epsilon})}{(\hbar\omega)^2} \mathrm{d}\hbar\omega$                     |
|--|--|---|
| $\hat{\mathbf{k}} = \begin{pmatrix} 0\\0\\1 \end{pmatrix}$ | $oldsymbol{\epsilon}_1 = rac{1}{\sqrt{2}} egin{pmatrix} 1 \ -i \ 0 \end{pmatrix}$ | $\frac{-4\pi^2}{3\sqrt{5}}\frac{\alpha}{\hbar c}\operatorname{Im}[\langle G a_{1-1}a_{2-1}^{\dagger} G\rangle\mathcal{Q}^{\star}\mathcal{D}]$ |
|  | $\epsilon_2 = rac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$         | $\frac{-4\pi^2}{3\sqrt{5}}\frac{\alpha}{\hbar c}\operatorname{Im}[\langle G a_{11}a_{21}^{\dagger} G\rangle\mathcal{Q}^{\star}\mathcal{D}]$   |

#### XNCD sum-rule

As for the Q-Q term, the result presented here will be limited to the case **k** along z, therefore, no study on the angular dependence of the spectra is possible. The sum-rule for a general wave vector is treated in [Natoli et al., 1998] in the case where the wave functions  $|F\rangle$  and  $|G\rangle$  are real. Here,  $|F\rangle$  and  $|G\rangle$  are any wave functions but it is important to remind that, if they cannot be chosen real, circular dichroism can be a combination of XMCD and XNCD [Sessoli et al., 2015].

The sum-rule for  $\sigma_{D-Q}^{\text{XNCD}} = \sigma_{D-Q}(\boldsymbol{\epsilon}_2) - \sigma_{D-Q}(\boldsymbol{\epsilon}_1)$  can be directly deduced from the results presented in Table 3.5:

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-Q}}^{\mathrm{XNCD}}}{(\hbar\omega)^2} \,\mathrm{d}\hbar\omega = \frac{-4\pi^2}{3\sqrt{5}} \frac{\alpha}{\hbar c} \operatorname{Im}[\langle G|a_{11}a_{21}^{\dagger} - a_{1-1}a_{2-1}^{\dagger}|G\rangle \mathcal{Q}^{\star}\mathcal{D}]$$

If the wave functions can be chosen real, it can be seen that:  $\boldsymbol{\zeta}_{\mathrm{D-Q}}^{-\lambda-\mu,-\nu} = \left(\boldsymbol{\zeta}_{\mathrm{D-Q}}^{\lambda+\mu,\nu}\right)^{\star}$  so that the quantities  $\langle G|a_{1-1}a_{2-1}^{\dagger}|G\rangle$  and  $\langle G|a_{11}a_{21}^{\dagger}|G\rangle$  are complex conjugate. The sum-rule for XNCD becomes:<sup>11</sup>

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-Q}}^{\mathrm{XNCD}}}{(\hbar\omega)^2} \,\mathrm{d}\hbar\omega = i \frac{4\pi^2}{3\sqrt{5}} \frac{\alpha}{\hbar c} \langle G|a_{11}a_{21}^{\dagger} - a_{1-1}a_{2-1}^{\dagger}|G\rangle \mathcal{QD} \,.$$
(3.18)

Therefore, XNCD probes the difference of occupation between the mixed p-d states with m = 1 and the mixed p-d states with m = -1 in the ground state which is a pure imaginary quantity. Contrary to the case of the Q-Q term, this sum-rule for the D-Q term can be applied to experimental XNCD spectra in which this term is the only contribution.

<sup>&</sup>lt;sup>11</sup> If  $x^*$  is the complex conjugate of x,  $\text{Im}[x - x^*] = -i(x - x^*)$ .

# 3.4.4 Sum-rules for the D-SP term

Here we derive the sum-rule for the new relativistic term of the cross section electric dipolespin-position in the case of collinear spins and diagonal spin orbit coupling. In that case, the spin is conserved during the transition (see subsection SP operator p.56) and we can use the collinear spin-position operator:

$$SP_{col}(\boldsymbol{\epsilon}) = \sigma_z(\epsilon_x y - \epsilon_y x).$$

#### XAS sum-rule

Using the expression for  $SP_{col}$  given in Appendix D.2,<sup>12</sup>

$$\sigma_{\rm D-SP}(\boldsymbol{\epsilon}) = \frac{2\pi^2 \hbar^2 \alpha_0 \omega^2}{mc^2} \sum_{\nu=-1}^{1} \operatorname{Re}[Y_1^{-\nu}(\boldsymbol{\epsilon}^*) \left(Y_1^1(\boldsymbol{\epsilon})\zeta_{\rm D-SP}^{1,\nu} - Y_1^{-1}(\boldsymbol{\epsilon})\zeta_{\rm D-SP}^{-1,\nu}\right)]$$

with

$$\zeta_{\rm D-SP}^{\lambda,\nu} = (-1)^{\nu} \left(\frac{4\pi}{3}\right)^2 \sum_F \langle F| \sum_i (\sigma_{zi} r_i Y_1^{\lambda}(\mathbf{u_{r_i}}))^{\star} |G\rangle \langle G| \sum_i r_i Y_1^{\nu}(\mathbf{u_{r_i}}) |F\rangle \delta(E_f - E_g - \hbar\omega).$$

With the spin made explicit, Eq. (3.2) rewrites:

$$\langle G|\sum_{i} r_{i} Y_{1}^{\lambda}(\mathbf{u}_{\mathbf{r}_{i}})|F\rangle = (-1)^{\lambda} \sqrt{\frac{1}{4\pi}} \sum_{s,s_{0}=\uparrow,\downarrow} \langle G|a_{0,0,s}^{\dagger}a_{1,-\lambda,s}|F\rangle \delta_{s,s_{0}} \mathcal{D}^{s}$$

Similarly to what has been done to obtain Eq. (3.2), we obtain:

$$\langle G|\sum_{i}\sigma_{zi}r_{i}Y_{1}^{\lambda}(\mathbf{u}_{\mathbf{r}_{i}})|F\rangle = \sum_{s',s_{0}=\uparrow,\downarrow} \langle s_{0}'|\sigma_{z}|s'\rangle(-1)^{\lambda}\langle G|a_{0,0,s_{0}'}^{\dagger}a_{1,-\lambda,s'}|F\rangle \mathcal{D}^{s_{0}',s'}.$$
(3.19)

Now,  $\langle s'_0 | \sigma_z | s' \rangle = \langle s' | \sigma_z | s' \rangle \delta_{s'_0, s'}$  where  $\langle s' | \sigma_z | s' \rangle = 1$  if  $s' = \uparrow$  and  $\langle s' | \sigma_z | s' \rangle = -1$  if  $s' = \downarrow$ .

Using the same method to integrate over the energy as for the D-D term, the sum-rule for the term D-SP is:<sup>13</sup>

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon})}{(\hbar\omega)^2} \,\mathrm{d}\hbar\omega = \frac{-8\pi^3\alpha}{9mc^2} \sum_{\nu=-1}^{1} \sum_{s=\uparrow,\downarrow} \langle s|\sigma_z|s\rangle \operatorname{Re}\left[Y_1^{-\nu}(\boldsymbol{\epsilon}^{\star})\left(Y_1^1(\boldsymbol{\epsilon})\langle G|a_{1-\nu\sigma}a_{1-1\sigma}^{\dagger}|G\rangle - Y_1^{-1}(\boldsymbol{\epsilon})\langle G|a_{1-\nu\sigma}a_{11\sigma}^{\dagger}|G\rangle\right)|\mathcal{D}^s|^2\right]$$

With this formula, it is possible to determine the sum-rule for any polarization. The results for pure circularly polarized light are reported in table 3.6.

<sup>&</sup>lt;sup>12</sup>Here, we used the fact that  $\text{Im}[(ix)^{\star}] = -\text{Re}[x^{\star}].$ 

<sup>&</sup>lt;sup>13</sup>Note that a  $\delta_{s,s'}$  appears in the calculation because  $a_{0,0,s}^{\dagger}a_{0,0,s'} = \delta_{s,s'} - a_{0,0,s'}a_{0,0,s}^{\dagger}$ 

| Table 3.6:              | Sum-rule for | the cross | $_{ m s}~{ m term}~{ m I}$ | D-SP | at the | <i>K</i> - | $\operatorname{and}$ | $L_1$ -edges | for fu | lly c | $\operatorname{circular}$ | polarized | light |
|-------------------------|--------------|-----------|----------------------------|------|--------|------------|----------------------|--------------|--------|-------|---------------------------|-----------|-------|
| with $\mathbf{k}$ along | g <i>z</i> . |           |                            |      |        |            |                      |              |        |       |                           |           |       |

| Polarization vector  | $\int_{E_0}^{\infty} \frac{\sigma_{\rm D-SP}(\boldsymbol{\epsilon})}{(\hbar\omega)^2} \mathrm{d}\hbar\omega$   |
|--|--|
| $egin{aligned} egin{aligned} egin{aligned} egin{aligned} egin{aligned} egin{aligned} 1 \ -i \ 0 \end{aligned} \end{aligned} \end{aligned}$ | $-\frac{2\pi^{2}\alpha}{3mc^{2}}(\langle G a_{1-1\uparrow}a_{1-1\uparrow}^{\dagger} G\rangle \mathcal{D}^{\uparrow} ^{2}-\langle G a_{1-1\downarrow}a_{1-1\downarrow}^{\dagger} G\rangle \mathcal{D}^{\downarrow} ^{2})$ |
| $oldsymbol{\epsilon}_2 = rac{1}{\sqrt{2}} egin{pmatrix} 1 \ i \ 0 \end{pmatrix}$  | $\frac{2\pi^2\alpha}{3mc^2}(\langle G a_{11\uparrow}a_{11\uparrow}^{\dagger} G\rangle \mathcal{D}^{\uparrow} ^2 - \langle G a_{11\downarrow}a_{11\downarrow}^{\dagger} G\rangle \mathcal{D}^{\downarrow} ^2)$            |

#### XMCD sum-rule

The sum-rule for  $\sigma_{D-SP}^{XMCD} = \sigma_{D-SP}(\epsilon_2) - \sigma_{D-Q}(\epsilon_1)$  can be directly deduced from the results presented in table 3.6:

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-SP}}^{XMCD}}{(\hbar\omega)^2} \,\mathrm{d}\hbar\omega = \frac{2\pi^2\alpha}{3mc^2} (\langle G|a_{11\uparrow}a_{11\uparrow}^{\dagger} + a_{1-1\uparrow}a_{1-1\uparrow}^{\dagger}|G\rangle |\mathcal{D}^{\uparrow}|^2 - \langle G|a_{11\downarrow}a_{11\downarrow}^{\dagger} + a_{1-1\downarrow}a_{1-1\downarrow}^{\dagger}|G\rangle |\mathcal{D}^{\downarrow}|^2)$$

Therefore, the XMCD spectrum due to this term is related to the difference of density of states between up and down spins.

If one defines the operator partial spin magnetization of the p states as:

$$Sz_{l=1}^{1,-1} = \sum_{m=-1,1} a_{1m\downarrow}^{\dagger} a_{1m\downarrow} - a_{1m\uparrow}^{\dagger} a_{1m\uparrow}.$$
 (3.20)

Under the hypothesis that  $|\mathcal{D}^{\uparrow}| = |\mathcal{D}^{\downarrow}|$ :

$$\int_{E_0}^{\infty} \frac{\sigma_{\mathrm{D-SP}}^{CD}(\boldsymbol{\epsilon})}{(\hbar\omega)^2} \,\mathrm{d}\hbar\omega = -\frac{2\pi^2\alpha}{3mc^2} \langle G|Sz_{l=1}^{1,-1}|G\rangle |\mathcal{D}|^2.$$
(3.21)

From this sum-rule, we can understand why despite its small prefactor, the D-SP term contributes significantly to XMCD:  $Sz_{l=1}^{1,-1}$  has a priori an average on the ground state that is notably larger than  $\langle Lz_{l=1} \rangle$  as it is the case for Fe, Co and Ni (see p.84).

This sum-rule is impossible to apply to experimental spectra, not only because of uncertainty on the upper limit of the integral (due to the fact that the 4p states are in the continuum), but mainly because experimentally this term is always accompanied by the D-D contribution.

#### Total dipole XMCD sum-rule

If we consider the sum of the D-D and the D-SP contribution to XMCD:  $\sigma_{\rm D}^{XMCD} = \sigma_{\rm D-D}^{XMCD} + \sigma_{\rm D-SP}^{XMCD}$ , then,

$$\int_{E_0}^{\infty} \frac{\sigma_{\rm D}^{XMCD}(\boldsymbol{\epsilon})}{(\hbar\omega)} \,\mathrm{d}\hbar\omega = -\frac{4\pi^2\alpha}{3} \left( L_{zl=1} + \frac{\hbar\omega}{2mc^2} S_{zl=1}^{1,-1} \right) |\mathcal{D}|^2 \tag{3.22}$$

This total sum-rule could be applied to measured XMCD spectra at the K-edge if the following hypothesis was realized:

• The contribution of the Q-Q term is negligible (else its contribution must be taken into account)

- $\hbar\omega$  is approximately constant over the energy range we integrate on (this is the case when we consider a few 10 eV over the edge and the edge energy is a few keV).
- The radial integral is spin-independent  $(|\mathcal{D}^{\uparrow}| = |\mathcal{D}^{\downarrow}|)$ .
- The transition occurs from a given  $n_0$  to a given n.

It is particularly difficult to make a general statement on the validity of the two last hypothesis.

#### 3.4.5 Partial density of states and spectra in the case of Fe and Ni

As explained and illustrated in [Cabaret, 2006, section 4], the comparison of the spectra with calculated projected density of states must be done with caution. Indeed, the definition of the partial density of states is not unique and the choice of the method used to perform the projection of the density of states on atomic orbitals has a big influence on the final result.



**Fig. 3.24:** *p*-projected density of states and *K*-edge D-D contribution to the XAS in Fe (on the left) and Ni (on the right) without core-hole for both spin channels (spin up: solid line, spin down: dashed line). This figure is an illustration of Eq. 3.6.

In the PAW formalism presented in section 2.3, it is easy to identify the partial density of states: if one consider that there is only one projector per channel the cross-section writes as a sum over n where the index n refers to quantum numbers for the angular momentum (l, m):

$$\sigma(\omega) = 4\pi^2 \alpha_0 \hbar \omega \sum_n \sum_f |\langle \tilde{f} | \tilde{p}_n^{\mathbf{R}_0} \rangle|^2 \delta(E_f - E_i - \hbar \omega) |\langle \phi_n^{\mathbf{R}_0} | O | i \rangle|^2.$$

In this expression, we can identify  $\sum_{f} |\langle \tilde{f} | \tilde{p}_n^{\mathbf{R}_0} \rangle|^2 \delta(E_f - E_i - \hbar \omega)$  as being the *n*-partial density of empty states. The remaining matrix element  $|\langle \phi_n^{\mathbf{R}_0} | O | i \rangle|^2$  is weakly energy-dependent as shown in [Cabaret, 2006]. For this reason, the DOS calculated with QUANTUM ESPRESSO are very similar to the spectra. This is confirmed by the comparison of the *p*-density of states with the calculated contribution to the XAS due to the D-D term depicted in Fig. 3.24 (the full width at half maximum of the broadening applied to the DOS is 1.6 eV and whereas it is 0.8 eV for the the spectra).

In Fig. 3.25, the D-D contribution to XMCD is compared to the density of orbital polarization of the empty p states and the D-SP contribution to XMCD is compared to the partial spin polarization  $Sz_{l=1}^{1,-1}(E)$ . It illustrates the validity of the sum rules (Eq. 3.9 and 3.21) at energies



Fig. 3.25: Comparison of  $Lz_p(E)$  (top) and  $Sz_p^{1,-1}(E)$  (bottom) obtained from *p*-projected densities of states with the K-edge D-D (top) and D-SP (bottom) contributions to the XAS in Fe (on the left) and Ni (on the right) without core-hole. These figures are illustrations of Eq. 3.9 and 3.21.  $Lz_p(E)$ have been multiplied by the same factor as the *p* density of states in Fig. 3.24 and  $Sz_p^{1,-1}(E)$  spectra by this factor times  $\frac{\hbar\omega}{2mc^2}$ .

higher than 10 eV. However, near the edge, the D-D contribution to XMCD and  $Lz_{l=1}(E)$  are completely different. This leads to the conclusion that, even with a method to compute the projected density of states that induces weak energy-dependence of the matrix elements, a calculation of the orbital polarization is not a good way to obtain the signature of the XMCD spectra at K-edge. The spectrum calculation must be performed. With the method presented in the previous chapter, the spectra calculations require less computer resources than DOS calculations.

## 3.5 Conclusion

In this chapter we presented the results of DFT *ab-initio* calculations of K-edge XAS and XMCD in Fe, Co and Ni and  $L_1$ -edge XAS and XNCD in LiIO<sub>3</sub>.

The main result of this chapter is the significance of the new D-SP term in XMCD. We explained it by deriving a sum-rule for this new term. We showed that it probes the spin

polarization of the p states whereas the D-D term probes their orbital polarization.

The overall agreement of the calculated spectra with the experimental one is fair. For XMCD, however, the secondary peaks are not always well reproduced. Given the complexity of the XMCD phenomenon (compared to the magnetic properties that are usually studied by DFT), it is a unique way to evaluate the precision of the modeling of magnetic systems. We tested the effect on the spectra of a large number of assumptions and approximations and came to the conclusion that Kohn-Sham DFT is unable to capture the complexity of the spin and orbital polarization of empty states.

The fact that the overall shape of the spectra is well described still opens perspectives for the interpretation of various experiments. This is the subject of the next chapter.

# Bibliography

- [Alagna et al., 1998] L. Alagna, T. Prosperi, S. Turchini, J. Goulon, A. Rogalev, C. Goulon-Ginet, C. R. Natoli, R. D. Peacock, and B. Stewart (1998). X-Ray Natural Circular Dichroism. *Phys. Rev. Lett.*, 80:4799–4802.
- [Altarelli, 1993] M. Altarelli (1993). Orbital-magnetization sum rule for x-ray circular dichroism: A simple proof. Phys. Rev. B, 47(2):597–598.
- [Altarelli, 1998] M. Altarelli (1998). Sum rules for X-ray magnetic circular dichroism. Il Nuovo Cimento D, 20(7-8):1067–1073.
- [Ankudinov and Rehr, 1995] A. Ankudinov and J. J. Rehr (1995). Sum rules for polarizationdependent x-ray absorption. *Phys. Rev. B*, 51(2):1282–1285.
- [Ankudinov and Rehr, 2000] A. L. Ankudinov and J. J. Rehr (2000). Theory of solid-state contributions to the x-ray elastic scattering amplitude. *Phys. Rev. B*, 62(4):2437–2445.
- [Arrio et al., 2010] M.-A. Arrio, J. Long, C. C. Moulin, A. Bachschmidt, V. Marvaud, A. Rogalev, C. Mathoniere, F. Wilhelm, and P. Sainctavit (2010). Photoinduced magnetization on Mo ion in copper octacyanomolybdate: an X-ray magnetic circular dichroism investigation. J. Phys. Chem, 100(114):593-600.
- [Bordage et al., 2010] A. Bordage, C. Brouder, E. Balan, D. Cabaret, A. Juhin, M.-A. Arrio, P. Sainctavit, G. Calas, and P. Glatzel (2010). Electronic structure and local environment of substitutional V3+ in grossular garnet Ca3Al2(SiO4)3: K-edge X-ray absorption spectroscopy and first-principles modeling. Am. Mineral., 95(8-9):1161-1171.
- [Brouder, 1990] C. Brouder (1990). Angular dependence of X-ray absorption spectra. J. Phys.: Condens. Matter, 2(3):701.
- [Brouder, 1994] C. Brouder (1994). Contribution quadripolaire au dichroïsme circulaire magnétique des seuils LII et LIII des terres rares (http://www-ext.impmc.upmc.fr/~brouder/ grandest94.ps). In Compte-rendu des Journées Grand-Est, pages 53-57. J.-P. Kappler, Strasbourg.
- [Brouder et al., 1996] C. Brouder, M. Alouani, and K. H. Bennemann (1996). Multiplescattering theory of x-ray magnetic circular dichroism: Implementation and results for the iron K edge. Phys. Rev. B, 54(10):7334-7349.
- [Bunău and Joly, 2009] O. Bunău and Y. Joly (2009). Self-consistent aspects of x-ray absorption calculations. J. Phys.: Condens. Matter, 21(34):345501.
- [Cabaret, 2006] D. Cabaret (2006). XANES : approche monoélectronique. http://www-ext. impmc.upmc.fr/~cabaret/cours\_aussois06.pdf

- [Cabaret et al., 2010] D. Cabaret, A. Bordage, A. Juhin, M. Arfaoui, and E. Gaudry (2010). First-principles calculations of X-ray absorption spectra at the K-edge of 3d transition metals: an electronic structure analysis of the pre-edge. *Phys. Chem. Chem. Phys.*, 12(21):5619.
- [Carra et al., 1993] P. Carra, B. T. Thole, M. Altarelli, and X. Wang (1993). X-ray circular dichroism and local magnetic fields. *Phys. Rev. Lett.*, 70(5):694–697.
- [Chen et al., 1995] C. T. Chen, Y. U. Idzerda, H.-J. Lin, N. V. Smith, G. Meigs, E. Chaban, G. H. Ho, E. Pellegrin, and F. Sette (1995). Experimental Confirmation of the X-Ray Magnetic Circular Dichroism Sum Rules for Iron and Cobalt. *Phys. Rev. Lett.*, 75:152–155.
- [Cowan, 1981] R. D. Cowan (1981). The Theory of Atomic Structure and Spectra. University of California Press.
- [Dixit and Alouani, 2016] A. Dixit and M. Alouani (2016). Ab initio calculations of X-ray magnetic circular dichroism spectra within the projector augmented wave method: An implementation into the {VASP} code. Comput. Phys. Commun., 207:136 144.
- [Ducher et al., 2016] M. Ducher, M. Blanchard, D. Vantelon, R. Nemausat, and D. Cabaret (2016). Probing the local environment of substitutional Al<sup>3+</sup> in goethite using X-ray absorption spectroscopy and first-principles calculations. *Physics and Chemistry of Minerals*, 43(3):217-227.
- [Ebert, 1996] H. Ebert (1996). Influence of the orbital polarization on the magnetic X-ray dichroism of transition metals. *Solid State Commun.*, 100(10):677–681.
- [Fuggle and Inglesfield, 1992] J. C. Fuggle and J. E. Inglesfield, editors (1992). Unoccupied Electronic States, volume 69 of Topics in Applied Physics. Springer Berlin Heidelberg.
- [Fujikawa and Nagamatsu, 2003] T. Fujikawa and S. Nagamatsu (2003). Relativistic multiple scattering theory of K-edge X-ray magnetic circular dichroism. J. Electron. Spectrosc. Relat. Phenom., 129(1):55–69.
- [Gotsis and Strange, 1994] H. J. Gotsis and P. Strange (1994). A first-principles theory of X-ray Faraday effects. J. Phys.: Condens. Matter, 6(7):1409.
- [Goulon et al., 1998] J. Goulon, C. Goulon-Ginet, A. Rogalev, V. Gotte, C. Malgrange, C. Brouder, and C. R. Natoli (1998). X-ray natural circular dichroism in a uniaxial gyrotropic single crystal of LiIO3. J. Chem. Phys., 108(15):6394-6403.
- [Group, ] P. D. Group. Clebsch-Gordan coeff., sph. harmonics, and d functions.
- [Guo, 1996] G. Y. Guo (1996). What does the K-edge x-ray magnetic circular dichroism spectrum tell us? J. Phys.: Condens. Matter, 8(49):L747.
- [Igarashi and Hirai, 1994] J. Igarashi and K. Hirai (1994). Magnetic circular dichroism at the K edge of nickel and iron. Phys. Rev. B, 50(24):17820-17829.
- [Igarashi and Hirai, 1996] J. Igarashi and K. Hirai (1996). Orbital moment and magnetic circular dichroism at the K edge in ferromagnetic cobalt. *Phys. Rev. B*, 53:6442–6450.
- [Kas et al., 2007] J. J. Kas, A. P. Sorini, M. P. Prange, L. W. Cambell, J. A. Soininen, and J. J. Rehr (2007). Many-pole model of inelastic losses in x-ray absorption spectra. *Phys. Rev. B*, 76:195116.

- [Kas et al., 2009] J. J. Kas, J. Vinson, N. Trcera, D. Cabaret, E. L. Shirley, and J. J. Rehr (2009). Many-pole model of inelastic losses applied to calculations of XANES. JPCS, 190(1):012009.
- [Kohn and Sham, 1965] W. Kohn and L. J. Sham (1965). Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133–A1138.
- [Martin, 2004] R. Martin (2004). *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press.
- [Materlik et al., 1983] G. Materlik, J. E. Müller, and J. W. Wilkins (1983). L-Edge Absorption Spectra of the Rare Earths: Assessment of the Single-Particle Picture. *Phys. Rev. Lett.*, 50:267–270.
- [Moruzzi et al., 1978] V. Moruzzi, J. Janak, and A. Williams (1978). V Effects of spin polarization. In V. Moruzzi, J. Janak, and A. Williams, editors, *Calculated Electronic Properties* of Metals. Pergamon.
- [Müller et al., 1982] J. E. Müller, O. Jepsen, and J. W. Wilkins (1982). X-ray absorption spectra: K-edges of 3d transition metals, L-edges of 3d and 4d metals, and M-edges of palladium. Solid State Commun., 42(5):365–368.
- [Mustre de Leon et al., 1991] J. Leon, J. J. Rehr, S. I. Zabinsky, and R. C. Albers (1991). Ab initio curved-wave x-ray-absorption fine structure. Phys. Rev. B, 44(9):4146-4156.
- [Natoli et al., 1998] C. R. Natoli, C. Brouder, P. Sainctavit, J. Goulon, C. Goulon-Ginet, and A. Rogalev (1998). Calculation of X-ray natural circular dichroism. *Eur. Phys. J. B*, 4(1):1– 11.
- [O'Handley, 1999] R. O'Handley (1999). Modern Magnetic Materials: Principles and Applications. Wiley.
- [Perdew et al., 1996] J. P. Perdew, K. Burke, and M. Ernzerhof (1996). Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77(18):3865–3868.
- [Rehr and Albers, 2000] J. J. Rehr and R. C. Albers (2000). Theoretical approaches to x-ray absorption fine structure. *Rev. Mod. Phys.*, 72(3):621–654.
- [Rehr and Ankudinov, 2005] J. J. Rehr and A. L. Ankudinov (2005). Progress in the theory and interpretation of XANES. *Coord. Chem. Rev.*, 249(1-2):131-140.
- [Rogalev et al., 2010] A. Rogalev, J. Goulon, F. Wilhelm, and A. Bosak (2010). X-Ray Detected Optical Activity. In E. Beaurepaire, H. Bulou, F. Scheurer, and J.-P. Kappler, editors, *Magnetism and Synchrotron Radiation*, number 133 in Springer Proceedings in Physics, pages 169–190. Springer Berlin Heidelberg.
- [Rogalev et al., 2006] A. Rogalev, F. Wilhelm, N. Jaouen, J. Goulon, and J.-P. Kappler (2006).
   X-ray Magnetic Circular Dichroism: Historical Perspective and Recent Highlights, pages 71–93.
   Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Sainctavit et al., 1989] P. Sainctavit, J. Petiau, M. Benfatto, and C. Natoli (1989). Comparison between XAFS experiments and multiple-scattering calculations in silicon and zincblende. *Physica B: Condensed Matter*, 158(1):347 - 350.

- [Schütz et al., 1987] G. Schütz, W. Wagner, W. Wilhelm, P. Kienle, R. Zeller, R. Frahm, and G. Materlik (1987). Absorption of circularly polarized x rays in iron. *Phys. Rev. Lett.*, 58(7):737-740.
- [Sessoli et al., 2015] R. Sessoli, M.-E. Boulon, A. Caneschi, M. Mannini, F. Poggini, and A. Rogalev (2015). Strong magneto-chiral dichroism in a paramagnetic molecular helix observed by hard X-ray. *Nature physics*, 11(1):69.
- [Stadnicka et al., 1985] K. Stadnicka, A. M. Glazer, and J. R. L. Moxon (1985). The structural chirality and optical activity of  $\alpha$ -LiIO<sub>3</sub>. Journal of Applied Crystallography, 18(4):237–240.
- [Stöhr and Siegmann, 2006] J. Stöhr and H. C. Siegmann (2006). Magnetism, From Fundamentals to Nanoscale Dynamics. Springer Series in Solid-State Sciences. Springer Berlin Heidelberg New York.
- [Svensson et al., 1983] C. Svensson, J. Albertsson, R. Liminga, A. Kvick, and S. C. Abrahams (1983). Structural temperature dependence in  $\alpha$ -lithium iodate: Neutron and x-ray study between 20 and 500 K. J. Chem. Phys., 78(12):7343-7352.
- [Thole et al., 1992] B. T. Thole, P. Carra, F. Sette, and G. Laan (1992). X-ray circular dichroism as a probe of orbital magnetization. *Phys. Rev. Lett.*, 68(12):1943–1946.
- [Torchio et al., 2011] R. Torchio, Y. O. Kvashnin, S. Pascarelli, O. Mathon, C. Marini, L. Genovese, P. Bruno, G. Garbarino, A. Dewaele, F. Occelli, and P. Loubeyre (2011). X-Ray Magnetic Circular Dichroism Measurements in Ni up to 200 GPa: Resistant Ferromagnetism. *Phys. Rev. Lett.*, 107(23):237202.
- [Train et al., 2008] C. Train, R. Gheorghe, V. Krstic, L.-M. Chamoreau, N. Ovanesyan, G. L. J. A. Rikken, M. Gruselle, and M. Verdaguer (2008). Strong magneto-chiral dichroism in enantiopure chiral ferromagnets. *Nature materials*, 7:729–734.
- [Sipr and Ebert, 2005] O. Sipr and H. Ebert (2005). Theoretical Fe  $L_{2,3}$  and K-edge x-ray magnetic circular dichroism spectra of free iron clusters. *Phys. Rev. B*, 72(13):134406.
- [Wu and Freeman, 1994] R. Wu and A. J. Freeman (1994). Limitation of the Magnetic-Circular-Dichroism Spin Sum Rule for Transition Metals and Importance of the Magnetic Dipole Term. *Phys. Rev. Lett.*, 73:1994–1997.
- [Wu et al., 1993] R. Wu, D. Wang, and A. J. Freeman (1993). First principles investigation of the validity and range of applicability of the x-ray magnetic circular dichroism sum rule. *Phys. Rev. Lett.*, 71:3581–3584.

# Chapter 4

# XAS and XMCD to study matter under pressure

## Contents

| 4.1 | $\mathbf{Intr}$ | oduction $\ldots \ldots 124$   |
|-----|-----------------|--|
| 4.2 | Exp<br>bear     | erimental method: XMCD under pressure on a dispersive<br>nline   |
|     | 4.2.1           | High-pressure cell   |
|     | 4.2.2           | XAS and XMCD on the dispersive beamline ODE at SOLEIL 127  |
|     | 4.2.3           | Data analysis  |
| 4.3 | $\mathbf{Pres}$ |  |
|     | 4.3.1           | Lattice parameter $\dots \dots \dots$  |
|     | 4.3.2           | Evolution of the spectra with pressure $\ldots \ldots 133$  |
|     | 4.3.3           | Effect of pressure on the density of states $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 135$  |
| 4.4 | Iron            | hydride  |
|     | 4.4.1           | Fe under a ${\rm H}_2$ atmosphere: experiment  |
|     | 4.4.2           | Structure of FeH $\ldots$  |
|     | 4.4.3           | FeH: Calculated XAS and XMCD spectra $\hdots \hdots \hdot$ |
|     |                 | Computational details  |
|     |                 | Results  |
|     |                 | Confirmation of the observed transition  |
|     | 4.4.4           | Effect of the core-hole  |
|     | 4.4.5           | Effect of the position of the H atoms  |
|     |                 | Relaxation of the position of H $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 145$   |
|     |                 | Effect on the XAS and XMCD spectra $\ldots$  |
|     |                 | Projected density of states and magnetic moment per site $\ldots \ldots \ldots 147$  |
|     | 4.4.6           | Magnetic structure of FeH  |
|     | 4.4.7           | Increasing pressure in the calculation   |
|     |                 | Disappearance of ferromagnetism under pressure $\ldots \ldots \ldots \ldots \ldots 149$  |
|     |                 | Calculated spectra under pressure $\ldots \ldots 150$  |

| 4.4.8   | Conclusion $\ldots \ldots 151$                           |
|---------|--|
| 4.5 Chr | omium Dioxide  |
| 4.5.1   | Electronic properties of $CrO_2$   |
|         | Lattice  |
|         | Simplified crystal field picture   |
|         | Partial densities of states  |
|         | Double exchange mechanism  |
|         | Correlations effect in $CrO_2$   |
| 4.5.2   | $\label{eq:pre-edge} {\rm Pre-edge\ structure\ } \ldots\ $ |
| 4.5.3   | Magnetism under pressure in ${\rm CrO}_2$ studied by XMCD at the Cr $K\text{-edge159}$   |
|         | Experimental details $\ldots \ldots 159$          |
|         | XAS and XMCD at the Cr $K$ -edge   |
|         | ${\rm Preliminary\ calculations\ }\ldots\ \ldots\ .\ 161$                  |
| 4.5.4   | Conclusion   |

# 4.1 Introduction

Experimental studies under pressure are very important in the field of earth sciences and they also are of several fundamental interests. The main effect of pressure is to reduce the interatomic distances which can lead to new properties (for example photomagnetic properties in Prussian blue analogues [Cafun et al., 2013]). Pressure is widely used in condensed matter physics to uncover new properties (a well known example is high  $T_c$  superconductivity). Also, in the context of combined experimental and numerical studies, pressure is a very good way to put calculations to the test. Indeed, pressure modifies the interactions between the atoms and only a robust method can model it correctly.

In Section 4.2, we present the experimental set-up for acquisition of XAS and XMCD spectra under pressure on the dispersive beamline ODE in SOLEIL synchrotron (synchrotron facility situated in Gif-sur-Yvette, France). In Section 4.3, we discuss the way of introducing pressure in our DFT calculations. Sections 4.4 and 4.5 are dedicated respectively to the study of the iron hydride FeH and of the transition metal oxide  $CrO_2$ .

Interest for iron hydrides is important because they are model objects for the study of transition metal hydrides [Antonov et al., 2002]. XAS and XMCD experiments on Fe under  $H_2$  atmosphere are presented in Section 4.4. Calculations give a rich insight into the transitions that occur under pressure and into the magnetic structure of the formed compound, FeH.

Transition metal oxides are strongly correlated materials which exhibit a wealth of phenomena, so that they have a lot of potential (see for example [Tokura, 2003]). Their theoretical study is one of the challenges to modern condensed matter physics. Among them Chromium dioxide is a particularly interesting case because it belongs to the class of half-metallic ferromagnets. A RIXS experiment and an XAS and XMCD study under pressure of  $CrO_2$  are presented in Section 4.5. Preliminary calculations show promising results.

# 4.2 Experimental method: XMCD under pressure on a dispersive beamline

# 4.2.1 High-pressure cell

**Principle** Diamond anvil cells (DAC) are the most used devices to generate static high pressures. They are very portable devices (a whole DAC can be held in the palm of a hand) that allow to reach very high pressures (studies up to 320 GPa [Loubeyre1 et al., 2002] have been reported ). The working principle is illustrated in Fig. 4.1: the sample is positioned between two diamond anvils on which a force is applied. Each diamond anvil has two flat surfaces: a surface on which the force is applied and a smaller surface (typically several 100  $\mu$ m in diameter) where the anvil puts force on the sample. The pressure transferred to the sample is proportional to the applied force and also to the ratio of the two surfaces of the anvils (the smaller the surface in contact with the sample, the larger the pressure). A pressure transmitting medium is added with the sample that is positioned in the hole of a gasket. This gasket is a thin metal foil (typically 300  $\mu$ m thick). The gasket is prepared by pressing it between the two anvils, reducing locally its thickness to 15-50  $\mu$ m. Then, a hole is drilled by electrical discharge or laser machining at the center of the obtained indentation.



Fig. 4.1: Principle of a diamond anvil cell. The sample is positioned in the hole in the gasket between the two anvils.

The role of the transmitting medium, as its name suggests, is to transmit pressure to the sample. The aim is that the sample stays in hydrostatic conditions, that is to say that it is not submitted to a pressure gradient. In the Refs. [Klotz et al., 2009, Miletich et al., 2000], the maximum pressure that ensures hydrostaticity is given for several usual media. The paper by Miletich et al. also contains a lot of information on the practical aspects of diamond anvil cells.

Application of the force The two diamond anvils are only a small part of the entire DAC: the cell design must allow a precise positioning of the diamonds and include a pressurizing mechanism. For the experiments presented in this thesis, we used membrane diamond anvil cells designed by J.C. Chervin [Chervin et al., 1995]: the force on the diamonds is generated by deformation of a membrane by effect of a Helium gas (one diamond is fixed and the other is on a piston that is pushed by the membrane). The magnitude of the applied force can be tuned by changing the gas pressure on the membrane,  $P_m$ . During an experiment,  $P_m$  can be safely increased up to the value that was used to indent the gasket. Beyond this value, a continuous monitoring of the diamonds on each other, which would result in their destruction. It is also needed to plot the curve representing the pressure inside the cell *vs* the applied pressure on the membrane, because a change of its slope is an indication that the gasket hole is about to open.

**Pressure measuring** It is not possible to compute the pressure inside the cell as a function of the force applied on the external surface of the anvil because some force is dissipated, for example in the gasket, and it is impossible to know *a priori* how much. Therefore a pressure probe must be positioned in the gasket hole in order to monitor the pressure on the sample during the experiment. The first possible method is to add a material whose equation of states is known and to rely on the measurement of its lattice parameter to determine the pressure. It is done for example with gold [Pépin et al., 2014], copper or silver [Mao et al., 1986]. This method can, however, only be used during diffraction experiments. The second possible method is to rely on the measurement of the fluorescence line of a chemically inert luminescent compounds. The most widely used pressure calibrant is ruby even if others exist [Hess and Exarhos, 1989, Datchi et al., 1997]. Small rubies (5-10  $\mu$ m) for the purpose of pressure monitoring are commercially available. Before the experiments, a few rubies are positioned along with the sample (if possible, two rubies in separate positions in order to be able to evaluate if there is a pressure gradient).



Fig. 4.2: Example of the ruby fluorescence spectrum at several pressures (figure taken from [Jephcoat et al., 1986]). Note that the spectrum are shifted in the horizontal direction. The wavelength position of peak  $R_1$  is indicated for each pressure.

The pressure measurement is performed by irradiating the sample with a blue laser and by analyzing the fluorescence of a ruby with a spectrometer. The position of the main peak of this fluorescence shifts to higher wavelengths (see Fig 4.2) when pressure increases. The relation between this peak position and pressure is calibrated [Syassen, 2008] which allows to deduce the pressure from the measured peak position. As shown in Fig 4.2 the fluorescence spectrum is broadened at high pressure, which reduces the accuracy of the measurement.

In practice, the sample pressure is measured before (just after the membrane pressure is changed) and after the recording of the spectra because pressure sometimes changes during the acquisition time. Despite the precise calibration of the ruby line, the uncertainty on the pressure is quite high because of this drift and also because the sample and the ruby may not be exposed to the same stress field. The precision depends on many factors (size of the diamonds, type of

gasket, size of the hole...) but it seems reasonable to consider that the uncertainty is  $\approx 0.5$  GPa for pressures up to 10 GPa and  $\approx 1$  GPa for pressures up to 30 GPa. For the experiments presented in this chapter, the pressure is never higher than 30 GPa but much higher pressures can be reached in DAC: pressures up to 100 GPa are available on ODE beamline at SOLEIL.

# 4.2.2 XAS and XMCD on the dispersive beamline ODE at SOLEIL

SOLEIL is a synchrotron radiation facility situated 30 kilometers south of Paris. The principle of such facility is to use the radiation (called synchrotron radiation) emitted by electrons moving at relativistic velocity in magnetic fields that curve their trajectory. The main components of the facility are depicted in Fig. 4.3.



Fig. 4.3: Illustration of the main components of SOLEIL synchrotron radiation facility: 1) the linac, a 16 meter long linear accelator 2) the booster, a circular accelerator which brings the energy level up to 2.75 GeV (SOLEIL operating value) 3) the storage ring (354 meter circumference) in which the electrons circulate for several hours after being injected from the booster 4) magnetic devices that control the trajectory of the electrons or make them oscillate (bending magnets or insertion devices) which results in the emission of synchrotron light 5) radiofrequency cavities to compensate for the energy lost by radiation 6) optic systems to select and shape the synchrotron radiation for the use of the beamlines.

Picture is taken from [Soleil, site]

The storage ring is composed of a combination of magnetic field devices (to change the electron beam's direction and produce the radiation) and electric field devices (to compensate for energy losses of the electrons and accelerate them). The beamlines are situated tangentially of this storage ring at the level of a magnetic device in order to collect the emitted light.

In SOLEIL the beamline ODE is a hard X-ray (3.5-25 keV) beamline for XAS and XMCD. The acronym ODE stands for "Optique Dispersive EXAFS" (Dispersive optics EXAFS).

**Source** The source of ODE is a bending magnet. It is a dipole magnet whose purpose is to bend the trajectory of the electrons (see Fig. 4.4).



**Fig. 4.4:** Bending magnet. Picture is taken from [Desy, site].

On its horizontal plane, the emitted radiation is linearly polarized. Above and below this plane, it is respectively elliptically right and left polarized [Wiedemann, 2007, p.852]. To perform XMCD experiments, a portion of this beam above or below the plane is selected and the

rate of circular polarization depends on the position and on the size of the slit used for this purpose.

**Dispersive setup** ODE is a dispersive beamline, it means that polychromatic X-rays are used which allows for the instant recording of the whole spectra. The principle is illustrated in Fig. 4.5. The X-rays provided by the source are focused by a bent monochromator so that all the diffracted beam converges to the same point (focal point, where the sample is positioned). This crystal is called polychromator because as the incident angle varies along the crystal, the diffracted wavelength also varies. This causes a spatial separation of the wavelength (hence the designation *dispersive*) and all the diffracted X-rays converge to the sample but with different angles depending on their energy.



Fig. 4.5: Principle of the energy dispersive EXAFS setup. The sample is positioned between the polychromator and the detector (red circle). Picture is taken from [Diamond, site].

The spatial separation of the energies after the light is transmitted through the sample permits to measure simultaneously the transmitted intensity I for all energies using a CCD detector that is position sensitive. The reference intensity  $I_0$  is measured in the absence of sample and the absorption spectrum is:

XAS(pixel number) 
$$\propto -\log \frac{I(\text{pixel number})}{I_0(\text{pixel number})}$$
.

To calibrate the link between the pixel number and the energy of the X-rays the spectrum XAS(pixel number) of a metal foil or another reference sample is measured at the beginning of each experiment. Then this spectrum is fitted with tabulated data.

In addition to the elements depicted in Fig. 4.5, there is also a bent mirror before the polychromator and another plane mirror between the sample and the CCD for vertical focalization and harmonic rejections respectively. To set-up the beamline to the energy edge of a specific element, the whole bench of experiment (that carries the sample and the detector) is rotated to be placed in the approximate angle corresponding to the chosen edge. After that, fine adjustments of the mirror and polychromator curvatures are required.

On ODE beamline, after a precise setting up of the beamline, the size of the focal spot is approximately 35  $\mu$ m in diameter (full width at half maximum). As there are no mechanical movements of the polychromator, the position of the focal point is very stable in time so that very small samples (70  $\mu$ m) can be studied. This is particularly interesting for high-pressure measurements because the smaller samples, the higher the pressure that can be reached.

Another advantage of the dispersive set-up is that the acquisition time of a XAS is very short, which offers perspective to study fast processes over time. Such time resolved experiments (with a few dozen  $\mu$ s resolution) are possible on ODE beamline.

Magnetic field and XMCD experiments The polarization of the light cannot be changed from right- to left-circularly polarized during the experiment: it would require a complete set up of the beamline each time it is done. For this reason, the applied magnetic field is flipped instead.

The magnetic field is generated by a big electromagnet (weighing several 100 kilograms) and can reach 2.1 T with a small air gap. When the sample is in the cryostat the air gap cannot be smaller than 46 mm and the maximum magnetic field is 1.3 T. The direction of the field is flipped by changing the sense of the current in the coils.

XAS spectra are alternatively measured with a magnetic field applied in one direction or the other:

| magnetic field direction | +       | -       | +       | -       | <br>+       |
|--------------------------|---------|---------|---------|---------|-------------|
| spectrum                 | $I_0^+$ | $I_1^-$ | $I_1^+$ | $I_2^-$ | <br>$I_N^+$ |

The aim of the iteration 0 is to cancel the possible linear derivation  $\epsilon$  due to beam variations :

After N iterations, the saved XMCD is:

$$\text{XMCD}_{N} = \frac{1}{2N} \log \frac{I_{0}^{+}(I_{1}^{+})^{2} \dots (I_{N-1}^{+})^{2} I_{N}^{+}}{(I_{1}^{-})^{2} \dots (I_{N}^{-})^{2}}$$

It has become less useful in SOLEIL (compared to previous generation synchrotron facilities) because the beam does not drift.

**Low temperature** A cryostat designed to receive diamond anvil cells is available on ODE beamline. The sample is cooled thanks to a flow of Helium and the temperature is controlled by a heater. Temperatures down to 5K are accessible.

#### 4.2.3 Data analysis

One drawback of performing experiments with a dispersive set-up is that the data processing can be complicated by the fact that the baseline of the data is not necessarily flat.



**Fig. 4.6:** Comparison of two types of processing for an experimental XAS with a non-straight baseline. On the left-hand side, the base line is determined manually (by comparison with another spectrum whose baseline was straight). The obtained spectrum is depicted in black in the last figure. On the right hand side, the software Athena [Ravel and Newville, 2005] is used with an order 2 polynomial for the post-edge and application of the flattening algorithm. In both cases, the pre-edge baseline is set to a straight line and the edge step is determined to be 1.71. The spectrum used as example is one of the spectra acquired during the experiments described section 4.4.1.

The required steps for data analysis are however the same as for any XAS and XMCD experiments:

- set the pre-edge of the XAS spectrum to zero (substract the pre-edge baseline)
- flatten the post-edge region of the XAS
- normalize the XAS so that the edge jump far from the edge is equal to 1
- multiply the XMCD signal by the same factor to obtain the normalized XMCD.

A software called Athena [Ravel and Newville, 2005] is available that facilitates the analysis of XAS spectra. The pre-edge baseline is treated as a straight line (see Fig. 4.6 top right). The order of the polynomial used to describe the post-edge baseline can be chosen between 1 and 3. Athena includes a flattening algorithm that aims at pushing the oscillation of the post-edge part along the y = 1 line.

Fig. 4.6 illustrates the processing of a XAS spectrum of which the post-edge baseline is not flat. Two methods are compared. The first method ("Manual processing") consists in determining the baseline by visual evaluation and substrating it with a data processing software (here, OriginPro was used). A comparison with a similar spectrum that serves as reference can be useful to set the baseline. The second method relies on Athena algorithm. The main advantage of this second method is its reproducibility associated to its automatic nature. However, it does not exempt from visually inspecting the spectrum as the algorithm might, in some situations, lead to inconsistent baselines.

The processing of XMCD spectra acquired on dispersive beamlines also requires to subtract the baseline. A software that would allow an automatic subtraction of the baseline for XMCD is currently being developped by Karine Provost at ICMPE. The great difficulty of the task is linked to the fact that the XMCD spectra can not be treated as oscillations of zero mean-value. It prevents the use of the algorithms that were developed to treat XAS spectra. We must therefore rely on the manual processing of the spectra. An example of such a processing is shown Fig 4.7. Such a treatment was applied to the XMCD spectra acquired on ODE beamline that are presented later in this chapter.



Fig. 4.7: Example of the normalization and removal of a non-straight baseline of a XMCD spectrum (one of the spectra acquired during the experiments described section 4.4.1).
# 4.3 Introduction of pressure in the calculation: example of *bcc*-Fe

## 4.3.1 Lattice parameter

In the calculation, the effect of pressure is taken into account simply by reducing the lattice parameters of the crystal considered in the self-consistent calculation. There are two ways of determining the pressure for a given lattice parameter from the calculations. The first way is two compute the forces in the cell; there is a keyword of the input of pw.x that allows to do so (tprnfor). The second method consists in plotting the total energy at the end of the self-consistent calculation E as a function of the lattice parameter a. Then, if we denote V the volume of the cell (in the case of *bcc*-Fe  $V = a^3$ ), the pressure writes

$$P(V) = -\left(\frac{\mathrm{d}E}{\mathrm{d}V}\right)_S.$$

The equation of states depicted in Fig. 4.8 on the left-hand side in pink was obtained by performing *scf* calculations for each lattice parameter. Then, E(V) was fitted with Murnaghan [Murnaghan, 1944] equation of states E(V) using the tool ev.x of QUANTUM ESPRESSO:

$$E(V) = E_0 + \frac{B_0 V}{B'_0} \left[ \frac{1}{B'_0 - 1} \left( \frac{V_0}{V} \right)^{B'_0} + 1 \right] - \frac{B_0 V_0}{B'_0 - 1}$$
(4.1)

where  $V_0$  and  $E_0$  are the equilibrium volume and energy at zero pressure.  $B_0 = -V_0 \frac{\partial P}{\partial V}$  is the equilibrium bulk modulus and  $B'_0$  is its first derivative with respect to pressure. The fitted parameters are:  $V_0 = 22.5 \text{ Å}^3$  (which corresponds to  $a_0 = 2.82 \text{ Å}$ ),  $B_0 = 180$  GPa and  $B'_0 = 7$ . Then the pressure for a given lattice parameter can be computed with the formula:

$$P(V) = \frac{B_0}{B'_0} \left[ \left( \frac{V}{V_0} \right)^{-B'_0} - 1 \right].$$

The relation P(a) obtained with this method is reproduced in the right panel of Fig. 4.8 (purple line). On this same plot, the pressures obtained by force calculation (with tprnfor) are represented as diamond scatters. We observe that these two methods give slightly different results.

The blue dots in Fig. 4.8, right correspond to the points P(a) obtained experimentally [Mao et al., 1967, Jephcoat et al., 1986] by performing X-ray diffraction experiments in DAC. The experimental points are fitted with Vinet exponential equation of state:<sup>1</sup>

$$P(V) = 3B_0 \left(\frac{V}{V_0}\right)^{-2/3} \left[1 - \left(\frac{V}{V_0}\right)^{1/3}\right] \exp\left[-\frac{3}{2}(B'_0 - 1)\left[1 - \left(\frac{V}{V_0}\right)^{1/3}\right]\right]$$
(4.2)

and the obtained parameters are:  $V_0 = 23.6\text{\AA}^3$   $(a_0 = 2.87 \text{\AA}), B_0 = 150 \text{ GPa}, B'_0 = 6.5$ . We observe a 2 % underestimation of the equilibrium lattice parameter by the calculation which is

<sup>&</sup>lt;sup>1</sup>Several forms (Murnaghan, third-order Birch-Murnaghan, Vinet...) have been proposed as universal equations of state. The choices made here of Murnaghan and Vinet E.of S. were only driven by the fact that they are commonly used.



**Fig. 4.8:** On the left: Calculated equation of states (E. of S.) energy *vs* lattice parameter for *bcc*-Fe and pressure deduced from this E. of S. On the right: Comparison of the curves pressure *vs* lattice parameters obtained experimentally ( [Jephcoat et al., 1986, Mao et al., 1967] fitted with Vinet equation) in blue, from the calculated E. of S. in purple and from the calculation of the force in the cell (diamond scatters).

consistent with the usual order of magnitude of the error on the  $a_0$  obtained by DFT calculations with PBE functionnal (see for example [Haas et al., 2009]).

From Fig. 4.8 we see that, to describe a given pressure P we have two choices of lattice parameter:  $a_{\exp}(P)$  that corresponds to the experimental lattice parameter (blue curve) and  $a_{calc}(P)$  that corresponds to the calculated one (purple curve).

In the following, for all the systems, we will use the experimental equation of state  $a_{\exp}(P)$ . The main reason of this choice is heuristic: as we want to reproduce the experimental spectra, it seems reasonable to consider the experimental lattice parameter and not a calculated lattice parameter that depends on the used functional. According to our observations, the effect of this choice on the spectra consists mainly in a slight change of the amplitude of the main XMCD peaks (of the order of 10%).

#### 4.3.2 Evolution of the spectra with pressure

As a pressure-induced bcc to hcp transition occurs in iron between 14 and 16 GPa, the range in which the effect of pressure of the bcc phase can be experimentally studied is limited.

When pressure is applied on an iron foil in a diamond anvil cell, the experiments by Mathon et al. [Mathon et al., 2004, Baudelet et al., 2005] illustrate that the XAS spectrum shows little changes up to 14 GPa: its features are only slightly shifted to higher energies (see Fig. 4.9) due to the reduction of interatomic distances. In the range of pressure where the structure is still *bcc* (information provided by the XAS) the amplitude of the XMCD spectra is almost constant with pressure (see Fig. 4.10 circles).

In Fig. 4.10, the calculated total magnetization in a unit cell of *bcc*-Fe is also plotted as a function of pressure (the computational details of the self-consistent calculations are given in the next section p.135). This total magnetization does not vary much in the range 3 GPa-14 GPa in accordance with the conclusion drawn in [Baudelet et al., 2005] from the fact that XMCD amplitude is almost constant.

There is no point for the experimental XMCD amplitude at 0 GPa in the figure from which



Fig. 4.9: Experimental XAS (left) and XMCD (right) spectra for Fe at ambient pressure (red curve) and 10 GPa (blue curve) from [Baudelet et al., 2005].

the circles of Fig. 4.10 are extracted [Mathon et al., 2004]. It is somewhat unfortunate because the main variation of the calculated total magnetization occurs between 0 and 3 GPa but it is probably linked to a problem inherent to experiments with diamond anvil cells: when a DAC is loaded with a pressure medium, merely closing it is sufficient for the pressure applied on the sample to be non zero. To record the spectra below 3 GPa, experiments with larger diamonds or an other type of pressure cell should be undertaken.



Fig. 4.10: Calculated total magnetic moment in a unit bcc Fe cell (squares) and amplitude of the experimental XMCD for bcc-Fe (circles). The points for the amplitude of XMCD are taken from [Mathon et al., 2004, Fig. 2] in the range in which the fraction of the bcc phase is 100% (i.e. before the beginning of the bcc/hcp phase transition). Error bars for the calculated total magnetization are not displayed because they are difficult to estimate.

By the mean of calculation, larger pressures can be set by decreasing the lattice parameter to values that are not accessible experimentally. The only limit is the point where the scf

calculation can not converge because the forces in the cell are too large. At 42 GPa, the calculated total magnetization is 94% of the total magnetization at ambient pressure (pink squares in Fig. 4.10).

The calculated XAS and XMCD spectra at the Fe K-edge (with a 1s core hole in a supercell) for several pressures are shown in Fig. 4.11. The consequence of the increase in pressure on the XAS is a shift of the features above the edge toward higher energies. The same trend is seen on the XMCD spectra. Except for P=10 GPa, the main other effect of pressure on XMCD is a decrease of the amplitude of the main negative peak. The total amplitude of the XMCD spectra at 42 GPa is 89 % of the amplitude at ambient pressure. At P=10 GPa, a small negative peak near the edge appears that seems to correspond to a negative feature present on the experimental spectra (Fig. 4.9- right part). The level of noise on the experimental spectra is high and this negative peak never appeared on any other experimental spectra, so the peak that appears in the calculations is probably a numerical artifact.



Fig. 4.11: Calculated XMCD and XAS spectra at *K*-edge in *bcc*-Fe with different lattice parameters: 2.87 Å (ambient pressure), 2.814 Å (P=10 GPa) and 2.71 Å (P=42 GPa from extrapolation of the equation of state). The calculation has been performed in the presence of a core-hole with technical parameters identical to the one used for Fe in the previous chapter.

## 4.3.3 Effect of pressure on the density of states

**Computational details** The densities of states (DOS) were computed with the regular code pw.x of QUANTUM ESPRESSO [Giannozzi et al., 2009] in a unit cell of *bcc*-Fe with the same norm-conserving PBE pseudopotentials used for spectra calculation. The cut-off energy was set to  $E_{cutoff} = 180$  Ry. The self-consistent calculation was performed using a  $\Gamma$ -centered  $16 \times 16 \times 16$  k-point grid for the unit cell that contains one atom. To compute the densities of states a second calculation including the unoccupied bands is performed starting from the already converged electron density. This non-self-consistent calculation was performed using a  $\Gamma$ -centered  $40 \times 40 \times 40$  k-point grid. The partial density of states are obtained by projecting wave functions onto orthogonalized atomic wave functions [Giannozzi et al., 2009]. The lattice parameter was set to a = 2.87 Å for ambient pressure, a = 2.814 Å for P=10 GPa and a = 2.71 Å for P=42 GPa. This last value was obtained from extrapolation of the experimental equation of state (in reality, at P=42 GPa, *bcc*-Fe is not stable). Collinear spins along [001]

were considered. Spin-up  $(\uparrow)$  denotes the majority spin and spin-down  $(\downarrow)$  the minority spin. The zero energy is taken at the Fermi level.

**Stoner criterion** The simplest model of itinerant magnetism is Hubbard model within the mean field approximation. The Hubbard Hamiltonian in second quantization formalism writes:

$$H = -t \sum_{i,\delta,s=\uparrow,\downarrow} c_{i,s}^{\dagger} c_{i+\delta,s} + U \sum_{i} n_{i}^{\uparrow} n_{i}^{\downarrow}$$

$$\tag{4.3}$$

with  $n_i^s = c_{i,s}^{\dagger} c_{i,s}$ . *i* labels the sites and  $i + \delta$  are the sites surrounding *i*,  $n_i$  is the occupation of a site. *t* is called the hopping integral and describes the ability of an electron to hop from one site to another one and *U* describes the intra-site Coulomb interaction. The local electronic density is  $n_i = n_i^{\uparrow} + n_i^{\downarrow}$  and the local spin density is  $m_i = \frac{n_i^{\uparrow} - n_i^{\downarrow}}{2}$ . The Stoner criterion is a criterion for the ferromagnetic state to be stable that is to say

The Stoner criterion is a criterion for the ferromagnetic state to be stable that is to say for an unbalance between spin-up and spin-down to be energetically favourable. As illustrated Fig. 4.12 a, in the case where U = 0, if  $n^{\uparrow} > n^{\downarrow}$ , it has an energetic cost because higher energy levels are occupied. For example, if  $n^{\uparrow} = n^{\downarrow} + \delta n$ , the cost in energy  $\Delta K = \frac{\rho(E_f)}{2} (\delta E)^2$  where  $\rho(E_f)$  is the total density at the Fermi level and  $\delta E = \frac{\delta n}{\rho(E_f)}$ . In the case U = 0, the material remains paramagnetic.



Fig. 4.12: Band ferromagnetism: schematic representation of the spin-up and spin-down density of states in a) the paramagnetic case, b the ferromagnetic case.

Now, if  $U \neq 0$ , an imbalance between spin-up and spin-down induces a Coulomb gain in the Hamiltonian of Eq.(4.3)

$$\Delta H_U = U\left(\frac{n}{2} + \frac{\delta n}{2}\right) \left(\frac{n}{2} - \frac{\delta n}{2}\right) - U\frac{n}{2}\frac{n}{2} = -U\left(\frac{\delta n}{2}\right)^2.$$

The criterion for ferromagnetism to exist in a given material at T = 0K is therefore:

$$|\Delta H_U| \ge \Delta K \Rightarrow U \ge \frac{2}{\rho(E_f)}.$$

A ferromagnetic instability is therefore expected in materials with a high density of states at Fermi level and large electron interactions. Due to the Coulomb energy, the minority-spin band shifts upward whereas the majority-spin band shifts downward as illustrated in Fig. 4.12 b).



Fig. 4.13: Partial densities of state of *bcc*-Fe with different lattice parameter: 2.87 Å (ambient pressure), 2.814 Å (P=10 GPa) and 2.71 Å (P=42 GPa from extrapolation of the equation of state). Top half of the graphs: spin-up DOS; lower half of the graphs: the opposite of the spin-down DOS.

The partial DOS of *bcc*-Fe with several lattice parameters are depicted Density of states in Fig. 4.13. The DOS of bcc-Fe is dominated by d electrons. There is a spontaneous spin split of the bands that is associated to the existence of a spin density. When the lattice parameter is decreased, the DOS widens which results in a decrease of the density at the Fermi level  $\rho(E_F)$ . This decrease is consistent with the decrease of the magnetic moment described in the previous subsection according to the Stoner criterion (see below). This qualitative interpretation is to be pursued with caution because U can a priori depend on the lattice parameters but it gives an idea of the mechanism at work when pressure increases: due to the DOS broadening the product  $U\rho(E_f)$  decreases which leads to a progressive cancellation of the Curie temperature which tends to zero when  $\frac{1}{2}U\rho(E_f) = 1$  [du Trémolet de Lacheisserie et al., 2005]. According to this mechanism, the magnetic moment of all materials tends to vanish at high pressure. In practice, as in the case of Fe, a structural transition often happens before the magnetic moment reaches zero. In Ni no structural transition occurs and ferromagnetism has been shown to maintain up to 200 GPa [Torchio et al., 2011], so an even higher pressure would be required to reach the disappearance of magnetism.

## 4.4 Iron hydride

The study of the combination of iron and hydrogen under pressure presents a strong geophysical interest. Iron is the main constituent of the Earth's core but, according to seismic models, the density of the core is several percent lower than the density of pure iron. This density deficit is attributed to the dissolution of light elements such as silicon, sulfur, oxygen, hydrogen, and carbon [Birch, 1952]. Hydrogen has become one of the major candidates for the light element in the Earth's core with the observation of its solubility in Fe under high pressure conditions [Sakamaki et al., 2009].

The understanding of iron hydrides is also essential for studying the behaviour of hydrogen in transition metals. Interest for this question has grown because the absorption of hydrogen by metals or alloys is a promising solution to address the problem of storage of hydrogen. This is crucial to allow its use as fuel [Schlapbach and Zuttel, 2001].

### 4.4.1 Fe under a $H_2$ atmosphere: experiment

A tiny foil of Fe was loaded in a diamond anvil cell with hydrogen  $(H_2)$  as a pressure medium.<sup>2</sup> The XAS and XMCD spectra at the Fe *K*-edge were measured simultaneously for several pressures from 2.7 GPa to 28 GPa and back to 2.9 GPa. Pressure was measured using the fluorescence line of a ruby.

**Evolution of the spectrum up to 4 GPa** Fig. 4.14 shows the measured absorption spectra up to 4 GPa. The low pressure spectrum corresponds to the XAS of *bcc*-Fe. We observe a drastic change of the spectrum around 3 GPa. The spectra for P > 3 GPa are completely different from the spectra for P < 3 GPa at high energy (above 7.14 keV i.e. 30 eV above the edge) and also near the edge. The main changes near the edge are a decrease of shoulders at 7115 eV and 7125 eV (a and b in Fig.4.14) and a shift to lower energy ( $\approx 1eV$ ) of the maximum at 7131 eV accompanied by an enhancement of this maximum (c in Fig.4.14).



Fig. 4.14: Experimental XAS spectra at the Fe K-edge for Fe under H<sub>2</sub> atmosphere of increasing pressure up to 4 GPa.

The measured XMCD (Fig. 4.15) also exhibits a transition around 3GPa with more gradual changes than the XAS. The main positive peak ( $\alpha$  in Fig. 4.15) decreases in amplitude. The amplitude of the main negative peak ( $\beta$  in Fig. 4.15), on the other hand, increases and it becomes sharper. Peak  $\beta$  is also clearly shifted to lower energies (from 7118.5 eV at 2.7 GPa to 7116 eV at 4 GPa) while the energy position of peak  $\alpha$  barely varies (from 7113 eV to 7112.5 eV).

<sup>&</sup>lt;sup>2</sup>When  $H_2$  is compressed, it remains fluid only up to P=5.7 GPa, yet hydrostaticity is preserved up to 177 GPa [Miletich et al., 2000, Table 1]



Fig. 4.15: Experimental XMCD spectra at the Fe K-edge for Fe under H<sub>2</sub> atmosphere of increasing pressure up to 4 GPa. Note that the energy range is not the same as on Fig. 4.14 in order for the region of interest to be more visible (the experimental spectra were not corrected for the circular polarization rate of the light  $P_c < 1$ ).

The pressure at which the changes are observed corresponds to the pressure of a transition reported in the literature from *bcc*-Fe to iron hydride FeH<sub>x</sub> [Badding et al., 1991, Choe et al., 1991]: hydrogen can be absorbed into the iron lattice by application of a 3.5 GPa pressure in a H<sub>2</sub> environment. In the experiment presented here, the H<sub>2</sub> fluid is, therefore, a source of hydrogen in addition to be a pressure-transmitting medium. The fraction of H atoms in the obtained hydride was estimated by outgassing to be  $x \approx 1$  [Schneider et al., 1991]. Magnetometry on samples of FeH<sub>x</sub> obtained by hydrogenation of Fe at high temperature (T=523 K) and high H<sub>2</sub> pressure (P=6.7 GPa) quenched to 95 K showed that, as *bcc*-Fe, it is a ferromagnet with a magnetic moment of approximately 2.2  $\mu_B$  per Fe atom [Antonov et al., 1981].<sup>3</sup>

Evolution of the spectra in the range 4 GPa - 28 GPa After the transition was achieved, the pressure on the sample was further increased in order to study the effect of pressure on iron hydride. Above 4 GPa, the only change on the XAS spectra is a progressive shift of the features to higher energies as the sample is compressed (see Fig. 4.16). The spectra at 28 GPa is different: the shoulders a and b are enhanced and the maximum c is reduced. In the literature, no change of the H content nor of the structure of the iron hydride was observed by XRD with compression up to 136 GPa [Pépin et al., 2014] of Fe under H<sub>2</sub> atmosphere.

The main effect of pressure on the XMCD spectrum is a slight decrease in amplitude (see Fig. 4.16). The maximum of the effect is also progressively shifted to higher energies. XMCD vanishes between 23 and 28 GPa, which is consistent with the observation made by Mössbauer spectroscopy [Mitsui and Hi, 2010] and previous XMCD experiments [Ishimatsu et al., 2012] that magnetic ordering vanishes at pressures higher than approximately 27 GPa.

**Reversibility of the transition** When the pressure is decreased, the spectra corresponding to *bcc*-Fe are recovered between 4 and 2.5 GPa. In other words, the pressure-induced transition that has been observed is reversible and the transition pressure is approximately the same in both ways (when the pressure is decreased as when the pressure is increased).

<sup>&</sup>lt;sup>3</sup>Note that the stoichiometry of the quenched samples in the study [Antonov et al., 1981] -  $x \approx 0.8$  - is different from the stoichiometry evaluated for the sample under pressure probably due to hydrogen losses during manipulations.



Fig. 4.16: Experimental XAS and XMCD spectra for FeH under H<sub>2</sub> atmosphere for several values of increasing pressure (the experimental XMCD spectra were not corrected for the circular polarization rate of the light  $P_c < 1$ ).

## 4.4.2 Structure of FeH



Fig. 4.17: Crystal structure of *dhcp*-FeH [Antonov et al., 1998]. The iron atoms are depicted in two different colors to emphasize the fact that they occupy non-equivalent sites. The arrows on the H atoms represent the direction of their displacement with respect to the center of the octahedral site. The letters A, B and C are the usual notations for close-packed layers in compact structures.

FeH has a double *hcp* lattice (*dhcp*) [Badding et al., 1991, Antonov et al., 1998]. The stacking sequence of closest packed layers follows the pattern ABAC (see Fig. 4.17) where A, B and C are the usual letters to represent the three possible orientations of the layers. There are two non equivalent crystallographic sites that contain Fe atoms. The symmetry of one of these sites is trigonal ((2*a*):  $\overline{3}$ m) and the symmetry of the other site is hexagonal ((2*c*):  $\overline{6}$ m2). H atoms are positioned in the interlayer. There is no symmetry constraint on the *z*-coordinates of the H atoms and, because of the low scattering cross section of H, it is not possible to determine the exact position of the H atoms using XRD. Neutron diffraction experiments on quenched iron deuterides (FeD) samples [Antonov et al., 1998] showed that the D atoms are not exactly positioned in the octahedral sites. Their displacement with respect to to the center of the site is represented by arrows in Fig. 4.17. The lattice parameters, atomic coordinates and bond length of dhcp-FeH at low pressure (3 GPa) are summarized in Table 4.1.

**Table 4.1:** *dhcp*-FeH lattice parameters and atomic positions. The position of the hydrogen atoms is given in three cases: exactly at the center of the octahedral site ("Oct. site"), at the position determined experimentally in FeD samples ("Exp. [Antonov et al., 1998]") and at the position obtained from structural relaxation ("Calc.", see section 4.4.5). The corresponding Fe-H bond length is given in these three cases.

| $dhcp$ -FeH P6 $_3/mmc$ , 3 GPa                   |              |           |            |                            |       |  |  |  |
|---|--------------|-----------|------------|----------------------------|-------|--|--|--|
| Lattice parar                                     | neters [Bade | ding et a | al., 1991] |                            |       |  |  |  |
| a (Å)   | 2.68         |           |            |                            |       |  |  |  |
| c (Å)   | 8.76         |           |            |                            |       |  |  |  |
| Volume $(\text{\AA}^3)$                           | 54.50        |           |            |                            |       |  |  |  |
| Atomic coord                                      | linates      |           |            |                            |       |  |  |  |
|   | x            | y         |            | z                          |       |  |  |  |
| Fe $(2a)$   | 0            | 0         |            | 0                          |       |  |  |  |
| Fe $(2c)$   | 1/3          | 2/3       |            | 1/4                        |       |  |  |  |
|   |              |           | Oct. site  | Exp.[Antonov et al., 1998] | Calc. |  |  |  |
| Н   | 1/3          | 2/3       | 0.875      | 0.882                      | 0.880 |  |  |  |
| Bond length (Å)                                   |              |           |            |                            |       |  |  |  |
|   | Oct. site    | Exp.      | Calc.      |                            |       |  |  |  |
| Fe(2a)-H  | 1.89         | 1.86      | 1.87       |                            |       |  |  |  |
| $\frac{\operatorname{Fe}(2c)\operatorname{-H}}{}$ | 1.89         | 1.93      | 1.92       |                            |       |  |  |  |

The volume of the cell as a function of pressure was determined experimentally in 1991 [Badding et al., 1991] and fitted with the Vinet equation of state (Eq.4.2 p.132). The obtained parameters were:  $V_0 = 55.6 \text{ Å}^3$ ,  $B_0 = 121$  GPa and  $B'_0 = 5.31$ . This relationship is plotted in Fig. 4.18 in comparison with the equation of state of *bcc*-Fe. The volume expansion due to the internalization of the H atoms is clearly visible on this figure. It amounts to 17% of the initial volume of iron.



Fig. 4.18: Relation between volume and pressure for *dhcp*-FeH (black curve) determined from the fit of experimental point [Badding et al., 1991]. For comparison the volume of *bcc*-Fe (see Fig. 4.8) was reproduced multiplied by two because the number of Fe atoms in the *bcc* cell is two while it is four in the *dhcp* cell. Inset: Lattice parameter *a* as a function of pressure in *dhcp*-FeH determined from the experimental equation of state [Badding et al., 1991] by setting the ratio c/a to  $2 \times 1.637$  [Antonov et al., 1998]. From the equation of state of [Badding et al., 1991] and by setting the ratio c/a to  $2 \times 1.637$  [Antonov et al., 1998], we obtain the lattice parameter a as a function of the pressure (the volume of the dhcp cell  $V = \frac{\sqrt{3}a^2c}{2}$ ) which is plotted in inset in Fig. 4.18.

## 4.4.3 FeH: Calculated XAS and XMCD spectra

#### **Computational details**

For the calculation, the experimental lattice parameters listed in Table 4.1 were used. The H atoms were positioned exactly at the center of the octahedral sites (see Section 4.4.5 for a discussion on this approximation).

A 1s static core hole was included in the pseudopotential for the absorbing Fe atom. A  $3\times3\times1$  supercell was built (72 atoms in the supercell among which 36 Fe atoms) so that the distance between periodically repeated core-holes was 8.0 Å. PBE norm-conserving pseudopotentials (parameters detailed in Appendix E) were used with a cutoff energy of 180 Ry. A Methfessel-Paxton smearing of 0.14 eV (0.01 Ry) and a  $\Gamma$ -centered  $4\times4\times4$  k-point grid were used for the self-consistent charge density calculation. A  $9\times9\times9$  k-point grid was used for spectra calculations. The spectra for the absorbing atom in site (2a) and (2c) were calculated.

Calculations were performed with collinear spins along the axis [001] (z in Fig. 4.17) within the diagonal spin-orbit coupling approximation (see p. 54). The wavevector  $\mathbf{k}$  was also set along the axis z.

The spectra were convolved with the same Lorentzian broadening function that was used for metal Fe (depicted in Fig. 3.4). No additional Gaussian broadening was added to account for the experimental resolution. The calculated spectra were normalized such that the edge jump in absorption is equal to 1.

#### Results

The self-consistent calculation converges to a ferromagnetic structure with a total magnetization of 80  $\mu$ B per supercell, i.e. 2.2  $\mu$ B per Fe atom, which is consistent with the magnetometry experiments [Antonov et al., 1981].

In Fig 4.19, the calculated XAS and XMCD spectra in FeH at the K-edge of Fe with the absorbing atom either on site (2a) or on site (2c) are depicted. The spectra for the two sites are very different. In particular the XMCD signal is mainly composed of two negative peaks for site (2a) and of a positive peak followed by a negative peak for the site (2c). On the right side of Fig 4.19, the contributions to XMCD of the terms listed in Section 2.2 are plotted. In both cases, the electric dipole-electric dipole term (D-D) dominates and the electric quadrupole-electric quadrupole term (Q-Q) is negligible. The electric dipole-spin position term (D-SP) is significant though way smaller than the D-D term. In XAS, as for the metals presented in the previous chapter, only the D-D term is significant. The total spectrum to be compared with experiment is the average of the spectra for each site (average of the two curves of Fig. 4.19 corresponding of each site ).

#### Confirmation of the observed transition

In the right panel of Fig. 4.20, the calculated XAS and XMCD spectra at the Fe K-edge in dhcp-FeH is plotted along with the spectra obtained for bcc-Fe in the previous chapter. Experimental XAS and XMCD spectra chosen before and after the observed pressure-induced transition are depicted on the left part of this figure.



**Fig. 4.19:** Left: Calculated contributions of the two crystallographic sites to the XAS and XMCD spectra at the K-edge of Fe in dhcp-FeH (sum of the three contributions D-D, Q-Q and D-SP with the absorbing atom in one site or the other). Right: Detail of the contribution of each term of the cross section (listed in Section 2.2) to the XMCD of each site.



Fig. 4.20: Left: Experimental XAS and XMCD spectra at the Fe K-edge of a Fe foil under  $H_2$  atmosphere before (blue) and after (green) the observed pressure induced transition. Right: Calculated XAS and XMCD spectra at the Fe K-edge in *bcc*-Fe (blue) and *dhcp*-FeH (green). The experimental XMCD spectra were not corrected for the circular polarization rate of the light which only partially explain the difference of amplitude between calculation and experiment (see the discussion on the amplitude Sec. 3.2.3).

The experimental spectra before and after the transition are very similar respectively to the spectra for bcc-Fe and for dhcp-FeH. The changes described in Section 4.4.1 are also all present when comparing the calculated spectra for Fe and FeH. We can therefore safely assert, that during the experiment, hydrogen went into the Fe lattice and a double hexagonal iron hydride formed. Consequently, the further increase of pressure was, as intended, a way of studying the compression of dhcp-FeH.

We see that the absorption of H atoms into the Fe lattice has a big impact on the XMCD spectra. It is very much likely because of the new structure that it permits to stabilize. In other words, it is possible that a dhcp Fe crystal with Fe atoms positioned as they are in the dhcp-FeH structure would yield the same XMCD. Such a structure is not stable so that it is difficult to achieve convergence of the self-consistent charge calculation that would allow to test this hypothesis.

It is striking that despite the similarity of the magnetic moment of Fe and FeH, their XMCD spectra are very different. It illustrates the fact that the link between XMCD at K-edge and magnetic moment is not straightforward.

**Table 4.2:** Comparison of the integral of the experimental XMCD spectrum for *bcc*-Fe and *dhcp*-FeH in the range 7105 eV - 7145eV. These integrals are very different despite the fact that the total magnetization in the cell divided by the number of Fe atoms is the same.

|                           | total magnetization per Fe atom | Integral of the exp. XMCD spectrum |
|---------------------------|---------------------------------|------------------------------------|
| $dhcp	ext{-}\mathrm{FeH}$ | $2.2 \ \mu_B$                   | $-6.6 \times 10^{-3}$              |
| <i>bcc</i> -Fe            | $2.2 \ \mu_B$                   | $-6.5 \times 10^{-4}$              |



## 4.4.4 Effect of the core-hole

Fig. 4.21: Comparison of the calculated XAS and XMCD spectra obtained in the presence of a 1s core hole (solid line) and of the spectra obtained without core-hole (dashed line) at the Fe K-edge in dhcp-FeH. The experimental spectra are plotted in light gray, for XMCD it is multiplied by 3 (see the discussion on the amplitude Sec. 3.2.3, the experimental spectrum was not corrected for the circular polarization rate of the light).

To study the effect of the core hole on the calculated spectra, we computed the spectra without core hole in the simple 8 atom cell with a  $\Gamma$ -centered  $12 \times 12 \times 4$  k-point grid for the scf calculation and a  $20 \times 20 \times 7$  grid for the spectra calculation.

The presence of the core hole in the calculation does not affect the calculated absolute and total magnetization at the end of the *scf* loop. However, as illustrated in Fig. 4.21, both the XAS and XMCD spectra are impacted. The energy positions and the sign of the features of the spectra are the same with and without core hole but there relative intensities are different.

Calculations without core-hole were used for the studies presented below because they require far less computer resources. The spectra are not intended to be compared with experimental spectra and the absence of core hole does not prevent to discuss variations of the spectra with respect to parameters of interest.

### 4.4.5 Effect of the position of the H atoms

As mentioned above, due to the low X-ray scattering power of H, it is not possible to determine the exact position of the H atoms by XRD. For a matter of simplicity, calculations presented in Fig. 4.20 and Fig. 4.19 were performed with the H atoms positioned exactly at the center of the octahedral sites. Yet, neutron powder diffraction experiments performed on quenched FeD samples showed that the D atoms were positioned slightly off the site centers (their vertical displacement are illustrated by arrows on Fig. 4.17 p. 140). We see below that the impact of this vertical displacement on XAS and XMCD spectra is tiny. The main interest for numerically moving the H atoms is to better understand the magnetic structure of FeH.

#### Relaxation of the position of H

As the vertical position of the atoms H was only measured in quenched FeD samples, we decided to perform an atomic relaxation in order to determine numerically this position for each pressure.



Fig. 4.22: Vertical position of the H atom in *dhcp*-FeH obtained by structural relaxation of the atomic positions with the code pw of QUANTUM ESPRESSO.

The principle of this relaxation is simply to allow for the atoms to move in the cell without changing the lattice parameter and to look for the lowest energy structure obtained at the end of a self consistent calculation. The choice of keeping the lattice parameters fixed was driven by the need to model FeH under high pressure. As explained in Section 4.3, the experimental volume is used to model a given pressure.

In practice, the relaxation of the atomic positions was performed by setting relax as type of calculation in input of the code pw.x of QUANTUM ESPRESSO. As a starting point, the H atoms were positioned at the center of the octahedral sites in a simple *dhcp* cell (8 atoms) using a  $10 \times 10 \times 10$  k-point grid. PBE norm-conserving pseudopotentials with a cutoff energy of 180 Ry (parameters detailed in Appendix E) were used. The calculations were spin polarized with spins along the axis [001]. Damped dynamics was used for structural relaxation [Giannozzi et al., 2009].

The relative position of the atom H for several lattice parameters, i.e. for several pressures, is presented in Fig. 4.22. The reduced coordinate is always in the interval  $0.880 \pm 0.015$ . The result of the relaxation is, therefore, very close to the experimental value that was 0.882 [Antonov et al., 1998].

#### Effect on the XAS and XMCD spectra

We performed calculations of the XAS and XMCD spectra without core-hole with the H atom either in the center of the octahedral site or in the relaxed position (positions denominated Oct. site and Calc. in Table 4.1). The lattice parameters were set to a=2.68Å and  $c/a=2 \times 1.637$ , which corresponds experimentally to a pressure of 3 GPa.

PBE norm-conserving pseudopotentials (parameters detailed in Appendix E) were used with cutoff energies 160 Ry. The calculations were spin polarized with spins along the axis [001]. A Methfessel-Paxton smearing of 0.14 eV (0.01 Ry) and a  $\Gamma$ -centered  $12 \times 12 \times 4$  k-point grid were used for the self-consistent charge density calculation. A  $20 \times 20 \times 7$  grid was used for spectra calculations and a constant Lorentzian broadening, with full width at half maximum set to 1.6 eV was applied. The wavevector **k** was set along the axis [001].

The XMCD spectra obtained for both positions of H atoms are plotted in Fig. 4.23. The top line of this figure shows the electric dipole-electric dipole (D-D) contribution and the bottom line the electric dipole-spin-position (D-SP) contribution. The electric quadrupole-electric quadrupole term, the contribution of which to XMCD is negligible, has not been studied here. Note that the scale used to display the different terms is not the same and the D-SP term is notably smaller than the D-D term.

Visually, the spectra seem almost unaffected by the vertical displacement of the H atoms: it is very difficult to distinguish the dashed lines from the solid lines. Yet, if one computes the integrals of the spectra (shown in the right in Fig. 4.23), an interesting trend emerges. The integrals of the D-SP spectra are almost exactly equal for both Fe sites when the H atoms are positioned in the center of the octahedral site whereas the integrals of the D-D spectra are different for both sites. For both terms, the absolute value of the integral of the contributions due to the site (2a) decreases when H is moved vertically and for the contributions due to the site (2c), it increases. In the previous chapter (Section 3.4), we demonstrated that the D-SP spectra is related to the partial spin magnetization of the occupied p states  $Sz_{l=1}^{1,-1}$  and that the corresponding sum-rule is reliable. We can therefore deduce that the spin magnetization of the site but it is different when, as observed experimentally, they are vertically shifted.



Fig. 4.23: Left: Contribution to the XMCD due to the D-D term (top) and to the D-SP term (bottom) of the cross section to the total calculated XMCD of dhcp-FeH with either the H atoms in the center of the octahedral site (solid lines) or in the position obtained by atomic relaxation (dashed lines). Right: Integral of this spectra plotted as a function of the z reduced coordinate of the H atom. The integral of the D-D term has been multiplied by -1 in order to have it positive.

The sign of the integral of the D-D term is opposite to the one of the D-SP term. This is consistent with Hund's rules that implies that for an orbital that is less than half-filled (that is the case of the 4p shell), the spin and orbital angular momentum are anti parallel.

The trend (increasing/decreasing of the absolute value) is the same for the D-D term as for the D-SP term which suggests that the sum-rule for the D-D term could be reliable for qualitative interpretation of small variations of the spectra.

#### Projected density of states and magnetic moment per site

The magnetic moment of FeH is mainly due to the d electrons of Fe.<sup>4</sup> The magnetic moment carried by the d electrons can be obtained from the d-projected densities of states (DOS) that are plotted on the left side of Fig. 4.24. These DOS present the expected spin split associated to ferromagnetism.

The curves corresponding to the two different positions for the hydrogen atoms are visually almost indistinguishable. The *d*-projected spin polarization  $2S_z^d$  (right side of Fig. 4.24) is obtained by computing the difference of the DOS for spin up with the DOS for spin down and integrating over unoccupied states from -10 eV to the Fermi level.

The magnetic moment associated to this projected spin momentum is given by  $-gS_z^d\mu_B$ where g is the gyromagnetic factor. As  $g \approx 2$  for the electron, the absolute value of the magnetic moment per atom due to the d electrons can be read directly on the right pannel of

<sup>&</sup>lt;sup>4</sup>By integration of the projected DOS, we evaluated the magnetic moment per atom due to the Fe p states to -0.05  $\mu_B$ , that due to the Fe s states to -0.002  $\mu_B$  and that due to H s states to -0.03  $\mu_B$ . We do not have an evaluation of the orbital magnetic moment.



Fig. 4.24: Effect of the position of H on the d projected density of states. Left: Projected d DOS with the H atoms either in the center of the octahedral site (solid lines) or in the position obtained by atomic relaxation (dashed lines) for both sites. The upper part of each plot corresponds to the DOS for spin up (majority spins) and the lower part to the DOS for spin down (minority spins). Right: d-projected spin polarization per Fe atom.

Fig. 4.24. We see that the position of the hydrogen atoms impacts the division of the magnetic moment between the two sites without changing its total value. The magnetic moment of the Fe atoms closer to the H interlayers is smaller while the magnetic moment of the Fe atoms further to the H interlayer is larger.

## 4.4.6 Magnetic structure of FeH

Using the tool pp.x of QUANTUM ESPRESSO, it is possible to extract the electron density from the output of the *scf* calculation and to plot it. We used this tool and the software XCrySDen [Kokalj, 2003] to plot the spin density depicted in Fig. 4.25.

We see that, unsurprisingly, most of the spin density is located on the Fe atoms: the region between the Fe atoms seems uniformly orange on the plane in the left side of Fig. 4.25. With a more appropriate scale is used (right side of Fig. 4.25), we see that the H atoms carry a negative spin polarization. It explains why the magnetic moment of the Fe atoms closer to the H interlayers is smaller.

It also sheds light on the reason why the magnetic moment per Fe atom in dhcp-FeH is the same as in bcc-Fe despite the fact that the Fe atoms are more distant one from each other in dhcp-FeH (2.68 Å whereas they are distant of 2.486 Å in bcc-Fe). This difference in distance suggests that the magnetic moment could be larger in dhcp-FeH (the magnetic moment decreases under pressure). From Fig. 4.25, we understand that the spin-density around the Fe atoms is, indeed, larger than in bcc-Fe but it is compensated by a negative moment on the H atoms which leads to the same total magnetic moment.



Fig. 4.25: Spin density  $(\rho^{\uparrow} - \rho^{\downarrow})$  calculated on the plane [110] which contains Fe(2*a*), Fe(2*c*) and H atoms. Left: Plane [110] depicted in the 3D structure. The full range of the spin density (-0.015 to 0.697) is included in the color scale. The region between the Fe atoms seems uniformly orange on the plane in the left side because most of the spin density is located on the Fe atoms. Right: Spin density in the plane [110]. A limited range of spin density (-0.015 to 0.015) is plotted in order to have the contributions of the H atoms to be visible.

## 4.4.7 Increasing pressure in the calculation

Application of pressure is a way of experimentally changing the interatomic distances.

#### Disappearance of ferromagnetism under pressure

Experimentally, we observe that XMCD vanishes between 23 and 28 GPa. This indicates that ferromagnetism disappears under pressure. As the Mössbauer spectrum of FeH (that is a sextet at 3 GPa) becomes a single line between 25 and 28 GPa (at room temperature) [Mitsui and Hi, 2010]), we deduce that the transition occurs toward a paramagnetic state. Indeed, Mössbauer spectroscopy probes the energy levels of the nuclei which are sensitive to its electronic environment. Under a magnetic field, the nuclear levels split into (2I+1) components and the allowed transitions ( $\Delta m_I = 0, \pm 1$ ) between the ground and excited levels lead to a sextet whereas in the absence of magnetic field this hyperfine spectrum collapses.

It is possible to run non magnetic self consistent field calculations which means that the DFT calculations are performed without spin (the density for spin up and for spin down are the same). The states obtained this way are usually referred to as *non magnetic states*. We will use this denomination in the following because it corresponds to the way the system is treated numerically. The proper way of modelling a paramagnetic state would require a random distribution of localized magnetic moments as it is done in [Mankovsky et al., 2013] using SPR-KKR method. This is however, to the best of our knowledge, computationally insurmountable with plane-waves DFT codes of the kind of QUANTUM ESPRESSO. It is therefore very common to describe paramagnetic states with non magnetic calculations.



Fig. 4.26: Calculated total energy as a function of the volume of the *dhcp*-FeH cell. Red squares correspond to non magnetic calculations and black circles to spin-polarized calculations that converged into a ferromagnetic structure. Both curves are fitted with Murnaghan equation of state (Eq. (4.1) p. 132). The fitted parameters are  $V_0 = 49$  Å<sup>3</sup>,  $B_0 = 263$  GPa and  $B'_0 = 3.8$  for the non magnetic state and  $V_0 = 53$  Å<sup>3</sup>,  $B_0 = 166$  GPa and  $B'_0 = 4.4$  for the ferromagnetic state.

The total energy of the non magnetic and ferromagnetic states of *dhcp*-FeH is plotted in Fig. 4.26. For a volume V=40 Å<sup>3</sup>, the non magnetic state becomes more stable. This is close to the result obtained by LMTO-ASA calculations [Elsässer et al., 1998] that showed a magnetic transition between 75 and 80 Bohr<sup>3</sup> (44 - 47 Å<sup>3</sup>). This volume corresponds experimentally to the pressure P=92 GPa [Badding et al., 1991]. Using Murnagham equation of state with the calculated parameters we can also obtain the calculated pressure (derivative of the energy with respect to the volume): P=94 GPa. The calculations are performed at low temperature (0 K) so it is not surprising to find a transition pressure larger than the one measured at room temperature. The uncertainty on the calculated transition pressure is very large (several tens of GPa) due to the fact that pressure varies quickly with volume in the range 40 - 45 Å<sup>3</sup> (the volume V=45 Å<sup>3</sup> corresponds to a pressure P=40 GPa) and a small error on the lattice parameter *a* results in a significant error on the volume.<sup>5</sup> For this reason, it is likely that, if experiments at low temperature were to be performed, the calculated value would also be off .

#### Calculated spectra under pressure

The XAS and XMCD spectra without core hole of the ferromagnetic state of FeH were computed with several lattice parameters and the XAS spectra of the non magnetic state was computed with the lattice parameter corresponding to P=40 GPa (see inset of Fig 4.18 for the relationship between the lattice parameter a and the pressure). PBE norm-conserving pseudopotentials (parameters detailed in Appendix E) were used with a cutoff energy 160 Ry. Except for 40 GPa, the calculations were spin polarized with spins along the axis [001]. A Methfessel-Paxton smearing of 0.14 eV (0.01 Ry) and a  $\Gamma$ -centered  $12 \times 12 \times 4$  k-point grid were used for the self-consistent charge density calculation. A  $20 \times 20 \times 7$  grid was used for spectra

<sup>&</sup>lt;sup>5</sup>For example, for a = 2.5Å with the uncertainty da=0.1Å, the uncertainty on the volume is dV=2.3Å<sup>3</sup>.



Fig. 4.27: Calculated XAS and XMCD spectra at the Fe K-edge in FeH at several pressures (*i.e.* for several lattice parameters). All calculations except for 40 GPa were performed from the ferromagnetic state of dhcp-FeH.

calculations and a constant Lorentzian broadening, with full width at half maximum set to 1.6 eV was applied. The wavevector **k** was set along the axis [001].

By comparison with Fig 4.16 p. 140, we see that our calculations reproduce very well the effect of pressure on the spectra. The effect of pressure on the XAS spectra of ferromagnetic FeH is to push the peaks at higher energies as the distances are reduced. The amplitude of XMCD decreases slowly without change of its shape. The similarity between the calculated XAS spectra for the non magnetic state (in purple in Fig. 4.27) and the XAS spectra at 28 GPa (in red in Fig. 4.16) that corresponds to the pressure when XMCD vanishes is also striking: the shoulders a and b are enhanced while the maximum c decreases in amplitude. We can therefore deduce that the evolution of the shape of the experimental XAS spectra at 28 GPa is a marker of the transition from ferromagnetic FeH to paramagnetic FeH that is not accompanied by any structural transition. We see that spectra calculations can also be a useful tool to determine if changes observed on spectra under pressure are accompanied by a structural transition. This not obvious when, as in the case presented here, pressure induces not only the expected stretch of the energies due to distances contraction, but also a variation of the intensities of the peaks.

## 4.4.8 Conclusion

In this study, calculations helped identifying two transitions that occurred when Fe was compressed under  $H_2$  atmosphere.

The first transition corresponds to the absorption of H atoms into the bcc-Fe lattice which leads to the creation of dhcp-FeH. The calculated spectra for bcc-Fe and dhcp-FeH are in good agreement with experimental ones. The example of Fe and FeH is perfect to illustrate that a quick interpretation of XMCD at K-edge saying that the integral of the spectra is proportional to the magnetic moment is not valid. Here, the magnetic moment in the cell divided by the number of Fe atoms is the same in *bcc*-Fe and *dhcp*-FeH but the integral of the XMCD spectra at the K-edge of Fe is completely different. The fact that the XAS and XMCD spectra are well reproduced by the calculations means that the materials under study are well simulated by our numerical method which allowed to draw conclusions on the magnetic structure of FeH based on the output of the calculations.

The second transition is a purely magnetic transition from a ferromagnetic state to paramagnetic state. The existence of this pressure-induced transition, that was inferred from the observation of the XMCD spectra, was confirmed by the calculation. The pressure of transition can, however, not be determined accurately from the calculation because of the enormous uncertainty on the obtained pressure.

From a methodological point of view, this study is an example of the possibilities that our method opens as soon as the spectra are well reproduced which ensures that the material is correctly simulated.

## 4.5 Chromium Dioxide

The time of glory of chromium dioxide seemed over with the gradual decline of the magnetic recording cassette tapes in which it had been used along with  $Fe_2O_3$ . Yet, it has been attracting a renewed interest in recent years. It is a particularly interesting transition metal oxide because it belongs to the class of half-metallic ferromagnets: the Fermi level crosses a metallic band for one spin channel while it lies in a gap for the other spin channel. This property makes it a good candidate for spintronic applications (for example as a source of spin polarized electrons) especially because it has a rutile structure like several other MO<sub>2</sub> oxides, where M is a transition metal (e.g. Ti, Mo, Ru). These oxides have various properties (e.g. TiO<sub>2</sub> is a diamagnetic semiconductor and RuO<sub>2</sub> is a paramagnetic metal [Rao and Gopalakrishnan, 1997, Table 6.4]) so their combination in layered devices is promising.

## 4.5.1 Electronic properties of CrO<sub>2</sub>

The interest for the electronic and magnetic properties of  $\text{CrO}_2$  from a fundamental point of view is strong. Despite a large number of studies, it is clearly stated in the review by Katsnelson et al. [Katsnelson et al., 2008], that the intrinsic correlated electronic structure that induces both ferromagnetism and metallicity remains to be understood. Yet, interesting lines of explanation can be found in the literature.

#### Lattice

 $CrO_2$  crystallizes in the tetragonal  $P4_2/mnm$  (136) space group [Thamer et al., 1957]. It has a rutile structure, that is represented in Fig. 4.28. The  $Cr^{4+}$  ions form a body-center tetragonal lattice and are surrounded by octahedra of oxygen atoms. The structure can be viewed as ribbons along **c**-axis of edge sharing octahedra joined together by corners (each oxygen is shared by three octahedra). The orientation of the octahedra in adjacent ribbons differs by a 90° rotation about the **c**-axis.

In  $CrO_2$ , the octahedra are axially compressed: the Cr-O length in the apical direction is 1.89 Å whereas in the equatorial directions it is 1.91 Å. Moreover, the arrangement of the four O is rectangular (instead of square in a perfect octahedra) so that when we define the local



**Fig. 4.28:** Left:  $CrO_2$  structure (Figure from [Yamasaki et al., 2006]). We define the local coordinate systems with  $\mathbf{z}_{1,2}$  pointing exactly to an O atom and  $\mathbf{x}_{1,2}$  almost pointing to an O atom. Right: the unit cell in the local coordinates of the octahedron around Cr1.

coordinates in the octahedra,  $\mathbf{x}$  and  $\mathbf{y}$  can not both point toward a ligand (see Fig 4.28, right). The lattice parameters of CrO<sub>2</sub> are listed in Table 4.3.

Table 4.3:  $CrO_2$  lattice parameters and atomic positions. As u < 0.304, the  $CrO_6$  octahedra are axially compressed.

| $CrO_2 P4_2/mnm$        |          |       |      |  |  |  |  |
|-------------------------|----------|-------|------|--|--|--|--|
| Lattice paramet         | ers [P.F | Porta | et a | al., 1972]                             |  |  |  |
| a (Å)                   | 4.42     |       |      |  |  |  |  |
| c (Å)                   | 2.92     |       |      |  |  |  |  |
| Volume $(\text{\AA}^3)$ | 56.9     |       |      |  |  |  |  |
| Atomic coordina         | ntes     |       |      |  |  |  |  |
|                         | x        | y     | z    |  |  |  |  |
| $\operatorname{Cr}(2a)$ | 0        | 0     | 0    |  |  |  |  |
| O(4f)                   | u        | u     | 0    | where $u=0.303$ [P.Porta et al., 1972] |  |  |  |
| Bond length (Å)         | )        |       |      |  |  |  |  |
| Cr-O apical             | 1.89     |       |      | (two bonds)                            |  |  |  |
| Cr-O equatorial         | 1.91     |       |      | (four bonds)                           |  |  |  |

#### Simplified crystal field picture

The half metallic behaviour of  $\text{CrO}_2$  can be understood in simplified terms using the crystalfield model [Katsnelson et al., 2008]. In an octahedron, the *d*-orbitals are split into two sets of levels: the  $t_{2g}$  orbitals and the  $e_g$  orbitals. In  $\text{CrO}_2$ , the minority spin states are at significant higher energy compared to the majority spin states (see diagram Fig. 4.29). In  $\text{Cr}^{4+}$ , the  $t_{2g}$  of up spins are 2/3 filled. For spins down, on the other hand, the Fermi level lies in a band gap between the occupied O p and the unoccupied Cr d states so that  $\text{CrO}_2$  is a half metal.

Yet, generally, two 3d electrons in the  $t_{2g}$  orbitals would make a Mott insulator with S=1 local moments and antiferromagnetic ordering but this seems to be about as far from the actual observed properties as one can get [Korotin et al., 1998].



Fig. 4.29: Simplified diagram representing the energies of the orbitals of the ion  $Cr^{4+}$  in ferromagnetic  $CrO_2$ . The  $t_{2g} - e_g$  splitting is due to the ligand in octahedral symmetry. In a ferromagnetic material, the minority spin states are at higher energy compared to the majority spin states. The convention chosen here is that up spins are majority spins.

#### Partial densities of states

To better understand the electronic properties of  $CrO_2$ , the projected densities of states in  $CrO_2$  shown in Fig. 4.30 can be useful.

These DOS illustrate the half-metallic property  $\text{CrO}_2$ : the Fermi level intersects the majority spin bands (spin up in Fig. 4.30) and it is in a gap of the minority spin states. It results in a complete spin polarization at the Fermi level which is consistent with observations made by Andreev spectroscopy [Soulen et al., 1998]. The total spin moment is precisely equal to 2.0  $\mu_B$ per Cr atom due to this gap.

The Cr d density of states (third panel) is consistent with the simplified crystal-field picture of Fig. 4.29. Due to the distortions of the CrO<sub>6</sub> octahedra, this crystal-field description is however more complicated than depicted in Fig. 4.29. As was pointed out by Korotin et al. [Korotin et al., 1998], the  $t_{2g}$  orbital is split because of the compression of the octahedra between the  $d_{xy}$  orbital at lower energy and the  $d_{zx}$  and  $d_{zy}$  orbitals at higher energy. We see at the bottom of Fig. 4.30, that the  $d_{xy}$  orbital is localized in energy. We deduce that it is not part of the continuum which means that it is also localized in real space contrary to the  $d_{zx}$ and  $d_{zy}$  orbitals which are more itinerant. The latter strongly hybridize with O p states.

#### Double exchange mechanism

Ferromagnetism in  $CrO_2$  has been attributed to a specific double exchange mechanism between Cr atoms in which the splitting of the  $t_{2g}$  orbital plays a crucial role [Schlottmann, 2003, Schlottmann, 2004, Korotin et al., 1998].

All Cr atoms in the structure have the same oxidation state so ferromagnetism can not be described exactly within Zener double exchange model but it has been explained in the same terms. According to this model, one electron on each site occupies the localized  $d_{xy}$  orbital (which is lower in energy than  $d_{yz}$  and  $d_{zx}$  due to the compression of the octahedra) and the other electron is able to hop from one site to the other (t term in the Hubbard model). Hund's rule couples the spins of the localized electrons with the spin of the itinerant electrons so that the hopping depends on the relative z projections of the spins. The presence of the hopping electrons lower the total energy of the system so that, in the ground state, the system favours ferromagnetic coupling. Note that the hopping between the Cr atoms is mediated by the O p-states and is not direct. It has been shown that, if the three orbitals were itinerant, an antiferromagnetic coupling would be favoured [Schlottmann, 2004] which led to the conclusion that the distortion of the octahedra is crucial to explain ferromagnetism.



Fig. 4.30: Projected Cr p, O p and Cr d density of states of CrO<sub>2</sub> for both spins and  $t_{2g}$  density of states near the Fermi level. The z direction is the apical direction of the CrO<sub>6</sub> octahedron.PBE norm-conserving pseudopotentials (parameters detailed in Appendix E) were used with a cutoff energy 180 Ry. The calculations were spin polarized with spins along the axis [001]. A Methfessel-Paxton smearing of 0.14 eV (0.01 Ry) and a  $\Gamma$ -centered  $6 \times 6 \times 9$  k-point grid was used for the self consistent calculation in the conventional cell that contains two Cr and one O. The non-self-consistent part of the calculation was performed with a  $12 \times 12 \times 18$  k-point grid.

Yet, the double exchange model proposed by P. Schlottmann requires high Coulomb interaction and, therefore corresponds to an insulator (Mott-Hubard limit) [Schlottmann, 2004]. As he himself states, the fact that the ligands are not included in the model Hamiltonians that P. Schlottmann considered is likely the origin of this problem. Yet, it indicates that the success of the proposed exchange mechanism is questionable.

In this context, a study of the effect of pressure on magnetism in  $CrO_2$  could be very interesting to get new insight on the possible exchange between Cr atoms. Indeed, application of pressure changes the local environment of  $CrO_2$  [Maddox et al., 2006] which is of great importance in the proposed mechanism.

#### Correlations effect in $CrO_2$

No Hubbard U term [Cococcioni and de Gironcoli, 2005] was included in the DOS calculation presented above. The importance of the on-site Coulomb repulsion in  $CrO_2$  is still a subject of active research. Mazin et al. [Mazin et al., 1999] showed that LDA calculations of optical conductivity in  $\text{CrO}_2$  are sufficient to interpret the experimental data and deduced that there is no experimental smoking gun in regard to strong correlations related exotic phenomena in  $\text{CrO}_2$ . Contrary to this result, photoemission spectra are better fitted if the Hubbard U correction is included [Laad et al., 2001] which seems to indicate that electronic localization is important. Another study by Toropova et al. [Toropova et al., 2005] compared results from LDA and LDA+U calculations to various experimental data, including  $L_{2,3}$  x-ray absorption spectra. They observed that LDA calculations better explain experimental data. The question of the existence of strong correlations in  $\text{CrO}_2$  remains rather unclear.

We saw above that simple PBE (GGA) calculations are sufficient to obtain half-metallic ferromagnetism in chromium dioxide. It is a little bit surprising because we could have expected that strong correlations would be required for the double exchange described above.

To pursue further the investigation of the properties of  $CrO_2$ , it is important to check if we can model correctly its electronic and magnetic structure. For this purpose, confrontation with XAS and XMCD experiments can be a valuable asset.

## 4.5.2 Pre-edge structure

The X-ray absorption spectra of  $\text{CrO}_2$  at K-edge of Cr measured on ODE beamline is shown in Fig. 4.31. There is a peak in the pre-edge but the signal to noise ratio is not sufficient to study its structure. Therefore, for the pre-edge structure, we are going to use the integrated data from a 1s2p Resonant Inelastic X-ray Scattering (RIXS) experiment that we performed on Galaxies beamline at SOLEIL. All the details of this RIXS-MCD experiments are given in the paper by Patric Zimmermann et al. [Zimmermann et al., 2017] attached in Appendix F of this thesis. By integration of the RIXS map over the full  $K_{\alpha}$  emission energy, we can obtain a fluorescence spectrum. It is similar to a TFY spectrum <sup>6</sup> as can be measured on certain XAS beamlines but with a slightly narrower emission energy range.



Fig. 4.31: XAS of  $CrO_2$  at the Cr *K*-edge acquired at room temperature and ambient pressure on ODE beamline. This spectra was acquired during the experiment discussed Section 4.5.3.

To interpret the experimental spectra, we computed the densities of state in the presence of a 1s core hole (see Section 2.5 p.69) which required the use of a supercell. It is possible to perform such calculations for the interpretation of the K pre-edge because the needed energy range is small. When the energy range increases, it becomes computationally more and more expensive and therefore intractable (hence, the use of the method presented in Chap. 2 to compute XAS spectra).

<sup>&</sup>lt;sup>6</sup>TFY spectra are usually assimilated to absorption spectra despite the fact that they can be distorted compared to XAS. [de Groot and Kotani, 2008, de Groot et al., 1994, Eisebitt et al., 1993]

A  $2 \times 2 \times 3$  supercell was used so that the distance between core-holes is 8.4 Å and the *k*-point grid was reduced accordingly. The densities of states were computed up to 6 eV above the Fermi level which already requires the inclusion of 220 bands in the non self consistent part of the calculation.

The p and d density of states projected on the Cr atom with a core-hole and on a neighboring Cr are plotted at the bottom of Figure 4.32.



Fig. 4.32: Top: Experimental absorption spectra (obtained from the integration of the RIXS plane over emission energy) and projected density of states in the presence of a 1s core hole on the Cr atom with a broadening of 0.3 eV. The p density of states is multiplied by 400 (the prefactors for terms D-D and Q-Q are different). The DOS have been shifted by 5996.5 eV in energy to match the experimental spectra. Bottom: Projected Cr p, O p and Cr d density of states of CrO<sub>2</sub> for both spins near the Fermi level with and without core-hole. Cr<sup>\*</sup> corresponds to projections around the chromium atom with a 1s core-hole and Cr around a neighboring chromium.

The main effect of the core hole on the DOS is to shift them to lower energies: it is visible on the *d*-projected DOS (on top of the bottom panel of Fig. 4.32). The energy position of pDOS, on the other hand, is almost unaffected by the core-hole. This leads to the conclusion that the p density of states in the pre-edge range is due to the hybridization of the p states of the absorbing atom with the d states of the neighboring Cr atoms. This hybridization results in several peaks in the Cr p density of states just above the Fermi level (note that, as the Cr site is centrosymmetric, on-site p-d hybridization is forbidden by symmetry so these peaks could not have been due to on-site hybridization). This hybridization is likely mediated by the O p states because the Cr-Cr distance is 3.45 Å which is too large to allow a direct hybridization. The energy position of these p features just above the Fermi level are not affected by the core hole because they lie at the energy of the d states of the neighboring Cr which has no core hole. The main effect of the core hole is, therefore, to separate the d and p features in the pre-edge range. The diagram Fig. 4.33 is an attempt to illustrate this point.





The experimental absorption spectra in the pre-edge is shown on top of Fig. 4.32. We know that the electric dipole - electric dipole (D-D) transition probes the empty p states whereas the electric quadrupole - electric quadrupole (Q-Q) transition probes the empty d states. The p and d densities of state are therefore plotted on the same graph as the experimental spectra. From this graph, we can infer that there is a significant D-D contribution in the pre-edge and a Q-Q contribution at lower energy that seems to appear as a shoulder in the experimental spectra.<sup>7</sup>

It is known that the energy position of the unoccupied d states near the Fermi level is very sensitive to core hole effects and that, when including a static 1s core-hole with full relaxation of the electronic structure as we do, it is usually calculated at a too high energy with respect to the edge (the 1s core hole is overscreened) [Cabaret et al., 2010]. The combined effect of the Hubbard U correction and of a core hole is, *a priori*, not predictable. In a different study at the Co K-edge in LiCoO<sub>2</sub> [Juhin et al., 2010] the effect of the Hubbard U correction in the presence of a core hole was to increase the splitting between the D-D and the Q-Q contributions. Here,

<sup>&</sup>lt;sup>7</sup> The attribution of features of the RIXS map to transitions toward d or p states is important for the further study of the RIXS and RIXS-MCD maps. Indeed, DFT-based calculation (which are monoelectronic calculations) can not accurately reproduce  $1s \rightarrow 3d$  transitions when there is a strong 3d-3d repulsion. An adequate theory to perform RIXS calculations in that case is ligand field multiplet (LFM) theory [Stavitski and de Groot, 2010]. In LFM calculations, only the absorbing atom with its ligands are considered so that only the localized states are accounted for but the big advantage is that it is a multielectronic method so that the 3d-3d repulsion can be included. LFM calculations are parametrized so that it is important to know a priori which region of the map is to be computed. Such calculations are presented in the paper [Zimmermann et al., 2017] reproduced in Appendix F of this thesis.

the splitting seems correctly computed so that the inclusion of the Hubbard U correction does not seem to be necessary. For this reason, we decided to pursue the interpretation of the edge features without U correction.

## 4.5.3 Magnetism under pressure in $CrO_2$ studied by XMCD at the Cr K-edge

We measured XAS and XMCD spectra under pressure at the Cr K-edge on ODE beamline at SOLEIL. The interest of performing experiments at the K-edge is twofold here: (i) the surface of the CrO<sub>2</sub> grains can be reduced to  $Cr_2O_3$  so it is important to probe the bulk and not the surface and (ii) it permits the application of pressure. The application of a hydrostatic pressure to  $CrO_2$  changes the local octahedral environment of Cr: X-ray diffraction measurements [Maddox et al., 2006] revealed that the degree of distortion of the octahedra increases upon application of pressure. Pressure is, therefore, expected to influence the possible doubleexchange mechanism which is related to the distortion of the octahedra and hence the magnetic state of  $CrO_2$ .

#### Experimental details

**Sample** We used a commercial  $CrO_2$  powder sold by Sigma-Aldrich under the name Magtrieve(TM). We grinded it mechanically to get grains of a size less than  $10\mu$ m. We performed X-ray diffraction of the obtained powder to check that  $CrO_2$  had not been reduced to  $Cr_2O_3$ . Indeed,  $CrO_2$  is metastable and the stable Chromium oxide at ambient conditions is  $Cr_2O_3$ . The diffraction pattern, shown in Fig. 4.34, shows no other peaks than the rutile  $CrO_2$  ones.



Fig. 4.34: XRD pattern of the finely grinded  $CrO_2$  powder fitted with the standard pattern of  $CrO_2$  in rutile phase.

Synchrotron experiments under pressure XAS and XMCD spectra under pressure at the Cr K-edge were acquired with a Si (111) polychromator. A portion of the beam below

the plane of the radiation emitted by the bending magnet was selected to obtain circularly polarized X-rays.

The K-edge of Cr lies relatively low in energy (5989 eV) so we had to use perforated diamond anvils [Dadashev et al., 2001] to apply pressure because full diamond would have absorbed too many X-rays. Silicon oil was used as pressure-transmitting medium. A hole with diameter around 100  $\mu$ m was drilled in a rhenium gasket of local thickness 20  $\mu$ m (obtained by indentation). Cr K-edge XAS spectra were measured in transmission within a 1.3 T magnetic field whether parallel or antiparallel to the direction of propagation of the magnetic field. XMCD was recorded according to the procedure described p. 129 and the sign was set to correspond to the conventional way of displaying XMCD spectra [Baudelet et al., 1991].

#### XAS and XMCD at the Cr K-edge

The XAS and XMCD spectra at the Cr K-edge as a function of pressure at two different temperatures (room temperature and T=15K) are shown in Fig. 4.35. The amplitude of the XMCD on the edge reaches  $3 \times 10^{-3}$  of the edge jump at low temperature. The XMCD spectra at ambient conditions that we measured is similar to the XMCD spectrum at the Cr K-edge in CrO<sub>2</sub> (recorded only up to 20 eV above the edge) that was published in 1997 [Attenkofer, K. and Schütz, G., 1997].

An isolated positive peak appears in the XMCD spectra approximately 46 eV above the edge. This feature is very similar to the evidences of multielectron excitations that were reported in XMCD of rare earth coumponds in [Dartyge et al., 1992]. This peak is likely due to the existence of a double excitation, that is to say an excitation of two electrons: a 1s electron and a 3p electron. Indeed, the binding energy of the 3p electrons ( $M_{2,3}$  edges) of Mn is 47.2 eV which corresponds to the energy position of the peak above the edge. We consider the binding energy in atom Mn (Z+1 compared to Cr) because the Z+1 atom qualitatively mimics the 1s core hole [Dartyge et al., 1992]. Dartyge et al. explain that the double excitation is due to electric dipole transitions because they checked the angular dependence. It is difficult to say if it is due to the D-D or the D-SP operator because our monoelectronic method for calculation is unable to model a double excitation.

At room temperature, on the right panel of Fig. 4.35, we observe that the XMCD amplitude decreases rapidly under pressure. In particular, there is a sudden decrease of the XMCD between 9 and 12 GPa accompanied by a slight deformation of the spectrum. The shape of the absorption spectrum also changes between 9 and 12 GPa. Surprisingly, almost no changes occur in the near-edge region but they start 40 eV above the edge. In the literature, a structural transition from rutile structure (P42/mnm) to an orthorhombic CaCl<sub>2</sub>-like structure (Pnnm, also a sixfold coordinated structure) is reported to occur at  $12\pm3$  GPa [Maddox et al., 2006]. It is likely that the observed changes of the spectra are due to this transition. The fact that the amplitude of the XMCD amplitude decreases rapidly with pressure is not in favor of the simplified picture in terms of double-exchange because the basis of this mechanism that is the distortion of the octahedra is enhanced by pressure. We can deduce that double exchange probably competes with one or several other mechanisms that lead to the progressive disappearance of magnetism under pressure.

On the left panel of Fig. 4.35, we see that the impact of pressure on the XMCD spectrum at low temperature is mainly a small decrease of its amplitude. Between 11 GPa and 15 GPa, the main positive peak of the XMCD spectrum is slightly shifted to higher energies. The corresponding XAS (15 GPa) exhibit several differences with the spectra at lower pressure but



Fig. 4.35: Experimental XAS and XMCD spectra for  $CrO_2$  for several values of increasing pressure at two temperatures: T=15K for the left panel and room temperature for the right panel. The experimental XMCD spectra were not corrected for the circular polarization rate of the light  $P_c < 1$ .

it is not similar to the room temperature-high pressure spectra. The fact that the magnetism at low temperature is more robust with pressure is interesting because it indicates that  $CrO_2$  remains ferromagnetic at high pressure but with a lower Curie temperature. This could be the sign of a competition between super-exchange (that favors antiferromagnetic ordering) and double exchange [Loh, 2011].

It is clear that calculations would be very useful to distinguish the changes due to a possible structural phase transitions from the changes related to the impact of pressure on the electronic and magnetic structure of  $CrO_2$ .

#### **Preliminary calculations**

For the preliminary calculations presented in this section, the experimental lattice parameters of Table 4.3 were used. The same pseudopotentials (parameters detailed in Appendix E) as for the DOS calculations were used. A static 1s core hole was included in the pseudopotential for the absorbing Cr atom. A  $4 \times 4 \times 4$  k-point grid was used for the *scf* calculation within a  $2 \times 2 \times 3$  supercell (72 atoms, distance between core holes: 8.42Å) with a Methfessel-Paxton smearing of 0.14 eV (0.01 Ry).

Calculations were performed with collinear spins along the axis [001] (**c**-axis in Fig. 4.28) within the diagonal spin-orbit coupling approximation (see p. 54). The wavevector **k** was also set along the **c**-axis.

The results of the calculation are shown in the right panel of Fig. 4.36 next to the experimental spectra. Visually, it is striking that, despite the fact that the X-ray absorption spectra is well reproduced by the calculation, the calculated XMCD spectrum seems completely different than the experimental XMCD. Yet, with a closer look, we see that, except in the pre-edge range,



Fig. 4.36: Left: Experimental XAS and XMCD spectra at the Cr K-edge in powder CrO<sub>2</sub> at ambient pressure and low temperature (T=4K). The experimental XMCD spectra were not corrected for the circular polarization rate of the light  $P_c \approx 0.7$ . Right: Calculated XAS and XMCD spectra at the Cr K-edge in CrO<sub>2</sub> with **k** along the **c**-axis. In both cases, the XMCD spectrum was multiplied by 200.

the peaks positions and sign are in good agreement with experiment. The relative intensity, on the other hand, is completely off which gives the visual impression of a totally different spectra.

The contributions to the calculated XMCD of the three terms D-D, D-SP and Q-Q are shown in Fig. 4.37. We see that, in the pre-edge region, all three contributions are more or less equally significant whereas at higher energies, the Q-Q term is negligible.



Fig. 4.37: Detail of the contribution of each term of the cross section (listed in Section 2.2) to the calculated XMCD at the Cr K-edge in CrO<sub>2</sub>.

The calculation was performed with only one direction of  $\mathbf{k}$ -vector so the calculated spectrum depicted in Fig. 4.36 is not the isotropic spectrum (which is in principle the spectrum to be compared with the experimental spectrum acquired on powder). This is expected to have a strong impact in the pre-edge range because the three terms are angular dependent.

The angular dependence of the the D-D and D-SP term are more simple than that of Q-Q and we checked by sampling three directions and calculating the spectra without core-hole, that angular dependency can not explain the discrepancy of the XMCD spectra in the range 6000 eV - 6040 eV.

The fact that the XMCD spectra is not well reproduced by the calculation is a warning sign that indicates that the method does not fully capture the complex magnetic structure of  $CrO_2$ . In particular, the D-SP contribution does not improve the agreement with experiment. It could be related to a bad modeling of the spin-density close to the nucleus.

## 4.5.4 Conclusion

 $CrO_2$  is a particularly interesting transition metal oxide because it is a half metallic ferromagnet. The fundamental electronic properties of  $CrO_2$  remain to be understood. The study of the impact of pressure on magnetism in  $CrO_2$  is a very promising way to shed a new light on the issues that are discussed and in particular on the double exchange mechanism that was proposed to explain ferromagnetism in  $CrO_2$ . Indeed, the distortion of the octahedra of the rutile structure plays a crucial role in this mechanism and the application of pressure changes the local environment of Cr. Combined XAS and XMCD experiments at the *K*-edge of Cr are the method of choice to undertake this study.

Experimentally we observed several variations of the XAS and XMCD spectra which still need to be interpreted. Preliminary calculations were not able to reproduce the experimental XMCD spectrum even if the calculated XAS spectrum is in good agreement with experiment. Further calculations, which include the Hubbard U correction and the intra-atomic exchange J (which are not implemented yet) should be performed in order to see if the agreement between calculated and experimental spectra is improved.

## Bibliography

- [Antonov et al., 2002] V. E. Antonov, M. Baier, B. Dorner, V. K. Fedotov, G. Grosse, A. I. Kolesnikov, E. G. Ponyatovsky, G. Schneider, and F. E. Wagner (2002). High-pressure hydrides of iron and its alloys. *Journal of Physics: Condensed Matter*, 14(25):6427.
- [Antonov et al., 1981] V. E. Antonov, I. T. Belash, E. G. Ponyatovskii, V. G. Thiessen, and V. I. Shiryaev (1981). Magnetization of iron hydride. *Physica Status Solidi (a)*, 65(1):K43– K48.
- [Antonov et al., 1998] V. E. Antonov, K. Cornell, V. K. Fedotov, A. I. Kolesnikov, E. G. Ponyatovsky, V. I. Shiryaev, and H. Wipf (1998). Neutron diffraction investigation of the dhcp and hcp iron hydrides and deuterides. *Journal of Alloys and Compounds*, 264(1-2):214-222.
- [Attenkofer, K. and Schütz, G., 1997] Attenkofer, K. and Schütz, G. (1997). Hard and Soft X-MCD Studies of CrO2. J. Phys. IV France, 7:C2-459-C2-460.
- [Badding et al., 1991] J. V. Badding, R. J. Hemley, and H. K. Mao (1991). High-Pressure Chemistry of Hydrogen in Metals: In Situ Study of Iron Hydride. *Science*, 253(5018):421– 424.
- [Baudelet et al., 1991] F. Baudelet, E. Dartyge, A. Fontaine, C. Brouder, G. Krill, J. P. Kappler, and M. Piecuch (1991). Magnetic properties of neodynium atoms in Nd-Fe multilayers studied by magnetic x-ray dichroism on Nd L<sub>II</sub> and Fe K edges Phys. Rev. B, 43:5857.
- [Baudelet et al., 2005] F. Baudelet, S. Pascarelli, O. Mathon, J. P. Itié, A. Polian, M. Astuto, and J. C. Chervin (2005). X-ray absorption spectroscopy and x-ray magnetic circular dichroism simultaneous measurements under high pressure: the iron bcc-hcp transition case. *Journal of Physics: Condensed Matter*, 17(11):S957.
- [Birch, 1952] F. Birch (1952). Elasticity and constitution of the Earth's interior. Journal of Geophysical Research, 57(2):227–286.
- [Cabaret et al., 2010] D. Cabaret, A. Bordage, A. Juhin, M. Arfaoui, and E. Gaudry (2010). First-principles calculations of X-ray absorption spectra at the K-edge of 3d transition metals: an electronic structure analysis of the pre-edge. *Physical Chemistry Chemical Physics*, 12(21):5619.
- [Cafun et al., 2013] J.-D. Cafun, J. Lejeune, J.-P. Itié, F. Baudelet, and A. Bleuzen (2013). XMCD at the Transition Metal K-Edges as a Probe of Small Pressure-Induced Structural Distortions in Prussian Blue Analogues. *The Journal of Physical Chemistry C*, 117(38):19645– 19655.
- [Chervin et al., 1995] J. C. Chervin, B. Canny, J. M. Besson, and P. Pruzan (1995). A diamond anvil cell for IR microspectroscopy. *Review of Scientific Instruments*, 66(3):2595–2598.

- [Choe et al., 1991] I. Choe, R. Ingalls, J. M. Brown, Y. Sato-Sorensen, and R. Mills (1991). Mössbauer studies of iron hydride at high pressure. *Physical Review B*, 44(1):1–4.
- [Cococcioni and de Gironcoli, 2005] M. Cococcioni and S. Gironcoli (2005). Linear response approach to the calculation of the effective interaction parameters in the LDA + U method. *Phys. Rev. B*, 71:035105.
- [Dadashev et al., 2001] A. Dadashev, M. P. Pasternak, G. K. Rozenberg, and R. D. Taylor (2001). Applications of perforated diamond anvils for very high-pressure research. *Review of Scientific Instruments*, 72(6):2633-2637.
- [Dartyge et al., 1992] E. Dartyge, A. Fontaine, C. Giorgetti, S. Pizzini, F. Baudelet, G. Krill, C. Brouder, and J.-P. Kappler (1992). Multielectron excitations in rare-earth compounds revealed by magnetic circular x-ray dichroism. *Phys. Rev. B*, 46:3155–3158.
- [Datchi et al., 1997] F. Datchi, R. LeToullec, and P. Loubeyre (1997). Improved calibration of the  $SrB_4O_7:Sm^{2+}$  optical pressure gauge: Advantages at very high pressures and high temperatures. *Journal of Applied Physics*, 81(8):3333–3339.
- [de Groot et al., 1994] F. Groot, M. Arrio, P. Sainctavit, C. Cartier, and C. Chen (1994). Fluorescence yield detection: Why it does not measure the X-ray absorption cross section. Solid State Communications, 92(12):991 – 995.
- [de Groot and Kotani, 2008] F. Groot and A. Kotani (2008). Core Level Spectroscopy of Solids. Advances in Condensed Matter Science. CRC Press.
- [Desy, site] Desy (website). Storage Rings as Synchrotron Radiation Sources accessed May 07, 2017. http://photon-science.desy.de/research/studentsteaching/primers/ storage\_rings\_beamlines/index\_eng.html.
- [Diamond, site] Diamond (website). Energy Dispersive EXAFS accessed Feb. 08, 2017. http: //www.diamond.ac.uk/Beamlines/Spectroscopy/Techniques/EDE.html.
- [du Trémolet de Lacheisserie et al., 2005] E. Lacheisserie, D. Gignoux, and M. Schlenker (2005). *Magnetism.* Number vol. 1. Springer.
- [Eisebitt et al., 1993] S. Eisebitt, T. Böske, J.-E. Rubensson, and W. Eberhardt (1993). Determination of absorption coefficients for concentrated samples by fluorescence detection. *Phys. Rev. B*, 47:14103–14109.
- [Elsässer et al., 1998] C. Elsässer, J. Zhu, S. G. Louie, B. Meyer, M. Fähnle, and C. T. Chan (1998). Ab initio study of iron and iron hydride: II. Structural and magnetic properties of close-packed Fe and FeH. Journal of Physics: Condensed Matter, 10(23):5113.
- [Giannozzi et al., 2009] P. Giannozzi et al. (2009). QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. J. Phys.: Condens. Matter, 21(39):395502.
- [Haas et al., 2009] P. Haas, F. Tran, and P. Blaha (2009). Calculation of the lattice constant of solids with semilocal functionals. *Phys. Rev. B*, 79:085104
- [Hess and Exarhos, 1989] N. J. Hess and G. J. Exarhos (1989). Temperature and pressure dependence of laser induced fluorescence in Sm:YAG-a new pressure calibrant. *High Pressure Research*, 2(1):57–64.

- [Ishimatsu et al., 2012] N. Ishimatsu, T. Shichijo, Y. Matsushima, H. Maruyama, N. Kawamura, M. Mizumaki, T. Matsuoka, and K. Takemura (2012). Magnetic States in Fe, Co, Ni hydrides under High Pressure Probed by X-ray Magnetic circular dichroism. *Journal of Physics: Conference Series*, 377(1):012041.
- [Jephcoat et al., 1986] A. P. Jephcoat, H. K. Mao, and P. M. Bell (1986). Static compression of iron to 78 GPa with rare gas solids as pressure-transmitting media. *Journal of Geophysical Research: Solid Earth*, 91(B5):4677–4684.
- [Juhin et al., 2010] A. Juhin, F. Groot, G. Vankó, M. Calandra, and C. Brouder (2010). Angular dependence of core hole screening in  $LiCoO_2$ : A DFT + U calculation of the oxygen and cobalt K-edge x-ray absorption spectra. *Phys. Rev. B*, 81:115115.
- [Katsnelson et al., 2008] M. I. Katsnelson, V. Y. Irkhin, L. Chioncel, A. I. Lichtenstein, and R. A. Groot (2008). Half-metallic ferromagnets: From band structure to many-body effects. *Reviews of Modern Physics*, 80(2):315–378.
- [Klotz et al., 2009] S. Klotz, J.-C. Chervin, P. Munsch, and G. L. Marchand (2009). Hydrostatic limits of 11 pressure transmitting media. *Journal of Physics D: Applied Physics*, 42(7):075413.
- [Kokalj, 2003] A. Kokalj (2003). Computer graphics and graphical user interfaces as tools in simulations of matter at the atomic scale. *Comp. Mater. Sci.*, 28:155–168.
- [Korotin et al., 1998] M. A. Korotin, V. I. Anisimov, D. I. Khomskii, and G. A. Sawatzky (1998). CrO<sub>2</sub>: A Self-Doped Double Exchange Ferromagnet. *Physical Review Letters*, 80(19):4305-4308.
- [Laad et al., 2001] M. S. Laad, L. Craco, and E. Müller-Hartmann (2001). Orbital correlations in the ferromagnetic half-metal CrO<sub>2</sub>. *Phys. Rev. B*, 64(21):214421.
- [Loh, 2011] N. Loh (2011). Magnetic exchange interactions in compressed CrO<sub>2</sub> and Co/CuMn multilayers studied with polarised x-rays and neutrons. PhD thesis, The University of Austern Autralia.
- [Loubeyre1 et al., 2002] P. Loubeyre1, F. Occelli1, and R. LeToullec (2002). Optical studies of solid hydrogen to 320 GPa and evidence for black hydrogen. *Nature*, 416:613–617.
- [Maddox et al., 2006] B. R. Maddox, C. S. Yoo, D. Kasinathan, W. E. Pickett, and R. T. Scalettar (2006). High-pressure structure of half-metallic CrO<sub>2</sub>. *Physical Review B*, 73(14):144111.
- [Mankovsky et al., 2013] S. Mankovsky, S. Polesya, H. Ebert, W. Bensch, O. Mathon, S. Pascarelli, and J. Minar (2013). Pressure induced bcc to hcp transition in Fe: Magnetism-driven structure transformation. *Physical Review B*, 88.
- [Mao et al., 1967] H. Mao, W. A. Bassett, and T. Takahashi (1967). Effect of Pressure on Crystal Structure and Lattice Parameters of Iron up to 300 kbar. Journal of Applied Physics, 38(1):272–276.
- [Mao et al., 1986] H. K. Mao, J. Xu, and P. M. Bell (1986). Calibration of the ruby pressure gauge to 800 kbar under quasi-hydrostatic conditions. *Journal of Geophysical Research: Solid Earth*, 91(B5):4673-4676.

- [Mathon et al., 2004] O. Mathon, F. Baudelet, J. P. Itié, A. Polian, M. d'Astuto, J. C. Chervin, and S. Pascarelli (2004). Dynamics of the Magnetic and Structural  $\alpha$   $\epsilon$  Phase Transition in Iron. *Phys. Rev. Lett.*, 93:255503.
- [Mazin et al., 1999] I. I. Mazin, D. J. Singh, and C. Ambrosch-Draxl (1999). Transport, optical, and electronic properties of the half-metal CrO<sub>2</sub>. *Physical Review B*, 59(1):411-418.
- [Miletich et al., 2000] R. Miletich, D. R. Allan, and W. F. Kuhs (2000). High-Pressure Single-Crystal Techniques. *Reviews in Mineralogy and Geochemistry*, 41(1):445–519.
- [Mitsui and Hi, 2010] T. Mitsui and N. Hi (2010). Ultrahigh-pressure study on the magnetic state of iron hydride using an energy domain synchrotron radiation <sup>57</sup>Fe Mössbauer spectrometer. MRS Proceedings, 1262.
- [Murnaghan, 1944] F. Murnaghan (1944). The compressibility of media under extreme pressures. *Proceedings of the National Academy of Sciences*, 30(9):244–247.
- [Pépin et al., 2014] C. M. Pépin, A. Dewaele, G. Geneste, P. Loubeyre, and M. Mezouar (2014). New Iron Hydrides under High Pressure. *Phys. Rev. Lett.*, 113:265504.
- [P.Porta et al., 1972] P.Porta, M.Marezio, J.P.Remeika, and P.D.Dernier (1972). Chromium dioxide: High pressure synthesis and bond lengths. *Materials Research Bulletin*, 7(2):157 – 161.
- [Rao and Gopalakrishnan, 1997] C. Rao and J. Gopalakrishnan (1997). New directions in solid state chemistry. 2nd edition, Edition en anglais. Cambridge solid state science series. Cambridge University Press.
- [Ravel and Newville, 2005] B. Ravel and M. Newville (2005). ATHENA, ARTEMIS, HEP-HAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. Journal of Synchrotron Radiation, 12(4):537-541.
- [Sakamaki et al., 2009] K. Sakamaki, E. Takahashi, Y. Nakajima, Y. Nishihara, K. Funakoshi, T. Suzuki, and Y. Fukai (2009). Melting phase relation of FeHx up to 20 GPa: Implication for the temperature of the Earth's core. *Physics of the Earth and Planetary Interiors*, 174(1-4):192-201.
- [Schlapbach and Zuttel, 2001] L. Schlapbach and A. Zuttel (2001). Hydrogen-storage materials for mobile applications. *Nature*, 414(6861):353–358.
- [Schlottmann, 2003] P. Schlottmann (2003). Double-exchange mechanism for CrO<sub>2</sub>. Physical Review B, 67(17):174419.
- [Schlottmann, 2004] P. Schlottmann (2004). Double exchange and charge fluctuations in CrO2. Journal of Applied Physics, 95(11):7471–7473.
- [Schneider et al., 1991] G. Schneider, M. Baier, R. Wordel, F. Wagner, V. Antonov, E. Ponyatovsky, Y. Kopilovskii, and E. Makarov (1991). Mössbauer study of hydrides and deuterides of iron and cobalt. *Journal of the Less Common Metals*, 172:333 – 342.
- [Soleil, site] Soleil (website). Soleil Synchrotron accessed May 07, 2017. https://www.synchrotron-soleil.fr/en.
- [Soulen et al., 1998] R. J. Soulen, J. M. Byers, M. S. Osofsky, B. Nadgorny, T. Ambrose, S. F. Cheng, P. R. Broussard, C. T. Tanaka, J. Nowak, J. S. Moodera, A. Barry, and J. M. D. Coey (1998). Measuring the Spin Polarization of a Metal with a Superconducting Point Contact. Science, 282(5386):85–88.
- [Stavitski and de Groot, 2010] E. Stavitski and F. M. Groot (2010). The CTM4XAS program for EELS and XAS spectral shape analysis of transition metal L edges. *Micron*, 41(7):687 694.
- [Syassen, 2008] K. Syassen (2008). Ruby under pressure. High Pressure Research, 28(2):75–126.
- [Thamer et al., 1957] B. J. Thamer, R. M. Douglass, and E. Staritzky (1957). The Thermal Decomposition of Aqueous Chromic Acid and Some Properties of the Resulting Solid Phases1. Journal of the American Chemical Society, 79(3):547-550.
- [Tokura, 2003] Y. Tokura (2003). Correlated electron physics in transition metal oxides. *Phys. Today*, 56(7):50.
- [Torchio et al., 2011] R. Torchio, Y. O. Kvashnin, S. Pascarelli, O. Mathon, C. Marini, L. Genovese, P. Bruno, G. Garbarino, A. Dewaele, F. Occelli, and P. Loubeyre (2011). X-Ray Magnetic Circular Dichroism Measurements in Ni up to 200 GPa: Resistant Ferromagnetism. *Phys. Rev. Lett.*, 107(23):237202.
- [Toropova et al., 2005] A. Toropova, G. Kotliar, S. Savrasov, and V. Oudovenko (2005). Electronic structure and magnetic anisotropy of CrO2. *Physical Review B*, 71(17).
- [Wiedemann, 2007] H. Wiedemann (2007). Particle Accelerator Physics. Springer-Verlag Berlin Heidelberg.
- [Yamasaki et al., 2006] A. Yamasaki, L. Chioncel, A. I. Lichtenstein, and O. K. Andersen (2006). Model Hamiltonian parameters for half-metallic ferromagnets NiMnSb and CrO<sub>2</sub>. *Physical Review B*, 74(2):024419.
- [Zimmermann et al., 2017] P. Zimmermann, N. Bouldi, M. O. Hunault, M. Sikora, J. M. Ablett, J.-P. Rueff, B. Lebert, P. Sainctavit, F. M. Groot, and A. Juhin (2017). 1s2p Resonant Inelastic X-ray Scattering Magnetic Circular Dichroism as a probe for the local and nonlocal orbitals in CrO2. Journal of Electron Spectroscopy and Related Phenomena.

## **Conclusion and Outlook**

The main purpose of this thesis was to compute X-ray magnetic circular dichroism spectra at K-edge in order to provide a tool to interpret the, so far very puzzling, experimental spectra.

In order to be able to compute the XMCD signal, we first derived the semi-relativistic absorption cross section. For this, we determined the relativistic cross section from the quantum electrodynamic photon-matter Hamiltonian and then we performed the semi-relativistic expansion of the fully relativistic states that appear in this cross section. We obtained a result with only even states (which are 2-components states in a mono-electronic system). This is methodologically very important because if the derivation would be performed the other way around (semi-relativistic expansion of the Hamiltonian to deduce the cross section), the obtained result would not be correct. We applied this method to derive the semi-relativistic X-ray absorption and scattering cross sections at order  $c^{-2}$  but the same derivation could be extended up to any order.

The most noticeable novelties of our theoretical results are a transition operator that we named spin-position and a two-body term in the scattering cross section that has, to our knowledge, never been reported. The spin-position operator has been studied at length in the rest of this thesis but the importance of the two-body term in scattering is still an open question.

We then implemented the significant terms of the absorption cross section into the existing code XSpectra of the DFT suite QUANTUM ESPRESSO. The package uses a basis of plane waves and pseudopotentials, and calculations are performed for a crystal with periodic boundary conditions. Spectra calculations are performed after a full self consistent relaxation of the valence electrons in the presence of a core hole. To compute XMCD spectra, spin-orbit coupling must be included in the DFT Hamiltonian used to compute the sum over unoccupied states. We used an existing implementation which relies on the approximation of a diagonal spin-orbit coupling, which seems quite rough at first sight. Yet, by comparison with results from the code FDMNES in which spin-orbit coupling is fully implemented, we verified that this approximation does not impact significantly the calculated spectra. Also we showed that the inclusion of spinorbit coupling in the self-consistent calculation is not needed to obtain XMCD spectra. The spin-orbit coupling should be accounted in the self-consistent part of the calculation only in materials exhibiting strong spin-orbit coupling (e.g. iridates). It would not be computationally expensive to use full spin-orbit coupling in the second part of the calculation (that consists in determining the sum over unoccupied states) so that it would be a useful implementation to make.

With this code, we computed X-ray absorption and X-ray natural circular dichroism at the  $L_1$  edge of LiIO<sub>3</sub>. The calculated spectra are in good agreement with experiment and we numerically recovered the expected angular dependence. The study of LiIO<sub>3</sub> made us appreciate

the considerable impact of the core hole on the XNCD phenomenon: the calculated spectra with or without core hole are completely different. Hence, in addition to be a way to study the chirality of materials, XNCD could be used as a benchmark to understand the process of relaxation in the presence of a core hole.

To pursue the test of the code, we also calculated XAS and XMCD spectra at the K-edge in Fe, Co and Ni. These calculations showed that the electric dipole - spin position term is very important in XMCD. We explained it by deriving a sum-rule for this new term: it probes the spin polarization of the p states whereas the electric dipole-electric dipole term probes their orbital polarization. The overall agreement of the calculated spectra with the experimental one is fair but the secondary peaks are not always well reproduced. Also, the electric dipole-spin position term, whose existence is prescribed by the theory, does not improve the agreement between the calculated and the experimental spectra. We tested the effect on the spectra of a large number of assumptions and approximations of our method to better understand the origin of this discrepancy. We could narrow it down to intrinsic limitations of the spin-polarized Kohn-Sham DFT which is unable to capture the complexity of the spin and orbital polarization of the unoccupied states near the nucleus. Despite the fact that Kohn-Sham DFT is deemed not to be appropriate for excited states, it was surprisingly successful to compute XAS spectra. It is, therefore, possible that improvements of the way spin-polarized systems are treated could improve the modeling of XMCD.

The fact that the overall shape of the spectra is well described by our numerical method still makes it a useful tool for the interpretation of experiments. In this thesis we presented two examples that concern materials of technological interest: FeH and  $CrO_2$ .

The study on FeH turned out to be very illustrative of the point that motivated this PhD work, namely that a quick interpretation of XMCD spectra by computing their integral and interpreting it as being proportional to the total magnetic moment is not valid at K-edge. The XMCD spectra at the Fe K-edge in Fe and FeH are very different despite the similarity of the total magnetic moment in the cells. Our method was successful to reproduce the experimental spectra and we were able to investigate the magnetic structure of FeH with the various tools included in QUANTUM ESPRESSO.

The study on  $CrO_2$  is an example of the possibility of enrichment of theory by experiment. The literature on the electronic properties of  $CrO_2$  is very rich but also rather contradictory. A mechanism has been proposed to explain the origin of ferromagnetism but no general agreement has been reached on the way the material can be numerically modeled. In this context, confrontation of the numerical results with XAS and XMCD experiments under pressure is a valuable asset. Indeed, pressure is a unique way to change the local environment of the Cr atoms and the understanding of the impact of pressure on the electronic and magnetic structure of  $CrO_2$  could give precious clues on the interactions involved. Calculations without the Hubbard U correction were sufficient to reproduce the half-metallic behaviour of  $CrO_2$  and to infer the presence of a strong electric dipole - electric dipole contribution in the XAS pre-edge. However, the XMCD spectra was not accurately reproduced by these preliminary calculations. Further calculations must be pursued before drawing any conclusion but it is possible that DFT+U is required to model  $CrO_2$  correctly. My PhD work raised many questions and we answered some important ones but, unfortunately, there was not enough time to address them all.

In particular, we did not study further the consequences of Chapter 1 on the X-ray scattering cross section. Yet, the result that we obtained could have lots of ramifications. First of all, the spin-position operator, whose significance in XMCD has been proven in this thesis, could also have an impact on magnetic scattering. This needs to be examined with the support of numerical calculations. Secondly, the fact that Thomson scattering and non-resonant scattering arise from the existence of negative energy intermediate states, despite not being a novelty, was not investigated in detail. It changes the physical picture of these processes. And, finally, the new two-body term, also arising from the negative energy states, remains entirely to be investigated. The evaluation of this term is being undertaken in collaboration with Yves Joly (Institut Néel, Grenoble). In particular, we are investigating the possibility of an experimental situation in which this term could be distinguished from the other contributions to the cross section.

In this thesis, I demonstrated that the calculation of XMCD is a challenge to theoreticians. It is far more complex to compute a spectra than to obtain the total magnetic moment of a material. In particular, it is absolutely impossible to apply the same procedure that is often used to get the right magnetic moment with DFT-based calculations, that is to say change the input until the convergence occurs toward the expected moment. We have seen that there is still a long way to go before being able to capture all the complexity of the XMCD spectra by Kohn-Sham DFT. Among the large number of existing functionals, we tested only LDA and PBE. It could be interesting to try using other functionals but it is difficult to know *a priori* if one could be appropriate to model correctly XMCD. Full modeling of XMCD is an ambitious objective that requires further development of spin-polarized DFT.

Now, being able to reproduce the experiment is not an end in itself. It is very important that new information can be extracted from the calculation. It is, for example, very common to relate peaks in the absorption spectra to transitions toward specific orbitals as was done in this thesis for  $\text{CrO}_2$ . We can also determine if changes of the spectra observed experimentally are related to a given structural phase transition or to a magnetic phase transition as was done in this thesis for FeH. It is also practicable, but not easy, to extract crystal field parameters from a calculated band structure by projecting the Bloch states onto Wannier functions.<sup>8</sup> These parameters can then be used in multiplet calculations which allow to include multielectronic effects. However, in regards to long-range magnetism, DFT calculations do not provide the amplitude of the interactions involved. Yet, the knowledge of the relative importance of the different exchange interactions in a given magnetic system is very important to get a picture of the physics involved. For this purpose, model Hamiltonians, in which one can tune the interactions, are precious. It could be useful to develop bridges between both kinds of approaches.

The numerical method presented in this thesis could be used to study a large variety of systems. For example, we performed experiments of XAS and XMCD in two Co<sup>II</sup> based molecular compounds that undergo a pressure-induced spin cross over (SCO) transition at room temperature. One of them also undergoes a temperature-induced SCO transition at ambient pressure. The most common SCO systems are Fe<sup>II</sup> based and the transition occurs between a

<sup>&</sup>lt;sup>8</sup>M. W. Haverkort, M. Zwierzycki, and O. K. Andersen (2012). Multiplet ligand-field theory using Wannier orbitals. *Phys. Rev. B*, 85:165113

diamagnetic state (Fe<sup>II</sup> in low spin configuration i.e. with S=0) and a paramagnetic state (Fe<sup>II</sup> in high spin configuration i.e. with S=2). In that case, the magnetic signature of the transition can be addressed unambiguously by conventional magnetic measurements (SQUID or VSM magnetometers). In the case of divalent cobalt, the SCO transition corresponds to a transition between two paramagnetic states so that spectroscopic methods are required to determine the exact nature of the electronic states involved in the SCO transition. We followed the pressureinduced transition in both compounds and the temperature-induced transition of the compound that transit at low temperature during two beamtimes. It is necessary to separate the study of the pre-edge (that requires a multielectronic treatment) from the study of the edge. On the edge, preliminary calculations of XAS with our DFT based method are promising and we will pursue this study in order to compare the electronic structure of the low-spin and high-spin configurations. The multiplet calculations of the pre-edge spectra are tricky because the symmetry of the Co site in the structures is low and, therefore, many crystal-field parameters must be set in input. To obtain *ab initio* computed values for these parameters CASSSCF/CASPT2 calculations have been undertaken by a collaborating team in Toulouse (Laboratoire de chimie et physique quantique). We see that, a combination of several theoretical methods is required for the quantitative interpretation of the electronic and magnetic properties of the compounds.

The numerical method presented in this thesis could also be applied, with only minor adjustments, to another advanced spectroscopy: X-ray magneto chiral dichroism (XM $\chi$ D) that occurs when inversion symmetry and time-reversal symmetry are both broken. XM $\chi$ D is the difference of absorption, by a material that is both chiral and magnetic, of non-polarized light depending on the direction of the magnetization. XM $\chi$ D is a quite recent method: the first ever spectra was recorded in 2002 and, until recent years it was limited to a very few number of compounds. The development of new systems that exhibit molecular magnetism and which are both chiral and magnetic has recently raised a new interest for this technique.<sup>9</sup> Indeed, it exists only when magnetic properties are combined with the breaking of inversion symmetry and this combination is expected to play an important role in many phenomena as multiferroicity. The fact that both XNCD and XMCD can be computed with our method suggests that magneto chiral calculations could be possible.

<sup>&</sup>lt;sup>9</sup>R. Sessoli, M.-E. Boulon, A. Caneschi, M. Mannini, L. Poggini, F. Wilhelm and Andrei Rogalev (2015). Strong magneto-chiral dichroism in a paramagnetic molecular helix observed by hard X-rays. *Nature Physics*, 11:69-74

# Appendix A

# Abbreviations and Notations

### Abbreviations

| bcc                  | body-centered cubic (crystal system)                                |
|----------------------|---|
| CD                   | Circular Dichroism  |
| $^{\mathrm{ch}}$     | core hole   |
| D-D                  | electric Dipole - electric Dipole                                   |
| $\mathrm{DFT}$       | Density Functional Theory   |
| D-Q                  | electric Dipole - electric Quadrupole                               |
| D-SP                 | electric Dipole - Spin Position                                     |
| dhcp                 | double hexagonal close packed (crystal system)                      |
| DOS                  | Density Of States   |
| $\mathbf{ESRF}$      | European Synchrotron Radiation Facility                             |
| EXAFS                | Extended X-ray Absorption Fine Structure                            |
| fcc                  | face-centered cubic (crystal system)                                |
| FDMNES               | Finite Difference Method for Near-Edge Structure (code)             |
| FW                   | Foldy-Wouthuysen  |
| GGA                  | Generalized Gradient Approximation                                  |
| hcp                  | hexagonal close packed (crystal system)                             |
| IMPMC                | Institut de Minéralogie de Physique des matériaux et de cosmochimie |
| KS                   | Kohn–Sham   |
| LDA                  | Local-Density Approximation (DFT functional)                        |
| LSDA                 | Local-Spin-Density Approximation (DFT functional)                   |
| LFM                  | Ligand Field Multiplet  |
| ODE                  | Optique Dispersive EXAFS (beamline)                                 |
| PAW                  | Projector Augmented Wave (method)                                   |
| PBE                  | Perdew Burke Ernzerhof (DFT functional)                             |
| Q-Q                  | electric Quadrupole - electric Quadrupole                           |
| QED                  | Quantum ElectroDynamics   |
| RIXS                 | Resonant Inelastic X-ray Scattering                                 |
| $\operatorname{scf}$ | self consistent field   |
| SOC                  | Spin-Orbit Coupling   |
| SPR-KKR              | Spin Polarized Relativistic Korringa-Kohn-Rostoker                  |
| TFY                  | Total Fluorescence Yield  |
| XAS                  | X-ray Absorption Spectroscopy                                       |
| XANES                | X-ray Absorption Near-Edge Structure                                |
| XMCD                 | X-ray Magnetic Circular Dichroism                                   |
| XNCD                 | X-ray Natural Circular Dichroism                                    |
| XRD                  | X-ray Diffraction   |

### Notations

### Matrix and operators

| р                          | momentum operator $\mathbf{p} = -i\hbar \nabla$ |                    |
|----------------------------|---|--------------------|
| $\pi$                      | $\mathbf{p} - e\mathbf{A}$ mechanical momentum  |                    |
| $\mathbf{L}$               | angular momentum                                |                    |
| $\tilde{I}$                | Identity matrix                                 |                    |
| $\epsilon_{jkl}$           | Levi-Civita symbols                             |                    |
| $\sigma_i$                 | Pauli matrices                                  | (see Appendix B.1) |
| $\beta$ and $\pmb{\alpha}$ | Standard Dirac matrices                         | (see Appendix B.3) |

### Constants

| e | charge of the electron   |
|---|--------------------------|
| с | speed of light in vacuum |
| m | electron mass            |

- $\hbar$  reduced Planck constant
- $\epsilon_0$  vacuum permittivity
- $\alpha_0$  fine structure constant
- $r_e$  classical electron radius
- $a_0$  Bohr radius
- $\mu_B$  Bohr magneton

$$\begin{array}{l} \alpha_0 = \frac{e^2}{4\pi\hbar c\epsilon_0} \\ r_e = \frac{e^2}{4\pi\epsilon_0 mc^2} \\ a_0 = \frac{\hbar^2 4\pi\epsilon_0}{me^2} \\ \mu_B = \frac{e\hbar}{2m} \end{array}$$

e < 0

### Electromagnetic wave

- **B** magnetic field
- E electric field
- **A** vector potential
- V scalar potential
- k wave vector
- $\epsilon$  polarization vector

## Appendix B

## Definitions and useful formula

### **B.1** Pauli Matrices

The Pauli matrices are dimensionless operators:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
(B.1)

They possess the following property, which can easily be verified from their explicit form:

$$\sigma_j \sigma_k = i \sum_l \epsilon_{jkl} \sigma_l + \delta_{jk} \tilde{I}. \tag{B.2}$$

The vector  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  is often used to designate the set of three Pauli matrices. For a spin 1/2 particle, the spin operator writes

$$\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}.\tag{B.3}$$

The eigenvectors of  $\sigma_z$  are noted  $|\uparrow\rangle$  and  $|\downarrow\rangle$  where the spin up spinor is  $|\uparrow\rangle = \begin{pmatrix} 1\\0 \end{pmatrix}$  and the spin down spinor is  $|\downarrow\rangle = \begin{pmatrix} 0\\1 \end{pmatrix}$ . The application of  $\sigma_x$  of  $\sigma_y$  on  $|\uparrow\rangle$  changes it in  $|\downarrow\rangle$  and vice versa:  $\sigma_z|\uparrow\rangle = |\uparrow\rangle \quad \sigma_y|\uparrow\rangle = i|\downarrow\rangle \quad \sigma_x|\uparrow\rangle = |\downarrow\rangle$ (D.4)

$$\begin{array}{ccc}
\sigma_z |\downarrow\rangle = |\downarrow\rangle & \sigma_y |\downarrow\rangle = i|\downarrow\rangle & \sigma_x |\downarrow\rangle = |\uparrow\rangle \\
\sigma_z |\downarrow\rangle = -|\downarrow\rangle & \sigma_y |\downarrow\rangle = -i|\uparrow\rangle & \sigma_x |\downarrow\rangle = |\uparrow\rangle
\end{array}$$
(B.4)

So that if,  $|g\rangle = f_1|\uparrow\rangle + f_2|\downarrow\rangle$  then

$$\langle g | \boldsymbol{\sigma} | g \rangle = \begin{pmatrix} 2 \operatorname{Re}(f_1^* f_2) \\ 2 \operatorname{Im}(f_1^* f_2) \\ |f_1|^2 - |f_2|^2 \end{pmatrix}$$
(B.5)

### B.2 Angular Momentum matrix for l=1

In the basis  $|lm\rangle = |1-1\rangle, |10\rangle, |11\rangle$ , the components of the angular momentum operator write as  $3\times 3$  matrices. As  $\langle lm'|L_z|lm\rangle = m\hbar\delta_{m'm}$ ,

$$L_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

As  $\langle lm'|L_{\pm}|lm\rangle = \sqrt{l(l+1) - m(m\pm 1)}\hbar\delta_{m'm}$ ,

$$L_{+} = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \text{ and } L_{-} = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

This leads to

$$L_x = \frac{1}{2}(L_+ + L_-) = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0 \end{pmatrix}$$

and

$$L_y = \frac{1}{2i}(L_+ - L_-) = \frac{\hbar}{i\sqrt{2}} \begin{pmatrix} 0 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & 0 \end{pmatrix}.$$

### **B.3** Dirac Matrices

We note  $\beta$  and  $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$  the standard Dirac matrices:

$$\beta = \begin{pmatrix} \tilde{I} & 0\\ 0 & -\tilde{I} \end{pmatrix} \text{ and } \alpha_i = \begin{pmatrix} 0 & \sigma_i\\ \sigma_i & 0 \end{pmatrix}.$$
(B.6)

Some of their basic properties are:<sup>1</sup>

$$\beta^2 = 1 \tag{B.7}$$

$$\{\alpha_i,\beta\} = 0 \tag{B.8}$$

$$\{\alpha_j, \alpha_k\} = 2\delta_{jk} \tag{B.9}$$

$$[\alpha_j, \alpha_k] = 2(\alpha_j \alpha_k - \delta_{jk}) = 2i \sum_m \epsilon_{jkm} \Sigma_m$$
(B.10)

$$\alpha_j \alpha_k = \delta_{jk} + i \sum_m \epsilon_{jkm} \Sigma_m \tag{B.11}$$

In the two last equations  $\Sigma_m = \begin{pmatrix} \sigma_m & 0 \\ 0 & \sigma_m \end{pmatrix}$ . We also define the vector  $\Sigma$  as  $(\Sigma_x, \Sigma_y, \Sigma_z)$ . The product  $\beta \alpha$  is:

The product  $\beta \alpha$  is:

$$\beta \boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix} \tag{B.12}$$

and from (B.8), we see that  $\boldsymbol{\alpha}\beta = -\beta\boldsymbol{\alpha}$ .

### B.4 Commutators

• The canonical commutation relation writes  $[r_i, p_j] = i\hbar \delta_{ij}$ .

 $<sup>^1\</sup>mathrm{B.}$  Thaller, The Dirac Equation, Springer-Verlag Berlin Heidelger (1992) p. 37

• If **a** and **b** are vectors that commute with  $\boldsymbol{\alpha}$ :

$$\begin{aligned} [\boldsymbol{\alpha} \cdot \mathbf{a}, \boldsymbol{\alpha} \cdot \mathbf{b}] &= \sum_{ij} (\alpha_i \alpha_j a_i b_j - \alpha_j \alpha_i b_j a_i) \\ &= \sum_{ij} \delta_{ij} [a_i, b_j] + i \sum_{mij} \epsilon_{ijm} \Sigma_m (a_i b_j + b_j a_i) \\ &= \sum_i [a_i, b_i] + i \boldsymbol{\Sigma} \cdot (\mathbf{a} \times \mathbf{b} - \mathbf{b} \times \mathbf{a}) \end{aligned}$$

# Appendix C

## Point groups



**Fig. C.1:** Stereographic projections of the crystal classes. The centrosymetric groups are in orange, the chiral groups in green and the non-centrosymetric non-chiral groups in yellow. The name of the 13 crystal classes that can exhibit circular dichroism associated to the cross term electric dipole-electric quadrupole (D-Q) are boxed.

**References:** Sylvain Ravy - Cours Structure de la matière condensée- www.lps.u-psud.fr/spip.php? article531 (table of stereographic projections)

Rogalev, A., Goulon, J., Wilhelm, F., and Bosak, A. (2010). X-Ray Detected Optical Activity. In Beaurepaire, E., Bulou, H., Scheurer, F., and Kappler, J.- P., editors, Magnetism and Synchrotron Radiation, number 133 in Springer Proceedings in Physics (name of the crystal classes that exhibit D-Q XNCD)

# Appendix D

# Operators in XSpectra code

### D.1 Spherical harmonics

Below, a list of the spherical harmonics in spherical and cartesian coordinates assuming that

$$x = r \sin \theta \cos \varphi$$
$$y = r \sin \theta \sin \varphi$$
$$z = r \cos \theta$$

For l = 0

$$Y_0^0(\theta,\varphi) = \sqrt{\frac{1}{4\pi}}$$

For l = 1

$$Y_1^{-1}(\theta,\varphi) = \sqrt{\frac{3}{8\pi}} \cdot e^{-i\varphi} \cdot \sin\theta = \sqrt{\frac{3}{8\pi}} \cdot \frac{(x-iy)}{r}$$
$$Y_1^0(\theta,\varphi) = \sqrt{\frac{3}{4\pi}} \cdot \cos\theta = \sqrt{\frac{3}{4\pi}} \cdot \frac{z}{r}$$
$$Y_1^1(\theta,\varphi) = -\sqrt{\frac{3}{8\pi}} \cdot e^{i\varphi} \cdot \sin\theta = -\sqrt{\frac{3}{8\pi}} \cdot \frac{(x+iy)}{r}$$

For 
$$l = 2$$

$$Y_2^{-2}(\theta,\varphi) = \sqrt{\frac{15}{32\pi}} \cdot e^{-2i\varphi} \cdot \sin^2 \theta = \sqrt{\frac{15}{32\pi}} \cdot \frac{(x-iy)^2}{r^2}$$

$$Y_2^{-1}(\theta,\varphi) = \sqrt{\frac{15}{8\pi}} \cdot e^{-i\varphi} \cdot \sin \theta \cdot \cos \theta = \sqrt{\frac{15}{8\pi}} \cdot \frac{(x-iy)z}{r^2}$$

$$Y_2^0(\theta,\varphi) = \sqrt{\frac{15}{4\pi}} \cdot (3\cos^2 \theta - 1) = \sqrt{\frac{15}{4\pi}} \cdot \frac{(2z^2 - x^2 - y^2)}{r^2}$$

$$Y_2^1(\theta,\varphi) = -\sqrt{\frac{15}{8\pi}} \cdot e^{i\varphi} \cdot \sin \theta \cdot \cos \theta = -\sqrt{\frac{15}{8\pi}} \cdot \frac{(x+iy)z}{r^2}$$

$$Y_2^2(\theta,\varphi) = \sqrt{\frac{15}{32\pi}} \cdot e^{2i\varphi} \cdot \sin^2 \theta = \sqrt{\frac{15}{32\pi}} \cdot \frac{(x+iy)^2}{r^2}$$

The spherical harmonics verify the orthogonality relation:

$$\int Y_l^m(\mathbf{u}_{\mathbf{r}})Y_{l'}^{m'}(\mathbf{u}_{\mathbf{r}})\mathrm{d}\theta\mathrm{d}\phi\sin\theta = \delta_{l,l'}\delta_{m,m'}$$

The integral of the product of three spherical harmonics is given by:<sup>1</sup>

$$\int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} Y_{l}^{m*} Y_{l_{1}}^{m_{1}} Y_{l_{2}}^{m_{2}} d\Omega = \sqrt{\frac{(2l_{1}+1)(2l_{2}+1)}{4\pi(2l+1)}} (l_{1} 0 l_{2} 0 | l 0) (l_{1} m_{1} l_{2} m_{2} | l m)$$

### D.2 Operators

The three operators that enter the different terms of the absorption cross section are:

$$D = \boldsymbol{\epsilon} \cdot \mathbf{r},$$
  

$$SP = \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}),$$
  

$$Q = \frac{1}{2|\mathbf{k}|} (\boldsymbol{\epsilon} \cdot \mathbf{r}) (\mathbf{k} \cdot \mathbf{r})$$

Using the spherical harmonics above, we can rewrite

$$D = \frac{4\pi}{3} r \sum_{\lambda=-1}^{\lambda=1} (-1)^{\lambda} Y_1^{-\lambda}(\boldsymbol{\epsilon}) Y_1^{\lambda}(\mathbf{u_r})$$
(D.1)

In the case of collinear spins with diagonal spin-orbit coupling:

$$SP_{\rm col} = \sigma_z(\epsilon_x y - \epsilon_y x) = \sigma_z \frac{4i\pi}{3} r(Y_1^{-1}(\boldsymbol{\epsilon})Y_1^1(\mathbf{u_r}) - Y_1^1(\boldsymbol{\epsilon})Y_1^{-1}(\mathbf{u_r}))$$
(D.2)

The electric quadrupole operator writes:

$$Q(\mathbf{r}) = \left(\frac{4\pi}{3}\right)^2 \frac{r^2}{2} \sum_{\lambda=-1}^{\lambda=1} \sum_{\mu=-1}^{\mu=1} (-1)^{\lambda+\mu} Y_1^{-\lambda}(\boldsymbol{\epsilon}) Y_1^{\lambda}(\mathbf{u_r}) Y_1^{\mu}(\mathbf{u_r}) Y_1^{-\mu}(\mathbf{k})$$
  
As  $Y_1^{\lambda}(\mathbf{u_r}) Y_1^{\mu}(\mathbf{u_r}) = \sum_{\nu=0}^2 \sqrt{\frac{3*3}{4\pi(2\nu+1)}} Y_{\nu}^{\lambda+\mu} (1\lambda 1\mu | \nu\lambda + \mu) (1010 | \nu0)$  and  
 $(1010 | \nu0) = -\sqrt{\frac{1}{3}} \qquad \text{if } \nu = 0$   
 $= 0 \qquad \text{if } \nu = 1$   
 $= \sqrt{\frac{2}{3}} \qquad \text{if } \nu = 2.$ 

The product of two spherical harmonics with l = 1 writes:

$$Y_{1}^{\lambda}(\mathbf{u_{r}})Y_{1}^{\mu}(\mathbf{u_{r}}) = \sqrt{\frac{3}{4\pi}}Y_{0}^{\lambda+\mu}(1\lambda1\mu|0\lambda+\mu) + \sqrt{\frac{3}{10\pi}}Y_{2}^{\lambda+\mu}(1\lambda1\mu|2\lambda+\mu).$$

This formula involves two Clebsch–Gordan coefficients whose values are:

| (               | $1\lambda 1\mu 0$    | $(\lambda + \mu)$     |                      | (1                  | $\lambda 1\mu  2$    | $\lambda + \mu$      |                      |
|-----------------|----------------------|-----------------------|----------------------|---------------------|----------------------|----------------------|----------------------|
| $\lambda / \mu$ | -1                   | 0                     | 1                    | $\lambda \ / \ \mu$ | -1                   | 0                    | 1                    |
| -1              | 0                    | 0                     | $\sqrt{\frac{1}{3}}$ | -1                  | 1                    | $\sqrt{\frac{1}{2}}$ | $\sqrt{\frac{1}{6}}$ |
| 0               | 0                    | $-\sqrt{\frac{1}{3}}$ | 0                    | 0                   | $\sqrt{\frac{1}{2}}$ | $\sqrt{\frac{2}{3}}$ | $\sqrt{\frac{1}{2}}$ |
| 1               | $\sqrt{\frac{1}{3}}$ | 0                     | 0                    | 1                   | $\sqrt{\frac{1}{6}}$ | $\sqrt{\frac{1}{2}}$ | 1                    |

<sup>&</sup>lt;sup>1</sup>Sakurai and Napolitano-Modern quantum mechanics-Revised edition p.217

The fact that  $(1\lambda 1\mu|0\lambda + \mu)$  is zero except if  $\lambda + \mu = 0$  implies that there is a term proportional to  $Y_0^0 \sum_{\lambda=-1}^{\lambda=1} (-1)^{\lambda} Y_1^{\lambda}(\boldsymbol{\epsilon}) Y_1^{\lambda}(\mathbf{k})$  in Q. This is proportional to  $\boldsymbol{\epsilon} \cdot \mathbf{k}$ . So, as  $\boldsymbol{\epsilon} \cdot \mathbf{k} = 0$ , this term is zero. Finally,

$$Q = \left(\frac{4\pi}{3}\right)^2 \sqrt{\frac{3}{10\pi}} \frac{r^2}{2} \sum_{\lambda=-1}^{\lambda=1} \sum_{\mu=-1}^{\mu=1} (-1)^{\lambda+\mu} (1\lambda 1\mu | 2\lambda+\mu) Y_1^{-\lambda}(\boldsymbol{\epsilon}) Y_2^{\lambda+\mu}(\mathbf{u_r}) Y_1^{-\mu}(\mathbf{k})$$
(D.3)

### D.3 Initial vector $|\tilde{\varphi}_{\mathbf{R}_0}\rangle$

We must compute

$$|\tilde{\varphi}_{\mathbf{R}_{0}}\rangle = \sum_{n} |\tilde{p}_{n}^{\mathbf{R}_{0}}\rangle \langle \phi_{n}^{\mathbf{R}_{0}}|O|i\rangle$$

with n referring to (l, m) and an additional number  $\varsigma$ . We write:

$$\begin{split} \langle \mathbf{r} | i \rangle &= f_0^i(r) Y_0^0(\mathbf{u_r}) = \frac{f_0^i(r)}{\sqrt{4\pi}} \\ \langle \mathbf{r} | \tilde{p}_n^{\mathbf{R}_0} \rangle &= f_{\varsigma,l}^p(r) Y_l^m(\mathbf{u_r}) \\ \langle \mathbf{r} | \phi_n^{\mathbf{R}_0} \rangle &= f_{\varsigma,l}^\phi(r) Y_l^m(\mathbf{u_r}) \end{split}$$

The corresponding variables in the code are listed in Table D.1.

If one considers an operator O which is a linear combination of  $Y_l^m(\mathbf{u_r})Y_{l'}^{m'}(\boldsymbol{\epsilon})$ :

$$O = \sum_{a} C_{a}(r) Y_{l_{a}}^{m_{a}}(\mathbf{u}_{\mathbf{r}}) Y_{l_{a}'}^{m_{a}'}(\boldsymbol{\epsilon}),$$

Then,

$$\begin{aligned} \langle \mathbf{r} | \tilde{\varphi}_{\mathbf{R}_0} \rangle &= \sum_{\varsigma,l,m} f_{\varsigma,l}^p(r) Y_l^m(\mathbf{u}_{\mathbf{r}}) \int \mathrm{d}\mathbf{r}' f_{\varsigma,l}^\phi(r) Y_l^m(\mathbf{u}_{\mathbf{r}}')^* \left( \sum C_a(r') Y_{l_a}^{m_a}(\mathbf{u}_{\mathbf{r}}') Y_{l_a'}^{m_a'}(\boldsymbol{\epsilon}) \right) \frac{f_0^i(r')}{\sqrt{4\pi}} \\ &= \sum_a \frac{1}{\sqrt{4\pi}} f_{\varsigma,l_i}^p(r) Y_{l_a}^{m_a}(\mathbf{u}_{\mathbf{r}}) Y_{l_a'}^{m_a'}(\boldsymbol{\epsilon}) \int C_a(r') r'^2 f_{\varsigma,l_a}^\phi(r') f_0^i(r') \mathrm{d}r' \end{aligned}$$

This formula can be applied to the three operators of interest.

#### Electric-dipole operator

Using Eq. (D.1) we get:

$$\begin{aligned} \langle \mathbf{r} | \tilde{\varphi}_{\mathbf{R}_0} \rangle &= \frac{\sqrt{4\pi}}{3} \sum_{\varsigma} f_{\varsigma,l}^p(r) \left( \sqrt{\frac{3}{8\pi}} Y_1^{-1}(\mathbf{u}_{\mathbf{r}})(\epsilon_x + i\epsilon_y) \right. \\ &+ \sqrt{\frac{3}{4\pi}} Y_1^0(\mathbf{u}_{\mathbf{r}})\epsilon_z - \sqrt{\frac{3}{8\pi}} Y_1^1(\mathbf{u}_{\mathbf{r}})(\epsilon_x - i\epsilon_y) \right) \int r'^3 f_{\varsigma,l}^\phi(r') f_0^i(r') \mathrm{d}r' \end{aligned}$$

#### Electric-quadrupole operator

Using Eq. (D.3), with  $\hat{\mathbf{k}} = \frac{\mathbf{k}}{|k|}$ :

$$\begin{split} \langle \mathbf{r} | \tilde{\varphi}_{\mathbf{R}_{0}} \rangle &= \sum_{\varsigma} f_{\varsigma,l}^{p}(r) \left( Y_{2}^{-2}(\mathbf{u}_{\mathbf{r}}) \frac{1}{\sqrt{3 \star 40}} (\epsilon_{x} + i\epsilon_{y}) (\hat{k}_{x} + i\hat{k}_{y}) \right. \\ &+ Y_{2}^{2}(\mathbf{u}_{\mathbf{r}}) \frac{1}{\sqrt{3 \star 40}} (\epsilon_{x} - i\epsilon_{y}) (\hat{k}_{x} - i\hat{k}_{y}) \\ &+ Y_{2}^{0}(\mathbf{u}_{\mathbf{r}}) \frac{\sqrt{2}}{3\sqrt{40}} (2\epsilon_{z}\hat{k}_{z} - \epsilon_{x}\hat{k}_{x} - \epsilon_{y}\hat{k}_{y}) \\ &+ Y_{2}^{-1}(\mathbf{u}_{\mathbf{r}}) \frac{1}{\sqrt{3 \star 40}} \left[ (\epsilon_{x} + i\epsilon_{y})\hat{k}_{z} + \epsilon_{z}(\hat{k}_{x} + i\hat{k}_{y}) \right] \\ &- Y_{2}^{1}(\mathbf{u}_{\mathbf{r}}) \frac{1}{\sqrt{3 \star 40}} \left[ (\epsilon_{x} - i\epsilon_{y})\hat{k}_{z} + \epsilon_{z}(\hat{k}_{x} - i\hat{k}_{y}) \right] \right) \int r'^{4} f_{\varsigma,l}^{\phi}(r') f_{0}^{i}(r') dr'. \end{split}$$

#### Spin-Position operator

Using Eq. (D.2),

$$\begin{aligned} \langle \mathbf{r} | \tilde{\varphi}_{\mathbf{R}_0} \rangle &= \operatorname{sign}(\sigma_z) \frac{\sqrt{4\pi}}{3} i \sum_{\varsigma} f_{\varsigma,l}^p(r) \Big( \sqrt{\frac{3}{8\pi}} Y_1^1(\mathbf{u}_{\mathbf{r}})(\epsilon_x - i\epsilon_y) \\ &+ \sqrt{\frac{3}{8\pi}} Y_1^{-1}(\mathbf{u}_{\mathbf{r}})(\epsilon_x + i\epsilon_y) \Big) \int r'^3 f_{\varsigma,l}^\phi(r') f_0^i(r') \mathrm{d}r' \end{aligned}$$

where  $\operatorname{sign}(\sigma_z)$  is the sign of the spin associated to the current k.

| Table D.1: | Correspondence | between som | e variables | in XSpectra | and the | corresponding | quantities. |
|------------|----------------|-------------|-------------|-------------|---------|---------------|-------------|
|------------|----------------|-------------|-------------|-------------|---------|---------------|-------------|

| core_wfn                                | $rf_0^i(r)$   |
|---|---|
| <pre>paw_recon(xiabs) % aephi(ip)</pre> | %psi $rf_{arsigma,l}^{\phi}$  |
| $\texttt{paw\_vkb}^a$                   | $\sum_{arsigma} y_1^m(\mathbf{u_r}) f_{arsigma,l}^p(r)$                 |
| paw_vkb_cplx                            | $\sum_{arsigma} Y_1^m(\mathbf{u_r}) f_{arsigma,l}^p(r)$                 |
| psiwfc                                  | $\langle {f r}   	ilde{arphi}_{{f R}_0}  angle$                         |
| xnorm                                   | $\sqrt{\langle 	ilde{arphi}_{{f R}_0}   	ilde{arphi}_{{f R}_0}  angle}$ |
| $\texttt{xanes\_dip}^b$                 | $\int r'^3 f^\phi_{\varsigma,l}(r') f^i_0(r') \mathrm{d}r'$             |
| xanes_qua                               | $\int r'^4 f^\phi_{\varsigma,l}(r') f^i_0(r') \mathrm{d}r'$             |
| xepsilon(1:3)                           | $\epsilon$  |
| xkvec(1:3)                              | $\hat{\mathbf{k}}$  |
| sign_sigma                              | $\operatorname{sign}(\sigma_z)$   |

<sup>&</sup>lt;sup>*a*</sup> In Quantum Espresso, the real harmonics are defined such that  $y_1^1 \propto -x$ ,  $y_1^{-1} \propto -y$  and  $y_1^0 \propto z$ . <sup>*b*</sup> Note that the radial integrals are the same for all *k*-points. They are computed at the beginning of the calculation.

## D.4 Prefactors

In Rydberg atomic units,  $\hbar = 2m = 1$  and  $c = \frac{2}{\alpha}$ , so  $|\mathbf{k}| = \frac{\omega}{c} = \frac{\omega\alpha}{2}$ . So that,

$$\sigma_{\rm D-Q} = -8\pi^2 \omega \alpha \sum_f \operatorname{Im} \langle f | \frac{1}{2} (\boldsymbol{\epsilon} \cdot \mathbf{r}) (\mathbf{k} \cdot \mathbf{r}) | i \rangle \langle i | \boldsymbol{\epsilon}^* \cdot \mathbf{r} | f \rangle \delta(E_f - E_i - \hbar \omega)$$
  
$$\sigma_{\rm D-SP} = -\pi^2 \omega^2 \alpha^3 \sum_f \operatorname{Im} \langle f | \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}) | i \rangle \langle i | \boldsymbol{\epsilon}^* \cdot \mathbf{r} | f \rangle \delta(E_f - E_i - \hbar \omega)$$

If one considers  $I_1$  and  $I_2$  as they are defined in subsection 2.3.2 then,

$$\sigma_{\rm D-Q} = 2\pi\omega\alpha |\mathbf{k}| (I_1 - I_2) = \pi\omega^2 \alpha^2 (I_1 - I_2)$$
  
and  $\sigma_{\rm D-SP} = \frac{1}{4}\pi\omega^2 \alpha^3 (I_1 - I_2).$ 

## Appendix E

## Technical details of the pseudopotentials

The pseudopotentials have been generated with the code atomic (ld1.x) of Quantumespresso. The complete description of the input of ld1.x can be found in the directory atomic/Doc of Quantum Espresso or online.<sup>1</sup> A documentation is also available in the directory atomic/Doc: Notes on pseudopotential generation - Paolo Gianozzi .<sup>2</sup>

The reader is referred to these documents for details on the pseudopotentials generation.

## E.1 Norm conserving (NC) pseudopotentials

| Element             | functional | type of PP | lloc | generation    | cut-off | GIPAW                       | energy position    |
|---------------------|------------|------------|------|---------------|---------|-----------------------------|--------------------|
|                     |            |            |      | configuration | radius  | $\operatorname{projectors}$ | of projectors (Ry) |
| Fe                  | PBE        | TM SR      | 2    | $3s^{2}$      | 1.1     | 3p                          | 0.0                |
|                     |            |            |      | $3p^6$        | 1.1     | 4p                          | 0.0                |
|                     |            |            |      | $3d^6$        | 1.1     | 3d                          | 0.0                |
|                     |            |            |      |               |         | 4d                          | 0.0                |
| Co                  | PBE        | TM SR      | 1    | $3s^{2}$      | 1.1     | 3p                          | 0.0                |
|                     |            |            |      | $3p^6$        | 1.1     | 4p                          | 0.0                |
|                     |            |            |      | $3d^{7}$      | 1.1     | 3d                          | 0.0                |
|                     |            |            |      |               |         | 4d                          | 0.0                |
| Ni                  | PBE        | TM SR      | 2    | $3s^{2}$      | 1.1     | 3p                          | 0.0                |
|                     |            |            |      | $3p^6$        | 1.1     | 4p                          | 0.0                |
|                     |            |            |      | $3d^{8}$      | 1.1     | 3d                          | 0.0                |
|                     |            |            |      |               |         | 4d                          | 0.0                |
| Н                   | PBE        | TM SR      | 0    | $1s^1$        | 0.9     | -                           | -                  |
| $\operatorname{Cr}$ | PBE        | TM SR      | 1    | $3s^{2}$      | 1.0     | 3p                          | 0.0                |
|                     |            |            |      | $3p^6$        | 1.7     | 4p                          | 0.0                |
|                     |            |            |      | $3d^2$        | 1.4     | 3d                          | 0.0                |
|                     |            |            |      |               | 1.4     | 4d                          | 0.0                |
| Ο                   | PBE        | TM SR      | 2    | $2s^2$        | 1.0     | -                           | -                  |
|                     |            |            |      | $2p^4$        | 1.0     | -                           | -                  |

Table E.1: Technical details of the norm-conserving pseudopotentials used in this thesis.

<sup>&</sup>lt;sup>1</sup>URL for ld1 input: http://www.quantum-espresso.org/wp-content/uploads/Doc/INPUT\_LD1.html# idm140471497748848

<sup>&</sup>lt;sup>2</sup>URL for documentation: www.quantum-espresso.org/wp-content/uploads/Doc/pseudo-gen.pdf

- PBE stands for Perdew, Burke, and Ernzerhof exchange-correlation functional.<sup>3</sup>
- TM is the used pseudization algorithm: Troullier-Martins.<sup>4</sup>
- SR stands for scalar-relativistic
- lloc is the angular momentum of the local chanel.
- For K- and  $L_1$ -edge spectra calculations GIPAW p- (for electric dipole transitions) and dprojectors (for electric quadrupole transitions) must be defined for the absorbing atom. These projectors are used in XSpectra to compute the cross-section (see Chapter 2). If their energy position is set to 0.0, the one-electron energy of the all-electron state is used to pseudize the corresponding states.

### E.2 Ultrasoft pseudopotentials

| Element | functio-             | type of PP | lloc | generation                     | cut-off | US     | energy | pro-    | energy                    |
|---------|----------------------|------------|------|--------------------------------|---------|--------|--------|---------|---------------------------|
|         | $\operatorname{nal}$ |            |      | $\operatorname{configuration}$ | radius  | radius |        | jectors | $\operatorname{position}$ |
| Li      | PBE                  | TM SR      | -1   | $1s^2$                         | 0.7     | 1.2    | 0.0    | -       | -                         |
|         |                      |            |      | $2s^2$                         | 0.7     | 1.2    | 0.0    |         |                           |
|         |                      |            |      | $2p^0$                         | 2.0     | 2.2    | 0.3    |         |                           |
|         |                      |            |      | $2p^0$                         | 2.0     | 2.2    | 1.7    |         |                           |
| Ι       | PBE                  | TM SR      | -1   | $5s^2$                         | 1.3     | 1.7    | 0.0    | 5p      | 0.0                       |
|         |                      |            |      | $5s^0$                         | 1.3     | 1.7    | 6.0    | 6p      | 6.0                       |
|         |                      |            |      | $5p^5$                         | 1.4     | 1.8    | 0.0    | 4d      | 0.0                       |
|         |                      |            |      | $5p^0$                         | 1.4     | 1.8    | 6.0    | 5d      | -0.5                      |
|         |                      |            |      | $4d^{10}$                      | 1.3     | 1.7    | 0.0    |         |                           |
|         |                      |            |      | $4d^0$                         | 1.3     | 1.7    | -2.5   |         |                           |
| О       | PBE                  | TM SR      | 2    | $2s^2$                         | 1.0     | 1.35   | 0.0    | _       | _                         |
|         |                      |            |      | $2s^0$                         | 1.0     | 1.35   | 0.05   |         |                           |
|         |                      |            |      | $2p^4$                         | 1.0     | 1.35   | 0.0    |         |                           |
|         |                      |            |      | $2p^0$                         | 1.0     | 1.35   | 0.05   |         |                           |

Table E.2: Technical details of the ultrasoft pseudopotentials used in this thesis.

For Ultrasoft pseudopotentials, an additional cut-off radius for charge compensation must be defined in addition to the cut-off radius for the wave-functions. This Ultrasoft radius must be larger than the other one.

A negative lloc means that the local potential is generated by pseudizing the all electron potential.

 $<sup>^3</sup>$  J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett.  $~77,\,3865~(1996)$ 

<sup>&</sup>lt;sup>4</sup>N. Troullier and J.L. Martins, Phys. Rev. B 43, 1993 (1991).

# Appendix F

## Articles

This appendix contains three articles that are mentioned in the manuscript:

- N. Bouldi and Ch. Brouder (2017). Gauge invariance and relativistic effects in photon absorption and scattering by matter. Accepted in Euro. Phys. J. B
- N. Bouldi, N. J. Vollmers, C. G. Delpy-Laplanche, Y. Joly, A. Juhin, Ph. Sainctavit, Ch. Brouder, M. Calandra, L. Paulatto, F. Mauri, and U. Gerstmann (2017). X-ray magnetic and natural circular dichroism from first principles: Calculation of K- and L<sub>1</sub>-edge spectra. Phys. Rev. B 96, 085123
- P. Zimmermann, N. Bouldi, M. O.J.Y.Hunault, M. Sikora, J. M.Ablett, J.-P. Rueff, B. Lebert, Ph. Sainctavit, F. M.F.de Groot and A. Juhin. (2017) 1s2p Resonant Inelastic X-ray Scattering Magnetic Circular Dichroism as a probe for the local and non-local orbitals in CrO2 1s2p Resonant Inelastic X-ray Scattering Magnetic Circular Dichroism as a probe for the local and non-local orbitals in CrO2. J. Electron. Spectrosc. Relat. Phenom.

# Gauge invariance and relativistic effects in x-ray absorption and scattering by solids

Nadejda Bouldi<sup>12</sup> and Christian Brouder<sup>1</sup>

<sup>1</sup> Sorbonne Universités, UPMC Univ Paris 06, UMR CNRS 7590, Muséum National d'Histoire Naturelle, IRD UMR 206, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, A place Luccium, E 75007, Duris, France.

4 place Jussieu, F-75005 Paris, France

<sup>2</sup> Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, BP 48, 91192 Gif-sur-Yvette Cedex, France

Received: date / Revised version: date

**Abstract.** There is an incompatibility between gauge invariance and the semi-classical time-dependent perturbation theory commonly used to calculate light absorption and scattering cross-sections. There is an additional incompatibility between perturbation theory and the description of the electron dynamics by a semi-relativistic Hamiltonian. In this paper, the gauge-dependence problem of exact perturbation theory is described, the proposed solutions are reviewed and it is concluded that none of them seems fully satisfactory. The problem is finally solved by using the fully relativistic absorption and scattering cross-sections given by quantum electrodynamics. Then, a new general Foldy-Wouthuysen transformation is presented. It is applied to the many-body case to obtain correct semi-relativistic transition operators. This transformation considerably simplifies the calculation of relativistic corrections. In the process, a new light-matter interaction term emerges, called the spin-position interaction, that contributes significantly to the magnetic x-ray circular dichroism of transition metals. We compare our result with the ones obtained by using several semi-relativistic time-dependent Hamiltonians. In the case of absorption, the final formula agrees with the result obtained from one of them. However, the correct scattering cross-section is not given by any of the semi-relativistic Hamiltonians.

**PACS.** 78.70.Ck X-ray scattering - 78.70.Dm X-ray absorption spectra - 11.15.Bt Perturbation theory, applied to gauge field theories - 31.30.jx Nonrelativistic limits of Dirac-Fock calculations

#### 1 Introduction

The x-ray absorption and scattering cross-section is often derived by considering the incident photon as a perturbing external electromagnetic potential [1,2]. Since only gauge-invariant observables are physically meaningful, it is legitimate to investigate whether this approach leads to gauge-invariant cross-sections. This question is even more crucial in a context in which experimental improvements urge the community to compute the x-ray cross-sections with an increasing accuracy.

The short answer is that the absorption and scattering cross-sections derived by semi-classical arguments are not gauge invariant. This is a (not so well-known) consequence of a widely studied conflict between time-dependent perturbation theory and gauge invariance (see section 2 for a non exhaustive list of references). Indeed,  $|\langle \phi_n | \psi(t) \rangle|^2$  gives the probability to find the system described by the state  $|\psi(t)\rangle$  in the eigenstate  $|\phi_n\rangle$  at time t, where  $|\psi(t)\rangle$  is a solution of the time-dependent Schrödinger equation for the Hamiltonian  $H(t) = H_0 + H_1(t)$ , while  $|\phi_n\rangle$  is an eigenstate of the time-independent Hamiltonian  $H_0$ . A time-dependent gauge transformation of H(t) will be

assigned to  $H_1(t)$  but not to  $H_0$ , which must remain independent of time. The transition probability is then modified because the state  $|\psi(t)\rangle$  is gauge-transformed and not the state  $|\phi_n\rangle$  (see section 2 for a more detailed argument). We meet there a serious difficulty: "Until this problem is understood, therefore, it seems that no calculation can be trusted at all." [3]

There is another conflict between time-dependent perturbation theory and semi-relativistic physics [4]. The semirelativistic approximation of  $|\phi_n\rangle$  is obtained by applying to it the time-independent Foldy-Wouthuysen transformation:  $|\phi_n^{\text{FW}}\rangle = U_{H_0}|\phi_n\rangle$ . The semi-relativistic approximation of  $|\psi(t)\rangle$  is derived from the time-dependent Foldy-Woutuysen transformation  $|\psi^{\text{FW}}(t)\rangle = U_{H(t)}|\psi(t)\rangle$ . Since  $U_{H(t)} \neq U_{H_0}$ , the transformed transition probability  $|\langle \phi_n^{\text{FW}} | \psi^{\text{FW}}(t) \rangle|^2$  is not equal to  $|\langle \phi_n | \psi(t) \rangle|^2$ , even if the Foldy-Wouthuysen transformations  $U_{H_0}$  and  $U_{H(t)}$  are known to all orders.

In this paper we discuss and solve these two conflicts. In a nutshell, the gauge problem is solved by deriving relativistic absorption and scattering cross-sections from quantum electrodynamics instead of the usual semi-classical argument where the incident light wave is described by a time-dependent potential. The semi-relativistic problem is solved by applying a many-body Foldy-Wouthuysen transformation to the relativistic cross-sections instead of describing the dynamics of the system with a semirelativistic Hamiltonian. The final result is a semi-relativistic absorption and scattering cross-section involving a new term that couples the spin and the position operators. In a companion paper, we show that this new term contributes significantly to the x-ray absorption of magnetic materials [5].

We now describe the outline of this paper. Section 2 discusses the gauge transformation of transition probabilities and reviews the solutions to the gauge-dependence problem proposed in the literature. Since none of them was widely accepted, we turn to the quantum electrodynamics framework in section 3, where we derive the relativistic electric dipole, quadrupole and magnetic dipole relativistic transition operators. In section 4, we describe a many-body Foldy-Wouthuysen transformation that we apply to the transition operators. They are used to obtain semi-relativistic absorption and scattering cross-sections in sections 5 and 6, where a new spin-position term is derived. In section 7, the conflict between time-dependent perturbation theory and semi-relativistic methods is described and illustrated by presenting the explicit calculations using four different semi-relativistic Hamiltonians commonly used in the literature. The conclusion presents possible extensions of the present work.

#### 2 Gauge invariance

The gauge invariance of the absorption and scattering cross-sections of light is a long-standing problem. It started in 1952 when Willis Lamb calculated the spectrum of Hydrogen in two gauges and obtained different results [6]. This gave rise to a long series of papers up to this day [7–11,3,12–25,4,26–55]. In 1987, the same Lamb (then Nobel prize winner) still considered this as "one of the outstanding problems of modern quantum optics." [39]

We quickly describe the meaning of gauge invariance and then consider its failure in semi-classical perturbation theory.

#### 2.1 The principle of gauge invariance

The two homogeneous Maxwell equations  $\nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0$ and  $\nabla \cdot \mathbf{B} = 0$ , where the dot denotes time derivative, imply the local existence of a vector potential  $\mathbf{A}$  and a scalar potential  $\Phi$  such that  $\mathbf{B} = \nabla \times \mathbf{A}$  and  $\mathbf{E} = -\nabla \Phi - \dot{\mathbf{A}}$ . We denote  $A = (\Phi, \mathbf{A})$ . The same  $\mathbf{E}$  and  $\mathbf{B}$  are obtained from the potentials  $A' = (\Phi - \dot{A}, \mathbf{A} + \nabla A)$ , that we also denote  $A' = A - \partial A$ , where  $\Lambda$  is any smooth function of space and time. In classical electromagnetism, gauge invariance means that the physics described by A and A'is the same. In quantum mechanics, consider a non-relativistic Hamiltonian

$$H_A = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\Phi,$$

or a relativistic (Dirac) Hamiltonian

$$H_A = c\boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{A}) + mc^2\beta + e\Phi,$$

where  $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$  and  $\beta$  are the Dirac matrices. Both Hamiltonians are of the form  $H_A = f(\mathbf{p} - e\mathbf{A}) + e\Phi$ , where f is some function. For such Hamiltonians it can be checked that  $M_A^{\dagger}(i\hbar\partial_t - H_{A'})M_A = i\hbar\partial_t - H_A$ , where  $M_A = e^{ieA/\hbar}$ . As a consequence, if  $\psi$  is a solution of the time-dependent Schrödinger equation  $i\hbar\dot{\psi} = H_A\psi$ , then  $\psi' = M_A\psi$  is a solution of  $i\hbar\dot{\psi}' = H_{A'}\psi'$ .

In quantum mechanics, a gauge transformation consists in both a change of the potentials and a change in the phase of the wavefunctions. An observable  $O_A$  depending on the electromagnetic potential A is said to be gauge invariant if  $M_A^{\dagger}O_{A'}M_A = O_A$  for every function  $A(t, \mathbf{r})$ . An observable must be gauge invariant to be considered a true physical quantity [56].

The principle of gauge invariance has become a cornerstone of particle physics. Since general relativity may also be considered as a gauge theory [57], it may be safely said that gauge invariance was the guiding principle of most of the fundamental physics of the twentieth century. Therefore, we need to check that the cross-section formulas are gauge invariant to ensure their true physical nature.

Note that the time-dependent Dirac or Schrödinger equations are always gauge invariant but the time-independent ones are not because  $H_A$  is not gauge invariant due to the the scalar potential  $\Phi$ . Indeed, under a gauge transformation  $\Phi$  becomes  $\Phi - \dot{A}$  and the term  $\dot{A}$  cannot be compensated for in the absence of a time derivative.

#### 2.2 Gauge dependence of transition probabilities

In time-dependent perturbation theory, a system is assumed to be in the ground state  $|\phi_g\rangle$  of a time-independent Hamiltonian  $H_{a_0}$ . Then, at time  $t_0$ , an electromagnetic wave represented by the time-dependent potential a is added to the system (with total potential  $A = a_0 + a$ ), which is represented at time t by the state  $|\psi(t)\rangle$ . A good way to take both the initial state and the dynamics into account is to use the evolution operator  $U_A(t, t_0)$ , which is the solution of  $i\hbar\partial_t U_A(t, t_0) = H_A(t)U_A(t, t_0)$  with the boundary condition  $U_A(t_0, t_0) = 1$ . Thus,  $|\psi(t)\rangle = U_A(t, t_0)|\phi_g\rangle$ . The probability of a transition to the eigenstate  $|\phi_n\rangle$  of  $H_{a_0}$  at time t is

$$P_{ng}(t) = |\langle \phi_n | \psi(t) \rangle|^2 = |\langle \phi_n | U_A(t, t_0) | \phi_g \rangle|^2.$$
(1)

Since we want to ensure that the breakdown of gauge invariance is not due to an approximation, we work with exact (i.e. to all orders) perturbation theory. If we carry out a gauge transformation of the perturbation from a to  $a' = a - \partial A$ , then the evolution operator becomes [35]

$$U_{A'}(t,t_0) = M_A(t) U_A(t,t_0) M_A^{\dagger}(t_0), \qquad (2)$$

where  $A' = a_0 + a'$ .

Therefore  $\langle \phi_n | \psi'(t) \rangle = \langle \phi_n | U_{A'}(t, t_0) | \phi_g \rangle$  is

$$\int d\mathbf{r} d\mathbf{r}' \phi_n^*(\mathbf{r}) e^{ie\Lambda(\mathbf{r},t)} U_A(\mathbf{r},t;\mathbf{r}',t_0) e^{-ie\Lambda(\mathbf{r}',t_0)} \phi_g(\mathbf{r}')$$

which is generally different from  $\langle \phi_n | \psi(t) \rangle$  since  $\Lambda(\mathbf{r}, t)$  is an arbitrary function (take for example  $\Lambda(\mathbf{r}, t) = \mathbf{r} \cdot \mathbf{k}(t)$ , where  $\mathbf{k}(t)$  is an arbitrary function of time). Therefore, in general,  $P'_{ng}(t) \neq P_{ng}(t)$  and the transition probabilities calculated in the two gauges are different. Moreover, since the transition rate entering cross-sections is the derivative of the transition probability with respect to time [58], the arbitrariness of the transition rate is increased by the fact that an arbitrary function  $\Lambda(\mathbf{r}, t)$  enters the integrand. Indeed, several papers evaluate the discrepancy between the probability calculated with two different gauges, and they find that it is generally not small [6,41,53]. By properly choosing  $\Lambda$ , the discrepancy can even be made arbitrary large [51].

The absence of gauge invariance is due to the fact that the operator is transformed but not the eigenstates of  $a_0$ . This is called a *hybrid transformation* in the literature [15].

#### 2.3 Proposed solutions

The lack of gauge invariance of transition probabilities is an alarming problem to which several solutions have been proposed. Since no clear consensus appears to have emerged [48], we present a critical review of these solutions.

The first one, called the consistent procedure, was proposed by Forney and coll. and Epstein [11, 13]. It is based on the observation that, if instead of gauge-transforming a we transform the potential  $a_0$  of the initial Hamiltonian to get  $H_{a'_0}$ , where  $a'_0 = a_0 - \partial A$ , then the evolution operator becomes again  $U_{A'}(t, t_0)$  (because  $a'_0 + a = a_0 + a - \partial A = A'$ ) but the eigenstates  $|\phi_g\rangle$  and  $|\phi_n\rangle$  are also transformed into time-dependent states  $|\phi'_g\rangle = M_A(t_0)|\phi_g\rangle$  and  $|\phi'_n\rangle = M_A(t)|\phi_n\rangle$ . Therefore, the transition probability is now conserved. In other words, gauge invariance is lost if we subtract  $\partial A$  from the perturbation but not if we subtract it from the unperturbed Hamiltonian  $H_0 = H_{a_0}$ .

Since the reference states  $|\phi_n\rangle$  become time-dependent, we leave the standard framework of time-dependent perturbation theory where the initial Hamiltonian  $H_0$  does not depend on time. Moreover, it is not physically clear why the gauge transformation should be applied to  $H_0$ and not to the perturbation.

In spite of these difficulties, many authors proposed to use the consistent procedure. However, as noticed by Yang [4], this does not really solve the problem because, if we start the calculation with the initial potential  $H_{a_0}$  and the perturbation a', the transition probability is  $P'_{ng}(t)$ . If we then use the consistent procedure to come back to the perturbation a, then we still find  $P'_{ng}(t)$  and we do not recover the result  $P_{ng}(t)$ . In other words, the transition probability is now gauge invariant (in the sense that a change of gauge does not modify the result) but it is gauge-dependent (in the sense that the result depends on the gauge we use in the perturbation to start the calculation). This gauge dependence would be a serious problem because we would have to select the "true" physical gauge for the perturbation.

A second solution appeared in a series of papers starting in 1976 [9,17,18,25,4,26,23,31,35,37,42], where Yang and collaborators proposed to define a gauge invariant transition probability. His idea is to start from the gaugeinvariant (but time-dependent) initial Hamiltonian

$$H_0(t) = \frac{(\mathbf{p} - e\mathbf{a}_0 - e\mathbf{a}(t))^2}{2m} + \mathbf{e}V, \qquad (3)$$

where V describes the electron-electron and electron-nuclear interactions so that  $H = H_0 + e\phi$ : the perturbation is only the scalar potential  $\phi$ . Then, the Hamiltonian  $H_0(t)$  is diagonalized at every time t:  $H_0(t)|\phi_n(t)\rangle = E_n(t)|\phi_n(t)\rangle$ and the transitions are calculated between the time-dependent states  $|\phi_n(t)\rangle$ . The corresponding transition probabilities are indeed gauge invariant. This solution has been used up to this day [40, 52, 55], although it was also strongly criticized [12, 19, 20, 27–29, 32–34, 59, 60]. The main arguments against Yang's interpretation are: (i) the quantity  $E_n(t)$  is not physical because you cannot measure an energy at a given time with arbitrary precision; (ii) the timedependent states  $|\phi_n(t)\rangle$  can be neither prepared nor detected; (iii) the term V in Eq. (3) should be removed from  $H_0(t)$  because it is a scalar potential and, as such, not gauge invariant. But if V is removed, then  $H_0(t)$  is so far from the true Hamiltonian that perturbation theory is no longer valid.

Following Goldman [61], Feuchtwang, Kazes and coll. proposed the following alternative solution [27, 28, 62, 33, 34]. They started from the well-known fact that the equations of motion of a Lagrangian are not modified by the addition of the total time derivative of a function [56]. Thus, two Lagrangians that differ by a total time derivative are equivalent [63]. Then, they remark that the addition of a total time derivative  $e\dot{A}$  to the Lagrangian induces a gauge transformation  $A \to A - \partial A$  of the Hamiltonian [62, 56, 49]. Finally, they use such a total derivative to compensate for the electric potential that is the cause of the gauge variance of the Hamiltonian. However, it is difficult to distinguish this procedure from picking up a specific gauge, namely the Weyl or temporal gauge where the scalar potential vanishes. We can conclude this short review by stating that no solution was found fully satisfactory.

To determine when gauge invariance can be achieved at the first order of perturbation theory, we consider a Dirac Hamiltonian in two gauges A and  $A' = A - \partial \Lambda$  and we calculate the difference

$$\langle \psi | H_A - H_{A'} | \psi' \rangle = e \langle \psi | c \boldsymbol{\alpha} \cdot \nabla \Lambda + \dot{\Lambda} | \psi' \rangle.$$

The advantage of the Dirac Hamiltonian is that the difference  $H_A - H_{A'}$  does not depend on A, but a similar calculation can be carried out in the non-relativistic case [60]. Then, we notice that  $c \boldsymbol{\alpha} \cdot \nabla \boldsymbol{\Lambda} = (i/\hbar)[H_D, \boldsymbol{\Lambda}]$  for any Dirac Hamiltonian  $H_D$ . Thus, if  $|\psi\rangle$  and  $|\psi'\rangle$  are eigenstates of  $H_D$  with energy E and E', we obtain

$$\langle \psi | H_A - H_{A'} | \psi' \rangle = e \langle \psi | \dot{\Lambda} | \psi' \rangle + ie \frac{E - E'}{\hbar} \langle \psi | \Lambda | \psi' \rangle.$$
(4)

If we consider the absorption cross-section of a photon of energy  $\hbar\omega$ , then energy conservation implies that  $E' = E + \hbar\omega$ . Thus, if  $\Lambda$  satisfies  $\dot{\Lambda} = -i\omega\Lambda$ , then  $\langle \phi | H_A - H_{A'} | \phi' \rangle = 0$  [60,42]. In other words, by restricting the gauge transformations to those satisfying  $\dot{\Lambda} = -i\omega\Lambda$ , the absorption cross-section, calculated up to first order in perturbation theory, is gauge invariant. However, in the resonant scattering cross-section, energy conservation does not apply to the transition involving intermediate states, and the cross-section is not gauge invariant even for those gauges [10,42,53].

Equation (4) shows that the matrix elements are also gauge invariant for a time-independent gauge transformation and energy conserving processes (i.e. E' = E). However, the gauge invariance principle is not supposed to restrict to gauges satisfying specific constraints such as  $\dot{A} = -i\omega \Lambda$  or  $\dot{A} = 0$ .

This rapid overview shows that, in the published semiclassical approaches where the photon is represented by an external potential, the transition probabilities are not gauge-invariant and no proposed solution has reached general acceptance. Therefore, we turn now to a framework where both electrons and photons are quantized: quantum electrodynamics (QED).

#### 2.4 Quantum electrodynamics

In QED the incident light is no longer described by an external electromagnetic field but by a photon, i.e. a state in a bosonic Fock space. Therefore, a scattering experiment is now described by the transition from an initial state involving both the electronic system in its ground state and the incident photon, to a final state involving both the electronic system in its (possibly) excited state and the scattered photon. Thus, the energy of the initial and final states is the same and, in the Schrödinger picture, the gauge transformation is expressed in terms of timeindependent operators instead of a time-dependent function  $\Lambda$  [64]. Equation (4) suggests that transition probabilities, which are now described through the so-called S-matrix, could be gauge invariant.

This is indeed the case, although a review of the literature on the gauge invariance of QED might look ambiguous because the kind of gauge transformation considered in different works can vary. In standard textbooks, "the S-matrix is gauge invariant by construction" [65] because only the so-called  $\xi$ -term is modified. In the most general gauge transformation, the space of states change from one gauge to the other [66]. For example, in the Coulomb gauge, only the transverse degrees of freedom are quantized and the photon states form a Hilbert space built by acting on the vacuum with creation operators of left and right polarized photons, while in the Lorenz gauge four degrees of freedom are quantized and the states (built by acting on the vacuum with creation operators of the left, right, longitudinal and scalar photons) can have a negative norm. In the Lorenz gauge, the Lorenz condition cannot be satisfied as an operator equation [67], it becomes a subsidiary condition used to determine a subspace of physical states with positive norm.

In other words, the state spaces of the Coulomb and Lorenz gauges have a quite different nature and the relation between them is delicate. Haller managed to show that the usual gauges are equivalent by devising a common framework containing all of them [43]. Note also that the gauge-invariance can only be expected for the renormalized S-matrix [68–70].

The gauge invariance under a general *infinitesimal* gauge transformation is well established within the Becchi-Rouet-Stora-Tyutin (BRST) approach: matrix elements of gauge-invariant operators between physical states are independent of the choice of the gauge-fixing functional if and only if the physical states  $|\alpha\rangle$  satisfy  $Q|\alpha\rangle = 0$ , where Q is the BRST charge [71,72]. The case of finite BRST transformations is in progress [73,74].

To summarize the discussion, the gauge invariance of the renormalized S-matrix is established for infinitesimal gauge transformations and for a reasonably large classe of gauges [15,75,76,70,43,77–80]. In other words, it is proved at the physicist level of rigour.

The most studied gauges are the Lorenz and Coulomb gauges. Renormalization is perfectly established for the Lorenz gauge, but in most practical calculations the subsidiary condition (Gauss' law) is not enforced [81]. Although it was proved that the S-matrix elements are often the same with and without the subsidiary condition [82, 83,15,12,28], this fails when the Hamiltonian is suddenly changed [84], as in the sudden creation of a core hole in photoemission or x-ray absorption [85,86]. In that case, Gauss' law has to be imposed in the Lorentz gauge and the Coulomb gauge result is recovered [84].

We choose to use quantum electrodynamics in the Coulomb gauge because it is the most accurate gauge for low-energy many-body calculations [56,87].

#### 3 Relativistic matrix elements

Since we now have a gauge-invariant framework, we can calculate the relativistic matrix elements that will be used in x-ray scattering and absorption cross-sections.

#### 3.1 The Hamiltonian

The quantum field Hamiltonian describing the interaction of light with matter in the Coulomb gauge is [15,88,56, 89]:

$$H = H_e + H_\gamma + H_{e\gamma},$$

where

$$\begin{split} H_e &= \int d\mathbf{r} \psi^{\dagger}(\mathbf{r}) \big( c \boldsymbol{\alpha} \cdot (-i\hbar \nabla - e\mathbf{a}) + \beta m c^2 + e\phi \big) \psi(\mathbf{r}) \\ &+ \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{8\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}, \end{split}$$

where  $\phi$  is a time-independent scalar external potential (for instance the nuclear potential), **a** is a time-independent vector potential (describing an external magnetic field) and  $\psi$  are fermion field operators. Normal ordering is implicit in  $H_e$ . It is the QED form of the Dirac Hamiltonian in the Coulomb gauge. The many-body version of this Hamiltonian is

$$H_N = \sum_{n=1}^N c \boldsymbol{\alpha}_n \cdot (-i\hbar \nabla_n - e\mathbf{a}(\mathbf{r}_n)) + \beta_n m c^2 + e\phi(\mathbf{r}_n) + \sum_{m \neq n} \frac{e^2}{8\pi\epsilon_0} \frac{1}{|\mathbf{r}_m - \mathbf{r}_n|},$$

where  $\alpha_n$  and  $\beta_n$  act on the *n*th Dirac electron. It can be given a well-defined mathematical meaning if the electronic system is described with respect to the Dirac sea [90], although the physical validity of the Dirac sea is sometimes disputed [91].

The photon Hamiltonian is

$$H_{\gamma} = \frac{\epsilon_0}{2} \int d\mathbf{r} |\mathbf{E}^{\perp}|^2 + c^2 |\mathbf{B}|^2 = \sum_{\mathbf{k},l} \hbar \omega_{\mathbf{k},l} a_{\mathbf{k},l}^{\dagger} a_{\mathbf{k},l},$$

where l stands for the polarization of a mode (there are two independent directions for a given wavevector  $\mathbf{k}$ ) and

$$H_{e\gamma} = -ec \int d\mathbf{r} \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \cdot \mathbf{A}(\mathbf{r}) \psi(\mathbf{r}),$$

describes the photon-matter interaction in the Coulomb gauge. According to Bialynicki-Birula, the Hamiltonian H also describes the dynamics of gauge-invariant states in any gauge [89]. The many-body version of this interaction Hamiltonian is

$$H_I = -ec \sum_{n=1}^N \boldsymbol{\alpha}_n \cdot \mathbf{A}(\mathbf{r}_n).$$

#### 3.2 S-matrix elements

Since we saw that the S-matrix is gauge invariant, we calculate its matrix-elements. We recall that

$$S = \lim_{\epsilon \to 0} T(e^{-\frac{i}{\hbar} \int_{-\infty}^{\infty} H_{\epsilon}(t)dt}),$$
 (5)

where  $H_{\epsilon}(t) = e^{-\epsilon|t|} e^{iH_0 t} H_{e\gamma} e^{-iH_0 t}$ . The adiabatic switching factor  $e^{-\epsilon|t|}$  enables us to describe physical processes

as matrix elements of S between eigenstates of  $H_0 = H_e + H_{\gamma}$ . The limit  $\epsilon \to 0$  can be shown to exist up to technical assumptions [92]. Note that  $H_0$  is not quadratic because of the Coulomb interaction term in  $H_e$ . The eigenstates of  $H_e$  are correlated multi-electronic wavefunctions. As a consequence, we are not in the textbook framework, the time-dependence of  $H_{\epsilon}(t)$  cannot be calculated explicitly and the Feynman diagram technique is no longer available to describe electrons. We can bypass this problem with the so-called "non-covariant" approach [93], using matrix elements of  $H_{\epsilon}(t)$  between eigenstates of  $H_0$ . Then, cross-sections are expressed in terms of the S-matrix and T-matrix elements related by:

$$\langle m|S|n\rangle = \delta_{mn} - 2i\pi\delta(e_m - e_n)\langle m|T|n\rangle.$$

Up to second order,

$$\langle m|T|n\rangle = \langle m|H_{e\gamma}|n\rangle + \sum_{p} \frac{\langle m|H_{e\gamma}|p\rangle\langle p|H_{e\gamma}|n\rangle}{e_{p} - e_{n} + i\gamma}, \quad (6)$$

where  $|m\rangle$ ,  $|p\rangle$  and  $|n\rangle$  are eigenstates of  $H_0$  with energy  $e_m$ ,  $e_p$  and  $e_n$ , respectively. The term  $i\gamma$  was added as a heuristic way to avoid divergence at resonance (i.e. when the states  $|n\rangle$  and  $|p\rangle$  are degenerate). More sophisticated methods exist to deal with such degeneracies [94] but they would bring us too far. From the physical point of view,  $\gamma$  describes the life-time of the state  $|p\rangle$ , which can decay by radiative or non-radiative relaxation. The sign of the damping term  $\gamma$  has been the object of some controversy [95–98].

Let us stress again that, since  $H_e$  is not quadratic, we essentially work in the Schrödinger picture, where the operators are independent of time, instead of the standard interaction picture which is used in most textbooks. Both approaches are equivalent [99]. A modern version of the Schrödinger picture of QFT is given by Hatfield [100].

Our purpose is now to calculate the matrix elements  $\langle m|H_{e\gamma}|n\rangle$ , where  $H_{e\gamma}$  is independent of time. The second quantized expression for the photon field in the Schrödinger picture is [101]:

$$\mathbf{A}(\mathbf{r}) = \sum_{\mathbf{k},l} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega_{\mathbf{k}}}} \left( \boldsymbol{\epsilon}_{\mathbf{k},l} a_{\mathbf{k},l} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} + \boldsymbol{\epsilon}_{\mathbf{k},l}^{\star} a_{\mathbf{k},l}^{\dagger} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{r}} \right).$$

Note that we do not assume the polarization vectors  $\boldsymbol{\epsilon}_{\mathbf{k},l}$  to be real.

We denote  $|n\rangle = a_{\mathbf{k},l}^{\dagger}|0\rangle|\Psi_n\rangle$  an eigenstate of  $H_0$  where one photon is present in mode  $\mathbf{k}, l$  and the electrons are in state  $|\Psi_n\rangle$  with energy  $E_n$ . The energy of  $|n\rangle$  is  $e_n = \hbar\omega_{\mathbf{k},l} + E_n$ . The interaction Hamiltonian  $H_{e\gamma}$  is linear in  $\mathbf{A}$ which is linear in photon creation and annihilation operators so that only one-photon transitions are possible. The state  $|n\rangle$  can make transitions towards  $|a\rangle = |0\rangle|\Psi_m\rangle$  by absorption and  $|e\rangle = a_{\mathbf{k},l}^{\dagger}a_{\mathbf{k}',l'}^{\dagger}|0\rangle|\Psi_m\rangle$  by emission. From now on, we denote  $\omega = \omega_{\mathbf{k},l}$ ,  $\epsilon = \epsilon_{\mathbf{k},l}$ ,  $\omega' = \omega_{\mathbf{k}',l'}$  and  $\epsilon' = \epsilon_{\mathbf{k}',l'}$ . The corresponding matrix elements are:

$$\langle a|H_{e\gamma}|n\rangle = -ec\sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\boldsymbol{\epsilon} \cdot \langle \boldsymbol{\Psi}_m| \int \psi^{\dagger} \boldsymbol{\alpha} \psi \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} |\boldsymbol{\Psi}_n\rangle,$$

and

$$\langle e|H_{e\gamma}|n\rangle = -ec\sqrt{\frac{\hbar}{2\epsilon_0 V\omega'}} \epsilon'^{\star} \cdot \langle \Psi_m| \int \psi^{\dagger} \boldsymbol{\alpha} \psi \mathrm{e}^{-i\mathbf{k}'\cdot\mathbf{r}} |\Psi_n\rangle$$

where

$$\int \psi^{\dagger} \boldsymbol{\alpha} \psi \mathrm{e}^{\pm i \mathbf{k} \cdot \mathbf{r}} = \int \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) \mathrm{e}^{\pm i \mathbf{k} \cdot \mathbf{r}} d\mathbf{r}.$$

#### 3.3 Electric dipole and multipole transitions

Johnson and coll. calculated relativistic multipole transitions of arbitrary order for two-electron atoms [102]. Here we stop at the first order in  $\mathbf{k}$  but we do not assume spherical symmetry of the Hamiltonian.

To carry out a multipole expansion of the previous matrix elements, we shall continue working with quantum fields instead of the usual many-body expressions. In that framework, the expressions are simpler because there is no electron index and we can use the following well-known trick [103,67].

Let  $F = \int \psi^{\dagger}(\mathbf{r}) f(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r}$ , where f is some function of  $\mathbf{r}$ . To calculate the commutator of F with some Hamiltonian  $H_0$ , we go to the interaction picture and define  $F_I(t) = e^{iH_0t/\hbar}Fe^{-iH_0t/\hbar}$ . Then, the time-derivative  $\dot{F}_I$  of  $F_I$  is given by  $-i\hbar\dot{F}_I(t) = [H_0, F_I(t)]$ . Now, we notice that F is related to the density operator  $\rho(\mathbf{r}) = \psi^{\dagger}(\mathbf{r})\psi(\mathbf{r})$  by  $F = \int \rho(\mathbf{r})f(\mathbf{r})d\mathbf{r}$ . Thus,  $-i\hbar\dot{F}_I(t) = -i\hbar\int\dot{\rho}(\mathbf{r},t)f(\mathbf{r})d\mathbf{r} =$  $[H_0, F_I(t)]$ . If  $H_0$  conserves the electric charge, the continuity equation  $e\dot{\rho}(\mathbf{r}) = -\nabla \cdot \mathbf{j}$  holds, where  $\mathbf{j}$  is the electric current operator. By taking t = 0 to recover the operators in the Schrödinger picture, we obtain

$$[H_0, F] = \frac{i\hbar}{e} \int \nabla \cdot \mathbf{j}(\mathbf{r}) f(\mathbf{r}) d\mathbf{r} = -\frac{i\hbar}{e} \int \mathbf{j}(\mathbf{r}) \cdot \nabla f(\mathbf{r}) d\mathbf{r}$$
$$= -i\hbar c \int \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) \cdot \nabla f(\mathbf{r}) d\mathbf{r}.$$
(7)

To find the electric dipole transition term we apply Eq. (7) with  $f(\mathbf{r}) = \boldsymbol{\epsilon} \cdot \mathbf{r}$  and  $H_0 = H_e$ :

$$[H_e, \int \psi^{\dagger}(\mathbf{r}) \boldsymbol{\epsilon} \cdot \mathbf{r} \psi(\mathbf{r}) d\mathbf{r}] = -i\hbar c \int \psi^{\dagger}(\mathbf{r}) \boldsymbol{\alpha} \psi(\mathbf{r}) \cdot \boldsymbol{\epsilon} d\mathbf{r},$$

and we obtain in the dipole approximation  $e^{i\mathbf{k}\cdot\mathbf{r}} \simeq 1$ 

$$\langle a|H_{e\gamma}|n\rangle = \frac{e(E_m - E_n)}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V\omega}} \langle \Psi_m| \int \psi^{\dagger} \boldsymbol{\epsilon} \cdot \mathbf{r} \psi |\Psi_n\rangle.$$

To deal with electric quadrupole and magnetic dipole transitions, we expand to the first order:  $e^{i\mathbf{k}\cdot\mathbf{r}} \simeq 1 + i\mathbf{k}\cdot\mathbf{r}$ . We apply Eq. (7) with  $f(\mathbf{r}) = \boldsymbol{\epsilon} \cdot \mathbf{r}\mathbf{k} \cdot \mathbf{r}$  and  $H_0 = H_e$ :

$$[H_e, \psi^{\dagger} \boldsymbol{\epsilon} \cdot \mathbf{r} \mathbf{k} \cdot \mathbf{r} \psi] = -i\hbar c \psi^{\dagger} \boldsymbol{\alpha} \psi \cdot (\boldsymbol{\epsilon} \mathbf{k} \cdot \mathbf{r} + \mathbf{k} \boldsymbol{\epsilon} \cdot \mathbf{r}),$$

where we removed the integral sign for notational convenience. Thus,

$$\psi^{\dagger}\boldsymbol{\epsilon}\cdot\boldsymbol{\alpha}\mathbf{k}\cdot\mathbf{r}\psi=\frac{i}{\hbar c}[H_{e},\psi^{\dagger}\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r}\psi]-\psi^{\dagger}\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\boldsymbol{\alpha}\psi.$$

If we add  $\psi^{\dagger} \boldsymbol{\epsilon} \cdot \boldsymbol{\alpha} \mathbf{k} \cdot \mathbf{r} \psi$  to both terms we obtain

$$\begin{split} 2\psi^{\dagger}\boldsymbol{\epsilon}\cdot\boldsymbol{\alpha}\mathbf{k}\cdot\mathbf{r}\psi &= \frac{i}{\hbar c}[H_{e},\psi^{\dagger}\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r}\psi]\\ &-\psi^{\dagger}(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}\times\boldsymbol{\alpha})\psi \end{split}$$

Finally, up to electric quadrupole transitions

$$\langle a|H_{e\gamma}|n\rangle = \frac{e\Delta E}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V\omega}} \langle \Psi_m| \int \psi^{\dagger} T\psi |\Psi_n\rangle, \quad (8)$$

where  $\Delta E = E_m - E_n$  and

$$T = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \mathbf{k} \cdot \mathbf{r} - \frac{\hbar c}{2\Delta E} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \boldsymbol{\alpha}).$$
(9)

The first term of T is the usual electric-dipole operator, the second one is the electric-quadrupole operator and the third one will turn out to be the magnetic-dipole operator (see section 4.3). Similarly,

$$\langle e|H_{e\gamma}|n\rangle = \frac{e\Delta E}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega'}} \langle \Psi_m| \int \psi^{\dagger} T' \psi |\Psi_n\rangle, \quad (10)$$

where

$$T' = \boldsymbol{\epsilon}'^{\star} \cdot \mathbf{r} - \frac{i}{2} \boldsymbol{\epsilon}'^{\star} \cdot \mathbf{rk}' \cdot \mathbf{r} + \frac{\hbar c}{2\Delta E} (\boldsymbol{\epsilon}'^{\star} \times \mathbf{k}') \cdot (\mathbf{r} \times \boldsymbol{\alpha}).$$

## 4 Foldy-Wouthuysen transformation of cross-sections

In the previous sections, we have shown that gauge invariance is ensured by describing the interaction of light and matter with quantum electrodynamics, where photons are quantized and electrons are described by four-component Dirac spinor quantum fields.

Although fully relativistic spectroscopy packages do exist [104,105], in most solid-state calculations, we do not use Dirac spinors but two-component (Pauli) wavefunctions. Moreover, semi-relativistic expressions are often physically clearer. Therefore, we need to link the two representations by using a generalization of the Foldy-Wouthuysen transformation.

In this section, we first describe the Foldy-Wouthuysen transformation and its many-body extension. Then, we use this framework to calculate the relativistic corrections to the dipole and quadrupole transitions. The calculations are considerably simpler than the usual approach, where the relativistic corrections are derived from a semirelativistic Hamiltonian.

#### 4.1 The Foldy-Wouthuysen transformation

The idea of the Foldy-Wouthuysen transformation is the following. If  $H_D$  is a time-independent relativistic Hamiltonian, it has the form

$$H_D = H^0 + \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

6

where  $H^0 = mc^2\beta$  and each  $H_{ij}$  is a 2x2 matrix. We write  $H_D$  as the sum of even and odd parts  $H_D = H^0 + \mathcal{E} + \mathcal{O}$ , where

$$\mathcal{E} = \begin{pmatrix} H_{11} & 0\\ 0 & H_{22} \end{pmatrix}, \quad \mathcal{O} = \begin{pmatrix} 0 & H_{12}\\ H_{21} & 0 \end{pmatrix},$$

satisfy  $\beta \mathcal{E}\beta = \mathcal{E}$  and  $\beta \mathcal{O}\beta = -\mathcal{O}$ . Note that  $H^0$  is also even. If  $|\psi_D\rangle$  is a solution of the Dirac equation  $H_D |\psi_D\rangle = E |\psi_D\rangle$ , where  $H_D$  is the Dirac Hamiltonian, then the upper two components of  $|\psi_D\rangle$  are called the large components and the lower two the small components. The Dirac equation couples the large and small components of  $|\psi_D\rangle$  through the odd terms of  $H_D$ . Foldy and Wouthuysen [106] looked for a unitary operator U that decouples the large and small components of  $|\psi\rangle = U |\psi_D\rangle$ . In other words,  $H = U H_D U^{\dagger}$  has only even components:  $H = \beta H \beta$ . The method proposed by Foldy and Wouthuysen consists in successive transformations of the form  $U = e^{iS}$  [106, 107].

This transformation does not satisfy Eriksen's condition  $U = \beta U^{\dagger} \beta$  discussed in the Appendix. This is because the product  $U = e^{iS^{(2)}} e^{iS^{(1)}}$  does not satisfy this equation even if  $e^{iS^{(1)}}$  and  $e^{iS^{(2)}}$  do. Silenko recently derived the correction that must be applied to go from Foldy-Wouthuysen to Eriksen transformations [108], and he showed that the correction is at an order beyond the one we consider in this paper.

In the Appendix, we derive a very general Foldy-Wouthuysen transformation which only assumes the existence of a self-adjoint operator  $H_D$  and a self-adjoint involution  $\eta$  (i.e.  $\eta^{\dagger} = \eta$  and  $\eta^2 = 1$ ). We show that we can build from  $H_D$  and  $\eta$  a unitary matrix U such that  $\eta U H_D U^{\dagger} \eta = U H_D U^{\dagger}$  (which means that  $U H_D U^{\dagger}$  couples only even components to even components and odd components to odd components). Moreover, Eriksen's condition  $U = \eta U^{\dagger} \eta$  is satisfied.

This transformation is exact in the following sense. The positive energy Dirac wavefunctions  $|\Phi\rangle$  and  $|\Psi\rangle$  can be expressed in terms of the Foldy-Wouthuysen ones  $|\phi\rangle$  and  $|\psi\rangle$  by  $|\Phi\rangle = U^{\dagger}|\phi\rangle$  and  $|\Psi\rangle = U^{\dagger}|\psi\rangle$ , where  $|\phi\rangle$  and  $|\psi\rangle$  are even (i.e. they satisfy  $\eta|\phi\rangle = |\phi\rangle$  and  $\eta|\psi\rangle = |\psi\rangle$ ). For one-particle wavefunctions, this means that the small components of  $|\phi\rangle$  and  $|\psi\rangle$  is zero.

The relativistic x-ray absorption cross-section can then be written

$$\sigma = 4\pi\alpha_0 \hbar\omega \sum_{\phi} |\langle \phi | UTU^{\dagger} | \psi \rangle|^2 \delta(E_{\phi} - E_{\psi} - \hbar\omega), (11)$$

where  $\alpha_0$  is the fine-structure constant.

This relation is exact if U is calculated to all orders in  $1/c^2$ . In the next sections, we shall expand U in powers of  $1/c^2$ .

#### 4.2 Semi-relativistic dipole transitions

We consider the matrix elements  $D = \langle \Phi | \int \psi^{\dagger} \boldsymbol{\epsilon} \cdot \mathbf{r} \psi | \Psi \rangle$ , that we rewrite in terms of Foldy-Wouthuysen wavefunctions as in Eq. (11). Since U is written as a many-body operator, we translate the quantum field expression for D into the many-body formula  $D = \langle \Phi | \boldsymbol{\epsilon} \cdot \mathbf{R} | \Psi \rangle$ , where  $\mathbf{R} = \sum_{n=1}^{N} \mathbf{r}_n [109, 110]$ . We calculate  $D = \langle \phi | U \boldsymbol{\epsilon} \cdot \mathbf{R} U^{\dagger} | \psi \rangle$ , where  $U = e^{iS}$  by using the Baker-Campbell-Hausdorff formula

$$e^{iS}Te^{-iS} = T + i[S,T] + \sum_{n=2}^{\infty} i^n \frac{L^n(T)}{n!},$$

where L(T) = [S, T] and  $L^n(T) = L(L^{n-1}(T))$ . At the first order in  $1/c^2$  we have  $U = U_1 \otimes \cdots \otimes U_N$ , where  $U_i = e^{iS_i^{(1)}}$ , with  $S_i^{(1)} = \frac{-i}{2mc^2}\beta_i\mathcal{O}_i$ . Hence, we can calculate the action of U on each variable independently. Removing temporarily the constant  $-i/2mc^2$ , we take the one-body operator  $S = \beta \mathcal{O}$  and compute

$$\begin{split} L(\hat{\epsilon} \cdot \mathbf{r}) &= c[\beta \boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{a}_0), \hat{\epsilon} \cdot \mathbf{r}] = c \sum_{ij} \beta \alpha^i \epsilon^j [p_i, r_j] \\ &= -i\hbar c \sum_{ij} \beta \alpha^i \epsilon^j \delta_{ij} = -i\hbar c \beta \boldsymbol{\alpha} \cdot \hat{\epsilon}, \end{split}$$

and

$$\begin{split} L^{2}(\hat{\epsilon} \cdot \mathbf{r}) &= -i\hbar c^{2}[\beta \boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{a}_{0}), \beta \boldsymbol{\alpha} \cdot \hat{\epsilon}] \\ &= -i\hbar c^{2}\sum_{ij}(p_{i} - ea_{0i})\epsilon_{j}[\beta \alpha^{i}, \beta \alpha^{j}] \\ &= i\hbar c^{2}\sum_{ij}(p_{i} - ea_{0i})\epsilon_{j}[\alpha^{i}, \alpha^{j}], \end{split}$$

where we used  $\beta \alpha_i = -\alpha_i \beta$  and  $\beta^2 = 1$ . We compute

$$[\alpha^i, \alpha^j] = 2i \sum_k \epsilon_{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} = 2i \sum_k \epsilon_{ijk} \Sigma^k,$$

which defines  $\Sigma^k$  the components of  $\Sigma$ . Therefore,

$$L^{2}(\hat{\epsilon} \cdot \mathbf{r}) = -2\hbar c^{2}(\mathbf{p} - e\mathbf{a}_{0}) \cdot (\hat{\epsilon} \times \boldsymbol{\Sigma}).$$

So that, for each particle, and up to  $O(m^{-2})$ ,

$$U_{n}\boldsymbol{\epsilon}\cdot\mathbf{r}_{n}U_{n}^{\dagger} = \boldsymbol{\epsilon}\cdot\mathbf{r}_{n} - i\frac{\hbar}{2mc}\beta_{n}\boldsymbol{\alpha}_{n}\cdot\boldsymbol{\epsilon}$$
$$-\frac{\hbar}{4m^{2}c^{2}}\boldsymbol{\pi}_{n}\cdot(\boldsymbol{\epsilon}\times\boldsymbol{\Sigma}_{n})$$

The many-body version is obtained by summing the righthand side over n.

In the matrix elements  $D = \langle \phi | U \boldsymbol{\epsilon} \cdot \mathbf{R} U^{\dagger} | \psi \rangle$ , recall that  $|\psi\rangle = \eta |\psi\rangle$  and  $|\phi\rangle = \eta |\phi\rangle$  because  $|\Psi\rangle$  and  $|\Phi\rangle$  are positive energy states, as shown in the Appendix. Therefore,  $\langle \phi | U \boldsymbol{\epsilon} \cdot \mathbf{R} U^{\dagger} | \psi \rangle = \langle \phi | \eta U \boldsymbol{\epsilon} \cdot \mathbf{R} U^{\dagger} \eta | \psi \rangle$  and all the terms that are odd in  $U \boldsymbol{\epsilon} \cdot \mathbf{R} U^{\dagger}$  are eliminated by the matrix elements. This eliminates the term proportional to  $\beta_n \boldsymbol{\alpha}_n$  and we are left with

$$D = \sum_{n=1}^{N} \langle \phi | \boldsymbol{\epsilon} \cdot \mathbf{r}_n - \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}_n) | \psi \rangle.$$

#### 4.3 Semi-relativistic multipole transitions

From Eq. (9), we write the multipole transitions

$$M = \frac{i}{2}E_2 - \frac{\hbar c}{2\Delta E}M_1,$$

where

$$E_{2} = \sum_{n} \langle \phi | U \boldsymbol{\epsilon} \cdot \mathbf{r}_{n} \mathbf{k} \cdot \mathbf{r}_{n} U^{\dagger} | \psi \rangle,$$
  
$$M_{1} = \sum_{n} \langle \phi | U(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r}_{n} \times \boldsymbol{\alpha}_{n}) U^{\dagger} | \psi \rangle,$$

correspond to the electric quadrupole and magnetic dipole transitions, respectively. Since multipole transitions are smaller than dipole ones, it is enough to use the first two terms of the Baker-Campbell-Hausdorf formula.

The term  $[S_n, \boldsymbol{\epsilon} \cdot \mathbf{r}_n \mathbf{k}_n \cdot \mathbf{r}_n]$  is odd and disappears in the matrix element. Thus, at the order we consider,

$$E_2 = \sum_n \langle \phi | \boldsymbol{\epsilon} \cdot \mathbf{r}_n \mathbf{k} \cdot \mathbf{r}_n | \psi \rangle.$$

Let  $T_2 = (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \boldsymbol{\alpha})$ . We write

$$\begin{split} [\beta \mathcal{O}, T_2] &= c[\beta \boldsymbol{\alpha} \cdot \mathbf{p}, T_2] - ec[\beta \boldsymbol{\alpha} \cdot \mathbf{a}_0, T_2] \\ &= c\beta(\{\boldsymbol{\alpha} \cdot \mathbf{p}, T_2\} - e\{\boldsymbol{\alpha} \cdot \mathbf{a}_0, T_2\}). \end{split}$$

The anticommutators are

$$\{\boldsymbol{\alpha} \cdot \mathbf{p}, T_2\} = \sum_{ijkl} \epsilon_{jkl} (\boldsymbol{\epsilon} \times \mathbf{k})_j (\alpha_i \alpha_l p_i r_k + r_k p_i \alpha_l \alpha_i)$$
$$= 2(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma} + \mathbf{L}),$$

and

$$\{\boldsymbol{\alpha} \cdot \mathbf{a}_0, T_2\} = \sum_{ijkl} \epsilon_{ikl} (\boldsymbol{\epsilon} \times \mathbf{k})_i a_j r_k \{\alpha_j, \alpha_l\}$$
$$= 2(\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \mathbf{a}_0).$$

Note that  $\hbar \Sigma = g\mathbf{S}$  with g = 2 (because the spin operator is  $\mathbf{S} = \hbar \Sigma/2$ ). Thus, we recover the fact that the Dirac equation gives a gyromagnetic factor g = 2 to the electron. Moreover,  $\mathbf{L} + \hbar \Sigma = \mathbf{L} + 2\mathbf{S}$  is the total magnetic moment of the electron.

Finally, since  $\mathbf{r}_n \times \boldsymbol{\alpha}_n$  is odd,

$$M_1 = \sum_n \frac{\beta_n}{mc} \langle \phi | (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n) | \psi \rangle,$$

where  $\mathbf{\Lambda}_n = \mathbf{L}_n - e\mathbf{r}_n \times \mathbf{a}_0(\mathbf{r}_n)$  is the moment of the mechanical momentum as defined in Ref. [111]. The term  $M_1$  describes magnetic-dipole transitions. The multipole transitions are

$$M = \sum_{n} \langle \phi | \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r}_{n} \mathbf{k} \cdot \mathbf{r}_{n} \\ - \frac{\hbar \beta_{n}}{2m\Delta E} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma}_{n} + \boldsymbol{\Lambda}_{n}) | \psi \rangle.$$

#### 5 Absorption cross-section

The absorption cross section is calculated by assuming that initially the system of electrons is in state  $|I\rangle$  that can be transformed into Foldy-Wouthuysen eigenstate  $|i\rangle$ , with energy  $E_i$ , and that a photon  $\mathbf{k}, \boldsymbol{\epsilon}$  is present. In the final state there is no photon and the system is in state  $|F\rangle$  ( $|f\rangle$  after transformation).

The transition probability per unit time from state m to state n is related to the T-matrix elements by [112]:

$$w = \frac{2}{\hbar} \delta_{mn} \operatorname{Im} \langle m | T | m \rangle + \frac{2\pi}{\hbar} \delta(e_n - e_m) |\langle n | T | m \rangle|^2.$$
 (12)

and must be divided by c/V (rate at which the photon crosses a unit of surface) to obtain the cross section. Since we consider real transitions (i.e.  $m \neq n$ ), only the second term is present.

From (8) and using the result of transformation derived in the previous section:

$$\sigma = 4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle f | T_{\rm FW} | i \rangle|^2 \delta(E_f - E_i - \hbar \omega),$$

where  $T_{\rm FW}$  is:

$$T_{\rm FW} = \sum_{n} \boldsymbol{\epsilon} \cdot \mathbf{r}_{n} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r}_{n} \mathbf{k} \cdot \mathbf{r}_{n} - \frac{\hbar}{4m^{2}c^{2}} \boldsymbol{\pi}_{n} \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}_{n}) \\ - \frac{\beta_{n}}{2m\omega} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\hbar \boldsymbol{\Sigma}_{n} + \boldsymbol{\Lambda}_{n}),$$

with  $\alpha_0$  the fine structure constant and  $\Delta E = E_f - E_i = \hbar \omega$ .

It corresponds to the usual formula for the cross section [113] with two more terms: the third one and the last one.

The third term was already found by Christos Gougoussis in his PhD thesis [114], but his final result was not in agreement with ours because of his use of the commutation relation, as described in section 7.6. We rewrite it by using  $\boldsymbol{\pi} = (m/i\hbar)[\mathbf{r}, H_0^{\text{FW}}] + O(c^{-2})$ , where  $H_0^{\text{FW}}$  is the Foldy-Wouthuysen Hamiltonian, to get:

$$\begin{aligned} -\frac{\hbar}{4m^2c^2} \langle f | \boldsymbol{\pi} \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}) | i \rangle \\ &= \frac{i}{4mc^2} (E_i - E_f) \langle f | \mathbf{r} \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\Sigma}) | i \rangle \\ &= \frac{i\hbar\omega}{4mc^2} \langle f | (\boldsymbol{\epsilon} \times \mathbf{r}) \cdot \boldsymbol{\Sigma} | i \rangle. \end{aligned}$$

We call *spin-position operator* the operator  $(\boldsymbol{\epsilon} \times \mathbf{r}) \cdot \boldsymbol{\Sigma}$ . Its evaluation at the K-edge of materials will be presented in a companion paper [5].

The amplitude of the last term depends on the choice of the space origin in the Coulomb gauge for  $\mathbf{a}_0$ . It does not make the cross section gauge dependent because the states are changed accordingly when choosing the origin of the gauge. If the origin of the gauge is chosen at the atom position, fields larger than  $10^6$  T are required for this term to be significant. Such fields are way beyond laboratory accessible values.

#### 6 Scattering cross-section

The scattering cross section is calculated by assuming that initially the system of electrons is in state  $|I\rangle$  with a photon  $\mathbf{k}_i, \boldsymbol{\epsilon}_i$  and that in the final state the system is in state  $|F\rangle$  with a scattered photon  $\mathbf{k}_f, \boldsymbol{\epsilon}_f$ . We do not consider the special case when  $\mathbf{k}_i, \boldsymbol{\epsilon}_i = \mathbf{k}_f, \boldsymbol{\epsilon}_f$ .

Eqs. (8), (10) and (12) yield:

$$w = \frac{2\pi}{\hbar} \sum_{F} \delta(E_{f} + \hbar\omega_{f} - E_{i} - \hbar\omega_{i}) \Big| \sum_{L} \frac{e^{2}c^{2}\hbar}{2\epsilon_{0}V} \frac{1}{\sqrt{\omega_{i}\omega_{f}}} \\ \frac{\langle F|e_{-\mathbf{k}_{f}}\psi^{\dagger}\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_{f}^{\star}\psi|L\rangle\langle L|e_{\mathbf{k}_{i}}\psi^{\dagger}\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_{i}\psi|I\rangle}{E_{i} - E_{l} + \hbar\omega_{i} + i\gamma} \\ + \frac{\langle F|e_{\mathbf{k}_{i}}\psi^{\dagger}\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_{i}\psi|L\rangle\langle L|e_{-\mathbf{k}_{f}}\psi^{\dagger}\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_{f}^{\star}\psi|I\rangle}{E_{i} - E_{l} - \hbar\omega_{f}} \Big|^{2},$$

where  $\gamma > 0$  and

$$e_{\mathbf{k}}\psi^{\dagger}\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}\psi = \sum_{j=1}^{3}\int e^{i\mathbf{k}\cdot\mathbf{r}}\psi^{\dagger}(\mathbf{r})\alpha^{j}\psi(\mathbf{r})\epsilon^{j}d\mathbf{r}.$$

The scattering cross-section is related to w by [115]:

$$\frac{d^2\sigma}{d\Omega d\omega_f} = \frac{V^2}{(2\pi)^3} \omega_f^2 \frac{1}{\hbar c^4} w.$$

Since the electric charge is related to the classical electron radius  $r_e$  by  $e^2 = 4\pi\epsilon_0 r_e mc^2$ , we obtain the relativistic Kramers-Heisenberg scattering cross-section:

$$\begin{aligned} \frac{d^2\sigma}{d\Omega d\omega_f} &= (r_e m c^2)^2 \frac{\omega_f}{\omega_i} \sum_F \delta(E_f + \hbar \omega_f - E_i - \hbar \omega_i) \\ & \left| \sum_L \frac{\langle F | e_{-\mathbf{k}_f} \psi^{\dagger} \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_f^{\star} \psi | L \rangle \langle L | e_{\mathbf{k}_i} \psi^{\dagger} \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_i \psi | I \rangle}{E_i - E_l + \hbar \omega_i + i\gamma} \right. \\ & \left. + \frac{\langle F | e_{\mathbf{k}_i} \psi^{\dagger} \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_i \psi | L \rangle \langle L | e_{-\mathbf{k}_f} \psi^{\dagger} \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_f^{\star} \psi | I \rangle}{E_i - E_l - \hbar \omega_f} \right|^2. \end{aligned}$$

In this expression, the sum over  $|L\rangle$  involves a complete set of states, with positive and negative energies. Since  $E_i$  is usually the positive energy of the ground state including the electron rest energy, we have  $E_i = mc^2 + E'_i > 0$ , where  $E'_i$  is the usual (negative) ground state energy. If  $|L\rangle$  is a positive energy state, we have  $E_l = mc^2 + E'_l$  with  $E'_l > E'_i$  and the first term is resonant at  $\hbar\omega_i = E'_l - E'_i$ . If  $|L\rangle$  is a negative energy state, then  $E_l = -mc^2 - E'_l$  and  $E_i - E_l - \hbar\omega_f = 2mc^2 + E'_i - E'_l - \hbar\omega_f$  cannot be resonant in standard experimental conditions.

We show that the resonant scattering term has a semirelativistic expansion close to, but different from, the standard one [116]. If we are interested in the resonant part of the scattering cross section, then  $E_l > 0$  and

$$\frac{d^2\sigma}{d\Omega d\omega_f} = \left(\frac{r_e m}{\hbar^2}\right)^2 \frac{\omega_f}{\omega_i} \sum_f \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i) \\ \left| \sum_{L^>} (E_l - E_i)(E_f - E_l) \right| \\ \frac{\langle f | T_{\rm FW}^{\prime fl}(\boldsymbol{\epsilon}_f) | l \rangle \langle l | T_{\rm FW}^{li}(\boldsymbol{\epsilon}_i) | i \rangle}{E_i - E_l + \hbar\omega_i + i\gamma} \right|^2.$$

with

$$\begin{split} T_{\rm FW}^{ij}(\boldsymbol{\epsilon}_i) &= \sum_n \boldsymbol{\epsilon}_i \cdot \mathbf{r}_n + \frac{i}{2} \boldsymbol{\epsilon}_i \cdot \mathbf{r}_n \mathbf{k}_i \cdot \mathbf{r}_n \\ &- \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}_i \times \boldsymbol{\Sigma}_n) \\ &- \frac{\hbar \beta_n}{2m \Delta E^{ij}} (\boldsymbol{\epsilon}_i \times \mathbf{k}_i) \cdot (\hbar \boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n), \end{split}$$

and

$$\begin{split} T_{\rm FW}^{\prime ij}(\boldsymbol{\epsilon}_f) &= \sum_n \boldsymbol{\epsilon}_f^{\star} \cdot \mathbf{r}_n - \frac{\imath}{2} \boldsymbol{\epsilon}_f^{\star} \cdot \mathbf{r}_n \mathbf{k}_f \cdot \mathbf{r}_n \\ &- \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}_f^{\star} \times \boldsymbol{\Sigma}_n) \\ &+ \frac{\hbar \beta_n}{2m \Delta E^{ij}} (\boldsymbol{\epsilon}_f^{\star} \times \mathbf{k}_f) \cdot (\hbar \boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n) \end{split}$$

where  $\Delta E^{ij} = E_i - E_j$ .

As in the absorption case, the spin-position term in the transition operator is not present in the usual formula [116].

#### 7 Other methods

In this section, we compare our semi-relativistic transition matrix elements with the ones obtained by using timedependent perturbation theory where the time-evolution is described by several time-dependent semi-relativistic Hamiltonians: the one proposed by Blume, the "gaugeinvariant" Foldy-Wouthuysen one, the textbook Foldy-Wouthuysen one and the effective Hamiltonian derived in non-relativistic QED (NRQED). Before making this comparison, we first explain why using a time-dependent semirelativistic Hamiltonian in a perturbation calculation can lead to incorrect results.

#### 7.1 Foldy-Wouthuysen subtelties

In this section, we assume that the exact time-dependent Foldy-Wouthuysen operator U is known. Thus, the following difficulties are not related to the use of an approximation, but to the interplay of the Foldy-Wouthuysen method with perturbation theory.

The first subtelty was noticed by Nieto [117,61]: If  $|\Psi\rangle$  is a solution of the time-dependent Dirac equation  $(i\hbar\partial_t - H)|\Psi\rangle = 0$ , then the Foldy-Wouthuysen transformation turns it into  $|\psi\rangle = U|\Psi\rangle$ , where U is a unitary time-dependent operator. The time-dependent Dirac equation for  $|\Psi\rangle$  implies that  $|\psi\rangle$  is a solution of the time-dependent Schrödinger equation  $(i\hbar\partial_t - H')|\psi\rangle = 0$ , where  $H' = UHU^{-1} + i\hbar(\partial_t U)U^{-1}$  is the time-dependent Foldy-Wouthuysen Hamiltonian. In the following, an uppercase Greek letter  $(|\Phi\rangle$  or  $|\Psi\rangle)$  refers to a solution of the Dirac equation and the corresponding lowercase letter  $(|\phi\rangle$  or  $|\psi\rangle$ ) to its Foldy-Wouthuysen transformation.

As a consequence, a matrix element  $\langle \Phi | H | \Psi \rangle$  is not equal to  $\langle \phi | H' | \psi \rangle$ , but to  $\langle \phi | H' - i\hbar(\partial_t U)U^{-1} | \psi \rangle$ . In other

words, H' has to be used to calculate the states  $|\phi\rangle$  and  $|\psi\rangle$  but not to calculate the matrix elements of the Hamiltonian.

The second subtelty was observed by Yang [4] and concerns the most straightforward way to use the Foldy-Wouthuysen Hamiltonian H'(t), where the time dependence is now explicit, to compute transition probabilities. This Hamiltonian is split into a time-independent part  $H'_0$  and a time-dependent one  $H'_1(t)$ , so that H'(t) = $H'_0 + H'_1(t)$ . The scalar product  $\langle \phi'_n | \psi(t) \rangle$ , where  $| \phi'_n \rangle$  is an eigenstate of  $H'_0$ , cannot be equal to the relativistic scalar product  $\langle \Phi_n | \Psi(t) \rangle$ . Indeed  $| \psi(t) \rangle = U(H_0 + H_1(t)) | \Psi(t) \rangle$ but  $| \phi'_n \rangle \neq U(H_0 + H_1(t)) | \Phi_n \rangle$  because  $| \phi'_n \rangle$  and  $| \Phi_n \rangle$ are independent of time whereas  $U(H_0 + H_1(t))$  depends on time. Since only the QED relativistic matrix elements where found to be gauge invariant,  $\langle \phi'_n | \psi(t) \rangle$  is generally not physically meaningful.

The two problems combine if first-order perturbation theory is naively applied with Foldy-Wouthuysen eigenstates and Hamiltonian. The Foldy-Wouthuysen interaction Hamiltonian  $H'_1(t) = H'(t) - H'_0 \neq U(H_0)(H(t) - H_0)U^{\dagger}(H_0)$ . As a consequence,  $\langle \phi'_n | H'_1(t) | \phi'_g \rangle$  is not equal to  $\langle \Phi_n | H_1(t) | \Phi_q \rangle$ .

To illustrate the variety of results that can be obtained by using first-order perturbation theory with semirelativistic Hamiltonians, we now examine four Hamiltonians used in practice. To help comparing these Hamiltonians, we express them in a common one-particle framework.

#### 7.2 The Blume Hamiltonian

Blume discussed the interaction of light with magnetic matter by starting from the Hamiltonian [118, 116]:

$$H^{\rm B} = \frac{\boldsymbol{\pi}^2}{2m} + eV - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B} - \frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma} \cdot (\mathbf{E} \times \boldsymbol{\pi}),$$
(13)

where  $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$ . This Hamiltonian is the sum of four terms: (i) the kinetic energy of the electron, (ii) an external potential, (iii) the Zeeman interaction between the electron and a magnetic field and (iv) the spin-orbit interaction (because, for a spherical V and a static  $\mathbf{A}, \boldsymbol{\sigma} \cdot (\mathbf{E} \times \mathbf{p}) = \frac{-1}{r} \frac{dV}{dr} \boldsymbol{\sigma} \cdot (\mathbf{r} \times \mathbf{p}) = \frac{-1}{r} \frac{dV}{dr} \boldsymbol{\sigma} \cdot \mathbf{L}$ ). There are several differences between our notation and

There are several differences between our notation and Blume's: he considers a many-body Hamiltonian (involving sums over electrons) and writes  $\sum_{ij} V(r_{ij})$  for our eV, he adds the Hamiltonian  $H_{\gamma}$  of the free photons, he uses  $\mathbf{A}/c$ ,  $\nabla \times \mathbf{A}/c$  and  $\mathbf{s}$  where we use  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\sigma/2$ , finally, his Zeeman term is wrong by a factor of 2 in his first two papers on the subject [118,116], but this was corrected in the third one [119]. In this third paper, Blume also replaces  $\mathbf{E}$  by  $-\dot{\mathbf{A}}$ . This is not compatible with his quantized description of the photon field. Indeed, the time-derivative  $\dot{\mathbf{A}}$  is present in the Lagrangian but, after the Legendre transformation leading to the Hamiltonian,  $\dot{\mathbf{A}}$  is replaced by its canonical momentum  $-\mathbf{E}$ . Note that Blume does not sketch any derivation of his Hamiltonian.

#### 7.3 Foldy-Wouthuysen Hamiltonian

We consider now the so-called "gauge-invariant" Foldy-Wouthuysen Hamiltonian for positive-energy states up to order  $1/(mc)^2$  [120]:

$$H^{\rm FW} = H^{\rm B} + mc^2 - \frac{e\hbar^2}{8m^2c^2}\nabla\cdot\mathbf{E} - \frac{ie\hbar^2}{8m^2c^2}\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{E}).$$

The difference between the Foldy-Wouthuysen and the Blume Hamiltonians consists of three terms: the rest energy  $mc^2$  of positive-energy eigenstates, the Darwin term proportional to  $\nabla \cdot \mathbf{E}$  and a last term, proportional to  $\boldsymbol{\sigma} \cdot (\nabla \times \mathbf{E})$  and called the *curl-term*, that we discuss presently. A basic difference between  $H^{\rm B}$  and  $H^{\rm FW}$  must first be stressed: the former is a QED expression where the quantum fields A, B and E are independent of time because they are written in the Schrödinger representation, while the latter was derived under the assumption that  $\mathbf{A}$  and V are external time-dependent potentials. In particular, the curl-term disappears if the external field A is independent of time [101]. In the semi-classical treatment of light-matter interaction, the photons are represented by an external time-dependent potential and this term is present.

These Hamiltonians can be written  $H(\mathbf{A}, \Phi)$ , where the total vector potential  $\mathbf{A}$  and scalar potential  $\Phi$  are a sum  $\mathbf{A} = \mathbf{a}_0 + \mathbf{a}$ ,  $\Phi = \phi_0 + \phi$ , of static external potentials  $\mathbf{a}_0$  and  $\phi_0$  (representing the static internal and external fields) perturbed by dynamical potentials  $\mathbf{a}$  and  $\phi$  representing the incident electromagnetic wave. We write the interaction Hamiltonian as  $H_I = H(\mathbf{A}, \Phi) - H(\mathbf{a}_0, \phi_0)$ . The two Hamiltonians  $H^{\mathrm{B}}$  and  $H^{\mathrm{FW}}$  lead to two different interactions:  $H_I^{\mathrm{B}} = h_1 + h_2 + h_3 + h_4 + h_5 + h_6$  and  $H_I^{\mathrm{FW}} = H_I^{\mathrm{B}} + h_7$ , where

$$h_{1} = \frac{e^{2}}{2m} \mathbf{a}^{2},$$

$$h_{2} = -\frac{e}{m} \mathbf{a} \cdot \boldsymbol{\pi}_{0},$$

$$h_{3} = -\frac{e\hbar}{2m} \boldsymbol{\sigma} \cdot (\nabla \times \mathbf{a}),$$

$$h_{4} = \frac{e^{2}\hbar}{4m^{2}c^{2}} \boldsymbol{\sigma} \cdot (\mathbf{e} \times \mathbf{a}),$$

$$h_{5} = -\frac{e\hbar}{4m^{2}c^{2}} \boldsymbol{\sigma} \cdot (\mathbf{e} \times \boldsymbol{\pi}_{0}),$$

$$h_{6} = \frac{e^{2}\hbar}{4m^{2}c^{2}} \boldsymbol{\sigma} \cdot (\mathbf{e}_{0} \times \mathbf{a}),$$

$$h_{7} = -\frac{ie\hbar^{2}}{8m^{2}c^{2}} \boldsymbol{\sigma} \cdot (\nabla \times \mathbf{e}),$$

with  $\pi_0 = \mathbf{p} - e\mathbf{a}_0$ . The *curl-term* in  $H^{\text{FW}}$  is the origin of the presence of  $h_7$  in  $H_I^{\text{FW}}$ . It originates from the term  $i\hbar(\partial_t U)U^{-1}$  in the time-dependent Foldy-Wouthuysen Hamiltonian. The Darwin term gives no contribution to the interaction because  $\nabla \cdot \mathbf{e}$  is zero for the electromagnetic wave. The terms  $h_5$  and  $h_6$  were omitted by Blume, who considered them to be small [118]. We shall see that  $h_5$  is the source of a spin-position term which is not negligible in x-ray magnetic circular dichroism (XMCD) spectra [5].

#### 7.4 Textbook Foldy-Wouthuysen Hamiltonian

Standard textbooks often derive a Foldy-Wouthuysen Hamiltonian  $H^{\text{TFW}}$  which is the same as  $H^{\text{FW}}$ , except for the fact that  $\pi$  is replaced by **p** in the spin-orbit term [121, fact that  $\pi$  is replaced by  $\mathbf{p}$  in the spin orbit term  $(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added [121] For all the Hamiltonians presented here, are placed by  $\mathbf{p}$  in the spin orbit term  $(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added [121] For all the Hamiltonians presented here, are placed by  $\mathbf{p}$  in the spin orbit term  $(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added [121] For all the Hamiltonians presented here, are placed by  $\mathbf{p}$  in the spin orbit term  $(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added [121] For all the Hamiltonians presented here, are placed by  $\mathbf{p}$  in the spin orbit term  $(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added [121] I all the Hamiltonians presented here, are placed by  $\mathbf{p} \cdot \mathbf{p}$  is the spin orbit term  $(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added [121] I all the Hamiltonians presented here, are placed by  $\mathbf{p} \cdot \mathbf{p}$  is the spin orbit term  $\mathbf{p} \cdot \mathbf{p}$  is the spin orbit term. nians is a term in  $\boldsymbol{\sigma} \cdot \mathbf{E} \times \mathbf{A}$ . This results in the absence of  $h_4$  and  $h_6$  in the perturbation Hamiltonian, which changes the transition probabilities.

#### 7.5 NRQED

To deal with QED calculations involving bound states, Caswell and Lepage proposed an alternative approach to relativistic effects, called non-relativistic QED (NRQED), which turned out to be highly successful [123]. They wrote the most general gauge-invariant non-relativistic Lagrangian plicit calculations showed that this contribution can apterms and fitted the coefficients of these terms to known QED processes [124].

The corresponding NRQED Hamiltonian is the same as  $H^{\rm FW}$  up to order  $c^{-2}$ , but its interpretation is different [124]. Indeed, NRQED is a quantum field theory, and the fields are independent of time in the Schrödinger representation. However, the curl-term is present in timeindependent NRQED although it is generated by a timedependence in  $H^{\rm FW}$ . In particular, the curl-term must not be removed from the Hamiltonian to calculate matrix elements of the Hamiltonian operator, in contrast to the example of section 7.1.

Besides these four different Hamiltonians, we consider an additional source of discrepancies between authors: the commutators.

#### 7.6 Commutators

To derive the multipole expansion of the matrix element of  $H_I$ , it is useful to replace  $\pi$  by a commutator with  $H_0 = H(\mathbf{a}_0, \phi_0)$ . The derivations that start from Blume's interaction Hamiltonian usually use the relation [125,2].

$$\mathbf{p} = \frac{mi}{\hbar} [H_0, \mathbf{r}]. \tag{14}$$

However, if one considers the static Hamiltonian given by Blume (13), its commutator with  $\mathbf{r}$  is:

$$[H_0^{
m B},{f r}]=-rac{i\hbar}{m}m{\pi}_0+rac{e\hbar}{4m^2c^2}(i\hbar)(m{\sigma} imes{f e}_0),$$

which is different from Eq. (14) because **p** is replaced by  $\pi_0 = \mathbf{p} - e\mathbf{a}_0$  and because of the term proportional to  $c^{-2}$ . The commutator of  $\mathbf{r}$  with  $H_0^{\text{TFW}}$  and  $H_0^{\text{FW}}$  are the same. In  $H_I^{\text{FW}}$  and  $H_I^{\text{B}}$ , when  $\pi_0$  in  $h_2$  is rewritten as a function of the commutator, the extra relativistic term leads to the cancellation of  $h_6$ , which is important in XMCD. On the other hand, it leads to a contribution  $\frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot[\nabla v_0\times\mathbf{a}]$ in  $H_I^{\rm TFW}$ .

If the mass-velocity term  $-(\mathbf{p} \cdot \mathbf{p})^2$  is present in  $H_0$ , the additional contribution to the commutator,  $\frac{i\hbar(\mathbf{p}\cdot\mathbf{p})\mathbf{p}}{2m^3c^2}$  is small compared to  $\frac{i\hbar}{m}\mathbf{p}$  if the order of magnitude of the kinetic energy of the core state satisfies  $E_k \ll mc^2$ .

 $\pi_0$  with  $H_0$ :

$$\mathbf{e}_0 = -\nabla v_0 = \frac{-i}{e\hbar} [\boldsymbol{\pi}_0, H_0].$$

In the case of absorption, the commutator transforms into a factor  $\Delta E = -\hbar \omega$  in the cross section so that  $h_5$  and  $h_6$ lead to the same contribution to the matrix element:

$$\frac{-ie\hbar\omega}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{a}\times\boldsymbol{\pi}_0),$$

which corresponds to the *spin-position* interaction. Expear two times, one time or cancel completely, according to which Hamiltonian and which commutator was used. Starting from  $H^{\rm FW}$ , the same absorption cross section as in our new approach can be derived. However, in the case of scattering, even with  $H^{\rm FW}$ , there is a factor  $\Delta E/\hbar\omega$ which is not correct. The same kind of discrepancy was already observed in the literature [10, 42].

#### **8** Conclusion

This paper was written because of the gauge-dependence of transition probabilities in the semi-classical approach and because we observed, after other authors [125, 126], that different semi-relativistic Hamiltonians lead to different cross-sections.

Our solution makes essential use of quantum electrodynamics as the correct gauge-invariant framework to discuss the interaction of light with matter. It is well-known that the semi-classical and QED absorption cross-sections are identical in the Coulomb gauge [127]. This is compatible with our discussion because, to go from the Coulomb gauge to another gauge, the semi-classical approach only involves the operator  $M_A$ , while QED involves a redefinition of the space of states, including in an essential way non-physical polarizations and even ghost states in the BRST approach. This redefinition is able to maintain gauge invariance where the semi-classical  $M_A$  fails to do  $\mathbf{so.}$ 

In the present paper, the stationary states of the electronic system was taken to be eigenstates of  $H_e$ . The interaction Hamiltonian  $H_{e\gamma}$  can modify these states through various QED effects, for example the Breit interaction discussed by Bethe and Salpeter [128]. We expect these contributions to be small in x-ray spectroscopy.

The explicit calculation of the spin-position contribution at the K-edge of Fe, Co and Ni will presented in a forthcoming publication [5].

It was known since Heisenberg in 1928 [129], that the Thomson cross-section which is due to the  $A^2$  term in the

non-relativistic approach, can be derived from the relativistic framework by using a sum over negative-energy states [130,101]. We intend to provide a more accurate discussion of the contribution of negative-energy states to the scattering cross-section by using our many-body Foldy-Wouthuysen approach.

#### 9 Acknowledgments

We are very grateful to Uwe Gerstmann, Matteo Calandra and Nora Jenny Vollmers for encouraging us to work on the problem of the relativistic effects in x-ray absorption spectroscopy. We thank Alexander Silenko for this help concerning the Foldy-Wouthuysen transformations. Discussions with Amélie Juhin, Sergio Di Matteo, Yves Joly and Philippe Sainctavit are gratefully acknowledged. We are very grateful to Maria Esteban for her guidance through the mathematical literature on the many-body Dirac equation.

This work was supported by French state funds managed by the ANR within the Investissements d'Avenir programme under Reference No. ANR-11-IDEX-0004-02, and more specifically within the framework of the Cluster of Excellence MATISSE led by Sorbonne Universités.

#### A Foldy-Wouthuysen transformation

In this appendix we derive a new general Foldy-Wouthuysen transformation and we apply it to the many-body Hamiltonian.

#### A.1 General Foldy-Wouthuysen transformation

To derive a many-body Foldy-Wouthuysen transformation, we first notice that, in the one-body case,  $\beta$  endows the space of spinors with the structure of a Krein space, where  $\beta$  is then called a *fundamental symmetry* [131]. For quite a different purpose [132], we investigated the tensor product of such spaces and showed that the fundamentaly symmetry of the Nth tensor power is essentially  $\eta = \beta^{\otimes N}$ . The abstract Krein-space framework leads us naturally to the following theorem:

Assume that  $H_D$  and  $\eta$  are self-adjoint operators and  $\eta^2 = 1$ . Then, there is a unitary operator U such that  $U = \eta U^{\dagger} \eta$  and  $\eta U H_D U^{\dagger} \eta = U H_D U^{\dagger}$ . Moreover, if  $|\psi_D\rangle$  is an eigenstate of  $H_D$  with positive (resp. negative) eigenvalue, then  $|\psi\rangle = U |\psi_D\rangle$  satisfies  $|\psi\rangle = \eta |\psi\rangle$  (resp.  $|\psi\rangle = -\eta |\psi\rangle$ ).

The condition  $U = \eta U^{\dagger} \eta$  does not appear in Foldy and Wouthuysen works. It was added by Eriksen [133,134, 108]. It means that U is self-adjoint for the Krein-space structure.

Let us start with general considerations involving a self-adjoint operator  $\eta$  such that  $\eta^2 = 1$ . It can be used to define projectors  $B_{\pm} = (1 \pm \eta)/2$ . It is clear that  $B_+ + B_- = 1$ ,  $B_{\pm}^2 = B_{\pm}$ ,  $B_{\pm}^{\dagger} = B_{\pm}$  and  $B_+B_- = B_-B_+ =$ 

0. A vector  $|\psi\rangle$  is said to be even (odd) if  $\eta|\psi\rangle = |\psi\rangle$  $(\eta|\psi\rangle = -|\psi\rangle)$ . Then, any vector  $|\psi\rangle$  can be written as the sum of its even part  $B_+|\psi\rangle$  and its odd part  $B_-|\psi\rangle$ . An operator H is said to be even (odd) if it transforms an even state into an even (odd) state and an odd state into an odd (even) state. An operator H is even (odd) if and only if  $\eta H \eta = H$  ( $\eta H \eta = -H$ ). Thus, the theorem states that  $UH_DU$  is an even operator. Any operator H can be written as the sum of its even part  $B_+HB_+ + B_-HB_$ and its odd part  $B_+HB_- + B_-HB_+$ .

Our proof of the theorem is essentially a generalized and rigorous version of Eriksen's proof [134]. We use the fact that  $H_D$  is self-adjoint to define  $\lambda = \operatorname{sign} H_D$  by functional calculus. The operator  $\lambda$  is called the *flat band Hamiltonian* in topological insulator theory [135]. In physical terms, let  $|\psi_D\rangle$  be an eigenstate of  $H_D$  for the energy E, then  $\lambda |\psi_D\rangle = |\psi_D\rangle$  if  $E \ge 0$  and  $\lambda |\psi_D\rangle = -|\psi_D\rangle$  if E < 0. Since  $\eta$  and  $\lambda$  are self-adjoint and  $\eta^2 = \lambda^2 = 1$ , they are bounded and  $\eta\lambda$  is unitary:  $\eta\lambda(\eta\lambda)^{\dagger} = \eta\lambda\lambda\eta =$  $\eta^2 = 1$  and  $(\eta\lambda)^{\dagger}\eta\lambda = 1$ . By the spectral theorem for unitary operators [136], there is a unique family of orthogonal projections  $P_t$  such that

$$\eta \lambda = \int_{-\pi}^{\pi} e^{it} P_t dt$$

In the finite dimensional case we could write this [137]

$$\eta \lambda = \sum_{n} e^{it_n} |\phi_n\rangle \langle \phi_n|.$$

Thus,

$$\lambda \eta = (\eta \lambda)^{\dagger} = \int_{-\pi}^{\pi} e^{-it} P_t dt = \int_{-\pi}^{\pi} e^{it} P_{-t} dt$$

and, by unicity of  $P_t$ ,  $\eta \lambda = \eta (\eta \lambda)^{\dagger} \eta$  implies  $P_t = \eta P_{-t} \eta$ . We can now define a unitary square root U of  $\eta \lambda$  by functional calculus [138,139]:

$$U = \sqrt{\eta \lambda} = \int_{-\pi}^{\pi} e^{it/2} P_t dt,$$

which satisfies

$$\eta U^{\dagger} \eta = \int_{-\pi}^{\pi} e^{-it/2} \eta P_t \eta dt = \int_{-\pi}^{\pi} e^{-it/2} P_{-t} dt = U.$$

We now show that this U satisfies the intertwining relation  $\eta U = U\lambda$ . Indeed, the relation  $U^2 = \eta\lambda$  implies  $U = U^{\dagger}\eta\lambda$ . By multiplying from the left with  $\eta$  and using  $\eta U^{\dagger}\eta = U$  we find  $\eta U = U\lambda$ . This important relation implies that  $H = \eta H\eta$  and that  $|\psi\rangle = U|\psi_D\rangle$  is even if  $|\psi_D\rangle$  is a positive energy state and odd if  $|\psi_D\rangle$  is a negative energy state.

The first property is easy to show:

$$\eta H\eta = \eta U H_D U^{\dagger} \eta = U \lambda H_D \lambda U^{\dagger} = U H_D \lambda^2 U^{\dagger} = H,$$

because  $\lambda$  commutes with  $H_D$  since it is a function of  $H_D$ .

To show the second property, let  $\Gamma_{\pm} = (1 \pm \lambda)/2$ , so that  $\Gamma_{\pm}$  projects onto the space of positive energy and  $\Gamma_{\pm}$  of negative energy, and recall that  $B_{\pm} = (1 \pm \eta)/2$ . For a one-body system,  $B_{\pm}$  projects onto the large/small components. Then,  $U\Gamma_{\pm} = U/2\pm U\lambda/2 = U/2\pm \eta U/2 = B_{\pm}U$ , which can be used to show that the Foldy-Wouthuysen wavefunctions  $|\psi\rangle = U|\psi_D\rangle$  corresponding to positive energy have only even components. Indeed, let  $|\psi_D\rangle$  be an eigenstate of  $H_D$  corresponding to a positive energy. By definition of  $\lambda$  we have  $\Gamma_+|\psi_D\rangle = |\psi_D\rangle$  and  $\Gamma_-|\psi_D\rangle = 0$ . Thus,  $U\Gamma_+|\psi_D\rangle = U|\psi_D\rangle = |\psi\rangle$  and  $U\Gamma_+ = B_+U$  implies  $|\psi\rangle = B_+U|\psi_D\rangle = B_+|\psi\rangle$ . Thus  $\eta|\psi\rangle = \eta B_+|\psi\rangle = B_+|\psi\rangle$  and  $|\psi\rangle$  is even. Similarly  $0 = B_-|\psi\rangle$ , so that the odd part of  $|\psi\rangle$  is zero.

For a one-body system, even components and large components are identical. Indeed a Dirac one-body wavefunction can be written

$$|\psi_D\rangle = \begin{pmatrix} \phi\\\psi \end{pmatrix},$$

If  $\eta = \beta$ , then the even part and the odd parts of  $|\psi_D\rangle$  are, respectively,

$$\begin{pmatrix} \phi \\ 0 \end{pmatrix}$$
 and  $\begin{pmatrix} 0 \\ \psi \end{pmatrix}$ .

so that the small components of  $|\psi\rangle$  are zero for a positiveenergy state. This is not true for many-body systems. For example, if we neglect antisymmetrization for notational convenience, a two-body state can be obtained as the tensor product of one-body wavefunctions:

$$|\psi_D\rangle = \begin{pmatrix} \phi_1\\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} \phi_2\\ \psi_2 \end{pmatrix}.$$

Then, the even part of  $|\psi_D\rangle$  is

$$\begin{pmatrix} \phi_1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} \phi_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ \psi_2 \end{pmatrix},$$

while its odd part is

$$\begin{pmatrix} \phi_1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ \psi_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} \phi_2 \\ 0 \end{pmatrix}.$$

The characterization of U as the square root of  $\eta\lambda$  is not easy to handel. We give now a much simpler characterization:

Let U be a unitary operator continuously defined (outside zero) in terms of  $H_D$  such that: (i)  $U = \eta U^{\dagger} \eta$ ; (ii)  $\eta U H_D U^{\dagger} \eta = U H_D U^{\dagger}$ ; (iii)  $U^2(-H_D) = -U^2(H_D)$ . Then  $U^{\dagger} \eta U = \pm \operatorname{sign}(H_D)$ .

To prove this, define  $Z = U^{\dagger}\eta U$ . Clearly,  $Z^{\dagger} = Z$ . Moreover, Z is defined in terms  $H_D$  since U is. However, for Z to be a function of  $H_D$  in the sense of functional calculus, Z needs to commute with  $H_D$  [140]:if we multiply condition (*ii*) from the right by  $\eta U$  we find  $\eta U H_D =$  $U H_D U^{\dagger} \eta U = U H_D Z$ . Hence,

$$ZH_D = U^{\dagger}\eta UH_D = U^{\dagger}UH_DZ = H_DZ.$$

Thus, there is a real function f(t) and a family of orthogonal projections  $P_t$  corresponding to the eigenstates of  $H_D$  such that

$$Z = \int_{-\infty}^{\infty} f(t) dP_t.$$

Moreover,  $Z^2 = 1$  because  $Z^2 = U^{\dagger} \eta U U^{\dagger} \eta U = U^{\dagger} \eta^2 U = U^{\dagger} U = 1$ . Therefore,  $f^2(t) = 1$  for every t. Finally, observe that  $Z = \eta^2 U^{\dagger} \eta U = \eta U^2$ , and condition (*iii*) implies that Z is an odd function of  $H_D$ : f(-t) = -f(t). To conclude that  $f(t) = \pm \text{sign}t$ , we need to add the condition of continuity on f outside zero. Indeed, functional calculus is valid for measurable functions and we could build a noncontinuous odd function f such that  $f^2 = 1$  outside the origin. In practice this does not take place because U is smoothly defined in terms of  $H_D$ , except at zero. No odd continuous function can satisfy  $f^2 = 1$  over  $\mathbb{R}$ . It has to be discontinuous at zero. Since it is crucial that  $f^2 = 1$  everywhere, we can choose either sign0 = 1 or sign0 = -1. Both solutions are valid.

#### A.2 Many-body Foldy-Wouthuysen transformation

To apply our Foldy-Wouthuysen approach to the manybody Dirac Hamiltonian we face the following problem. The generalization of  $H^0$  is imposed by the many-body Dirac Hamiltonian:

$$H_N^0 = \sum_{n=1}^N \beta_n m c^2,$$

where  $\beta_n$  is the matrix  $\beta$  acting on the *n*th electron (i.e.  $\beta_n = 1^{\otimes (n-1)} \otimes \beta \otimes 1^{\otimes (N-n)}$ ). This definition is valid because  $H_N^0$  commutes with the projector  $P_N$  onto the space of antisymmetric *N*-body states.

We show in the Appendix that a Foldy-Wouthuysen transformation can be defined whenever we have a selfadjoint operator  $\eta$  (with  $\eta^2 = 1$ ) to define parity. In the one-body case,  $\beta^2 = 1$  and  $\eta = \beta$  defines parity. But in the many-body case the operator  $\sum_{n=1}^{N} \beta_n$  suggested by  $H_N^0$ cannot be used for that purpose because its square is not proportional to the identity (it contains products  $\beta_n \beta_m$ ). It turns out that  $\eta = \beta_1 \otimes \cdots \otimes \beta_n$  is the natural many-body generalization of  $\beta$ . Indeed,  $\eta^{\dagger} = \eta$  and  $\eta^2 = 1$ . Moreover,  $\eta$  commutes with  $P_N$ , which allows us to work with tensor products instead of antisymmetric tensor products.

In the literature, the Foldy-Wouthuysen transformation was studied for two-body Hamiltonians [141, 142, 133], but the results were rather complicated and not easy to extend to the many-body case. Pachucki calculated an alternative semi-relativistic Hamiltonian by carrying out a Foldy-Wouthuysen transformation of the one-particle Dirac Hamiltonian followed by a perturbative expansion of the many-body Green function [143].

The even and odd parts of  $H_N$  are then  $H_N^0 + \mathcal{E}$  $\mathcal{O}$ , respectively:

$$\mathcal{E} = e \sum_{n=1}^{N} \phi_0(\mathbf{r}_n) + e \sum_{m \neq n} V(\mathbf{r}_m - \mathbf{r}_n),$$
$$\mathcal{O} = \sum_{n=1}^{N} c \boldsymbol{\alpha}_n \cdot \boldsymbol{\pi}_n = \sum_{n=1}^{N} \mathcal{O}_n,$$

where  $V(\mathbf{r}) = \frac{e}{8\pi\epsilon_0|\mathbf{r}|}$  is the Coulomb potential and  $\boldsymbol{\pi}_n =$  $-i\hbar\nabla_n - e\mathbf{a}_0(\mathbf{r}_n).$ 

At first order in  $c^{-2}$ , the Foldy-Wouthuysen operator is  $U = e^{iS^{(1)}}$  where

$$S^{(1)} = -\frac{i}{2mc^2} \sum_n \beta_n \mathcal{O}_n.$$

Indeed, it can be checked that  $i[S^{(1)}, H_N^0] = -\mathcal{O}$  removes the odd term of  $H_D$ . At this order  $U = U_1 \otimes \cdots \otimes U_N$  is a tensor power of one-body Foldy-Wouthuysen operators, as proposed by Moshinksy and Nikitin [144].

However, this tensor-power form does not hold at higher orders. Indeed, we show now that at the next order, the many-body Foldy-Wouthuysen Hamiltonian is the sum of one-body and two-body contributions. The usual formal Foldy-Wouthuysen transformation  $U = e^{iS^{(1)}}e^{iS^{(2)}}$  can be carried out almost unchanged and we find, with m as expansion parameter, at order  $m^{-2}$ :

$$\begin{split} H_{\rm FW} &= H_N^0 + \mathcal{E} + \frac{1}{2mc^2} \sum_{n=1}^N \beta_n \mathcal{O}_n^2 \\ &- \frac{1}{8m^2 c^4} \sum_{n=1}^N \left[ \mathcal{O}_n, \left[ \mathcal{O}_n, e\varphi_n + eV \right] \right] \\ &+ \frac{1}{8m^2 c^4} \sum_{p \neq n} \beta_p \beta_n [\mathcal{O}_p, \left[ \mathcal{O}_n, V \right] ]. \end{split}$$

This Hamiltonian obeys  $\eta H_{\rm FW} \eta = H_{\rm FW}$  which makes it a Foldy-Wouthuysen Hamiltonian.

It rewrites

$$H_{\rm FW} = \sum_{n=1}^{N} H_{\rm FW}^{n} + H_{\rm FW}^{MB}.$$
 (15)

where each  $H_{\rm FW}^n$  is the usual one-body Foldy-Wouthuysen Hamiltonian:

$$\begin{aligned} H_{\rm FW}^n &= \beta_n m c^2 + e \phi_0(\mathbf{r}_n) + \sum_{p \neq n} e V(\mathbf{r}_n - \mathbf{r}_p) \\ &+ \frac{1}{2m} \beta_n \boldsymbol{\pi}_n^2 - e \hbar \boldsymbol{\Sigma}_n \cdot \mathbf{b}_0(\mathbf{r}_n) - \frac{\hbar^2 e}{8m^2 c^2} \nabla \cdot \mathbf{E}_n \\ &+ \frac{\hbar e}{8m^2 c^2} \boldsymbol{\Sigma}_n \cdot (\boldsymbol{\pi}_n \times \mathbf{E}_n - \mathbf{E}_n \times \boldsymbol{\pi}_n) \end{aligned}$$

where

$$\mathbf{b}_0(\mathbf{r}_n) = \nabla \times \mathbf{a}_0(\mathbf{r}_n)$$

$$\mathcal{E}$$
 and and

$$\mathbf{E}_n = -\nabla \phi_n(\mathbf{r}_p) - \sum_{p \neq n} \nabla V(\mathbf{r}_n - \mathbf{r}_p).$$

The mass-velocity term  $\frac{\beta_n}{8m^3c^2}(\mathbf{p}_n\cdot\mathbf{p}_n)^2$  would be obtained by expanding to higher order.

The two-body term  $H_{\rm FW}^{MB}$  arises because  $V(\mathbf{r}_m - \mathbf{r}_n) =$  $V(\mathbf{r}_{mn})$  is a two body operator:

$$\begin{aligned} H_{\rm FW}^{n,p} &= \frac{\hbar e}{8m^2c^2} \sum_{p\neq n}^{N} \left( \hbar \Delta V(\mathbf{r}_{np}) \right. \\ &- \boldsymbol{\Sigma}_n \cdot \left( \boldsymbol{\pi}_n \times \nabla V(\mathbf{r}_{np}) - \nabla V(\mathbf{r}_{np}) \times \boldsymbol{\pi}_n \right) \\ &+ 2\hbar \beta_n \beta_p (\boldsymbol{\alpha}_n \cdot \nabla_n) (\boldsymbol{\alpha}_p \cdot \nabla_p) V(\mathbf{r}_{np}) \right). \end{aligned}$$

By using [145]:

$$\partial_j \partial_k V(\mathbf{r}) = \frac{e^2}{8\pi\epsilon_0} \Big( -\delta_{jk} \frac{4\pi}{3} \delta(\mathbf{r}) - \delta_{jk} \frac{1}{r^3} + \frac{3r^j r^k}{r^5} \Big),$$

the derivatives in the last term can be rewritten

$$(\boldsymbol{\alpha}_{n} \cdot \nabla_{n})(\boldsymbol{\alpha}_{p} \cdot \nabla_{p})V(\mathbf{r}_{np}) = \sum_{jk} \boldsymbol{\alpha}_{n}^{j} \boldsymbol{\alpha}_{p}^{k} \partial_{j} \partial_{k}V(\mathbf{r}_{np})$$
$$= \frac{e^{2}}{8\pi\epsilon_{0}} \Big( -\frac{4\pi}{3} \boldsymbol{\alpha}_{n} \cdot \boldsymbol{\alpha}_{p} \delta(\mathbf{r}_{np}) - \frac{\boldsymbol{\alpha}_{n} \cdot \boldsymbol{\alpha}_{p}}{|\mathbf{r}_{np}|^{3}}$$
$$+ 3\frac{\boldsymbol{\alpha}_{n} \cdot \mathbf{r}_{np} \boldsymbol{\alpha}_{p} \cdot \mathbf{r}_{np}}{|\mathbf{r}_{np}|^{5}} \Big).$$

This expression looks superficially like some contributions to the Breit interaction as presented by Bethe and Salpeter [128]. However, they are different since the Breit interaction is due to the exchange of a photon and not to a semirelativistic effect. Note that the last two terms are singular. It is known that the expansion of the Foldy-Wouthuysen transformation as a power serie in  $1/c^2$  becomes more and more singular because of the presence of the Coulomb potential [146]. At order  $m^{-2}$ , the transformation writes

$$\begin{split} U &= 1 + \frac{1}{2mc^2} \sum_n \beta_n \mathcal{O}_n - \frac{1}{8m^2c^4} \Big(\sum_n \beta_n \mathcal{O}_n\Big)^2 \\ &+ \frac{1}{4m^2c^4} \sum_n \beta_n \Big[\sum_m \beta_m \mathcal{O}_m, \mathcal{E}\Big] \end{split}$$

and it obeys  $U = \eta U^{\dagger} \eta$ . We also checked that  $U^2$  is odd in  $H_D$  after paying attention to the discontinuity at zero discussed in the Appendix. Thus, the positive (negative) energy eigenstate of  $H_D$  are transformed into even (odd) states by the action of U.

#### **B** Authors contributions

All the authors were involved in the preparation of the manuscript. All the authors have read and approved the final manuscript.

#### References

- 1. C. Brouder, J. Phys.: Condens. Matter 2, 701 (1990)
- Y. Joly, S. DiMatteo, O. Bunau, Eur. Phys. J. Special Topics 208, 21 (2012)
- 3. D.H. Kobe, A.L. Smirl, Am. J. Phys. 46, 624 (1978)
- 4. K.H. Yang, J. Phys. A: Math. Gen. 15, 437 (1982)
- N. Bouldi, N.J. Vollmers, C.G. Delpy-Laplanche, Y. Joly, A. Juhin, P. Sainctavit, C. Brouder, M. Calandra, L. Paulatto, F. Mauri et al., arXiv:1705.02192
- 6. J. W. E. Lamb, Phys. Rev. 85, 259 (1952)
- 7. I.P. Grant, J. Phys. B: Atom. Molec. Phys. 7, 1458 (1974)
- I.P. Grant, A.F. Starace, J. Phys. B: Atom. Molec. Phys. 8, 1999 (1975)
- 9. K.H. Yang, Annals Phys. 101, 6 (1976)
- R. Zeyher, H. Bilz, M. Cardona, Solid State Commun. 19, 57 (1976)
- J.J. Forney, A. Quattropani, F. Bassani, Nuovo Cimento 37B, 78 (1977)
- 12. Y. Aharonov, C.K. Au, Phys. Rev. A 20, 1553 (1979)
- 13. S.T. Epstein, Chem. Phys. Lett. **65**, 417 (1979)
- G. Grynberg, E. Giacobino, J. Phys. B: Atom. Molec. Phys. 12, L93 (1979)
- 15. K. Haller, R.B. Sohn, Phys. Rev. A 20, 1541 (1979)
- S. Olariu, I. Popescu, C.B. Collins, Phys. Rev. D 20, 3095 (1979)
- D.H. Kobe, K.H. Yang, J. Phys. A: Math. Gen. 13, 3171 (1980)
- 18. D.H. Kobe, C.T. Wen, Phys. Lett. 80A, 121 (1980)
- C. Leubner, P. Zoller, J. Phys. B: Atom. Molec. Phys. 13, 3613 (1980)
- 20. Y. Aharonov, C.K. Au, Phys. Lett. 86A, 269 (1981)
- 21. M.I. Shirokov, Sov. Phys. JETP 54, 645 (1981)
- 22. D.H. Kobe, Int. J. Quant. Chem. **21**, 685 (1982)
- D.H. Kobe, E.C.T. Wen, K.H. Yang, Phys. Rev. D 26, 1927 (1982)
- 24. Y. Aharonov, C.K. Au, Phys. Lett. 88A, 491 (1982)
- 25. K.H. Yang, Phys. Lett. **92A**, 71 (1982)
- 26. K.H. Yang, J. Phys. A: Math. Gen. 15, 1201 (1982)
- 27. T.E. Feuchtwang, E. Kazes, H. Grotch, P.H. Cutler, Phys. Lett. **93A**, 4 (1982)
- E. Kazes, T.E. Feuchtwang, P.H. Cutler, H. Grotch, Annals Phys. **142**, 80 (1982)
- 29. Y. Aharonov, C.K. Au, Phys. Lett. 95A, 412 (1983)
- 30. D. Lee, A.C. Albrecht, J. Chem. Phys. 78, 5373 (1983)
- 31. K.H. Yang, J. Phys. A: Math. Gen. 16, 919 (1983)
- 32. C.K. Au, J. Phys. B: Mol. Phys. 17, L59 (1984)
- 33. T.E. Feuchtwang, E. Kazes, P.H. Cutler, J. Phys. A: Math. Gen. 17, 1157 (1984)
- 34. T.E. Feuchtwang, E. Kazes, P.H. Cutler, H. Grotch, J. Phys. A: Math. Gen. 17, 151 (1984)
- 35. D.H. Kobe, K.H. Yang, Phys. Rev. A 32, 952 (1985)
- 36. D. Adu-Gyamfi, J. Phys. A: Math. Gen. 19, 3443 (1986)
- 37. K.H. Yang, D.H. Kobe, Annals Phys. 168, 104 (1986)
- 38. D.H. Kobe, S.M. Golsham, J. Phys. A: Math. Gen. 20, 2813 (1987)
- 39. J. W. E. Lamb, R.R. Schlicher, M.O. Scully, Phys. Rev. A 36, 2763 (1987)
- G.P. Arrighini, C. Guidotti, N. Durante, Gazz. Chim. Ital. 118, 703 (1988)
- N. Durante, G.P. Arrighini, C. Guidotti, Z. Phys. D 8, 63 (1988)
- 42. K.H. Yang, Ann. Phys. 186, 209 (1988)

- K. Haller, E. Lim-Lombridas, Found. Phys. 24, 217 (1994)
- 44. R.G. Woolley, Molec. Phys. **94**, 409 (1998)
- 45. R.G. Woolley, Inter. J. Quant. Chem. **74**, 531 (1999)
- R.G. Woolley, Proc. R. Soc. London A **456**, 1803 (2000)
   I.M. Savukov, W.R. Johnson, Phys. Rev. A **62**, 052506
- (2000) (2000)
- A.M. Stewart, J. Phys. A: Math. Gen. 33, 9165 (2000)
   S.W. Qian, Z.Y. Gu, Commun. Theor. Phys. 38, 267
- (2002)50. R.G. Woolley, On gauge invariance and molecular electro-
- dynamics, in New Trends in Quantum Systems in Chemistry and Physics. Volume 2. Advanced Problems and Complex Systems. Paris, France 1999, edited by J. Maruani, C. Minot, R. McWeeny, Y. Smeyers, S. Wilson (Kluwer Academic Publishers, Dordrecht, 2002), Vol. 7 of Progress in Theoretical Chemistry and Physics
- K. Rzazewski, R.W. Boyd, J. Modern Opt. 51, 1137 (2004)
- 52. S.W. Qian, J.S. Wang, Commun. Theor. Phys. 49, 308 (2008)
- A. Stokes, J. Phys. B: At. Mol. Opt. Phys. 46, 145505 (2013)
- A.D. Bandrauk, F. Fillion-Gourdeau, E. Lorin, J. Phys. B: At. Mol. Opt. Phys. 46, 153001 (2013)
- A. Mandal, K.L.C. Hunt, J. Chem. Phys. 144, 044109 (2016)
- C. Cohen-Tannoudji, J. Dupont-Roc, G. Grynberg, *Pho*tons and Atoms. Introduction to Quantum Electrodynamics (Wiley-VCH, New York, 1997)
- M. Blagojević, F.W. Hehl, Gauge Theories of Gravitation (Imperial College Press, London, 2013)
- A. Bohm, Quantum Mechanics: Foundations and Applications, 3rd edn. (Springer, New York, 1993)
- T.E. Feuchtwang, E. Kazes, J. Phys. A: Math. Gen. 18, 2859 (1985)
- T.E. Feuchtwang, E. Kazes, P.H. Cutler, Found. Phys. 16, 1263 (1986)
- 61. T. Goldman, Phys. Rev. D 15, 1063 (1977)
- E. Kazes, T.E. Feuchtwang, H. Grotch, P.H. Cutler, Phys. Rev. D 27, 1388 (1983)
- S. Weinberg, The Quantum Theory of Fields. I Foundations (Cambridge University Press, Cambridge, 1995)
- N. Nakanishi, I. Ojima, Covariant Operator Formalism of Gauge Theories and Quantum Gravity (World Scientific, Singapore, 1990)
- M.E. Peskin, D.V. Schroeder, An Introduction to Quantum Field Theory (Addison-Wesley, Reading, 1995)
- O. Steinmann, Perturbative Quantum Electrodynamics and Axiomatic Field Theory (Springer, Berlin, 2000)
- F. Strocchi, An Introduction to Non-Perturbative Foundations of Quantum Field Theory (Oxford Science Publications, Oxford, 2013)
- 68. I. Bialynicki-Birula, Phys. Rev. D 2, 2877 (1970)
- R.E. Kallosh, I.V. Tyutin, Soviet J. Nucl. Phys. 17, 98 (1973)
- 70. E.B. Manoukian, Int. J. Theor. Phys. 27, 787 (1988)
- S. Weinberg, The Quantum Theory of Fields II. Modern Applications (Cambridge University Press, Cambridge, 1995)
- 72. S. Hollands, Rev. Math. Phys. 20, 1033 (2008)
- I.A. Batalin, P.M. Lavrov, I.V. Tyutin, Int. J. Mod. Phys. A29, 1450166 (2014)

- 74. P.Y. Moshin, A.A. Reshetnyak, Inter. J. Mod. Phys. A 30, 1550021 (2015)
- 75. H. Matsuda, R. Kubo, Prog. Theor. Phys. 63, 275 (1980)
- B.L. Voronov, P.M. Lavrov, I.V. Tyutin, Sov. J. Nucl. Phys. 36, 292 (1982)
- 77. F. Lenz, H.W.L. Naus, K. Ohta, M. Thies, Annals Phys. 233, 17 (1994)
- 78. T. Kashiwa, N. Tanimura, Fortschr. Phys. 45, 381 (1997)
- 79. D. Grigore, Ann. Phys. (Leipzig) **10**, 439 (2001)
- A.K. Das, R.R. Francisco, J. Frenkel, Phys. Rev. D 88, 085012 (2013)
- 81. K. Haller, L.F. Landovitz, Phys. Rev. 171, 1749 (1968)
- 82. K. Haller, L.F. Landovitz, Phys. Rev. D 2, 1498 (1970)
- 83. K. Haller, Acta Phys. Aust. 42, 163 (1975)
- 84. K. Haller, L.F. Landovitz, Phys. Rev. Lett. 22, 245 (1969)
- 85. K. Ohtaka, Y. Tanabe, Rev. Mod. Phys. **62**, 929 (1990)
- E. Klevak, J.J. Kas, J.J. Rehr, Phys. Rev. B 89, 085123 (2014)
- I.P. Grant, Relativistic Quantum Theory of Atoms and Molecules (Springer, New York, 2007)
- W. Greiner, J. Reinhardt, Field Quantization (Springer, Berlin, 1996)
- I. Bialynicki-Birula, The Hamiltonian of quantum electrodynamics, in Quantum Electrodynamics and Quantum Optics, edited by A.O. Barut (Plenum Press, New York, 1984), Vol. 110 of NATO ASI Series B, pp. 41–61
- 90. M.J. Esteban, M. Lewin, E. Séré, Bull. Amer. Math. Soc. 45, 535 (2008)
- 91. W. Kutzelnigg, Chem. Phys. 395, 16 (2012)
- 92. C. Brouder, G. Panati, G. Stoltz, Ann. Henri Poincaré 11, 1285 (2010)
- W. Heitler, *The Quantum Theory of Radiation*, 3rd edn. (Dover, New York, 1984)
- 94. C. Brouder, G. Panati, G. Stoltz, Phys. Rev. Lett. 103, 230401 (2009)
- 95. B.A. Veklenko, Sov. Phys. J. 30, 555 (1987)
- 96. S. Hassing, E.N. Svendsen, J. Raman Spectr. 35, 87 (2004)
- 97. P.W. Milonni, R. Loudon, P.R. Berman, S.M. Barnett, Phys. Rev. A 77, 043835 (2008)
- 98. S. Mukamel, Phys. Rev. A 76, 021803 (2007)
- 99. S.T. Epstein, Phys. Rev. 98, 196 (1955)
- B. Hatfield, Quantum Field Theory of Point Particles and Strings (Addison-Wesley, Reading, 1992)
- P. Strange, *Relativistic Quantum Mechanics* (Cambridge University Press, Cambridge, 1998)
- 102. W.R. Johnson, D.R. Plante, J. Sapirstein, Relativistic calculations of transition amplitudes in the helium isoelectronic sequence (1995), Vol. 35 of Advances in Atomic, Molecular, and Optical Physics, pp. 255–329
- 103. D. Baye, Phys. Rev. C 86, 034306 (2012)
- 104. H. Ebert, P. Strange, L. Gyorffy, J. Appl. Phys. 63, 3055 (1988)
- 105. O. Šipr, J. Minár, H. Wende, H. Ebert, Phys. Rev. B 84, 115102 (2011)
- 106. L.L. Foldy, S.A. Wouthuysen, Phys. Rev. 78, 29 (1950)
- 107. W. Greiner, *Relativistic Quantum Mechanics*, 3rd edn. (Springer, Berlin, 2000)
- 108. A.J. Silenko, Phys. Rev. A 93, 022108 (2016)
- A.L. Fetter, J.D. Walecka, Quantum Theory of Many-Particle Systems (McGraw-Hill, Boston, 1971)
- 110. E.K.U. Gross, E. Runge, O. Heinonen, Many-Particle Theory (Adam Hilger, Bristol, 1991)

- C. Cohen-Tannoudji, B. Diu, F. Laloë, Mécanique Quantique (Hermann, Paris, 1973)
- J.D. Walecka, Advanced Modern Physics. Theoretical Foundations (World Scientific, Singapore, 2010)
- C. Brouder, M. Alouani, K.H. Bennemann, Phys. Rev. B 54, 7334 (1996)
- 114. C. Gougoussis, Excitations électroniques et magnétisme des matériaux: calcul ab initio de l'absorption X et du dichroïsme circulaire magnétique au seuil K (Université Pierre et Marie Curie-Paris 6, 2009), Ph.D. thesis
- J. Als-Nielsen, D. McMorrow, *Elements of Modern X-Ray* Physics (Wiley, New York, 2000)
- 116. M. Blume, J. Appl. Phys. 51, 3615 (1985)
- 117. M.M. Nieto, Phys. Rev. Lett. 38, 1042 (1977)
- 118. M. Blume, SSRL Report 83/02, 126 (1983)
- 119. M. Blume, Magnetic effects in anomalous dispersion, in Resonant Anomalous X-Ray Scattering, edited by G. Materlik, C.J. Sparks, K. Fischer (North-Holland, Amsterdam, 1994), pp. 495–512
- 120. L.L. Foldy, Phys. Rev. 87, 688 (1952)
- 121. C. Itzykson, J.B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980)
- 122. J.D. Bjorken, S.D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965)
- 123. W.E. Caswell, G.P. Lepage, Phys. Lett. 167B, 437 (1986)
- 124. G. Paz, Mod. Phys. Lett. A **30**, 1550128 (2015)
- 125. M. Takahashi, N. Hiraoka, Phys. Rev. B **92**, 094441 (2015)
- 126. S.D. Matteo (2016), private communication
- 127. R. Loudon, *The Quantum Theory of Light*, 3rd edn. (Clarendon Press, Oxford, 2000)
- H.A. Bethe, E.E. Salpeter, Quantum Mechanics of Oneand Two-Electron Atoms (Plenum, New York, 1977)
- 129. L.M. Brown, Stud. Hist. Phil. Mod. Phys. **33**, 211 (2002)
- J.J. Sakurai, Advanced Quantum Mechanics (Addison-Wesley, Reading, 1967)
- H. Baum, Spin-Strukturen und Dirac-Operatoren über pseudoriemannschen Mannigfaltigkeiten (Teubner, Lepzig, 1981)
- 132. C. Brouder, N. Bizi, F. Besnard, arXiv:1504.03890
- 133. E. Eriksen, Phys. Rev. 111, 1011 (1958)
- 134. E. Eriksen, M. Kolsrud, Nuovo Cimento (Suppl.) 18, 1 (1960)
- E. Prodan, H. Schulz-Baldes, Bulk and Boundary Invariants for Complex Topological Insulators (Springer, Berlin, 2016)
- M. Pourahmadi, Foundations of Time Series Analysis and Prediction Theory (Wiley, New York, 2001)
- 137. S. Lang, *Linear Algebra*, Undergraduate Graduate Texts in Mathematics, 3rd edn. (Springer-Verlag, Berlin, 1987)
   130. L. L. W. W. Line, M. Line, J. 150 (1999)
- 138. L.J. Wallen, Michigan J. Math. 16, 153 (1969)
- 139. A.A. Joye, Dynamical localization of random quantum walks on the lattice, in XVIIth International Congress on Mathematical Physics, edited by A. Jensen (World Scientific, Singapore, 2014), pp. 486–94
- 140. F.A. Berezin, M.A. Shubin, *The Schrödinger Equation* (Kluwer, Dordrecht, 1991)
- 141. Z.V. Chraplyvy, Phys. Rev. 91, 388 (1953)
- 142. Z.V. Chraplyvy, Phys. Rev. 92, 1310 (1953)
- 143. K. Pachucki, Phys. Rev. A 71, 012503 (2005)
- 144. M. Moshinsky, A. Nikitin, Rev. Mex. Fis. Suppl. 2 50, 66 (2004)
- 145. W. Weiglhofer, Am. J. Phys. 57, 455 (1989)
- 146. J.D. Morrison, R. Moss, Molec. Phys. 41, 491 (1980)
## X-ray magnetic and natural circular dichroism from first principles: Calculation of K- and $L_1$ -edge spectra

N. Bouldi,<sup>1,2,\*</sup> N. J. Vollmers,<sup>3</sup> C. G. Delpy-Laplanche,<sup>1</sup> Y. Joly,<sup>4</sup> A. Juhin,<sup>1</sup> Ph. Sainctavit,<sup>1</sup> Ch. Brouder,<sup>1</sup> M. Calandra,<sup>1</sup> L. Paulatto,<sup>1</sup> F. Mauri,<sup>5</sup> and U. Gerstmann<sup>3</sup>

<sup>1</sup>UPMC Université Paris 6, CNRS, UMR 7590, IRD, MNHN, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, case 115, 4 place Jussieu, 75252 Paris Cedex 5, France

<sup>2</sup>Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, Boîte Postale 48, 91192 Gif-sur-Yvette Cedex, France

<sup>3</sup>Lehrstuhl für Theoretische Physik, Universität Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany

<sup>4</sup>Institut Néel, CNRS and Université Joseph Fourier, Boîte Postale 166, 38042 Grenoble Cedex 9, France

<sup>5</sup>Departimento di Fisica, Università di Roma La Sapienza, Piazzale Aldo Moro 5, 00185 Roma, Italy

(Received 4 April 2017; published 17 August 2017)

An efficient first-principles approach to calculate x-ray magnetic circular dichroism (XMCD) and x-ray natural circular dichroism (XNCD) is developed and applied in the near-edge region at the *K* and  $L_1$  edges in solids. Computation of circular dichroism requires precise calculations of x-ray absorption spectra (XAS) for circularly polarized light. For the derivation of the XAS cross section, we used a relativistic description of the photon-electron interaction that results in an additional term in the cross section that couples the electric dipole operator with an operator  $\sigma \cdot (\epsilon \times \mathbf{r})$  that we call the spin position operator. The numerical method relies on pseudopotentials, on the gauge including projected augmented-wave method, and on a collinear spin relativistic description of the electron is structure. We apply the method to calculations of *K*-edge XMCD spectra of ferromagnetic iron, cobalt, and nickel and of I  $L_1$ -edge XNCD spectra of  $\alpha$ -LiIO<sub>3</sub>, a compound with broken inversion symmetry. For XMCD spectra we find that, even if the electric dipole term is the dominant one, the electric quadrupole term is not negligible (8% in amplitude in the case of iron). The term coupling the electric dipole operator with the spin-position operator is significant (28% in amplitude in the case of iron). We obtain a sum rule relating this term to the spin magnetic moment of the *p* states. In  $\alpha$ -LiIO<sub>3</sub> we recover the expected angular dependence of the XNCD spectra.

DOI: 10.1103/PhysRevB.96.085123

#### I. INTRODUCTION

A dichroic ("two-colored" in Greek) material has the property to absorb light differently depending on its polarization. X-ray circular dichroism is the difference between x-ray absorption spectra (XAS) obtained from left- and right-circularly polarized light, so it describes the dependence of the absorption cross section on the state of circularly polarized light.

In a magnetic sample, the breaking of time-reversal symmetry permits x-ray magnetic circular dichroism (XMCD). This is a powerful tool for studying the magnetic structure of complex systems as it gives element-specific information. Almost all synchrotron facilities around the world have a beamline dedicated to XMCD [1]. The existence of well-established magneto-optical sum rules that allow us to obtain the spin and orbital contribution to the magnetic moment directly from the integral of the spectra [2-4] made it an essential technique to study the magnetic properties of matter. These sum rules are widely and successfully applied at spin-orbit split  $L_{2,3}$  edges of transition metals [5–8] and  $M_{4,5}$  edges of actinides [9]. On the other hand, in the absence of spin-orbit splitting of the core state (like for the K or  $L_1$  edge), only the orbital magnetization sum rule [2,4] applies, and a quantitative analysis of the spectra is far from being straightforward. Yet, for 3d transition elements, measurements of XMCD at the K edge are the main way to probe magnetism under pressure, and XMCD is a widely used technique despite the interpretation difficulties [10–12].

X-ray natural circular dichroism (XNCD) occurs in noncentrosymmetric materials (for which the inversion symmetry is not a symmetry of the system). Up to now, it has been less widely used than XMCD, but it presents a fundamental interest as it gives access to element-specific stereochemical information [13]. In the domain of molecular magnetism, renewed interest in this technique has recently grown [14] with the emergence of new materials that are both chiral and magnetic. Contrary to optical activity to which a large number of mechanisms contribute [15], XNCD is largely dominated by a single contribution [13]. At  $L_1$  and K edges, XNCD exists only if p and d orbitals are mixed [16], yielding a unique measure of the mixing of even and odd orbitals.

The starting point of our work is a density functional theory (DFT)-based pseudopotential method. Using projector augmented-wave (PAW) reconstruction, the Lanczos algorithm, and a continued-fraction calculation [17–19], this method has proved to be successful for the calculation of absorption (XAS) spectra at the *K* edge [17,18,20,21]. The  $L_1$  edge, which corresponds to a 2*s* core hole, is expected to have the same behavior. In this paper, we propose the same kind of DFT-based approach for the calculation of XMCD and XNCD spectra in the near-edge region.

Several calculations of XMCD at the *K* edge in the near-edge region can be found in the literature. Most of these calculations are based on fully relativistic [22-27] or semirelativistic [28,29] multiple-scattering approaches with muffin-tin potentials even if efforts have been made to go beyond this approximation [30,31].

2469-9950/2017/96(8)/085123(12)

<sup>\*</sup>nadejda.bouldi@impmc.upmc.fr

The technique presented in this paper allows the use of a free-shape potential. Relativistic perturbations were taken into account both in the band structure [32] and in the photon-matter interaction [33]. The method has been implemented within a highly efficient reciprocal space code that allows the modeling of a large range of systems [19].

In Sec. II, the terms that enter the absorption cross section up to the electric quadrupole approximation are listed. Section III is dedicated to the presentation of the computational method. Results obtained for XAS and XNCD at the  $L_1$  edge of iodine in both enantiomers of  $\alpha$ -LiIO<sub>3</sub> and for *K*-edge XAS and XMCD spectra in 3*d* ferromagnetic metals are presented in Sec. IV. Finally, in Sec. V, the relativistic operator is examined in detail within the collinear spin approximation. Its corresponding sum rule is derived and evaluated numerically, and an expression that allows for a simple implementation of this term is given.

#### **II. CONTRIBUTIONS TO THE CROSS SECTION**

In the case of a fully circularly polarized light with a wave vector  $\mathbf{k}$  along *z*, the circular dichroism (XMCD and XNCD) cross section is written

$$\sigma^{\rm CD} = \sigma(\boldsymbol{\epsilon}_2) - \sigma(\boldsymbol{\epsilon}_1), \tag{1}$$

where  $\epsilon_2 = 1/\sqrt{2}(1,i,0)$ ,  $\epsilon_1 = \epsilon_2^* = 1/\sqrt{2}(1,-i,0)$ , and  $\sigma(\epsilon)$  is the XAS cross section of the material. The XMCD effect at the *K* edge of 3*d* transition elements results at most in an asymmetry in absorption of the order of  $10^{-3}$ . For this study, it is therefore important to compute the absorption cross section very accurately.

In a monoelectronic semirelativistic framework the contribution to the XAS cross section from a given core state of energy  $E_i$  is given by (see the Appendix)

$$\sigma = 4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle f | T | i \rangle|^2 \delta(E_f - E_i - \hbar \omega), \quad (2)$$

where  $\alpha_0$  is the fine-structure constant,  $|i\rangle$  is the twocomponent wave function that corresponds to the large components of the Dirac wave function of the core state, and the sum runs over unoccupied final states with energy  $E_f$ . The wave functions  $|f\rangle$  are eigenstates of the time-independent Foldy-Wouthuysen (FW) Hamiltonian of the electron in the presence of an electromagnetic field  $\mathbf{E}_0$ ,  $\mathbf{B}_0$  [34,35]:

$$H^{\rm FW} = mc^2 + \frac{\mathbf{p}^2}{2m} + eV - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}_0$$
$$- \frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma} \cdot (\mathbf{E}_0 \times \mathbf{p}) - \frac{e\hbar^2}{8m^2c^2}\boldsymbol{\nabla} \cdot \mathbf{E}_0.$$
(3)

Finally, *T* is the sum of three operators: (i) the electric dipole operator, (ii) the electric quadrupole operator, and (iii) a new light-matter interaction term that we named the spin-position operator (see the Appendix):

$$T = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \mathbf{k} \cdot \mathbf{r} + \frac{i\hbar\omega}{4mc^2} \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}), \qquad (4)$$

where  $\sigma$  is the vector of the Pauli matrices.

The absorption cross section expands in six terms, among which four terms are significant (see the orders of magnitude in the Appendix). The dominant term is the electric dipole-electric dipole (D-D) term:

 $\sigma$ 

$$\sigma_{\rm D-D} = 4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle f | \boldsymbol{\epsilon} \cdot \mathbf{r} | i \rangle|^2 \delta(E_f - E_i - \hbar \omega).$$
 (5)

It is usually the only term that is taken into account in calculations of XAS and XMCD spectra at the  $L_{2,3}$  edges and sometimes at the *K* edge [29,36–38].

The electric quadrupole–electric-quadrupole (Q-Q) term is

$$\chi_{Q-Q} = \pi^2 \alpha_0 \hbar \omega$$

$$\times \sum_f |\langle f | (\mathbf{k} \cdot \mathbf{r}) (\boldsymbol{\epsilon} \cdot \mathbf{r}) | i \rangle|^2 \delta(E_f - E_i - \hbar \omega). \quad (6)$$

At the *K* edge, it can reach a few percent of  $\sigma_{D-D}$  in amplitude. It contributes mainly to the preedge region. It is sometimes included in x-ray absorption calculations [17,39].

When neglecting spin-orbit coupling and in the absence of an external magnetic field, it is possible to choose real wave functions. In that case, the D-D and Q-Q terms verify  $\sigma(\epsilon) = \sigma(\epsilon^*)$ , which leads to a zero contribution to circular dichroism. For this reason it is crucial to account for relativistic effects in the wave functions calculation in order to compute XMCD.

On the other hand, the two following terms can give a nonvanishing contribution to the circular dichroism cross section even when wave functions can be chosen real.

The electric dipole-electric quadrupole cross term (D-Q) is

$$\sigma_{\mathrm{D-Q}} = -4\pi^2 \alpha_0 \hbar \omega \sum_f \mathrm{Im}[\langle f | (\mathbf{k} \cdot \mathbf{r}) (\boldsymbol{\epsilon} \cdot \mathbf{r}) | i \rangle \\ \times \langle i | \boldsymbol{\epsilon}^* \cdot \mathbf{r} | f \rangle] \delta(E_f - E_i - \hbar \omega).$$
(7)

If  $|i\rangle$  and  $|f\rangle$  are parity invariant (i.e., if inversion  $\mathbf{r} \rightarrow -\mathbf{r}$  is a symmetry of the system), then  $\sigma_{D-Q} = 0$ . It is, however, this term that is responsible for XNCD [16] because the electric dipole–magnetic dipole term (which is responsible for optical activity in the optical range) is very small in the x-ray range.

The cross term between the electric dipole and the relativistic operator that we call spin position (D-SP) is

$$\sigma_{\rm D-SP} = -\frac{2\pi^2 \alpha_0 \hbar^2 \omega^2}{mc^2} \sum_f {\rm Im}[\langle f | \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}) | i \rangle \\ \times \langle i | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | f \rangle] \delta(E_f - E_i - \hbar \omega).$$
(8)

It exists only in magnetic materials. Like the spin-orbit coupling term in the FW Hamiltonian, it arises from the coupling of the small components of the Dirac wave functions. To our knowledge, it has never been evaluated before. We will show in the following that, despite the small prefactor of this term, its contribution to XMCD at the *K* edge of 3*d* metals can account for up to one third of the XMCD intensity near the edge.

#### **III. METHOD**

In the framework of the final-state rule [40] the absorption cross section is obtained from one-electron wave functions. Within the frozen-core approximation, the 1s (*K*-edge) or 2s ( $L_1$ -edge) unperturbed core states  $|i\rangle$  can be determined from an all-electron isolated atom calculation. The stationary final states  $|f\rangle$  are calculated self-consistently in the presence of a core hole. Here, they are calculated

within a semirelativistic pseudopotential-based DFT and PAW reconstruction framework [32]. The absorption cross section is then calculated in a continued-fraction scheme using the Lanczos algorithm [17,18].

#### A. Collinear semirelativistic self-consistent field calculation

Self-consistent field calculations in this study are based on DFT with a plane-wave basis set and pseudopotentials as implemented in QUANTUM ESPRESSO [19] including the spin-orbit coupling (SOC) term [32]. Since an accurate implementation of SOC plays a crucial role for the evaluation of XMCD spectra, we briefly describe the underlying approach in the following.

In pseudopotential-based methods the potential near the nuclei is replaced by a fictitious smooth potential. The valence-electron wave functions are replaced by pseudo-wave-functions that are exempt from the rapid oscillations near the core. The size of the plane-wave basis set needed to describe the system is therefore considerably lowered, which leads to a much better computational efficiency compared to an all-electron approach, making an *ab initio* description of large systems with thousands of electrons possible.

In the PAW formalism, as described by Blöchl [41], the physical valence wave functions  $|\Psi\rangle$  can be reconstructed from the pseudo-wave-functions  $|\tilde{\Psi}\rangle$  as they are related through a linear operator  $\mathcal{T}: |\Psi\rangle = \mathcal{T}|\tilde{\Psi}\rangle$ , with

$$\mathcal{T} = \mathbb{1} + \sum_{\mathbf{R},n} (|\phi_{\mathbf{R},n}\rangle - |\tilde{\phi}_{\mathbf{R},n}\rangle) \langle \tilde{p}_{\mathbf{R},n}|.$$
(9)

In our case, the set of all-electron partial waves centered on atomic site **R**,  $|\phi_{\mathbf{R},n}\rangle$ , contains solutions of the Dirac equation for the isolated atom within a scalar relativistic approximation [42],  $|\tilde{\phi}_{\mathbf{R},n}\rangle$  are the corresponding pseudo partial waves, and  $\langle \tilde{p}_{\mathbf{R},n} |$  form a complete set of projector functions. The operator  $\mathcal{T}$  acts only in augmentation regions enclosing the atoms. Outside the augmentation regions the all-electron and pseudo-wave-functions coincide.

The pseudo-Hamiltonian is given by  $\mathcal{T}^{\dagger} H^{\text{FW}} \mathcal{T}$  [32,43]

$$\tilde{\mathcal{H}} = E_{\rm kin} + e\tilde{V}^{\rm loc}(\mathbf{r}) + \sum_{\mathbf{R}} e\tilde{V}_{\mathbf{R}}^{\rm nl} + \tilde{\mathcal{H}}_{\rm SO}, \qquad (10)$$

where  $E_{\rm kin}$  is the kinetic energy as implemented in QUANTUM ESPRESSO and  $\tilde{V}^{\rm loc}$  and  $\tilde{V}^{\rm nl}_{\bf R}$  are the local part and the nonlocal part in the separable form of the pseudopotentials.  $\tilde{\mathcal{H}}_{\rm SO}$  is the pseudo-Hamiltonian corresponding to the time-independent spin-orbit term in the Foldy-Wouthuysen transformed Hamiltonian [43]:

$$\tilde{\mathcal{H}}_{\rm SO} = \mathcal{T}^{\dagger} \left( \frac{e\hbar}{4m^2 c^2} \boldsymbol{\sigma} \cdot [\boldsymbol{\nabla} V(\mathbf{r}) \times \mathbf{p}] \right) \mathcal{T}$$
$$= \frac{e\hbar}{4m^2 c^2} \left( \boldsymbol{\sigma} \cdot [\boldsymbol{\nabla}^{\rm loc}(\mathbf{r}) \times \mathbf{p}] + \sum_{\mathbf{R}} F_{\mathbf{R}}^{\rm nl} \right). \quad (11)$$

 $F_{\mathbf{R}}^{\text{nl}}$  at the atomic site **R** are [44]

$$F_{\mathbf{R}}^{\mathrm{nl}} = \sum_{n,m} |\tilde{p}_{n}^{\mathbf{R}}\rangle \boldsymbol{\sigma} \cdot \left[ \left\langle \phi_{\mathbf{R},n} \middle| \boldsymbol{\nabla} v_{\mathbf{R}}(\mathbf{r}) \times \mathbf{p} \middle| \phi_{\mathbf{R},n} \right\rangle - \left\langle \tilde{\phi}_{\mathbf{R},n} \middle| \boldsymbol{\nabla} \tilde{v}_{\mathbf{R}}^{\mathrm{loc}}(\mathbf{r}) \times \mathbf{p} \middle| \tilde{\phi}_{\mathbf{R},n} \right\rangle \right] \left\langle \tilde{p}_{m}^{\mathbf{R}} \middle|,$$
(12)

where  $v_{\mathbf{R}}$  and  $\tilde{v}_{\mathbf{R}}^{\text{loc}}$  are the atomic all-electron and localchannel pseudopotentials, respectively. As these potentials are spherical,  $F_{\mathbf{R}}^{\text{nl}}$  is rewritten as

$$F_{\mathbf{R}}^{\mathrm{nl}} = \sum_{n,m} \left| \tilde{p}_{n}^{\mathbf{R}} \right\rangle \boldsymbol{\sigma} \cdot \left( \langle \phi_{\mathbf{R},n} | \frac{1}{r} \frac{\partial v_{\mathbf{R}}}{\partial r} \mathbf{L} | \phi_{\mathbf{R},n} \rangle - \langle \tilde{\phi}_{\mathbf{R},n} | \frac{1}{r} \frac{\partial \tilde{v}_{\mathbf{R}}^{\mathrm{loc}}}{\partial r} \mathbf{L} | \tilde{\phi}_{\mathbf{R},n} \rangle \right) \langle \tilde{p}_{m}^{\mathbf{R}} |.$$
(13)

The local potential  $\tilde{V}^{\text{loc}}(\mathbf{r}) = \sum_{\mathbf{R}} \tilde{v}_{\mathbf{R}}^{\text{loc}}(\mathbf{r})$ , and the quantity  $\frac{1}{r} \frac{\partial \tilde{v}_{\mathbf{R}}^{\text{loc}}}{\partial r}$  decreases in  $1/r^3$ , so that the action of the operator  $\tilde{V}^{\text{loc}}(\mathbf{r}) \times \mathbf{p}$  in the augmentation region is, at first order, the same as the action of  $\nabla \tilde{v}_{\mathbf{R}}^{\text{loc}}(\mathbf{r}) \times \mathbf{p}$ . In the PAW framework any pseudo-wave-function in the augmentation region can be expanded according to  $|\tilde{\Psi}\rangle = \sum_n |\tilde{\phi}_{n,\mathbf{R}}\rangle \langle \tilde{p}_n^{\mathbf{R}} |\tilde{\Psi}\rangle$ . Therefore, the term proportional to  $\tilde{v}_{\mathbf{R}}^{\text{loc}}$  and the term proportional to  $\tilde{V}^{\text{loc}}(\mathbf{r})$  partially compensate each other, so that the dominant contribution arises from the term

$$\frac{e\hbar}{4m^2c^2}\sum_{n\mathbf{R}m}\boldsymbol{\sigma}\cdot\left|\tilde{p}_n^{\mathbf{R}}\right\rangle\langle\phi_{\mathbf{R},n}|\frac{1}{r}\frac{\partial v_{\mathbf{R}}}{\partial r}\mathbf{L}|\phi_{\mathbf{R},m}\rangle\left\langle\tilde{p}_m^{\mathbf{R}}\right|.$$
(14)

In this study, we consider collinear spin along z, and only the z Pauli matrix is considered (diagonal spin-orbit-coupling approximation):

$$\boldsymbol{\sigma} = \sigma_z \mathbf{e}_z. \tag{15}$$

In XMCD experiments a magnetic field is usually applied parallel to the beam [45], which justifies considering the quantization axis parallel to  $\mathbf{k}$ .

This semirelativistic approach, which includes spin-orbit coupling in a two-component approach, is computationally less expensive than a fully relativistic one. It has been shown to reproduce the fully relativistic band structure [32]. For heavy atoms, the formula can be generalized by substituting  $\nabla \tilde{V}^{\text{loc}}$  and  $\frac{\partial \tilde{v}_{R}^{\text{loc}}}{\partial r}$  with reduced gradients, resulting in a zeroth-order regular approximation (ZORA) type of Hamiltonian [32].

In this study, the calculations have been performed using Troullier-Martins norm-conserving pseudopotentials and are based on the generalized gradient approximation (GGA) with Perdew-Burke-Ernzerhof (PBE) functionals [46]. The charge density is evaluated self-consistently in the presence of a core hole which is described by removing a 1s or 2s electron in the pseudopotential of the absorbing atom. A jellium background charge is added in order to ensure charge neutrality. A large unit cell (supercell) must be built to minimize the interactions between periodically reproduced core holes, and the k-point grid can be reduced accordingly.

#### B. Cross-section calculation

We implemented XMCD and XNCD in the XSPECTRA code [18] of QUANTUM ESPRESSO [19] distribution. The first results of this implementation for the terms D-D and Q-Q can be found in Ref. [47].

In the PAW formalism it has been shown [17,18] that the contribution of the operator O to the absorption cross section,

$$\sigma(\omega) = 4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle f | O | i \rangle|^2 \delta(E_f - E_i - \hbar \omega), \quad (16)$$

can be rewritten, as the initial wave function is localized around the absorbing atoms  $\mathbf{R}_0$ ,

$$\sigma(\omega) = 4\pi^2 \alpha_0 \hbar \omega \sum_f |\langle \tilde{f} | \tilde{\varphi}_{\mathbf{R}_0} \rangle|^2 \delta(E_f - E_i - \hbar \omega), \quad (17)$$

with

$$\left|\tilde{\varphi}_{\mathbf{R}_{0}}\right\rangle = \sum_{n} \left|\tilde{p}_{n}^{\mathbf{R}_{0}}\right\rangle \left\langle \phi_{n}^{\mathbf{R}_{0}}\right| O \left|i\right\rangle.$$
(18)

This sum involves, in principle, an infinite number of projectors, but experience demonstrated that two or three linearly independent projectors are, in general, sufficient in order to achieve the convergence of the D-D term at the *K* edge in the near-edge region [48].

The determination of all empty states in Eq. (17) would require a lot of computing resources and, as a consequence, would limit the size of the manageable supercell. To increase the efficiency of the method, the cross section is evaluated as developed in Refs. [17,18] via the Green's function using the Lanczos algorithm [49], which avoids the heavy workload of a large matrix inversion. The cross terms D-SP and D-Q are not in the form of Eq. (16), but they can be determined from two calculations of this type using the relationship

$$\operatorname{Im}[DB^{\star}] = \frac{1}{4}(|D+iB|^2 - |D-iB|^2), \quad (19)$$

where *B* is either the electric quadrupole or the spin-position operator and *D* is the electric dipole operator. For the term D-SP within the diagonal spin-orbit-coupling approximation, we have checked that this approach yields the same result as the computational-time-sparing calculation from the D-D spin-polarized contributions presented in Sec. V [Eq. (26)].

The calculated spectra are broadened with a Lorentzian function. Furthermore, the occupied states, which do not contribute to the absorption cross section, are cut according to the method described in Sec. III B of Ref. [29].

For the selected examples below, the different contributions to the cross sections for left- and right-circularly polarized light  $\sigma(\epsilon_2)$  and  $\sigma(\epsilon_1)$  were computed accurately in order to obtain circular dichroism.

#### **IV. APPLICATIONS**

#### A. Technical details

For LiIO<sub>3</sub>, the experimental structure is used [50]: the  $\Delta$  enantiomer of  $\alpha$ -LiIO<sub>3</sub> belongs to the hexagonal space group *P*6<sub>3</sub> with lattice parameters *a* = 5.48 Å and *c* = 5.17 Å. The atomic positions [15] are Li 2(a) (0,0,0.076), I 2(b) (1/3,2/3,0), and O 6(c) (0.247,0.342,0.838). The  $\Lambda$ enantiomer is the mirror image of the  $\Delta$  one (see Fig. 1), and it belongs to the same space group. A 2 × 2 × 2 supercell (80 atoms) is used so that the smallest distance between a core hole and its periodic image is 10.344 Å.  $\Gamma$ -centered *k*-point grids, 3 × 3 × 3 for the self-consistent charge-density calculation and 9 × 9 × 9 for the spectra calculation, are used. A constant Lorentzian broadening, with FWHM set to the core-hole lifetime broadening of 3.46 eV [51], is applied. As XNCD is a structural effect and not a magnetic effect, the calculation is not spin polarized.



FIG. 1. Top: hexagonal  $\alpha$ -LiIO<sub>3</sub> unit cell for (a)  $\Delta$  and (b)  $\Lambda$  enantiomers [15]. Bottom: top view of the cells [projection on (001)].

The XMCD calculations for the 3*d* ferromagnetic metals are carried out using the following experimental lattice parameters: a = 2.87 Å for bcc Fe, a = 3.52 Å for fcc Ni, and a = 2.51 Å and c = 4.07 Å for hcp Co. The number of atoms per supercell is 64 atoms for Fe and Ni and 96 atoms for Co, so the smallest distance between the periodically repeated core holes is 9.84 Å in Fe, 9.97 Å in Ni, and 10.03 Å in Co. A Methfessel-Paxton cold smearing of 0.14 eV (0.01 Ry) and a centered  $2 \times 2 \times 2k$ -point grid are used for the self-consistent charge density calculation. The spectra calculation is performed with a  $6 \times 6 \times 6$  grid for Fe and Co and an  $8 \times 8 \times 8$  grid for Ni. These calculations are performed with collinear spins along the easy axis of the crystals, that is to say, [001] for bcc Fe and hcp Co and [111] for fcc Ni [52], and the wave vector **k** is set along the same axis.

The spectra are convolved with a Lorentzian broadening function to simulate the effect of the finite lifetime of the core hole (constant in energy) and of the inelastic scattering of the photoelectron (additional energy-dependent broadening) for which we use the curves published by Müller *et al.* [53]. Experimental and calculated spectra are normalized such that the edge jump is equal to 1.

During the calculation of the spectra the origin of the energy is set to the Fermi energy of the material  $E_F$ . For the spectra to be compared with experiment, a rigid shift in energy is applied to the calculated spectra to make the maxima of the calculated XAS correspond to the maxima of the experimental spectra. The same shift is applied to the XMCD spectra. In the plots, the origin of energy  $E_0$  is therefore the one chosen in the publications from which the experimental spectra are extracted.

#### **B.** XNCD at the $L_1$ edge of I in $\alpha$ -LiIO<sub>3</sub>

Natural circular dichroism in the inorganic noncentrosymmetric lithium iodate (LiIO<sub>3</sub>) crystal was measured in 1998 [15], and it was attributed to the interference of electric dipole and electric quadrupole transitions [15,16]. Previous calculations [15,16,54] were indeed able to reproduce the overall peak positions and intensities in this framework. The agreement is, however, not entirely satisfactory for the



FIG. 2. (a) Comparison of experimental [15] and calculated XAS and XNCD spectra at the I  $L_1$  edge in LiIO<sub>3</sub> for both enantiomers with  $\mathbf{k} \parallel \mathbf{c}$ . The XNCD spectra arise from the D-Q term exclusively. Here, in the calculation as in the experiment  $\sigma^{CD} = \sigma^R - \sigma^L$ . (b) Calculated contributions to the XAS at the I  $L_1$  edge in LiIO<sub>3</sub>. The D-Q term was multiplied by 100, and the Q-Q term was multiplied by 500. (c) Angular dependence of the XNCD at the I  $L_1$  edge in LiIO<sub>3</sub>. Inset: XNCD amplitude as a function of the angle following the law  $3 \cos^2 \theta - 1$ , where  $\theta$  is the angle between  $\mathbf{c}$  and the incident wave vector  $\mathbf{k}$ .

absorption spectra. These discrepancies have been attributed to the use of muffin-tin potentials [54].

The approach presented in this work, which does not rely on the muffin-tin approximation, was applied to compute the XAS and XNCD spectra for  $\alpha$ -LiIO<sub>3</sub>. The absorption is dominated by the D-D term, as shown in Fig. 2(b). The XNCD spectra, on the other hand, are entirely due to the D-Q cross term.

As illustrated by Fig. 2(a), both the calculated XAS and XNCD spectra at the I  $L_1$  edge are in good agreement with experiment. However, the amplitude of the calculated XNCD is  $4 \times 10^{-2}$  compared to the edge jump, while the amplitude of the experimental spectra from Ref. [15] is  $6.5 \times 10^{-2}$ . Such an underestimation was also observed in Ref. [16] within a multiple-scattering approach.

From Fig. 2(a) (bottom), it becomes obvious that the XNCD spectra for both enantiomers are opposite. Indeed, changing an enantiomer for the other ( $\Delta \leftrightarrow \Lambda$ ) has the same effect for XNCD as changing the sign of the magnetic field ( $\mathbf{B} \leftrightarrow -\mathbf{B}$ ) for XMCD.

The angular dependence of the calculated XNCD spectra is depicted in Fig. 2(c), and its amplitude is plotted in the inset as a function of  $\theta$ , the angle between **k** and the **c** axis of the crystal. This amplitude varies as  $3\cos^2 \theta - 1$ , so it is maximal in the case where **k** is parallel to the **c** axis. This dependence is consistent with the formula derived in Ref. [16] for point group  $C_6$  (the point group of the space group of the crystal). Note that, as  $\epsilon$  is kept perpendicular to **k** and  $C_6$  is a dichroic point group [55], the XAS spectra also present an angular dependence. This does not prevent a comparison of the amplitude of the XNCD spectra because the edge jump remains unchanged.

#### C. XMCD at the *K* edge of 3*d* transition metals

XMCD was recorded at the Fe *K* edge in magnetized Fe in 1987 [56]. Ever since, a large number of calculations for the electric dipole term of the XMCD spectra on the Fe *K* edge in bcc Fe in the near-edge region have been reported (see, for example, Refs. [23,25–27,29,38,57,58]). Calculations of XMCD at the *K* edge in fcc Ni and hcp Co are fewer in number [26,57,59,60] and are not really conclusive.

These calculations have been performed with various methods, often within the electric dipole approximation and with muffin-tin potentials. Here, we present the calculation of the three terms (D-D, Q-Q, and D-SP) that are likely to contribute to the XMCD cross section at the K edge of ferromagnetic 3d transition metals, showing the relevance of the D-SP term.

The contribution of the D-SP term to the absorption cross section is not shown here because it is negligible. On the other hand, its contribution to the XMCD spectra (Fig. 3) is significant: it reaches 28% of the D-D term in amplitude. This can be understood considering the sum rules that are made explicit in the next section: in the XMCD cross section, the D-SP term probes the spin polarization of the *p* states, whereas the D-D term probes their orbital polarization. In Ref. [59] the 4p orbital magnetic moment in Co, Fe, and Ni is evaluated to a few  $10^{-4}\mu_B$  (Fe:  $5 \times 10^{-4}\mu_B$ , Co:  $16 \times 10^{-4}\mu_B$ , Ni:  $6 \times 10^{-4} \mu_B$ ), and in Ref. [61] the 4p spin magnetic moment in Fe and Co is evaluated to several  $10^{-2} \mu_B$  (Fe: 5 ×  $10^{-2} \mu_B$ , Co:  $6 \times 10^{-2} \mu_B$ ) in the opposite direction. This difference in order of magnitude of both quantities compensates for the smallness of the prefactor  $(\hbar\omega/4mc^2)$  of the D-SP term (see Table I in the Appendix).



FIG. 3. Calculated contributions to the *K*-edge XMCD spectra in the ferromagnetic 3*d* metals Fe, Co, and Ni.

To check possible numerical problems, we also performed the calculations using the FDMNES code [39], where, for this purpose, the D-SP term was introduced in the same way. This code follows Wood and Boring [64] to eliminate the small component and obtain a couple of Schrödinger-like equations, including the spin-orbit effect, closely akin to but improving the Pauli equation. Despite the very different approach (no pseudopotential, calculation in real space and no diagonal spin-orbit-coupling approximation), we found very similar results for both the shape and relative amplitude of the D-SP contribution.

The agreement with the experimental spectra is fair, as illustrated in Fig. 4. As usual in independent-particle calculations, the energy axis is slightly compressed [65–67] due to the energy dependence of the real part of the self-energy [68] for which corrections to the calculated spectra could be applied [69]. Alternatively, the position of the calculated peaks could be improved by phenomenological rescaling [65,68].

For Fe, the main peaks of the experimental XMCD are reproduced by the calculation. As in calculations by others [27], the positive peak at 10 eV is overestimated, probably due to the approximate description of the exchange-correlation energy. Indeed, the comparison between the spectra calculated with PBE and local-density approximation (LDA) functionals (Fig. 5) shows that this peak would be even more enhanced with LDA.

For Ni and Co, a main negative peak is present near the main rising edge in the calculation as in the experiment, but the satellite peaks that appear in the calculation are difficult to link to the experiment.

In these calculations, the polarization rate of the light is taken to be 100%, and a single crystal with full 3d spin polarization is considered. In Fe, Ni, and Co, saturation is reached with the usual experimentally applied magnetic field, and the anisotropy is quite weak, so that the rate of circular

polarization of the light  $P_c$  is expected to account for most of the discrepancy in amplitude between the calculated and experimental XMCD spectra. The data for Fe and Co were recorded in a 5-T magnet at 5 K and within a setup that reaches a 90% circular polarization rate [70]. The correction to the amplitude of the calculated spectra to fit the experimental condition should therefore be of the order of 0.9. Here, it is approximately 0.6 in the case of Fe and 1.0 in the case of Co. The data for Ni were recorded at ambient temperature in a 0.7-T magnet within a dispersive setup with a diamond quarter-wave plate for which we can infer that  $P_c \approx 0.7$  [71]. However, no correction to the amplitude of the calculated spectra is needed to make it correspond to the amplitude of the experimental spectra. Thus, whereas our calculation overestimates the amplitude of the XMCD spectra in the case of Fe, it underestimates it in the case of Ni.

#### V. CONTRIBUTION OF THE D-SP TERM TO XMCD: THE CASE OF COLLINEAR SPINS

#### A. The SP operator

In this section, we study the *spin-position* operator SP( $\epsilon$ ) =  $\boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r})$ . We consider collinear spins along *z* with independent spin channels. The spin part of the wave functions  $|s\rangle$  can either be the spin-up spinor  $\binom{0}{0}$  or the spin-down spinor  $\binom{0}{1}$ .

The D-SP term is the cross term between the electric dipole and the spin-position operator. Spin does not appear in the electric dipole operator, so it is diagonal in spin:

$$\langle \phi_i s | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | \phi_f s' \rangle = \langle \phi_i | \boldsymbol{\epsilon}^{\star} \cdot \mathbf{r} | \phi_f \rangle \delta_{ss'}.$$
(20)

This imposes s' = s. On the other hand, the vector of Pauli matrices  $\sigma$  appears explicitly in the spin-position operator:

$$\langle \phi_i s | \boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}) | \phi_f s \rangle = \langle \phi_i | (\boldsymbol{\epsilon} \times \mathbf{r}) | \phi_f \rangle \cdot \langle s | \boldsymbol{\sigma} | s \rangle.$$
(21)

As  $\langle s | \sigma_x | s \rangle = \langle s | \sigma_y | s \rangle = 0$ , we can exclude *a priori* the terms that are proportional to  $\sigma_x$  and  $\sigma_y$  in the spin-position operator. In that case the spin-position operator is rewritten as

$$SP_{col}(\boldsymbol{\epsilon}) = \sigma_{z}(\boldsymbol{\epsilon}_{x}y - \boldsymbol{\epsilon}_{y}x)$$
  
=  $\sigma_{z}\frac{4i\pi}{3}r[Y_{1}^{-1}(\boldsymbol{\epsilon})Y_{1}^{1}(\mathbf{u}_{r}) - Y_{1}^{1}(\boldsymbol{\epsilon})Y_{1}^{-1}(\mathbf{u}_{r})].$ 
(22)

Its selection rules are almost the same as those for the electric dipole [72] one:  $\Delta l = \pm 1$ ,  $\Delta m = \pm 1$ .

As 
$$Y_1^{-1}(\epsilon_1) = 0, Y_1^{-1}(\epsilon_2) = \sqrt{3/4\pi}, Y_1^0(\epsilon_1) = Y_1^0(\epsilon_2) = 0, Y_1^1(\epsilon_1) = -\sqrt{3/4\pi}, \text{ and } Y_1^1(\epsilon_2) = 0,$$

$$SP_{col}(\boldsymbol{\epsilon}_1) = i\sqrt{\frac{4\pi}{3}}rY_1^{-1}(\mathbf{u}_r)\sigma_z = \sigma_z i\boldsymbol{\epsilon}_1 \cdot \mathbf{r}, \qquad (23)$$

$$SP_{col}(\boldsymbol{\epsilon}_2) = i\sqrt{\frac{4\pi}{3}}rY_1^1(\mathbf{u}_r)\sigma_z = -\sigma_z i\boldsymbol{\epsilon}_2 \cdot \mathbf{r}.$$
 (24)

Hence,

$$\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}_1) = -\frac{\hbar\omega}{2mc^2} [\sigma_{\mathrm{D-D}}^{\uparrow}(\boldsymbol{\epsilon}_1) - \sigma_{\mathrm{D-D}}^{\downarrow}(\boldsymbol{\epsilon}_1)], \qquad (25)$$

$$\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}_2) = \frac{\hbar\omega}{2mc^2} [\sigma_{\mathrm{D-D}}^{\uparrow}(\boldsymbol{\epsilon}_2) - \sigma_{\mathrm{D-D}}^{\downarrow}(\boldsymbol{\epsilon}_2)], \qquad (26)$$



FIG. 4. Comparison between the experimental XAS and XMCD spectra for Fe, Co [62,63], and Ni [60] and the total calculated spectra. The wave vector and the magnetization axis were set to the easy axis of the crystals, that is to say, [001] for bcc Fe and hcp Co and [111] for fcc Ni [52].

with

$$\sigma_{\rm D-D}^{s}(\boldsymbol{\epsilon}) = 4\pi^{2}\alpha_{0}\hbar\omega$$
$$\times \sum_{f} |\langle f^{s}|\boldsymbol{\epsilon}\cdot\mathbf{r}|i^{s}\rangle|^{2}\delta(E_{f}-E_{i}-\hbar\omega), \quad (27)$$

where  $s = \uparrow$  or  $\downarrow$ . Therefore, in the diagonal spin-orbitcoupling collinear spin case, the D-SP term can be computed from the D-D cross section for the up and down spin channels.

#### **B.** Sum rule at the *K* edge

A sum rule is a formula in which the integral of the circular dichroism spectra due to a given term of the cross section

is expressed as a function of the ground-state expectation value of some operator. The sum rules at  $L_{2,3}$  edges are well established [61,73] and are widely used to extract quantitative magnetic ground-state properties. Their derivation is based on several approximations, including the fact that the radial integrals are spin and energy independent [74]. At the *K* edge the sum rule for the electric dipole-electric dipole term [3,4,59,75] relates the integral of the XMCD spectra to the orbital magnetic moment of occupied *p* states that is proportional to  $\langle L_z \rangle_p$ . This sum rule is, however, almost impossible to apply in practice because the upper limit of the integral is not well defined and, in the case of 3*d* transition elements, the 4*p* states are almost unoccupied, so  $\langle L_z \rangle_p$  is very small and has a minor impact on the magnetic moment



FIG. 5. Total calculated bcc Fe *K*-edge XMCD spectra without a core hole using PBE and LDA functionals (all other technical parameters are identical). Here, the broadening was taken to be constant (0.8 eV) along the whole energy range.

of the material. Deriving a similar sum rule for the D-SP term is nevertheless very useful to understand why, despite its very small prefactor, this term is so large in XMCD. We derive it following the method of Thole *et al.* [2,4] with many-body wave functions and operators assuming that all spins are collinear and within the diagonal spin-orbit-coupling approximation.

In a many-body framework, using the expression for SP in terms of spherical harmonics (22),

$$\sigma_{\mathrm{D-SP}}(\boldsymbol{\epsilon}) = \frac{2\pi^2 \hbar^2 \alpha_0 \omega^2}{mc^2} \sum_{\nu=-1}^{1} \mathrm{Re} \{ Y_1^{-\nu}(\boldsymbol{\epsilon}^{\star}) \times [Y_1^{1}(\boldsymbol{\epsilon}) \zeta_{\mathrm{D-SP}}^{1\nu} - Y_1^{-1}(\boldsymbol{\epsilon}) \zeta_{\mathrm{D-SP}}^{-1\nu}] \}, \quad (28)$$

with

$$\zeta_{\rm D-SP}^{\lambda\nu} = (-1)^{\nu} \left(\frac{4\pi}{3}\right)^2 \sum_{f} \langle f | \sum_{i} \left[\sigma_{z_i} r_i Y_1^{\lambda} \left(\mathbf{u}_{\mathbf{r}_i}\right)\right]^{\star} | g \rangle \\ \times \langle g | \sum_{i} r_i Y_1^{\nu} \left(\mathbf{u}_{\mathbf{r}_i}\right) | f \rangle \delta(E_f - E_g - \hbar \omega).$$
(29)

In a second quantized form with l, m, and  $\sigma$  being the usual quantum numbers [16],

$$\langle g | \sum_{i} r_{i} Y_{1}^{\lambda} (\mathbf{u}_{\mathbf{r}_{i}}) | f \rangle$$

$$= \sum_{lm\sigma l_{0}m_{0}\sigma'_{0}} \sqrt{\frac{3(2l+1)}{4\pi (2l_{0}+1)}}$$

$$\times (10l0|l_{0}0)(1\lambda lm|l_{0}m_{0}) \langle g | a^{\dagger}_{l_{0}m_{0}\sigma} a_{lm\sigma} | f \rangle \mathcal{D}_{l_{0},l}, \quad (30)$$

where  $\mathcal{D}_{l_0,l} = \int dr r^3 R_{l_0}^{\star}(r) R_l(r)$  is assumed (as usual in sumrule derivations) to be spin independent. The experimental procedure enables us to obtain the signal corresponding to a specific  $l_0$ . At the *K* edge  $l_0 = 0$  and  $m_0 = 0$ , so that

$$\langle g|\sum_{i}r_{i}Y_{1}^{\nu}(\mathbf{u}_{\mathbf{r}_{i}})|f\rangle = \sqrt{\frac{1}{4\pi}}\sum_{\sigma}(-1)^{\nu}\langle g|a_{00\sigma}^{\dagger}a_{l-\nu\sigma}|f\rangle\mathcal{D},$$
(31)

where  $\mathcal{D} = \mathcal{D}_{0,1}$ .

Similarly, as 
$$\langle \sigma_0' | \sigma_z | \sigma' \rangle = \sigma' \delta_{\sigma_0',\sigma'},$$
  
 $\langle g | \sum_i r_i Y_1^{\lambda} (\mathbf{u}_{\mathbf{r}_i}) \sigma_{zi} | f \rangle = \sum_{\sigma'} \sigma' (-1)^{\lambda} \langle g | a_{00\sigma'}^{\dagger} a_{l-\lambda\sigma'} | f \rangle \mathcal{D}.$ 
(32)

Using the completeness relation  $\int dE \sum_{f} |f\rangle \langle f|\delta(E_{f} - E_{g} - E) = 1 - |g\rangle \langle g|$ , as the core shell is full and under the assumption that the radial integral  $\mathcal{D}$  does not depend on energy,

$$\int dE \,\zeta_{\rm D-SP}^{\lambda\nu} = \frac{4\pi}{9} \sum_{\sigma} (-1)^{\lambda} \sigma \,\langle g | a_{1-\nu\sigma} a_{1-\lambda\sigma}^{\dagger} | g \rangle |\mathcal{D}|^2.$$
(33)

The combination of Eqs. (33) and (28) leads to

$$\int d\hbar\omega \frac{\sigma_{\text{D-SP}}(\boldsymbol{\epsilon}_2)}{(\hbar\omega)^2} = \frac{\pm 2\pi^2 \alpha_0}{3mc^2} |\mathcal{D}|^2 \langle g|a_{1\pm1\uparrow}a_{1\pm1\uparrow}^{\dagger} - a_{1\pm1\downarrow}a_{1\pm1\downarrow}^{\dagger}|g\rangle. \quad (34)$$

The difference between the two integrals yields the XMCD sum rule for the D-SP term:

$$\int d\hbar\omega \, \frac{\sigma_{\rm D-SP}^{\rm XMCD}}{(\hbar\omega)^2} = -\frac{2\pi^2 \alpha_0}{3mc^2} \langle S z_{l=1}^{1,-1} \rangle |\mathcal{D}|^2, \qquad (35)$$

with the operator

$$Sz_{l=1}^{1,-1} = \sum_{m=-1,1} a_{1m\downarrow}^{\dagger} a_{1m\downarrow} - a_{1m\uparrow}^{\dagger} a_{1m\uparrow}$$
 (36)

corresponding to a partial spin polarization of the occupied p states.

If one considers the derivative of this sum rule, we see that the electric dipole–spin-position circular dichroism signal probes the spin polarization of the empty p states. Figure 6 illustrates the correspondence between both quantities. This proves the validity of the D-SP sum rule. Unfortunately, this sum rule cannot be applied directly to experimental spectra, mainly because of the superposition of the D-D contribution to the D-SP contribution.

#### VI. CONCLUSION

We have developed an efficient computational approach to determine accurate XMCD and XNCD spectra. The main result is that the contribution from the relativistic term D-SP in the transition operator is significant in XMCD spectra despite being negligible in XAS. This importance is explained by the fact that this term probes the spin of the p states that is two orders of magnitude larger than its orbital counterpart.

For XNCD, the calculated spectra are in good agreement with experiment, and the angular dependence corresponds to the expected one.

A big advantage of the method employed in this paper to perform XMCD and XNCD calculations is its large adaptability, which opens opportunities for applications to several kinds of systems such as strongly correlated materials or molecules absorbed on functionalized surfaces. The same method could be applied to compute x-ray magnetochiral



FIG. 6. Comparison between the calculated D-SP spectra without a core hole for Fe, Co, and Ni and the calculated projected densities  $Sz_{l=1}^{1,-1}(E)$ .  $Sz_{l=1}^{1,-1}(E)$  has been multiplied by the factor between the *p* density of states and the dipole XAS spectra times  $R = \frac{\hbar\omega}{2mc^2}$  in accordance with the sum rule (35).

dichroism (XM $\chi$ D) that has been observed in magnetized chiral systems [14]. The features of XM $\chi$ D differ from those of XMCD and XNCD, making it a promising probe of the interplay between chirality and magnetism.

#### ACKNOWLEDGMENTS

This work was supported by French state funds managed by the ANR within the Investissements d'Avenir program under Reference No. ANR-11-IDEX-0004-02 and more specifically within the framework of the Cluster of Excellence MATISSE led by Sorbonne Universités. We are grateful to D. Cabaret for very interesting and constructive feedback on this work. We also thank F. Baudelet and L. Nataf for providing reference spectra for the energy scaling. U.G. and N.J.V. acknowledge support from DFG (FOR 1405). The numerical calculations were performed using HPC resources from the Paderborn Center for Parallel Computing (PC<sup>2</sup>) and from GENCI-IDRIS (Grant No. 2016-100172).

#### APPENDIX: SEMIRELATIVISTIC TRANSFORMATION OF THE RELATIVISTIC CROSS SECTION

We start with the expression for the cross section in a relativistic framework [33], and we adapt it to the specific need of our numerical method that is the determination of large components of the Dirac wave function for the core state and of Foldy-Wouthuysen (FW) wave functions for the valence states.

#### PHYSICAL REVIEW B 96, 085123 (2017)

#### 1. Relativistic cross section

The contribution to the x-ray absorption (XAS) cross section from a given four-component Dirac core state  $|\Psi_i\rangle$  of energy  $E_i$  is given by [33]

$$\sigma(\hbar\omega) = 4\pi^2 \alpha_0 \hbar\omega \sum_f |\langle \Psi_f | \mathscr{T} | \Psi_i \rangle|^2 \delta(E_f - E_i - \hbar\omega),$$
(A1)

where the sum runs over unoccupied final states  $|\Psi_f\rangle$  with energy  $E_f, \alpha_0$  is the fine-structure constant, and  $\mathscr{T}$  is the transition operator, defined as

$$\mathscr{T} = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \mathbf{k} \cdot \mathbf{r} - \frac{\hbar c}{2\omega} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \boldsymbol{\alpha}), \qquad (A2)$$

where the polarization vector  $\boldsymbol{\epsilon}$ , the wave vector  $\mathbf{k}$ , and the energy  $\hbar\omega$  describe the incident electromagnetic wave;  $\mathbf{r}$  is the position operator; and  $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$  is the vector of Dirac matrices.

Here, as in our numerical calculations, a one-electron scheme is used. In a many-body framework the formula for the cross section would be the same but with *N*-electron wave functions and many-body operators that are written as sums over electrons.

In Ref. [33], the transformation into a two-component representation for  $|\Psi_i\rangle$  and  $|\Psi_f\rangle$  was performed by applying a time-independent FW transformation at order  $c^{-2}$ . The FW transformation of  $\Psi_l$  is obtained by applying a unitary operator:  $\psi_l^{\text{FW}} = U_{\text{FW}}\Psi_l$ , with [76]

$$U_{\rm FW} = 1 + \frac{\beta}{2mc^2}\mathcal{O} - \frac{1}{8m^2c^4}\mathcal{O}^2,$$
 (A3)

where  $\beta$  is the standard Dirac matrix. In this expression,  $\mathcal{O}$  is the odd operator entering the Dirac Hamiltonian:  $\mathcal{H}^{D} = \beta mc^{2} + \mathcal{O} + \mathcal{E}$ , where  $\mathcal{E}$  is even. It is defined as  $\mathcal{O} = c\boldsymbol{\alpha} \cdot (\mathbf{p} - e\mathbf{A}_{0})$ , where  $\mathbf{p}$  is the momentum operator and  $\mathbf{A}_{0}$  is the static vector potential.

Only the large components of  $\psi_l^{\text{FW}}$ , denoted  $\phi_l^{\text{FW}}$ , are nonzero up to order  $c^{-2}$ . The cross section can be written as a function of the large components of  $\psi_i^{\text{FW}}$  and  $\psi_f^{\text{FW}}$  [33]:

$$\sigma = 4\pi^2 \alpha_0 \hbar \omega \sum_f \left| \left\langle \phi_f^{\text{FW}} \middle| T_{\text{FW}} \middle| \phi_i^{\text{FW}} \right\rangle \right|^2 \delta(E_f - E_i - \hbar \omega).$$
(A4)

The operator  $T_{\rm FW}$  is the projection on the upper components of  $U_{\rm FW} \mathscr{T} U_{\rm FW}^{\dagger}$ :

$$T_{\rm FW} = T_{\rm D} + T_{\rm Q} + T_{\rm MD} + T_{A_0} + T_{\rm SP},$$
 (A5)

where

$$T_{\rm D} = \boldsymbol{\epsilon} \cdot \mathbf{r} \tag{A6}$$

and

$$T_{\rm Q} = \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r} \mathbf{k} \cdot \mathbf{r} \tag{A7}$$

are the standard *electric dipole* and *electric quadrupole* operators.

The magnetic dipole operator  $T_{MD}$  is written as

$$T_{\rm MD} = \frac{1}{2m\omega} (\mathbf{k} \times \boldsymbol{\epsilon}) \cdot (\hbar \boldsymbol{\sigma} + \mathbf{L}), \qquad (A8)$$

where  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$  and  $\boldsymbol{\sigma}$  is the vector of the Pauli matrices.  $T_{\text{MD}}$  is proportional to the total magnetic moment operator ( $\hbar \boldsymbol{\sigma} + \mathbf{L}$ ) = (2 $\mathbf{S} + \mathbf{L}$ ), where  $\mathbf{S}$  is the spin operator.  $T_{\text{MD}}$  is also present in common nonrelativistic derivations [55,77]. Another use of the name "magnetic dipole" is for the spinquadrupole coupling term, usually denoted  $T_z$ , that appears in the spin sum rule for XMCD at spin-orbit split edges. It is not the quantity that is discussed here. The selection rules of  $T_{\text{MD}}$ are  $l_i = l_f$  and  $n_i = n_f$  [55], so  $T_{\text{MD}}$  vanishes in the x-ray energy range because the states involved in the transition have different principal quantum numbers.

The correction to this term due to the static vector potential  $\mathbf{A}_0$  is

$$T_{A_0} = -\frac{e}{2m\omega} (\mathbf{k} \times \boldsymbol{\epsilon}) \cdot (\mathbf{r} \times \mathbf{A}_0). \tag{A9}$$

The last term in Eq. (A5) is present only when relativistic effects are included in the calculation of the transition operator:

$$T_{\rm SP} = -\frac{\hbar}{4m^2c^2}(\mathbf{p} - e\mathbf{A}_0) \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\sigma}). \tag{A10}$$

A similar term was already found in Ref. [47] but was derived from a semirelativistic Hamiltonian, and this approach presents a conflict with time-dependent perturbation theory [33]. It can be rewritten noticing that, in the nonrelativistic limit,  $|\phi_i^{FW}\rangle$  and  $|\phi_f^{FW}\rangle$  are eigenstates of

$$H_0^0 = \frac{(\mathbf{p} - e\mathbf{A}_0)^2}{2m} + eV(\mathbf{r}) - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B}_0, \qquad (A11)$$

where  $\mathbf{B}_0$  is the static external magnetic field. This Hamiltonian obeys  $\mathbf{p} - e\mathbf{A}_0 = (m/i\hbar)[\mathbf{r}, H_0^0]$ , so that

$$-\frac{\hbar}{4m^2c^2} \langle \phi_f^{\rm FW} | (\mathbf{p} - e\mathbf{A}_0) \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\sigma}) | \phi_i^{\rm FW} \rangle$$
  
$$= \frac{i}{4mc^2} (E_i - E_f) \langle \phi_f^{\rm FW} | \mathbf{r} \cdot (\boldsymbol{\epsilon} \times \boldsymbol{\sigma}) | \phi_i^{\rm FW} \rangle$$
  
$$= \frac{i\hbar\omega}{4mc^2} \langle \phi_f^{\rm FW} | (\boldsymbol{\epsilon} \times \mathbf{r}) \cdot \boldsymbol{\sigma} | \phi_i^{\rm FW} \rangle.$$

We name  $\boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r})$  the *spin-position* operator and define the associated transition operator:

$$T_{\rm SP} = \frac{i\hbar\omega}{4mc^2}\boldsymbol{\sigma} \cdot (\boldsymbol{\epsilon} \times \mathbf{r}). \tag{A12}$$

For technical reasons, in the present paper we consider a different situation than the one in Ref. [33]: we use the FW wave function for the final states and large components of the Dirac wave function for the initial (core) state. This difference in treatment is linked to the fact that the core wave function is determined from a relativistic atomic code, whereas the unoccupied states are calculated with a semirelativistic condensed-matter code.

#### 2. Rewriting the cross section with large components of the Dirac wave function for the core state

We denote  $\phi_i$  and  $\chi_i$  the large and small components of  $\Psi_i$ , respectively. The order of magnitude of the ratio between the small and large components is v/c, where v is the velocity of the particle [78]. Up to order  $c^{-1}$ , the small component is

written as [78,79]

$$\chi_i = \frac{1}{2mc} \boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{A}_0)\phi_i.$$
(A13)

Only the second term in  $U_{\rm FW}$  [Eq. (A3)] couples the small and large components. From Eqs. (A13) and (A3), the large component of the FW transformed wave function can be expressed as a function of the large components of the Dirac wave function up to order  $c^{-2}$ ,

$$\phi_i^{\text{FW}} = \left(1 - \frac{1}{8m^2c^4} [\mathcal{O}^2]_p\right) \phi_i + \frac{1}{4mc^3} \mathcal{O}_p \boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{A}_0) \phi_i.$$
(A14)

 $[\mathcal{O}^2]_p$  is the projection of  $\mathcal{O}^2$  onto large components [78]:

$$[\mathcal{O}^2]_p = c^2 (\mathbf{p} - e\mathbf{A}_0)^2 - c^2 e\hbar\boldsymbol{\sigma} \cdot \mathbf{B}_0$$
  
=  $2mc^2 [H_0^0 - eV(\mathbf{r})],$ 

and  $\mathcal{O}_p = c\boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{A}_0)$  is the projection of  $\beta \mathcal{O}$  on the top right components.

The identity  $c\mathcal{O}_p\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}_0)=[\mathcal{O}^2]_p$  leads to

$$\phi_i^{\text{FW}} = \left(1 + \frac{1}{8m^2c^4} [\mathcal{O}^2]_p\right) \phi_i.$$
 (A15)

From this relation, the cross section of Eq. (A4) can be adapted to the case that we consider here.

In Ref. [33] the expansion was made to order  $1/c^2$  for the dipole contribution and to order kr for multipole contributions. At the same order,

$$\sigma = 4\pi^2 \alpha_0 \hbar \omega \sum_f \left| \left\langle \phi_f^{\rm FW} \middle| T_{\rm FW}' \middle| \phi_i \right\rangle \right|^2 \delta(E_f - E_i - \hbar \omega),$$
(A16)

where  $T'_{\rm FW}$  is

$$T'_{\rm FW} = T_{\rm FW} \left( 1 + \frac{1}{8m^2c^4} [\mathcal{O}^2]_p \right)$$
  
=  $T_{\rm FW} + T^e$ .

There is one extra term in the cross section compared to  $T_{FW}$  that is related to the use of large components of the Dirac wave function instead of the Foldy-Wouthuysen wave function for the core state:

$$T^{e} = \frac{1}{4mc^{2}} \left[ \boldsymbol{\epsilon} \cdot \mathbf{r} H_{0}^{0} - \boldsymbol{\epsilon} \boldsymbol{\epsilon} \cdot \mathbf{r} V(\mathbf{r}) \right].$$
(A17)

We show in the next section that it is negligible for the core states considered in this work. As the magnetic dipole term is negligible in the x-ray range,  $T_{\rm FW}$  thus contains four operators [see Eq. (A5) and the subsequent comments], so  $T'_{\rm FW}$  is written as

$$T'_{\rm FW} = T_{\rm D} + T_{\rm Q} + T_{a_0} + T_{\rm SP} + T^e.$$
 (A18)

#### 3. Order of magnitude of the operators

As the core wave function is very localized, we obtain an idea of the relative order of magnitude of the operators in Eq. (A18) by evaluating them at the radius corresponding to the core state. In Table I these evaluations are given compared to the electric dipole operator.

#### X-RAY MAGNETIC AND NATURAL CIRCULAR DICHROISM ...

TABLE I. Order of magnitude of the operators in Eq. (A18) evaluated at the core-state radius  $r_c$  compared to the electric dipole operator. The mean radius of the core orbitals is deduced from the effective nuclear charge:  $r_c = \frac{3}{2} \frac{a_0}{Z_{\text{eff}}}$  [80,81].  $B_0$  has been fixed to  $2 \times 10^4$  T (1.2 eV), which is two orders of magnitude larger than the exchange splitting observed for the Fe K edge. The Coulomb potential is  $V = \frac{-Z_{\text{eff}}e}{4\pi\epsilon_0 r_c}$ , and the core-state energy  $E_i$  is evaluated in a planetary model,  $E_i = \frac{-Z_{\text{eff}}e^2}{8\pi\epsilon_0 r_c}$ .

|   |   | Edge   |   |  |  |  |   |
|---|---|--|---|--|--|--|---|
|   |   | $L_1$  | L_2   |  |  | K  |   |
|   | Order of Magnitude  | I  | Fe  | Gd   | Bi   | 0  | Fe  |
| Energy (keV)  |   | 5.19   | 0.72  | 7.898  | 15.71  | 0.53   | 7.11  |
| $Z_{\rm eff}$   |   | 39.067   | 22.089  | 29.8527  | 39.2335  | 7.6579   | 25.381  |
| $T_{\rm Q}$ (A7)<br>$T_{\rm SP}$ (A12)<br>$T_{A_0}$ (A9)<br>$T^e$ (A17) | $kr_c/2$<br>$\hbar\omega/4mc^2$<br>$ekr_c B_0/4m\omega$<br>$(E_i - eV)/4mc^2$ | $\begin{array}{c} 2.7\times10^{-2}\\ 2.6\times10^{-3}\\ 6.0\times10^{-6}\\ 6.7\times10^{-3} \end{array}$ | $\begin{array}{c} 6.6 \times 10^{-3} \\ 3.5 \times 10^{-4} \\ 1.1 \times 10^{-5} \\ 2.2 \times 10^{-3} \end{array}$ | $\begin{array}{l} 5.3\times10^{-2}\\ 3.9\times10^{-3}\\ 7.8\times10^{-6}\\ 3.9\times10^{-3} \end{array}$ | $\begin{array}{l} 8.1\times 10^{-2}\\ 7.7\times 10^{-3}\\ 6.0\times 10^{-6}\\ 6.8\times 10^{-3} \end{array}$ | $\begin{array}{c} 1.4\times 10^{-2}\\ 2.6\times 10^{-4}\\ 3.0\times 10^{-5}\\ 2.6\times 10^{-4} \end{array}$ | $5.7 \times 10^{-2} \\ 3.5 \times 10^{-3} \\ 9.2 \times 10^{-6} \\ 2.8 \times 10^{-3} \\ \end{array}$ |

When expanding the square modulus of the matrix elements in Eq. (A16), we keep the terms with contributions higher than  $10^{-3}$  compared to the dominant electric dipole term. We also neglect the term  $T^e$ : as  $V(\mathbf{r})$  is almost spherical at the core state radius, it concerns transitions to the same orbitals as the electric dipole term. It does not contain a spin operator, so that, even in XMCD, it yields only a negligible correction to the electric dipole contribution. Therefore, we are left with the four significant terms, D-D, Q-Q, D-Q, and D-SP, discussed in Sec. II.

- [1] A. Rogalev and F. Wilhelm, Phys. Met. Metallogr. 116, 1285 (2015).
- [2] B. T. Thole, P. Carra, F. Sette, and G. van der Laan, Phys. Rev. Lett. 68, 1943 (1992).
- [3] P. Carra, B. T. Thole, M. Altarelli, and X. Wang, Phys. Rev. Lett. 70, 694 (1993).
- [4] M. Altarelli, Phys. Rev. B 47, 597 (1993).
- [5] J. Vogel, A. Fontaine, V. Cros, F. Petroff, J.-P. Kappler, G. Krill, A. Rogalev, and J. Goulon, Phys. Rev. B 55, 3663 (1997).
- [6] J. Stöhr, J. Magn. Magn. Mater. 200, 470 (1999).
- [7] K. W. Edmonds, N. R. S. Farley, T. K. Johal, G. van der Laan, R. P. Campion, B. L. Gallagher, and C. T. Foxon, Phys. Rev. B 71, 064418 (2005).
- [8] Y. Prado, M.-A. Arrio, F. Volatron, E. Otero, C. Cartierdit-Moulin, P. Sainctavit, L. Catala, and T. Mallah, Chem. Eur. J. 19, 6685 (2013).
- [9] F. Wilhelm, R. Eloirdi, J. Rusz, R. Springell, E. Colineau, J.-C. Griveau, P. M. Oppeneer, R. Caciuffo, A. Rogalev, and G. H. Lander, Phys. Rev. B 88, 024424 (2013).
- [10] F. Baudelet, S. Pascarelli, O. Mathon, J. P. Iti, A. Polian, M. d'Astuto, and J. C. Chervin, J. Phys. Condens. Matter 17, S957 (2005).
- [11] R. Torchio, A. Monza, F. Baudelet, S. Pascarelli, O. Mathon, E. Pugh, D. Antonangeli, and J. P. Itié, Phys. Rev. B 84, 060403 (2011).
- [12] J.-D. Cafun, J. Lejeune, J.-P. Itié, F. Baudelet, and A. Bleuzen, J. Phys. Chem. C 117, 19645 (2013).
- [13] A. Rogalev, J. Goulon, F. Wilhelm, and A. Bosak, in *Magnetism and Synchrotron Radiation*, edited by E. Beaurepaire, H. Bulou, F. Scheurer, and J.-P. Kappler, Springer Proceedings in Physics Vol. 133 (Springer, Berlin, 2010), pp. 169–190.
- [14] R. Sessoli, M.-E. Boulon, A. Caneschi, M. Mannini, L. Poggini,
   F. Wilhelm, and A. Rogalev, Nat. Phys. 11, 69 (2015).

- [15] J. Goulon, C. Goulon-Ginet, A. Rogalev, V. Gotte, C. Malgrange, C. Brouder, and C. R. Natoli, J. Chem. Phys. **108**, 6394 (1998).
- [16] C. R. Natoli, C. Brouder, P. Sainctavit, J. Goulon, C. Goulon-Ginet, and A. Rogalev, Eur. Phys. J. B 4, 1 (1998).
- [17] M. Taillefumier, D. Cabaret, A.-M. Flank, and F. Mauri, Phys. Rev. B 66, 195107 (2002).
- [18] C. Gougoussis, M. Calandra, A. P. Seitsonen, and F. Mauri, Phys. Rev. B 80, 075102 (2009).
- [19] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys. Condens. Matter **21**, 395502 (2009).
- [20] D. Cabaret, A. Bordage, A. Juhin, M. Arfaoui, and E. Gaudry, Phys. Chem. Chem. Phys. 12, 5619 (2010).
- [21] A. Bordage, C. Brouder, E. Balan, D. Cabaret, A. Juhin, M.-A. Arrio, P. Sainctavit, G. Calas, and P. Glatzel, Am. Mineral. 95, 1161 (2010).
- [22] H. Ebert, P. Strange, and B. L. Gyorffy, Z. Phys. B 73, 77 (1988).
- [23] H. Ebert, Solid State Commun. 100, 677 (1996).
- [24] S. Stähler, G. Schütz, and H. Ebert, Phys. Rev. B 47, 818 (1993).
- [25] H. J. Gotsis and P. Strange, J. Phys. Condens. Matter 6, 1409 (1994).
- [26] G. Y. Guo, J. Phys. Condens. Matter 8, L747 (1996).
- [27] O. Šipr and H. Ebert, Phys. Rev. B 72, 134406 (2005).
- [28] C. Brouder and M. Hikam, Phys. Rev. B 43, 3809 (1991).
- [29] C. Brouder, M. Alouani, and K. H. Bennemann, Phys. Rev. B 54, 7334 (1996).
- [30] C. R. Natoli, M. Benfatto, and S. Doniach, Phys. Rev. A 34, 4682 (1986).
- [31] Y. Joly, O. Bunău, J. E. Lorenzo, R. M. Galra, S. Grenier, and B. Thompson, J. Phys. Conf. Ser. **190**, 012007 (2009).

- [32] U. Gerstmann, N. J. Vollmers, A. Lücke, M. Babilon, and W. G. Schmidt, Phys. Rev. B 89, 165431 (2014).
- [33] N. Bouldi and C. Brouder, arXiv:1610:05900.
- [34] C. Itzykson and J. B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).
- [35] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
- [36] J. J. Rehr and R. C. Albers, Rev. Mod. Phys. 72, 621 (2000).
- [37] C. R. Natoli, D. K. Misemer, S. Doniach, and F. W. Kutzler, Phys. Rev. A 22, 1104 (1980).
- [38] T. Fujikawa and S. Nagamatsu, J. Electron Spectrosc. Relat. Phenom. 129, 55 (2003).
- [39] O. Bunău and Y. Joly, J. Phys. Condens. Matter 21, 345501 (2009).
- [40] U. von Barth and G. Grossmann, Phys. Rev. B 25, 5150 (1982).
- [41] P. E. Blöchl, Phys. Rev. B 50, 17953 (1994).
- [42] D. D. Koelling and B. N. Harmon, J. Phys. C 10, 3107 (1977).
- [43] D. Ceresoli, U. Gerstmann, A. P. Seitsonen, and F. Mauri, Phys. Rev. B 81, 060409 (2010).
- [44] C. J. Pickard and F. Mauri, Phys. Rev. Lett. 88, 086403 (2002).
- [45] A. Rogalev, F. Wilhelm, N. Jaouen, J. Goulon, and J.-P. Kappler, X-ray Magnetic Circular Dichroism: Historical Perspective and Recent Highlights (Springer, Berlin, 2006), pp. 71–93.
- [46] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996).
- [47] C. Gougoussis, Ph.D. thesis, Université Pierre et Marie Curie– Paris 6, 2009.
- [48] O. Bunău and M. Calandra, Phys. Rev. B 87, 205105 (2013).
- [49] C. Lanczos, J. Res. Natl. Bur. Stand. 49, 33 (1952).
- [50] C. Svensson, J. Albertsson, R. Liminga, Å. Kvick, and S. C. Abrahams, J. Chem. Phys. 78, 7343 (1983).
- [51] Unoccupied Electronic States, edited by J. C. Fuggle and J. E. Inglesfield, Topics in Applied Physics Vol. 69 (Springer, Berlin, 1992).
- [52] R. O'Handley, *Modern Magnetic Materials: Principles and Applications* (Wiley, New York, 1999).
- [53] J. E. Müller, O. Jepsen, and J. W. Wilkins, Solid State Commun. 42, 365 (1982).
- [54] A. L. Ankudinov and J. J. Rehr, Phys. Rev. B 62, 2437 (2000).
- [55] C. Brouder, J. Phys. Condens. Matter 2, 701 (1990).
- [56] G. Schütz, W. Wagner, W. Wilhelm, P. Kienle, R. Zeller, R. Frahm, and G. Materlik, Phys. Rev. Lett. 58, 737 (1987).
- [57] J. I. Igarashi and K. Hirai, Phys. Rev. B 50, 17820 (1994).
- [58] A. Dixit and M. Alouani, Comput. Phys. Commun. 207, 136 (2016).

#### PHYSICAL REVIEW B 96, 085123 (2017)

- [59] J. I. Igarashi and K. Hirai, Phys. Rev. B 53, 6442 (1996).
- [60] R. Torchio, Y. O. Kvashnin, S. Pascarelli, O. Mathon, C. Marini, L. Genovese, P. Bruno, G. Garbarino, A. Dewaele, F. Occelli *et al.*, Phys. Rev. Lett. **107**, 237202 (2011).
- [61] C. T. Chen, Y. U. Idzerda, H.-J. Lin, N. V. Smith, G. Meigs, E. Chaban, G. H. Ho, E. Pellegrin, and F. Sette, Phys. Rev. Lett. 75, 152 (1995).
- [62] M. A. Laguna-Marco, C. Piquer, and J. Chaboy, Phys. Rev. B 80, 144419 (2009).
- [63] J. Wong, *Reference X-ray Spectra of Metal Foils*, EXAFS Materials, Inc., 871 El Cerro Blvd (Danville, CA, 1999).
- [64] J. H. Wood and A. M. Boring, Phys. Rev. B 18, 2701 (1978).
- [65] G. Materlik, J. E. Müller, and J. W. Wilkins, Phys. Rev. Lett. 50, 267 (1983).
- [66] L. A. Grunes, Phys. Rev. B 27, 2111 (1983).
- [67] B. Lengeler and R. Zeller, J. Less Common Met. 103, 337 (1984).
- [68] J. Mustre de Leon, J. J. Rehr, S. I. Zabinsky, and R. C. Albers, Phys. Rev. B 44, 4146 (1991).
- [69] J. J. Kas, J. Vinson, N. Trcera, D. Cabaret, E. L. Shirley, and J. J. Rehr, J. Phys. Conf. Ser. **190**, 012009 (2009).
- [70] N. Kawamura, N. Ishimatsu, and H. Maruyama, J. Synchrotron Rad. 16, 730 (2009).
- [71] C. Giles, C. Malgrange, J. Goulon, F. de Bergevin, C. Vettier, A. Fontaine, E. Dartyge, and S. Pizzini, Nucl. Instrum. Methods Phys. Res., Sect. A 349, 622 (1994).
- [72] D. Sébilleau, X-ray and Electron Spectroscopies: An Introduction (Springer, Berlin, 2006), pp. 15–57.
- [73] R. Wu, D. Wang, and A. J. Freeman, Phys. Rev. Lett. 71, 3581 (1993).
- [74] M. Altarelli, Nuovo Cimento Soc. Ital. Fis. D 20, 1067 (1998).
- [75] A. Ankudinov and J. J. Rehr, Phys. Rev. B 51, 1282 (1995).
- [76] E. Eriksen, Phys. Rev. 111, 1011 (1958).
- [77] S. Di Matteo, Y. Joly, and C. R. Natoli, Phys. Rev. B 72, 144406 (2005).
- [78] P. Strange, *Relativistic Quantum Mechanics* (Cambridge University Press, Cambridge, 1998).
- [79] E. van Lenthe, E. J. Baerends, and J. G. Snijders, J. Chem. Phys. 105, 2373 (1996).
- [80] E. Clementi and D. L. Raimondi, J. Chem. Phys. 38, 2686 (1963).
- [81] E. Clementi, D. L. Raimondi, and W. P. Reinhardt, J. Chem. Phys. 47, 1300 (1967).

# 1s2p Resonant Inelastic X-ray Scattering Magnetic Circular Dichroism as a probe for the local and non-local orbitals in CrO<sub>2</sub>

Patric Zimmermann<sup>\*</sup>, Nadejda Bouldi<sup>+,†</sup>, Myrtille O.J.Y. Hunault<sup>\*</sup>, Marcin Sikora<sup>‡</sup>, James M. Ablett<sup>+</sup>, Jean-Pascal Rueff<sup>+,⊕</sup>, Blair Lebert<sup>+,†</sup>, Philippe Sainctavit<sup>+,†</sup>, Frank M.F. de Groot<sup>\*</sup> and Amélie Juhin<sup>†</sup>

\*Debye Institute of Nanomaterial Science, Utrecht University, 3584 CA Utrecht, The Netherlands

<sup>+</sup>Synchrotron SOLEIL, L'Orme des Merisiers, BP48 Saint-Aubin, 91192 Gif-sur-Yvette, France

<sup>†</sup>Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC), Sorbonne Universités, UMR CNRS 7590, UPMC Univ Paris 06, Muséum National d'Histoire Naturelle, IRD UMR206, 4 Place Jussieu, F-75005 Paris, France

<sup>⊕</sup>Laboratoire de Chimie Physique-Matière et Rayonnement, Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7614, F-75005 Paris, France.

<sup>‡</sup>Academic Centre for Materials and Nanotechnology, AGH University of Science and Technology, 30-059 Kraków, Poland

 $Contacts: \ P.Zimmermann@uu.nl, \ F.M.F.deGroot@uu.nl$ 

RIXS-MCD, magnetic circular dichroism, non-local, 4p3d hybridisation, CrO<sub>2</sub>, distortion

August 21, 2017

We have determined the magnetic ground state of the half-metal  $CrO_2$  based on 1s2p Resonant Inelastic X-ray Scattering Magnetic Circular Dichroism (RIXS-MCD) experiments. The two-dimensional RIXS-MCD map displays the 1s X-ray absorption spectrum combined with the 1s2p X-ray emission decay, where there is a large MCD contrast in the final state involving the 2p core hole.

Our measurements show that the Cr K pre-edge structure is dominated by dipolar contributions and the quadrupole peak is invisible in direct K pre-edge absorption. Using RIXS-MCD, we reveal that the quadrupole 1s3d pre-edge has a large MCD contrast, which appears at lower energy with respect to the K pre-edge maximum.

We use crystal field multiplet calculations to model the excitonic RIXS-MCD spectral shape in tetragonal  $(D_{4h})$  symmetry. The RIXS-MCD is strongly sensitive to the ground state distortion of the Cr<sup>4+</sup> sites. The calculations of the RIXS-MCD maps suggest that the 3d spin-orbit interaction is fully quenched ( $\zeta_{3d} = 0 \text{ meV}$ ) and the ground state electron configuration must contain a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) contribution, which is required to explain the appearance of the Magnetic Circular Dichroism (MCD) in the Cr K pre-edge. This is in apparent contrast with the compressed tetragonal distortion.

## 1. Introduction

The detection of the X-ray magnetic circular dichroism (XMCD) has become a powerful tool for the element-specific study of the magnetic properties of complex systems. The MCD of 3d transition metal ions is usually studied at the spin-orbit split  $L_{2,3}$  absorption edges (2p  $\rightarrow$  3d) to allow the determination of the spin and orbital magnetic moments using the sum rules.<sup>15,44</sup>

The  $L_{2,3}$ -edges of 3d transition metals are in the soft X-ray range requiring vacuum conditions, implying that they are difficult for liquid or high-pressure cells. This limits the number of possible applications and the nature of the samples. The energy of the K-edge of 3d transition metals lies in the hard X-ray range, but the direct K-edge MCD signal is weak and the absence of spin-orbit splitting a priori prohibits a quantitative analysis using the spin sum rules.

Some of the above limitations can be addressed with the novel RIXS-MCD approach, in which one combines XMCD and *resonant inelastic X-ray scattering* (RIXS) at the K pre-edge of 3d transition metals according to the following two-step-model (FIG. 1):

$$1s^{2} 2p^{6} 3d^{N} \xrightarrow{0}_{XAS} 1s^{1} 2p^{6} 3d^{N+1} \xrightarrow{}_{XES} 1s^{2} 2p^{5} 3d^{N+1}$$

The excitation step (X-ray Absorption Spectrum, XAS) is performed with circular polarised light, being either left (lcp,  $\circlearrowleft$ ) or right circular polarised (rcp,  $\circlearrowright$ ). The detection of the subsequent X-ray emission spectrum (XES) can in principle also be polarisation dependent, however, for this study no polarisation analyser was used in the XES channel.



Figure 1: Atomic term scheme for the 1s2p RIXS-MCD photon-in photon-out process for a  $3d^2$  electron configuration.  $E_{in}$  and  $E_{out}$  denote the energies of the incoming and outgoing light, and  $E_T$  the energy transferred to the system.

RIXS measurements are established to study the electronic structure, while RIXS-MCD additionally offers magnetic information with the advantage that it enhances the contrast of resonant features. The first 1s2p RIXS-MCD experiments have shown that the RIXS-MCD signal of iron in magnetite, can be of the same order of magnitude as  $L_{2,3}$ -edge XMCD.<sup>27,44,51,52</sup> The RIXS-MCD approach can thus be considered as a high resolution magnetic spectroscopy, while hard X-rays yield bulk sensitivity.

Chromium dioxide (CrO<sub>2</sub>) is a half-metallic ferromagnet ( $T_{\text{Curie}} \approx 390 \text{ K}$ ) which means that one spin channel is conductive while the other one is insulating; in other words the electrons in the occupied Cr(3d) bands in CrO<sub>2</sub> show nearly 100% spin polarisation.<sup>35,40,53</sup> This makes it a promising candidate for future applications in the field of spintronics<sup>40</sup>, for example as a source for spin-polarised currents, magnetic tunnel junctions or other magneto-electronic devices that require a large spin polarisation.<sup>60</sup> The ferromagnetism is usually explained in terms of a specific double-exchange mechanism<sup>30,46</sup> related to the existence of the combination of a strongly localised state just below the Fermi level and another dispersed band at higher energy. As Schlottmann states Hund's rule couples the spin of the localised electron with that of the itinerant electrons and the hopping becomes correlated.<sup>46</sup> This implies a strong correlation between the spins of the localised and non-localised electrons.<sup>46</sup> The metallicity is due to the dispersed bands that hybridise with the O(2p) bands and cross the Fermi level. This mechanism and the importance of local and non-local correlations is a subject of active research.<sup>29</sup>

In spite of the large number of studies, the electronic structure that induces both ferromagnetism and metallicity in  $CrO_2$  remains to be understood and the origin of the half-metallic ferromagnetism is highly nontrivial.<sup>54</sup>

It is expected that RIXS-MCD probes specifically the local magnetic contribution yielding valuable information on the complex electronic structure that leads to metallic ferromagnetism in  $CrO_2$ . In RIXS-MCD, transitions involving localised d states are expected to give an enhanced intensity with respect to conventional XMCD, while those involving delocalised states are not.<sup>52</sup>

RIXS measurements are established to study the electronic structure, while RIXS-MCD additionally offers magnetic information with the advantage that local features appear on resonance with enhanced contrast. The bulk sensitivity of 1s2p RIXS-MCD addresses the common problem of a reduced surface ( $Cr^{4+} \rightarrow Cr^{3+}$ ) in surface sensitive measurements in the soft X-ray range, as in L<sub>2,3</sub>-edge XMCD<sup>2,8,17,18,22,23,28,40</sup> or L-edge RIXS.<sup>32</sup>

## 2. Technical details

#### 2.1. Experimental setup and measurements

The measured sample is a commercially available crystalline  $\text{CrO}_2$  powder (MAGTRIEVE<sup>TM</sup>) with a density of  $\rho = 4.85 \frac{\text{g}}{\text{cm}^3}$  at 25°C and a grain size of 44  $\mu$ m (mesh 325).<sup>50</sup> The powder was milled by hand with mortar and pestle and then pressed into a pill.

The measurements were performed at ambient conditions ( $T \approx 25^{\circ}$ C,  $p \approx 1$  bar) at the GALAX-IES inelastic-scattering beamline at the SOLEIL synchrotron radiation facility.<sup>45</sup> The synchrotron radiation was monochromatised using a Si(111) nitrogen-cooled fixed-exit double-crystal monochromator (DCM) with  $\Delta E/E \approx 1.4 \cdot 10^{-4}$ , followed by a Pd-coated spherical collimating mirror. The X-rays were then focused to a spot-size of 30  $\mu$ m (vertical) by 90  $\mu$ m (horizontal) full width at half maximum (FWHM) at the sample position by a 3:1 focusing toroidal Pd-coated mirror. A vertical Rowland circle geometry was implemented using a Ge(422) spherical-bent analyser crystal (R = 1m) which was used to energy-select and focus the emitted X-rays onto a silicon avalanche photodiode detector.

The overall resolution was found to be FWHM  $\approx 0.74 \,\text{eV}$  or  $E/\Delta E \approx 7700$  by measuring the quasi-elastic line at 5.4 keV, corresponding to the energy of the Cr K<sub> $\alpha$ </sub> fluorescence line. A diamond quarter-wave plate in (111)-orientation with a thickness of  $d = 500 \,\mu\text{m}$ , located immediately after the DCM, was used in order to select between left ( $\circlearrowleft$ ) and right circular polarised ( $\circlearrowright$ ) light. An electromagnet created a magnetic flux density  $\vec{B}$  of up to  $|\vec{B}| \approx 0.7 \,\text{T}$  on the sample.

The setup was aligned in longitudinal geometry<sup>31</sup> with  $\vec{k}||\vec{B}||\vec{z}|$  as shown in FIG. 2. The angles of the sample and the analyser with respect to the incident beam were  $\theta = 45^{\circ}$  and  $\theta = 90^{\circ}$  respectively.



Figure 2: Scheme of the experimental setup with the sample in orange between the two poles of the electromagnet. The circularly polarised incident beam in blue, magnetic field in red and the detected X-ray emission in green. The detector, an *avalanche photo diode* (APD), is not shown for clarity.

Each measurement was performed with alternating polarisation (lcp  $\leftrightarrow$  rcp) for each data point. All spectra were acquired as incident energy scans with the spectrometer fixed to detect a given emission energy  $E_{out}$ . For the spectra in FIG. 7 the emitted photons corresponding to the Cr K $\alpha_1$ emission at  $E_{out} = 5415.3 \,\text{eV}$  were detected. For the RIXS maps the detected emission energies were varied accordingly between  $E_{out} = 5406 \,\text{eV}$  and  $E_{out} = 5423 \,\text{eV}$ . After completion of a spectrum the direction of the magnetic field was reversed and the measurement repeated. This yields two inverse spectra which are added to minimise noise and reduce systematic errors.

The two-dimensional RIXS maps are displayed either in an emitted energy view with the intensity  $I(E_{in}, E_{out})$ , or in an energy transfer view with the intensity  $I(E_{in}, E_T)$ . The incident energy  $E_{in}$  is in both cases the horizontal x axis, and the y axis is either the emitted photon energy  $E_{out}$  or the energy transfer  $E_T = E_{in} - E_{out}$ . The incident energy  $E_{in}$  of the experimental spectra was calibrated against a CrO<sub>2</sub> reference<sup>38</sup>. The energies of the emitted photons  $E_{out}$  and energy transfer  $E_T$  were calibrated with literature values from the X-ray Data Booklet (http://xdb.lbl.gov/). The intensities of all spectra are normalised with respect to the sum maximum and given in arbitrary units (a.u.).

## 2.2. Calculations of the densities of states

The density of states (DOS) has been calculated using density-functional theory (DFT) as implemented in Quantum-Espresso<sup>16</sup>, i.e. using a plane-wave basis set, pseudo-potentials and periodic boundary conditions. We used Troullier-Martins<sup>61</sup> norm-conserving pseudo-potentials and the formulation of Perdew-Burke-Ernzerhof<sup>41</sup> (GGA) for the exchange and correlation density functional. The electric dipole and quadrupole contributions to the absorption cross-section are calculated with Xspectra.<sup>19,56</sup>

To interpret the experimental spectra, a 1s core hole is added in the calculation. It is described within a static approximation by including a core-hole in the pseudo-potential of the absorbing Cr atom. The Hubbard U correction, whose relevance in the case of  $CrO_2$  is discussed in Refs<sup>6,30,36,58</sup>, was not included in the calculation.

The lattice parameters and the dimensionless internal coordinate defining the positions of the atoms were set to their experimental values (a = b = 4.421Å, c = 2.916Å<sup>18,35</sup>,  $u = 0.303^{33}$ ). The self-consistent charge density calculation was performed using a centered  $6 \times 6 \times 9$  k-points grid

for the conventional cell that contains one Cr atom and two O atoms. The non-self-consistent calculation is performed with a  $12 \times 12 \times 18$  k-point grid and the projected density of states are obtained by projecting wave functions onto orthogonalised atomic wave functions. A  $2 \times 2 \times 3$  supercell is used to minimise interactions between neighbouring core-holes upon application of periodic boundary conditions and the k-point grid is reduced accordingly.

#### 2.3. Crystal-field multiplet calculations

The quadrupole contribution to the pre-edge structure cannot be accurately reproduced with the present DFT-based calculations due to the strong 3d3d correlation effects that determine the pre-edge spectral shape and also the X-ray emission matrix elements. Instead, a crystal-field multiplet (CFM) calculation of the  $1s^2 3d^2 \rightarrow 1s^1 3d^3$  quadrupole transition with subsequent dipole decay  $1s^1 3d^3 \rightarrow 1s^2 2p^5 3d^3$  is applied.

The multiplet calculations are based on the absorption and emission matrices created with the CTM4XAS program.<sup>55</sup> It takes into account all the 3d-3d, 1s-3d and 2p-3d electronic Coulomb interactions, as well as the spin-orbit coupling  $\zeta$  on every open shell of the absorbing atom and treats the geometrical environment through a crystal-field potential. All calculations are performed using the C<sub>4</sub> point group symmetry to take into account the presence of a magnetic field  $(\vec{B}||\vec{k})$ . The RIXS process is modelled with the Kramers-Heisenberg relation<sup>44</sup> using additional scripts written for this purpose.

The atomic Slater coefficients for the 3d3d Coulomb interaction  $F_{\rm dd}$  are scaled to 65% of the Hartree-Fock values and the 2p3d Coulomb  $F_{\rm pd}$  and the exchange interactions  $(G_{\rm pd}, G_{\rm sd})$  are scaled to 52%. This reduction of the Slater integrals is a result of the expansion of the wave function due to charge transfer effects and the actual values used are an empirical result as they show the best results presented here. The atomic values for the spin-orbit interaction are in the ground state  $\zeta_{3d} = 41 \text{ meV}$ , in the intermediate state  $\zeta_{3d} = 54 \text{ meV}$  and in the final state  $\zeta_{2p} = 5.668 \text{ eV}$  and  $\zeta_{3d} = 53 \text{ meV}$ .

For CrO<sub>2</sub>, there are some reported values for the crystal-field parameter 10Dq (Ikeno: 2.28 eV<sup>25</sup>, Lewis: 2.5 eV<sup>33</sup>), but we are not aware of any reports for the tetragonal distortion parameters Ds and Dt. Based on the magnitude of the spatial distortion we estimate the two distortion parameters, Ds and Dt, to be of the order of tens of millielectronvolt (meV). The crystal-field parameters (Dq,Ds,Dt) have been varied across a significant section of the parameter space. Only a few representative maps have been chosen to illustrate the general appearance for a given ground state. The displayed calculations refer to the crystal-field splitting parameter 10Dq = 2.347 eV and the distortion parameters Ds = -0.036 eV, Dt = -0.007 eV for  ${}^{3}\text{E}_{g}$  and Dt = -0.2 eV for the  ${}^{3}\text{B}_{2g}$ case. The molecular field M reflecting the interatomic exchange interactions is set to M = 30 meV, being estimated with the Curie temperature  $T_{c} \approx 390 \text{ K}$ .

The following Lorentzian broadenings are applied for the intermediate state (IS)  $L_{IS} = 1.2 \text{ eV}$ and the final state (FS)  $L_{FS} = 0.6 \text{ eV}$ . Here  $L_{FS}$  is an intermediate value between the two natural broadenings for the  $2p_{3/2}$  and  $2p_{1/2}$  final states, corresponding to the  $K\alpha_1$  and  $K\alpha_2$  emission respectively. As discussed elsewhere<sup>11,14</sup> the lifetime broadening of the  $2p_{1/2}$  shell ( $K\alpha_2$  line,  $L_2$ edge) can be up to five times larger with respect to the  $2p_{3/2}$  shell ( $K\alpha_1$  line,  $L_3$  edge). The experimental (Gaussian) broadening  $G_{IS} = 0.7 \text{ eV}$  is set to the experimentally acquired FWHM of the quasi-elastic scattering peak. For the emission, a resolution of  $G_{FS} = 0.4 \text{ eV}$  is used. All broadenings are given as full width at half maximum.

Finally, the energy calibration for the theoretical maps is not absolute. The calculated spectra were shifted in both directions for the best agreement with the experimental MCD.

## 3. Theory

#### 3.1. Crystal-field multiplet theory

The CrO<sub>2</sub> crystal has a rutile structure and belongs to the spacegroup P4<sub>2</sub>/mnm (136).<sup>35,53</sup> With the metal ion as the inversion center (inversion symmetry) local pd-mixing is forbidden, but mixing between different sites is possible. The Cr<sup>4+</sup> ions occupy the six-fold oxygen-coordinated sites with Wyckoff position 2*a* corresponding to D<sub>2*h*</sub> point group symmetry.<sup>1,18,35,40,53,59,60</sup> In the ground state, Cr<sup>4+</sup> in CrO<sub>2</sub> has a high spin (S = 1) 3d<sup>2</sup> electron configuration,<sup>8</sup> which corresponds to the multi-electronic ground state <sup>3</sup>F<sup>+</sup> in spherical symmetry (O<sub>3</sub>) as derived with Hund's rules. The splitting of the atomic multi-electronic state by the crystal-field, through the successive branchings O<sub>3</sub>  $\rightarrow$  O<sub>h</sub>  $\rightarrow$  D<sub>4h</sub>  $\rightarrow$  D<sub>2h</sub> is illustrated in FIG. 3.



Figure 3: Energy splittings of the multi-electronic state  ${}^{3}F^{+}$  for  $Cr^{4+}$  (3d<sup>2</sup> configuration in highspin S = 1), for each symmetry for  $O_{3} \rightarrow O_{h} \rightarrow D_{4h} \rightarrow D_{2h}$  (local exchange is included, but no spin-orbit or magnetic field is taken into account). The order of the energy levels shown above is not fixed. It can be altered in dependency on the specific crystal-field parameters chosen.

In  $O_h$  symmetry only one crystal-field parameter Dq (or 10Dq) is needed.  $D_{4h}$  symmetry requires in total three crystal-field parameters (Dq,Ds,Dt), and in  $D_{2h}$  symmetry five splitting parameters  $(Dq,Ds,Dt,D_{\alpha},D_{\beta})$  are required. In FIG. 4, we illustrate in the single-electron picture the splitting of the atomic  $3d^2$  orbital into the partially filled and spin-polarised  $t_{2g}$ , and the empty  $e_g$  level in  $O_h$  symmetry.



Figure 4: Crystal field splitting of the  $3d^2$  high-spin multi-electronic configuration  ${}^{3}F^{+}$  (O<sub>3</sub>), as derived with Hund's rules in spherical symmetry (O<sub>3</sub>), into the multi-electronic  ${}^{3}T_{1g}$  (O<sub>h</sub>) ground state that is dominated by the  $|t_{2g}^2 e_g^0\rangle$  single-electron configuration in octahedral symmetry (O<sub>h</sub>).

For a tetravalent Cr cation (Cr<sup>4+</sup>) the crystal-field splitting energy 10Dq separating the  $t_{2g}$  and the  $e_g$  states is estimated (0.6 eV per valency) to be approximately  $10Dq \approx 2.4 \text{ eV}$ , in agreement with other reported values (Ikeno:  $2.28 \text{ eV}^{25}$ , Lewis:  $2.5 \text{ eV}^{33}$ ). The other parameters introduced by lower point group symmetries describe further the splittings of the mono-electronic levels as the consequence from a distortion of the metal ion site.

#### 3.2. Importance of the distortion

For  $CrO_2$  there are publications dating back to the 1970s<sup>42</sup> reporting a *compression* of the octahedron. The metal-ligand distances in the equatorial plane  $d_e$  and the apical distances  $d_a$  on the local z-axis can be directly acquired from X-ray diffraction (XRD) data. The published values are summarised in Table 1.

| Reference                  | $d_e$ | $d_a$ | $d_e - d_a$ |
|----------------------------|-------|-------|-------------|
| Porta et $al^{42}$         | 1.910 | 1.891 | 0.019       |
| Deng et al <sup>13</sup>   | 1.914 | 1.911 | 0.003       |
| Burdett et al <sup>5</sup> | 1.911 | 1.888 | 0.023       |
| Baur et $al^3$             | 1.917 | 1.882 | 0.035       |
| Average                    | 1.913 | 1.893 | 0.020       |

Table 1: Metal-ligand distances in the equatorial plane  $d_e$  and apex direction  $d_a$  for CrO<sub>2</sub> from various sources (all values in angstrom  $(10^{-10} \text{ m})$  rounded to the last given digit).

Though the difference  $d_e - d_a$  varies between 0.3 pm (0.15%) and 3 pm (1.5%) all reports in Table 1 confirm an axial compression. We note that XRD measurements, due to relying on Bragg diffraction, are always an average across many coordination spheres, while the crystal-field model used here is strictly local.

As already pointed out by Korotin et al.<sup>30</sup>, the compression of the Cr octahedron induces the further splitting of the  $t_{2g}$  orbitals.<sup>33,47</sup> As shown in FIG. 3, a longitudinal compression of the z-axis results in the lower point group symmetry  $D_{4h}$  and splits the orbital triplet  ${}^{3}T_{1g}$  ( $O_{h}$ ) ground state into  ${}^{3}E_{g}$  and  ${}^{3}B_{2g}$  ( $D_{4h}$ ).



Figure 5: Distortions for a  $3d^2$  configuration yielding the transition from octahedral (O<sub>h</sub>) to tetragonal (D<sub>4h</sub>) symmetry. This changes the multi-electronic  ${}^{3}T_{1g}$  (O<sub>h</sub>) ground state into an orbital doublet  ${}^{3}E_{g}$  (D<sub>4h</sub>) in case of a compression, and it yields the orbital singlet  ${}^{3}B_{2g}$ (D<sub>4h</sub>) as the ground state for an axial elongation (T = 0K, no spin-orbit interaction). For the assignment of the energys see equations (1) to (4).

In the single particle picture as illustrated in FIG. 5, a compression of the octahedron along its z-axis splits the three-fold orbital degenerate  $t_{2g}^2$  level (O<sub>h</sub>, two spin-up electrons in xy, xz, yz) into

the  $b_{2g}^1$  (D<sub>4h</sub>) level (xy orbital) and the xz, yz orbitals subsequently form the formally half-filled and two-fold orbital degenerate  $e_g^1$  (D<sub>4h</sub>) level.<sup>46</sup> In addition, the empty two-fold degenerate  $e_g$  (O<sub>h</sub>) level splits into the  $a_{1g}$  (D<sub>4h</sub>) level (z<sup>2</sup> orbital) and the  $b_{1g}$  (D<sub>4h</sub>) level (x<sup>2</sup> - y<sup>2</sup> orbital). Altogether this yields the multi-electronic ground state  ${}^{3}E_{g}$  (D<sub>4h</sub>) in the compression case. This situation, with one strongly localised electron in the  $b_{2g}$  (D<sub>4h</sub>) level (xy orbital) and the second itinerant electron at a higher energy level in the  $e_g$  (D<sub>4h</sub>) level is crucial for ferromagnetic CrO<sub>2</sub>.<sup>13,30,33,46</sup>

This is of great importance because the **double-exchange mechanism** between the Cr ions requires one strongly localised electron in the xy orbital and another delocalised and dispersed electron in a higher level mediating the magnetic exchange information (ferromagnetic coupling) via *non-local* transitions. <sup>30,46,47,48</sup>

The spins of the strongly localised electrons in the xy orbitals of each Cr site are then coupled via the double-exchange mechanism.<sup>30,39,47</sup> It has been shown that if all electrons were itinerant, an antiferromagnetic coupling would be favoured<sup>48</sup> making the *compression* (meaning the  ${}^{3}E_{g}$ groundstate) a requirement to explain the ferromagnetism in  $CrO_{2}$ .<sup>47</sup>

On the contrary, in the case of an *elongation* of the octahedrons z-axis, the orbital singlet  ${}^{3}B_{2g}$ (D<sub>4h</sub>) becomes the ground state, where the two 3d electrons are in the  $e_{g}^{2}$  (D<sub>4h</sub>) level (xz, yz orbitals) as shown in FIG. 5. In this case, the two electrons are not treated differently, making one localised and one delocalised electron impossible in the elongation case. Hence the ferromagnetism in CrO<sub>2</sub> and the distortion of the octahedra are due to this connection closely related.

These considerations highlight that the nature of the Cr site distortion and the relative energies of the electronic levels of the Cr<sup>4+</sup> ion have a subtle, yet critical, relationship. To approximate the relative energies of the 3d orbitals in the single-particle picture using the three crystal-field parameters Dq, Ds and Dt in tetragonal (D<sub>4h</sub>) symmetry, we use the following set of formulae:<sup>9,43</sup>

$$E_{a_{1q}} = +6Dq - 2Ds - 6Dt \qquad (d_{z^2} \text{ orbital})$$
(1)

$$E_{b_{1g}} = +6\mathrm{D}q + 2\mathrm{D}s - \mathrm{D}t \qquad (\mathrm{d}_{x^2 - y^2} \text{ orbital})$$

$$\tag{2}$$

$$E_{e_g} = -4Dq - Ds + 4Dt \qquad (d_{xz}, d_{yz} \text{ orbital})$$
(3)

$$E_{b_{2q}} = -4\mathrm{D}q + 2\mathrm{D}s - \mathrm{D}t \qquad (\mathrm{d}_{xy} \text{ orbital}) \tag{4}$$

Usually  $d_a$  and  $d_e$  are of the order of angstroms  $(10^{-10} \text{ m})$  yielding for  $|\mathbf{D}q|$  values of the order of ~ 100 meV, and  $|\mathbf{D}s|$  and  $|\mathbf{D}t|$  are usually < 100 meV. The expressions (1) to (4) are an approximation in the sense that they consider the crystal-field effects due to the Coulomb interaction between the ion and the ligands, but they entirely disregard local spin-orbit and non-local exchange interactions. However, they are useful to derive the general behaviour in a single-particle picture for a given parameter set ( $\mathbf{D}q, \mathbf{D}s, \mathbf{D}t$ ).

Although the exact point group symmetry of the Cr ions is  $D_{2h}$ , the  $D_{2h}$  splittings are by far too small to be resolvable in the present experimental data. Furthermore the two additional parameters in  $D_{2h}$  are essentially unknown, making an evaluation of the crystal-field parameters difficult. Therefore, in the present paper we use  $D_{4h}$  symmetry as an approximation in all calculations. In other words, the two  $D_{2h}$  parameters  $D_{\alpha}$ ,  $D_{\beta}$  are implicitly set to zero. For an accurate assessment of the two additional  $D_{2h}$  parameters based on experimental data a high resolution of a few millielectronvolt (meV) is required.

Often apical compressions and elongations are discussed as a so called *Jahn-Teller distortion* describing the change of symmetry from  $O_h$  to  $D_{4h}$ . Jahn-Teller distortions arise when the ground state energy can be reduced due to the distortion and hence it has a stabilising effect. However, for  $CrO_2$  we note that the point group symmetry is not due to a Jahn-Teller distortion, because

this would induce an axial elongation for a  $3d^2$  system, where in fact an axial compression is observed. An axial elongation due to Jahn-Teller would lead to the energy stabilised  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state corresponding to a double occupied  $e_g$  level (see FIG. 5). In other words the axial compression in CrO<sub>2</sub> is not a Jahn-Teller distortion, instead another mechanism is required to justify the stabilisation of the observed compression.

## 3.3. $Cr^{4+}$ ground states ${}^{3}E_{g}$ (D<sub>4h</sub>) and ${}^{3}B_{2g}$ (D<sub>4h</sub>)

The two multi-electronic states, the 3-fold (spin) degenerate  ${}^{3}B_{2g}$  (D<sub>4h</sub>) and the 6-fold degenerate  ${}^{3}E_{g}$  (D<sub>4h</sub>), have formally the lowest ground state energy, for an elongation and compression respectively. Hence, at absolute zero T = 0 K these are therefore the corresponding ground states (see FIG. 5). However, using only the lowest multi-electronic ground state  ${}^{3}E_{g}$  (D<sub>4h</sub>) for the expected compression implies a temperature of T = 0 K. The partial term scheme in FIG. 6 illustrates that the multi-electronic state  ${}^{3}B_{2g}$  (D<sub>4h</sub>), corresponding in the single electron picture to the two electrons in a  $|e_{g}^{2}\rangle$  configuration, is also one of the excited states in the compression case.



Figure 6: Examples illustrating possible  ${}^{3}B_{2g}(D_{4h})$  and  ${}^{3}E_{g}(D_{4h})$  contributions in the compression case. The picture is restricted to the three  $t_{2g}(O_{h})$  orbitals (xy,xz,yz) leaving the  $e_{g}(O_{h})$  orbitals empty yielding a total of 9 micro states: 6 for  ${}^{3}E_{g}(D_{4h})$  and 3 for  ${}^{3}B_{2g}(D_{4h})$ .

Comparing FIG. 5 and FIG. 6 also shows that possible  ${}^{3}B_{2g}$  (D<sub>4h</sub>) contributions from an excited state in the compression case can be possible. The multi-electronic ground state GS is in this example an unknown linear combination of the  ${}^{3}B_{2g}$  (D<sub>4h</sub>) and  ${}^{3}E_{g}$  (D<sub>4h</sub>) configurations as schematically formalised in expression (5):

$$GS = \alpha \cdot {}^{3}B_{2g} + \beta \cdot {}^{3}E_{g} \tag{5}$$

However, it shall be clear that this is yet simplified and the real linear combination directly relates to the exact multiplet structure, and the result therefore depends on the specific distortion and on the magnetic splitting energies as well as spin-orbit and exchange interactions. In other words, the linear combination depends on the exact order and energy splittings of the individual states due to the various interactions.

To consider temperatures T > 0 K the effective spectrum is commonly approximated as a linear combination of all contributing ground states according to the

**Boltzmann distribution** 
$$p_i \propto \exp\left(\frac{-E_i}{k_B T}\right)$$
 (6)

where the occupation probability  $p_i$  for the multi-electronic state *i* depends, with the Boltzmann constant  $k_B$  and energy  $E_i$ , on the absolute temperature *T*.

Due to the fact that in  $CrO_2$  several energy splittings in the tens of meV lead to a complex multiplet structure, it is without additional information impossible to determine the exact linear

combination of the actual ground state. Additionally local and non-local mixing add further difficulties when identifying the ground state character. This can be illustrated with energy level diagrams showing the dependency of the energy levels on the crystal-field parameters Dq, Ds, Dtand M (see FIG. 21 in the appendix).

In this paper we discuss, based on crystal-field multiplet calculations, the different characters of the RIXS and RIXS-MCD maps for two characteristic contributions for the  ${}^{3}B_{2g}$  (D<sub>4h</sub>) and  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground states. The crystal-field multiplet approach without additional charge transfer is a *local* theory, that we use here to model the effect of the local point group symmetry on the splittings of the multi-electronic energy levels of Cr<sup>4+</sup> ions (3d<sup>2</sup>). In other words, we calculate only the local contributions explicitly, while non-local hybridisation (charge transfer) is approximated by the reduction of the Slater integrals.<sup>10</sup>

Hence it is expected that the calculated sum pre-edge spectrum (Sum = (5 + (5))) will lack any intersite transitions or metal-to-ligand charge transfer and will therefore not be able to reproduce the experimental RIXS map for the sum correctly. However, it is assumed that RIXS-MCD probes a local property, the local magnetic moment, and should therefore be well reproducible within a crystal-field multiplet approach.

## 4. Results and Discussion

#### 4.1. Chromium K-edge

The X-ray Absorption Near Edge Structure (XANES) measured with High Energy Resolution Fluorescence Detection (HERFD) at the Cr K-edge are shown in FIG. 7. It covers the Cr K pre-edge region and the main edge, corresponding to a 1s X-ray absorption scan while the spectrometer (analyser angle) remains fixed in order to detect the Cr K $\alpha_1$  ( $E_{out} = 5415.3 \,\text{eV}$ ) emission that has been calibrated far from resonance.



Figure 7: Experimental High Energy Resolution Fluorescence Detection Magnetic Circular Dichroism (HERFD-MCD) spectra at the Cr K-edge on  $\text{CrO}_2$ , for  $E_{out} = 5415.3 \text{ eV}$ , illustrating the low MCD intensity in the pre-edge ( $E_{in} \approx 5992 \text{ eV}$ ) and in the dipole main edge ( $E_{in} \geq 5998 \text{ eV}$ ).

The dichroism (FIG. 7 bottom) confirms the general trend of two inverted MCD signals for the

positive and negative magnetic field direction and it illustrates how small the MCD signal in the pre-edge region really is. The asymmetry between the MCD signals for the two magnetic field directions relates mostly to the relative high noise level as only three scans were averaged. In the following we will look only at the pre-edge structure with incident energies  $E_{in} < 6 \text{ keV}$ .

## 4.2. Density of states

The electric quadrupole and dipole absorption cross-section of the Cr K-edge calculated by firstprinciple methods (solid lines) is shown in FIG. 8 (top) and compared to the experimental spectrum (dotted line).



Figure 8: Electric dipole and quadrupole contributions to the absorption cross-section for incident  $\vec{k} = (0, 0, 1)$  and polarisation  $\vec{\varepsilon} = (1, 0, 0)$ , and calculated densities of states (DOS) in CrO<sub>2</sub> for the ground state and with a core hole in the Cr 1s shell (indicated as Cr<sup>•</sup>). Relative contributions of the electric dipole and quadrupole transitions to the pre-edge structure (top), and projected densities of states (DOS) in CrO<sub>2</sub> for both spin directions. The Fermi level lies in a gap of the minority spin states and it is fully spin-polarised. Cr(4p),Cr(3d) and O(2p) are overlapping in energy and considered as contributions to the intersite Cr(4p)-O(2p)-Cr(3d) band.

In 1s X-ray Absorption Spectroscopy (XAS), the electric dipole transition probes the empty Cr p states whereas the electric quadrupole transition probes the empty Cr d states. From the calculated p and d density of states projected on the Cr atom with a core-hole (FIG. 8, lower 3 plots), we can infer that there is a strong electric dipole contribution between 2 and 5 eV, that is due to the hybridisation of the Cr(4p) states of the absorbing atom with the Cr(3d) states of the neighbouring Cr atoms. Furthermore we find an electric quadrupole contribution at lower energy that appears as a shoulder in the experimental spectra. The main effect of accounting for the onsite Coulomb repulsion through the U term would be an upward shift of the unoccupied d-band.<sup>36</sup> It is known that the energy position of the unoccupied d states near the Fermi level is very sensitive to core hole effects and that, when including a full static 1s core-hole, it is usually calculated at a too high energy with respect to the edge.<sup>6</sup> Hence, the energy separation between the d peak and the p states may not be accurately computed.

In Figure 9 the Cr(3d) densities are displayed as partial density of states projected onto the corresponding orbitals. The xy orbital is localised in energy and thus not part of the continuum, hence it is also localised in real space contrary to the xz,yz orbitals which are more itinerant, in agreement with the effect of the crystal-field in the compression case. The latter two orbitals (Cr xz,yz) strongly hybridise with O(2p) states related to the double-exchange mechanism.



Figure 9: Partial projected Cr(3d) densities of states near the Fermi level. The xy orbital appears very localised and lowest in energy being in agreement with the expected compression.

Because Cr is centro-symmetric, on-site pd-hybridisation is forbidden by symmetry. Nevertheless intersite hybridisation between the *local* Cr(4p) orbitals and the *non-local* Cr(3d) orbitals of the neighbouring ions is possible (cf. FIG. 7 in <sup>12</sup>). This *non-local* or *intersite* hybridisation can arise in solids where the many metal sites form bands adding some local Cr(4p) character to the mostly Cr(3d) character of the pre-edge. In CrO<sub>2</sub> the O(2p) orbitals mediate the hybridisation between the local Cr(4p) character and the Cr(3d) bands of all Cr sites in the solid to form a "non-local band" Cr(4p)-O(2p)-Cr(3d) due to intersite 4p3d hybridisation.<sup>12,26,62</sup>

In the ground state, this non-local band overlaps in energy with the Cr(3d) and the Cr(4p) character in the pre-edge. When a 1s core hole, indicated as Cr<sup>•</sup>, is created the local electronic structure is effectively shifted to lower energy due to the reduced Coulomb repulsion of the core level. The shift of the local Cr<sup>•</sup>(3d) level with respect to the unchanged Cr(3d) level is expected to be approximately  $\Delta E_{in} \approx 2$  to 2.5 eV while the Cr<sup>•</sup>(4p) level is, due to a different screening, only slightly shifted by  $\Delta E_{in} \approx -0.5$  eV to lower energy.<sup>26,62</sup>

All other non-local Cr sites and the non-local band Cr(4p)-O(2p)-Cr(3d) remain essentially unaffected and the pre-edge consists of two types of transitions:

- i) Purely local or onsite transitions i.e. the native quadrupole  $1s \rightarrow 3d$  peak that will appear lowest in energy.
- ii) Non-local or intersite transitions i.e. from the local 1s shell into the non-local Cr(4p)-O(2p)-Cr(3d) band which is approximately 2.5 eV above the native  $Cr^{\bullet}(3d)$  peak.

This non-local mixing combined with the above described complex order of the multi-electronic ground states makes a detailed analysis of the real electronic configuration difficult.

#### 4.3. 1s2p RIXS-MCD map of CrO<sub>2</sub>

The sum of both circular polarisations  $(\circlearrowleft + \circlearrowright)$  yields the 1s2p RIXS map shown in FIG. 10.a, and the difference  $(\circlearrowright - \circlearrowright)$  results in the experimental 1s2p RIXS-MCD map of a crystalline CrO<sub>2</sub> powder displayed in FIG. 10.b.



Figure 10: a.) Experimental 1s2p RIXS sum ( $\circlearrowleft + \circlearrowright$ ) and b.) the corresponding RIXS-MCD map ( $\circlearrowright - \circlearrowright$ ) of the CrO<sub>2</sub> powder covering the K $\alpha_{1,2}$  doublet, namely the  $2p_{\frac{3}{2}} \rightarrow 1s$  (K $\alpha_1$ ) and  $2p_{\frac{1}{2}} \rightarrow 1s$  (K $\alpha_2$ ) decays after absorption in the Cr K pre-edge. The vertical and diagonal lines highlight the CIE and CEE slices discussed in this paper (see FIG. 12, FIG. 13 and FIG. 14).

The sum (FIG. 10.a) shows the typical band-like diagonal orientation for the K $\alpha_{1,2}$  fluorescence decays. The experimental 1s2p RIXS-MCD map (FIG. 10.b) clearly displays an intense magnetic circular dichroism for the  $2p_{3/2} \rightarrow 1s$  decay channel (K $\alpha_1$ ) around  $E_T \approx 577 \,\text{eV}$ . Its vertical orientation indicates a resonant final state (FS) effect, as opposed to a MCD in horizontal direction would mean resonance on an intermediate state (IS).

The MCD of the  $2p_{1/2} \rightarrow 1s$  decay channel (K $\alpha_2$ ) in the upper half of the map is an order of magnitude less intense with respect to the K $\alpha_1$  dichroism. The K $\alpha_1$  and K $\alpha_2$  regions both display a weak diagonal dichroic background in the MCD map (FIG. 10.b).

The K $\alpha_1$  pre-peak maximum in the RIXS map (FIG. 10.a) is at  $E_{in} \approx 5995 \,\text{eV}$  and  $E_T \approx 580 \,\text{eV}$ while the center of the K $\alpha_1$  MCD appears approximately at  $E_{in} \approx 5992.5 \,\text{eV}$  and  $E_T \approx 577.5 \,\text{eV}$  (FIG. 10.b). In other words, the position of the intense  $K\alpha_1$  dichroism, marked with a circle in both maps, is located at approximately  $\Delta E_{in} \approx 2.5 \text{ eV}$  lower incident energy with respect to the pre-peak maximum of the  $K\alpha_1$  emission.

### 4.4. Constant Emitted Energy (CEE) slices

The Constant Emitted Energy (CEE) slices are in the following used to investigate the Cr K $\alpha_1$  pre-peak structure. They are comparable to HERFD-XAS and appear as diagonals in a RIXS map in an energy transfer view  $I(E_{in}, E_T)$ . They are identical to the corresponding horizontal slices in maps in an emitted energy view  $I(E_{in}, E_{out})$ . Subsequently, the CEE slices shown as diagonals in FIG. 10 appear as horizontal lines in an emitted energy view as displayed in FIG. 11.



Figure 11: RIXS and RIXS-MCD maps of the  $K\alpha_1$  region in an emitted energy view showing the positions of the extracted CEE slices as horizontal lines.

These CEE slices correspond to the emitted energies  $E_{out} \in \{5414.5, 5414.75, 5415.0, 5415.25\}$  eV. They are shown in FIG. 12 and immediately reveal that the pre-edge structure consists of at least two visible peaks. One peak at  $E_{in} \approx 5992$  eV which is assigned to the quadrupole peak  $(1s \rightarrow 3d)$ giving rise to the MCD, and another structure centred around  $E_{in} \approx 5995$  eV is interpreted as a non-local peak, whose contribution to the MCD signal is small.



Figure 12: Constant Emitted Energy (CEE) slices for  $E_{out} \in \{5414.5, 5414.75, 5415.0, 5415.25\}$  eV top: CEE sum slices showing that the K $\alpha_1$  pre-peak clearly consists of multiple peaks; bottom: CEE MCD slices revealing the relative position of the dichroism.

Therefore, the transitions into the spin-polarised  $t_{2g}$  will dominate the MCD.<sup>13,24,33</sup> The empty  $e_g$  level however can contribute to the MCD only via the exchange splitting between the spin-up and spin-down  $e_g$  states. This seems to be consistent with the MCD appearing at a lower incident energy with respect to the pre-peak maximum as visible in FIG.12. Here we will use crystal-field multiplet theory in an attempt to calculate the local quadrupole contributions.

## 4.5. Constant Incident Energy (CIE) slices

A vertical slice extracted from a two dimensional RIXS map is a *Constant Incident Energy* (CIE) slice. The CIE slices across the 1s pre-peak are expected to be similar to conventional  $L_{2,3}$ -edge spectra, because the final state  $1s^2 2p^5 3d^{N+1}$  is identical in 1s2p RIXS and in 2p XAS. However, the spectra are not identical because the 1s2p RIXS process involves the matrix elements for the  $1s \rightarrow 3d$  excitation and the  $2p \rightarrow 1s$  decay, which are different from the direct 2p XAS matrix elements. Thus the MCD CIE slices across the intense dichroism extracted from the MCD map are expected to be similar but not identical to conventional  $L_{2,3}$ -edge XMCD spectra.<sup>7</sup>

In the following the CIE slices corresponding to the vertical lines in the two maps in FIG. 10 are discussed. The first set of three CIE slices intersecting the intense  $K\alpha_1$  dichroism in the **onsite** region at the incident energies  $E_{in} \in \{5992.25, 5992.50, 5992.75\}$  eV is displayed in FIG. 13.



Figure 13: Constant Incident Energy (CIE) slices across the onsite region (vgl. Figure 12) for  $E_{in} \in \{5992.25, 5992.50, 5992.75\}$  eV derived from the two maps in Figure 10. top: CIE sum slices showing the  $K\alpha_{1,2}$  doublet; bottom: CIE MCD slices with the strong dichroism for the  $K\alpha_1$  emission, while the  $K\alpha_2$  region appears only weakly dichroic.

As expected, the CIE sum slices in FIG. 13 show the typical band-like shift with increased incident energy  $E_{in}$  being consistent with the diagonal appearance of the K $\alpha$  pre-edge emission in FIG. 10.a.

The MCD slices on the other hand show only minor differences with no energy shift indicating a resonant ( $\rightarrow$  excitonic) effect. Only the minima and maxima of the MCD intensity, at  $E_T \approx 576.5 \,\text{eV}$  and  $E_T \approx 578 \,\text{eV}$  respectively, vary a bit. This can be explained with the overlapping diagonal band character which is also visible as weak diagonal dichroism in the RIXS-MCD map (FIG. 10.b). Four other CIE slices across the **intersite** region in the maximum of the Cr K $\alpha_1$  pre-edge structure, extracted at the incident energies  $E_{in} \in \{5994.50, 5994.75, 5995.00, 5995.25\}$  eV, are displayed in FIG. 14.



Figure 14: Constant Incident Energy (CIE) slices across the intersite region (vgl. Figure 12) for  $E_{in} \in \{5994.50, 5994.75, 5995.00, 5995.25\}$  eV extracted from Figure 10. top: CIE sum slices showing the K $\alpha_{1,2}$  doublet; bottom: CIE MCD slices with a relatively weak dichroism for the K $\alpha_{1,2}$  region.

The CIE sum slices (FIG. 14 top) are again similar with the typical band-like shift with increased incident energy  $E_{in}$ . The CIE MCD slices (FIG. 14 bottom) show some energy dependent shift, which indicates that the dichroism in this region, is not purely quadrupolar ( $\rightarrow$ excitonic). The MCD intensities are in this case almost an order of magnitude smaller with respect to the dichroism measured in the low-energy tail of the pre-peak around  $E_{in} \approx 5992 \,\text{eV}$  as shown in FIG. 13

The observed energy shift is also consistent with the partially diagonal appearance of the dichroism in the 1s2p RIXS-MCD map (FIG. 10.b) around  $E_T \approx [580 \pm 1] \text{ eV}$ . This may be assigned to the non-local mixing of the 3d states inducing some band-like character visible as dipole band transitions into the spin-polarised 3d-band. A detailed quantitative analysis of the dichroism in this intersite region is difficult due to the involved intersite 4p3d hybridisation: this aspect goes beyond the local CFM approach, which does not account for the band structure of the solid and thus it prevents the calculation of transitions to delocalised levels.

It is however noteworthy that overall the  $K\alpha_2$  dichroism shows in this experiment no indication of a resonating behaviour as opposed to the  $K\alpha_1$  dichroism.

### 4.6. RIXS-MCD ground state character of CrO<sub>2</sub>

The detailed analysis of the CEE and CIE slices have revealed that the intense dichroism in the experimental RIXS-MCD map in FIG. 10.b is dominated by *local* excitonic transitions. Therefore, we assume that the local contributions to the pre-edge sum, and that most of the MCD intensity can be calculated within the local crystal-field multiplet framework.

As discussed above, for ferromagnetic  $CrO_2$  an electronic structure corresponding to an axial compression with one strongly localised and one itinerant electron is required for the proposed double-exchange mechanism. The energy level diagram in FIG. 15 shows the behaviour of the multi-electronic states in the case of the compressed distortion of the Cr site for a floating magnetic exchange interaction M and without spin-orbit coupling.



Figure 15: Multi-electronic energy level diagram floating the magnetic exchange parameter M from 0 to 50 meV. The threefold spin degenerate states  ${}^{3}E_{g}$  (D<sub>4h</sub>) and  ${}^{3}B_{2g}$  (D<sub>4h</sub>) are separated at M = 0 meV by  $\Delta E = |-3Ds + 5Dt| = 73$  meV. The spin degeneracy is then removed due to the magnetic field ( $\rightarrow$  Zeeman effect). The circles highlight intersections between the states indicating a possible mixing between the states and a change of the ground states order. The dashed line marks M = 30 meV used in the calculations.

As shown by this diagram, the super-exchange interaction M removes the threefold spin degeneracy of the  ${}^{3}E_{g}(D_{4h})$  and  ${}^{3}B_{2g}(D_{4h})$  terms. The lowest multi-electronic ground state arises from  ${}^{3}E_{g}(D_{4h})$  and should therefore dominate the spectrum. The energy  $\Delta E$  separating the  ${}^{3}E_{g}(D_{4h})$ and  ${}^{3}B_{2g}(D_{4h})$  at M = 0 meV corresponds in this case to  $\Delta E = |-3Ds + 5Dt| = 73$  meV and is therefore also dependent on the specific parameter set chosen. In this case the splittings are too large to justify a population of  ${}^{3}B_{2g}(D_{4h})$  states only due to thermal excitation (T = 300K  $\rightsquigarrow E = kT \approx 25$  meV).

However, it is important to note that the twofold orbital degeneracy of the  ${}^{3}E_{g}$  (D<sub>4h</sub>) state, leading in total to six levels, is not removed here by the magnetic field scaled with M. Hence, only the threefold spin degeneracy is removed. This is due to the fact that for ferromagnets one has to use the *Heisenberg exchange coupling* in the Hamiltonian in which the magnetic operator  $\hat{O}$  is defined without the orbital momentum L:

$$\widehat{\mathcal{O}} = B_x \cdot (2\,\widehat{S}_x) + B_y \cdot (2\,\widehat{S}_y) + B_z \cdot (2\,\widehat{S}_z) \tag{7}$$

Whereas for paramagnets one has to use the usual definition including the orbital momentum L:

$$\widehat{O} = B_x \cdot (2\,\hat{S}_x + \hat{L}_x) + B_y \cdot (2\,\hat{S}_y + \hat{L}_y) + B_z \cdot (2\,\hat{S}_z + \hat{L}_z) \tag{8}$$

For both cases,  $B_x$ ,  $B_y$ ,  $B_z$  are the components of the magnetic field with  $B_x = (1, 0, 0) \cdot M$  (and  $B_y, B_z$  analog), and  $\hat{S}_x$ ,  $\hat{S}_y$ ,  $\hat{S}_z$  are the operators for the corresponding spin components. For the definition in (8)  $\hat{L}_x$ ,  $\hat{L}_y$ ,  $\hat{L}_z$  are then analogous the operators for the corresponding orbital components.

As we have to use for the ferromagnet  $\operatorname{CrO}_2$  the definition in (7), this explains why the twofold orbital degeneracy of the  ${}^{3}\mathrm{E}_{g}$  (D<sub>4h</sub>) state is not removed in the energy level diagram shown in Fig. 15 (ferromagnet and spin-orbit interaction off). The  ${}^{3}\mathrm{E}_{g}$  (D<sub>4h</sub>) state will only split into six separate levels (3 for spin, 2 for orbit) when equation (8) is used (paramagnets), or for a ferromagnet with the definition in (7) when spin-orbit interaction is on as shown in the appendix in Fig. 21 top row. It would go beyond the scope of this paper to discuss all calculated RIXS-MCD maps that were made to find acceptable agreement with the experimental data. Thus here we present the calculated RIXS-MCD map for a chosen parameter set that corresponds to the compressed distortion, as a benchmark example for the interpretation of the excitonic RIXS and RIXS-MCD features.

## 4.6.1. RIXS-MCD for ${}^{3}E_{g}$ (D<sub>4h</sub>) ground states

One example calculation for a RIXS and RIXS-MCD map for  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state is shown in FIG. 16 and shortly discussed in the following. The general appearance of the RIXS-MCD map is not similar to the experimental MCD map.



Figure 16: Calculated 1s2p RIXS sum and MCD map for a  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state with spin-orbit coupling  $\zeta_{3d} = 100\%$ .

Most notable is that the minimum and maximum of the MCD are more separated on both axis when compared to the experimental data. Even under the consideration of the absence of any non-local contributions at higher energies it appears clear that the dichroism is not shifted to the low-energy tail of the pre-edge structure as marked with a circle in the RIXS and RIXS-MCD maps in FIG. 10 and as it is better visible in the CEE slices in FIG. 12.

The CIE slices show the  $K\alpha_1$  well reproduced, while the  $K\alpha_2$  dichroism apparently has the opposite tendency in the calculation  $(\downarrow\uparrow\downarrow\uparrow)$  with respect to the experimental data  $(\downarrow\uparrow\uparrow\downarrow\downarrow)$ . The red and blue arrows correspond to the local minima and local maximal in the MCD slice and they emphasise the opposite behaviour of the minima and maxima in the MCD of the  $K\alpha_2$  region for the comparison between the experimental and calculation data. The effect is also illustrated with arrows in Figure 17 comparing the calculated CIE slice, which is highlighted by the red line in the RIXS MCD map in Figure 16.b, with the experimental slice already shown in Figure 13.

This behaviour is visible for essentially all calculated  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground states, but because the experimental K $\alpha_{2}$  MCD appears to be mostly due to non-local mixing, an accurate modelling of the K $\alpha_{2}$  MCD is beyond the scope of the present multiplet model.

We have performed many calculations across a significant section of the parameter space, and none of the parameter sets for a  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state yields an overall satisfactory result comparable to the experimental data.



Figure 17: Comparison the MCD CIE slices to visualise the inverted behaviour of  $K\alpha_2$  MCD with respect to the  $K\alpha_1$  emphasised with arrows indicating the local minima and maxima. The calculated MCD CIE slice is scaled to match the experimental intensity.

## 4.6.2. RIXS-MCD for ${}^{3}B_{2q}$ (D<sub>4h</sub>) ground states

As discussed above  ${}^{3}B_{2g}$  (D<sub>4h</sub>) corresponds in the single electron picture for 3d<sup>2</sup> having both electrons in the  $e_{g}^{2}$  level (*xz,yz* orbitals) as displayed in FIG. 5 and FIG. 6. Even though a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state appears to contradict the requirements for the proposed double-exchange mechanism as cited above,  ${}^{30,46}$  and it is known that the ground state at absolute zero temperature should be a  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state, our calculations show that we find multiple solutions yielding a RIXS-MCD map comparable to the experimental data for  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground states. Furthermore we find that for  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state calculations the RIXS and RIXS-MCD maps are in many cases essentially the same for 3d spin-orbit interaction  $\zeta_{3d}$  either on or off.

One example for a calculation for a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state with 3d spin-orbit interaction set to  $\zeta_{3d} = 0.0 \text{ eV}$  is displayed in FIG 18.



Figure 18: Calculated 1s2p RIXS sum and MCD map for a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state without spinorbit interaction  $\zeta_{3d} = 0\%$ .

This is in agreement with other reports finding a reasonably good agreement with spin-orbit interaction effectively switched off due to the vanishing influence of an orbital momentum and correlated spin-orbit effects.<sup>49</sup> Most notable is that the RIXS sum for a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state in FIG. 18.a now consists of two single-peak structures, as opposed to the double-peak structures as visible for the  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state.

A comparison of the calculated RIXS-MCD map in FIG. 18.b with the experimental map (FIG. 10.b) shows that the weak diagonal dichroism related to the band-character is absent in the calculation. But this is expected and the general appearance of the RIXS-MCD map is quite well reproduced. In fact, under assumption of a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state we find many solutions similar to the experimental RIXS-MCD map.

In addition we now also find a different behaviour in the CIE slices for the calculation  $(\downarrow\uparrow\uparrow\downarrow)$  with respect to the calculation for the  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state. The results for  ${}^{3}B_{2g}$  (D<sub>4h</sub>) do in fact also agree better with the experimental slices  $(\downarrow\uparrow\uparrow\downarrow)$  as shown in Fig. 19.



Figure 19: Comparison of the experimental and the calculated CIE slices extracted from FIG. 10.b and FIG. 18.b. (The calculated slice is scaled to match experimental data.)

Again, it would go beyond the scope of this paper to discuss all calculated RIXS and RIXS-MCD maps for  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground states, thus it shall be clear that the shown maps are not supposed to be taken as proposed solutions. Instead the calculations shown here serve only as example to illustrate the general appearance of the RIXS and RIXS-MCD map with a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) ground state character.

From these calculations, we conclude that the RIXS-MCD signal appears to be dominated by ground state contributions with a  ${}^{3}B_{2g}$  (D<sub>4h</sub>) character. This would therefore suggest that the ground state is a mixed one, which further triggers the counter-intuitive distortion with respect to pure Jahn-Teller considerations. Such a complex ground state has already been suggested for various systems and in some cases is mediated by both double-exchange mechanism and vibronic couplings, which were not accounted in our present model.<sup>4,57</sup>.

## 5. Summary

The RIXS-MCD approach delivers valuable information to experimentally disentangle the quadrupole from the dipole contributions, where the quadrupole part of the spectrum reveals magnetic information. As such, hard X-ray RIXS-MCD is a bulk-sensitive high resolution magnetic spectroscopy.

In the case of  $CrO_2$  the pre-edge is dominated by dipole transitions into the non-local Cr(4p)-O(2p)-Cr(3d) band originating from intersite 4p3d hybridisation and the quadrupole pre-edge not detectable. The 1s2p RIXS-MCD unravels the strong resonant features in the  $K\alpha_1$  dichroism assigned to excitonic states. The non-local states only show a weak non-resonant dichroism.

The assignment of the resonant features in the K $\alpha_1$  dichroism was interpreted with the CFM theory that describes the 3d3d and 2p3d multiplet interactions. Although the compressed axial distortion of the Cr<sup>4+</sup> octahedron predicts a  ${}^{3}E_{g}$  ( $D_{4h}$ ) ground state, the RIXS-MCD calculation of this ground state only fails to reproduce the observed dichroism. Our calculations suggest that i) the spin-orbit coupling of the 3d electrons is quenched, and ii) there is non-negligible contribution from the excited multi-electronic state  ${}^{3}B_{2g}$  ( $D_{4h}$ ) in the ground state.

We conclude that we cannot give an unambiguous description of the ground state from a local crystal-field approach. Both, the existence of a double-exchange mechanism and the strong non-local Cr(4p)-O(2p)-Cr(3d) hybridisation indicate strong Cr-Cr interactions.

Recent computational developments that enable the coupling between first principle densityfunctional theory and multi-electronic calculations would enable to go beyond the crystal-field model and provide valuable insights into the nature of the ground state of  $CrO_2$ .

## 6. Acknowledgments

We wish to acknowledge SOLEIL for provision of the facilities at GALAXIES; explicitly we appreciate the exceptional support from the staff at the GALAXIES beamline at SOLEIL. We thank Pieter Glatzel (ESRF) for lending us the analyser crystal used in this experiment and for sharing his Matlab scripts which were very helpful to write the scripts needed for our calculations.

For financial support we further acknowledge the European Research Council (ERC). Marcin Sikora (AGH University, Kraków, Poland) acknowledges support from EUSpec COST Action and National Science Centre, Poland.

DFT calculations were done using HPC resources from GENCI-IDRIS (Grant i2016096863). The energy level diagrams have been calculated with Quanty (www.quanty.org).<sup>20,21,34</sup> All molecule structures were created with Vesta.<sup>37</sup>

## A. RIXS-MCD for ${}^{3}E_{q}$ (D<sub>4h</sub>) without spin-orbit interaction

Another set of calculated maps for a  ${}^{3}E_{g}$  (D<sub>4h</sub>) ground state with the 3d spin-orbit interaction turned off ( $\zeta_{3d} = 0.0 \text{ eV}$ ) is shown in Figure 20. This could be explained with a quenching of the 3d spin-orbit interaction because the atomic value  $\zeta_{3d} = 54 \text{ meV}$  is small.



Figure 20: Calculated 1s2p RIXS sum and MCD map for a  ${}^{3}E(D_{4h})$  ground state with 3d spin-orbit interaction  $\zeta_{3d}$  off.

Furthermore the orbital magnetic moments of 3d transition metals are generally quenched because of the crystal-field  $^{18,22,23,28}$ . The metallicity of CrO<sub>2</sub>, crystal-field or symmetry effects as well as angular averaging for a powder, all can reduce the spin-orbit interaction.

The general appearance of the RIXS-MCD map is remotely similar to the experimental data, but the MCD intensity is by several orders of magnitude too low  $(I_{MCD} \sim 10^{-7})$  with respect to the quadrupole maximum in the sum. Furthermore the dominant MCD of the K $\alpha_1$  region again does not appear in the low energy tail of the pre-edge structure of the RIXS sum, which is the case for the experimental data (see circle in Figure 10).

Usually it is expected that the 3d spin-orbit interaction  $\zeta_{3d}$  separates the MCD minima and maxima in the incident energy  $E_{in}$  direction. However, there is clearly a separation of the MCD minimum and maximum also in the incident direction visible in this example proving that it is not purely be the 3d spin-orbit interaction  $\zeta_{3d}$ .

## B. Values used for the calculations

In this part we summarise the values used for the calculations shown in this paper for the *compression case* with the ground state  ${}^{3}E_{g}$  (Table 2) and for the *elongation case* for the ground state  ${}^{3}B_{2g}$  (Table 3).

|                  | GS                 | IS                 | FS                 |
|------------------|--------------------|--------------------|--------------------|
| $10\mathrm{D}q$  | $2.347\mathrm{eV}$ | $2.347\mathrm{eV}$ | $2.347\mathrm{eV}$ |
| $\mathrm{D}s$    | $-36\mathrm{meV}$  | $-36\mathrm{meV}$  | $-36\mathrm{meV}$  |
| Dt               | $-7\mathrm{meV}$   | $-7\mathrm{meV}$   | $-7\mathrm{meV}$   |
| M                | $30\mathrm{meV}$   | $30\mathrm{meV}$   | $30\mathrm{meV}$   |
| $F^2(2p3d)$      | -                  | -                  | $4.687\mathrm{eV}$ |
| $G^1(2p3d)$      | -                  | -                  | $3.508\mathrm{eV}$ |
| $G^2(1s3d)$      | -                  | $38\mathrm{meV}$   | -                  |
| $G^3(2p3d)$      | -                  | -                  | $1.996\mathrm{eV}$ |
| $F^2(3d3d)$      | $6.133\mathrm{eV}$ | $6.496\mathrm{eV}$ | $6.538\mathrm{eV}$ |
| $F^4(3d3d)$      | $3.868\mathrm{eV}$ | $4.092\mathrm{eV}$ | $4.123\mathrm{eV}$ |
| SOC $\zeta_{2p}$ | -                  | -                  | $5.668\mathrm{eV}$ |
| SOC $\zeta_{3d}$ | $41\mathrm{meV}$   | $54\mathrm{meV}$   | $53\mathrm{meV}$   |

Table 2: Values for the ground state (GS), intermediate state (IS) and final state (FS) for the  ${}^{3}E_{g}$  state. Note, the values for spin-orbit interaction  $\zeta_{3d}$  are set to 0 meV for the spectra shown in Figure 20. The results of the calculation are shown in section 4.6.1 and appendix A.

|                  | GS                 | IS                 | FS                 |
|------------------|--------------------|--------------------|--------------------|
| $10\mathrm{D}q$  | $2.347\mathrm{eV}$ | $2.347\mathrm{eV}$ | $2.347\mathrm{eV}$ |
| $\mathrm{D}s$    | $0\mathrm{meV}$    | $0\mathrm{meV}$    | $0\mathrm{meV}$    |
| Dt               | $200\mathrm{meV}$  | $200\mathrm{meV}$  | $200\mathrm{meV}$  |
| M                | $30\mathrm{meV}$   | $30\mathrm{meV}$   | $30\mathrm{meV}$   |
| $F^2(2p3d)$      | -                  | -                  | $4.687\mathrm{eV}$ |
| $G^1(2p3d)$      | -                  | -                  | $3.508\mathrm{eV}$ |
| $G^2(1s3d)$      | -                  | $38\mathrm{meV}$   | -                  |
| $G^3(2p3d)$      | -                  | -                  | $1.996\mathrm{eV}$ |
| $F^2(3d3d)$      | $6.133\mathrm{eV}$ | $6.496\mathrm{eV}$ | $6.538\mathrm{eV}$ |
| $F^4(3d3d)$      | $3.868\mathrm{eV}$ | $4.092\mathrm{eV}$ | $4.123\mathrm{eV}$ |
| SOC $\zeta_{2p}$ | -                  | -                  | $5.668\mathrm{eV}$ |
| SOC $\zeta_{3d}$ | $0\mathrm{meV}$    | $0\mathrm{meV}$    | $0\mathrm{meV}$    |

Table 3: Values for the ground state (GS), intermediate state (IS) and final state (FS) for the  ${}^{3}B_{2g}$  state. Note, to reflect a fully quenched spin-orbit interaction  $\zeta_{3d}$  it is set to 0 meV. The results of the calculation are shown in section 4.6.2.

## C. Energy Level Diagrams



Energy level diagrams for 10Dq, Ds, Dt and M with 3d spin-orbit interaction off



Figure 21: Energy level diagrams for  $Cr^{4+}$  with 3d spin-orbit interaction  $\zeta_{3d}$  on (top row) and off (bottom row). The vertical red dashed marker lines highlight the approximate value used in the calculation (10Dq = 2.347 eV, Ds = -0.036 eV, Dt = -0.007 eV, M = 30 meV). The intersections of the dashed marker line with the energy levels do agree between the four corresponding diagrams and hence enable to illustrate the behaviour in the parameter space in the proximity of the chosen crystal-field and magnetic splitting parameter values. Furthermore it is noteworthy that there are several crossings between the individual states which on the one hand implies a change of the order of the ground states, and on the other hand can indicate a mixing of the ground state characters.
## References

- [1] D. Ahlers, K. Attenkofer, and G. Schütz. Spin-dependent extended x-ray absorption fine structure in magnetic oxides. Journal of Applied Physics, 83(11):7085, 1998.
- [2] K. Attenkofer and G. Schuetz. Hard and Soft X-MCD Studies of CrO<sub>2</sub>. Journal de Physique IV France, 7, 1997.
- [3] W. H. Baur and A. A. Khan. Rutile-type compounds. IV. SiO<sub>2</sub>, GeO<sub>2</sub> and a comparison with other rutile-type structures. Acta Crystallographica Section B, 27(11):2133-2139, 1971. .cif file: http://www.crystallography.net/cod/9007540.html.
- [4] S. A. Borshch, E. L. Bominaar, G. Blondin, and J. J. Girerd. Double exchange and vibronic coupling in mixed valence systems. Origin of the broken-symmetry ground state of [Fe3S4] 0 cores in proteins and models. Journal of the American Chemical Society, 115(12):5155-5168, 1993.
- [5] J. K. Burdett, G. J. Miller, J. W. Richardson, and J. V. Smith. Low-Temperature Neutron Powder Diffraction Study of CrO<sub>2</sub> and the Validity of the Jahn-Teller Viewpoint. J. Am. Chem. Soc., 110:8064–8071, 1988.
- [6] D. Cabaret, A. Bordage, A. Juhin, M. Arfaoui, and E. Gaudry. First-principles calculations of X-ray absorption spectra at the K-edge of 3d transition metals: an electronic structure analysis of the pre-edge. Physical Chemistry Chemical Physics, 12:5619–5633, 2010.
- [7] W. A. Caliebe, C. C. Kao, J. B. Hastings, M. Taguchi, A. Kotani, T. Uozumi, and F. M. F. de Groot. 1s2p resonant inelastic X-ray scattering in α-Fe<sub>2</sub>O<sub>3</sub>. Physical Review B, 58, 1998.
- [8] C. Chang, D. Huang, A. Tanaka, G. Guo, S. Chung, S. Kao, S. Shyu, and C. Chen. Electronic structure of CrO<sub>2</sub> studied by magnetic circular dichroism in resonant photoemission. Physical Review B, 71(5), 2005.
- [9] F. de Groot and A. Kotani. Core Level Spectroscopy of Solids. Advances in Condensed Matter Science. CRC Press, 2008. ISBN 9781420008425.
- [10] F. M. de Groot. X-ray absorption of transition metal oxides. PhD thesis, University of Nijmegen, 1991. p. 84ff.
- [11] F. M. de Groot, J. C. Fuggle, B. T. Thole, and G. A. Sawatzky.  $L_{2,3}$  Xray-absorption edges of  $d^0$  compounds:  $K^+$ ,  $Ca^{2+}$ ,  $Sc^{3+}$ , and  $Ti^{4+}$  in  $O_h$  (octahedral) symmetry. Physical Review B, 41(2):928–937, 1990.
- [12] F. M. F. de Groot, G. Vankó, and P. Glatzel. The 1s x-ray absorption pre-edge structures in transition metal oxides. Journal of Physics: Condensed Matter, 21(10):104207–7, 2009.
- [13] Z.-Y. Deng, J.-M. Zhang, and K.-W. Xu. The electronic and magnetic properties of the Fdoped CrO<sub>2</sub> from first-principles study. Journal of Magnetism and Magnetic Materials, 379: 196–201, 2014.
- [14] J. C. Fuggle, B. T. Thole, G. A. Sawatzky, and F. M. de Groot. 2p x-ray absorption of 3d transition-metal compounds: An atomic multiplet description including the crystal field. Physical Review B, 42:5459–5468, 1990.
- [15] T. Funk, A. Deb, Simon J. George, H. Wang, and S. P. Cramer. X-ray magnetic circular dichroism—a high energy probe of magnetic properties. Coordination Chemistry Reviews, 249 (1-2), 2005.

- [16] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. Journal of Physics: Condensed Matter, 21(39):395502, 2009.
- [17] E. Goering, A. Bayer, S. Gold, G. Schütz, M. Rabe, U. Ruediger, and G. Guentherodt. Strong Anisotropy of Projected 3d Moments in Epitaxial CrO<sub>2</sub> Films. Physical Review Letters, 88, 2002.
- [18] E. Goering, M. Justen, J. Geissler, U. Ruediger, M. Rabe, G. Guentherodt, and G. Schütz. Magnetic anisotropy of textured CrO<sub>2</sub> thin films investigated by X-ray magnetic circular dichroism. Applied Physics A: Materials Science & Processing, 74(6), 2002.
- [19] C. Gougoussis, M. Calandra, A. P. Seitsonen, and F. Mauri. First-principles calculations of x-ray absorption in a scheme based on ultrasoft pseudopotentials: From α-quartz to high-Tc compounds. Physical Review B, 80(7):075102, 2009.
- [20] M. W. Haverkort, M. Zwierzycki, and O. K. Andersen. Multiplet ligand-field theory using wannier orbitals. *Physical Review B*, 85:165113, 2012.
- [21] M. W. Haverkort, G. Sangiovanni, P. Hansmann, A. Toschi, Y. Lu, and S. Macke. Bands, resonances, edge singularities and excitons in core level spectroscopy investigated within the dynamical mean-field theory. *EPL (Europhysics Letters)*, 108(5):57004, 2014.
- [22] D. Huang, H. Jeng, C. Chang, G. Guo, J. Chen, W. Wu, S. Chung, S. Shyu, C. Wu, H. Lin, and C. Chen. Orbital magnetic moments of oxygen and chromium in CrO<sub>2</sub>. Physical Review B, 66(17), 2002.
- [23] D. Huang, C. Chang, J. Chen, H. Lin, S. Chung, H. Jeng, G. Guo, W. Wu, S. Shyu, and C. Chen. Orbital moments of CrO<sub>2</sub> and Fe<sub>3</sub>O<sub>4</sub> studied by MCD in soft X-ray absorption. Journal of Electron Spectroscopy and Related Phenomena, 137-140:633-639, 2004.
- [24] D. J. Huang, L. Tjeng, J. Chen, C. F. Chang, W. Wu, S. C. Chung, A. Tanaka, G. Y. Guo, H. J. Lin, S. G. Shyu, C. Wu, and C. T. Chen. Anomalous spin polarization and dualistic electronic nature of CrO2. Physical Review B, 67(21), 2003.
- [25] H. Ikeno, T. Mizoguchi, and I. Tanaka. Ab initio charge transfer multiplet calculations on the L<sub>2,3</sub> XANES and ELNES of 3d transition metal oxides. Physical Review B, 83(15):155107–14, 2011.
- [26] A. Juhin, F. M. de Groot, G. Vankó, M. Calandra, and C. Brouder. Angular dependence of core hole screening in LiCoO<sub>2</sub>: A DFT+U calculation of the oxygen and cobalt K-edge x-ray absorption spectra. Physical Review B, 81(11):115115-11, 2010.
- [27] A. Juhin, A. López-Ortega, M. Sikora, C. Carvallo, M. Estrader, S. Estradé, F. Peiró, M. D. Baró, P. Sainctavit, P. Glatzel, and J. Nogués. Direct evidence for an interdiffused intermediate layer in bi-magnetic core-shell nanoparticles. Nanoscale, 6:11911–11920, 2014.
- [28] V. Kanchana, G. Vaitheeswaran, and M. Alouani. Electronic structure and X-ray magnetic circular dichroism of CrO<sub>2</sub>. Journal of Physics: Condensed Matter, 18(22):5155–5162, 2006.

- [29] M. I. Katsnelson, V. Y. Irkhin, L. Chioncel, A. I. Lichtenstein, and R. A. de Groot. Halfmetallic ferromagnets: From band structure to many-body effects. Review of Modern Physics, 80(2):315–378, 2008.
- [30] M. A. Korotin, V. I. Anisimov, D. I. Khomskii, and G. A. Sawatzky. CrO2: A Self-Doped Double Exchange Ferromagnet. Physical Review Letter, 80(19):4305–4308, 1998.
- [31] A. Kotani. Resonant inelastic X-ray scattering and its magnetic circular dichroism. Journal of Physics and Chemistry of Solids, pages 2150–2156, 2005.
- [32] E. Kurmaev, A. Moewes, S. Butorin, M. Katsnelson, L. Finkelstein, J. Nordgren, and P. Tedrow. *Half-metallic electronic structure of CrO2in resonant scattering*. *Physical Review* B, 67(15):155105–4, 2003.
- [33] S. P. Lewis, P. B. Allen, and T. Sasaki. Band structure and transport properties of CrO<sub>2</sub>. Physical Review B, 55:10253–10260, 1997. also: arXiv:mtrl-th/9608006.
- [34] Y. Lu, M. Höppner, O. Gunnarsson, and M. W. Haverkort. Efficient real-frequency solver for dynamical mean-field theory. *Physical Review B*, 90:085102, 2014.
- [35] B. Maddox, C. Yoo, D. Kasinathan, W. Pickett, and R. Scalettar. High-pressure structure of half-metallic CrO<sub>2</sub>. Physical Review B, 73:144111, 2006.
- [36] I. I. Mazin, D. J. Singh, and C. Ambrosch-Draxl. Transport, optical, and electronic properties of the half-metal CrO<sub>2</sub>. Physical Review B, 59:411–418, 1999.
- [37] K. Momma and F. Izumi. VESTA3 for three-dimensional visualization of crystal, volumetric and morphology data. Journal of Applied Crystallography, 44(6):1272–1276, 2011.
- [38] S. Ould-Chikh, O. Proux, P. Afanasiev, L. Khrouz, M. N. Hedhili, D. H. Anjum, M. Harb, C. Geantet, J.-M. Basset, and E. Puzenat. *Photocatalysis with Chromium-Doped TiO<sub>2</sub>: Bulk* and Surface Doping. ChemSusChem, 7(5):1361–1371, 2014.
- [39] P. W. Anderson and H. Hasegawa. Considerations on Double Exchange. Physical Review, 100: 675–681, 1955.
- [40] M. Pathak. Half-metallic CrO<sub>2</sub> thin films for spintronic applications. PhD thesis, University of Alabama, 2011.
- [41] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. Physical Review Letters, 77:3865–3868, 1996.
- [42] P. Porta, M. Marezio, J. Remeika, and P. Dernier. Chromium dioxide: High pressure synthesis and bond lengths. Materials Research Bulletin, 7(2):157–161, 1972.
- [43] S. L. Reddy, T. Endo, and G. S. Reddy. Electronic (Absorption) Spectra of 3d Transition Metal Complexes, 2012. INTECH Open Access Publisher http: //cdn.intechopen.com/pdfs/38537/InTech-Electronic\_absorption\_spectra\_of\_3d\_ transition\_metal\_complexes.pdf.
- [44] M. Rovezzi and P. Glatzel. Hard X-ray emission spectroscopy: a powerful tool for the characterization of magnetic semiconductors. Semicond. Sci. Technol., 29(2), 2014.
- [45] J.-P. Rueff, J. M. Ablett, D. Céolin, D. Prieur, T. Moreno, V. Balédent, B. Lassalle-Kaiser, J. E. Rault, M. Simon, and A. Shukla. *The GALAXIES beamline at the SOLEIL synchrotron: inelastic X-ray scattering and photoelectron spectroscopy in the hard X-ray range. Journal of Synchrotron Radiation*, 22(1):175–179, 2015.

- [46] P. Schlottmann. Double-exchange mechanism for CrO2. Physical Review B, 67(17):174419–7, 2003.
- [47] P. Schlottmann. Spin exchange in CrO<sub>2</sub>: the role of a localized level. Journal of Magnetism and Magnetic Materials, 272-276:553–554, 2004.
- [48] P. Schlottmann. Double exchange and charge fluctuations in CrO2. Journal of Applied Physics, 95(11):7471–7473, 2004.
- [49] Schütz, G and Fischer, P and Attenkofer, K and Knülle, M and Ahlers, D and Stähler, S and Detlefs, C and Ebert, H and de Groot, Frank M F. X-ray magnetic circular dichroism in the near and extended absorption edge structure (invited). Journal of Applied Physics, 76(10): 6453-8, 1994.
- [50] Sigma-Aldrich. Magtrieve supplementary information, 2015. http://www.sigmaaldrich. com/catalog/product/aldrich/480037.
- [51] M. Sikora, A. Juhin, T.-C. Weng, P. Sainctavit, C. Detlefs, F. de Groot, and P. Glatzel. Strong K-edge Magnetic Circular Dichroism Observed in Photon-in-Photon-out Spectroscopy. Physical Review Letters, 105(3), 2010.
- [52] M. Sikora, A. Juhin, G. Simon, M. Zajac, C. Kapusta, L. Morellon, M. Ibarra, and P. Glatzel. 1s2p Resonant inelastic X-ray scattering-magnetic circular dichroism: A sensitive probe of 3d magnetic moments using hard x-ray photons. Journal of Applied Physics, 111(7), 2012.
- [53] G. P. Singh, S. Ram, J. Eckert, and H.-J. Fecht. Synthesis and morphological stability in CrO<sub>2</sub> single crystals of a half-metallic ferromagnetic compound. Journal of Physics: Conference Series, 144:012110–012116, 2009.
- [54] I. Solovyev, I. Kashin, and V. Mazurenko. Mechanisms and origins of half-metallic ferromagnetism in CrO2. Physical Review B, 92(14):144407–15, 2015.
- [55] E. Stavitski and F. M. de Groot. The CTM4XAS program for EELS and XAS spectral shape analysis of transition metal L edges. Micron, 41(7):687–694, 2010.
- [56] M. Taillefumier, D. Cabaret, A.-M. Flank, and F. Mauri. X-ray Absorption Near-Edge Structure calculations with pseudopotentials. Application to K-edge in diamond and alpha-quartz. Physical Review B, 66(19), 2002.
- [57] B. T. Thole, G. Van Der Laan, and P. H. Butler. Spin-mixed ground state of Fe phthalocyanine and the temperature-dependent branching ratio in X-ray absorption spectroscopy. Chemical Physics Letters, 149(3):295–299, 1988.
- [58] A. Toropova, G. Kotliar, S. Y. Savrasov, and V. S. Oudovenko. Electronic structure and magnetic anisotropy of CrO2. Physical Review B, 71(17):172403-4, 2005.
- [59] M. Torres Deluigi, F. M. de Groot, G. López-Díaz, G. Tirao, G. Stutz, and J. Riveros de la Vega. Core and Valence Structures in K<sub>β</sub> X-ray Emission Spectra of Chromium Materials. Journal of Physical Chemistry C, 118(38):22202–22210, 2014.
- [60] D. Tripathy and A. Adeyeye. Electronic properties of field aligned CrO2 powders. Physica B: Condensed Matter, 368(1-4):131–138, 2005.
- [61] N. Troullier and J. L. Martins. Efficient pseudopotentials for plane-wave calculations. Physical Review B, 43:1993–2006, 1991.

[62] G. Vankó, F. M. de Groot, S. Huotari, R. Cava, T. Lorenz, and M. Reuther. Intersite 4p-3d hybridization in cobalt oxides: a resonant x-ray emission spectroscopy study. arXiv:0802.2744, pages 1–7, 2008.

## Abstract Résumé

**Keywords:** X-ray absorption near-edge spectrocopy, X-ray circular dichroism, DFT calculations, Magnetism, Transition metals, High-pressure

The main purpose of this thesis was to compute X-ray magnetic circular dichroism spectra at the K-edge in order to provide a tool to interpret the, so far very puzzling, experimental spectra. Computation of circular dichroism requires precise calculations of X-ray absorption spectra (XAS) for circularly polarized light. We have found that there is an incompatibility of the semi-classical time-dependent perturbation theory commonly used to calculate light absorption and scattering cross-sections with both gauge invariance and semi-relativistic descriptions of the electron dynamics. The problems are solved by applying a Foldy-Wouthuysen transformation to the fully relativistic cross-sections given by quantum electrodynamics. In the process, a new light-matter interaction term emerges, that we named the "spin-position" interaction. An efficient first-principles approach was developed to compute the absorption cross-section in order to obtain X-ray magnetic circular dichroism (XMCD) and X-ray natural circular dichroism (XNCD). The numerical method relies on density-functional theory with plane waves and pseudopotentials. We find that the term coupling the electric dipole operator with the spinposition operator contributes significantly to the XMCD at the K-edge of ferromagnetic iron, cobalt, and nickel. We obtain a sum rule relating this term to the spin magnetic moment of the p states. We also applied the method to calculations of K-edge XMCD in FeH and  $CrO_2$ . In both cases, the combination of experiment and theory leads to mutual enrichment.

**Mots-clés:** Spectroscopie d'absorption près du seuil d'absorption de rayons X, Dichroïsme circulaire de rayons X, Calculs DFT, Magnétisme, Métaux de transition, Hautes-pressions

Le but principal de cette thèse était de calculer les spectres de dichroïsme circulaire magnétique de rayons X au seuil K afin de fournir un outil pour interpréter les spectres expérimentaux, jusqu'ici très déroutants. La détermination du dichroïsme circulaire nécessite le calcul précis des spectres d'absorption des rayons X polarisés circulairement. Nous avons constaté que la théorie des perturbations semi-classique dépendante du temps, communément utilisée pour calculer les sections efficaces d'absorption et de diffusion, est incompatible à la fois, avec l'invariance de jauge et avec les descriptions semi-relativistes de la dynamique des électrons. Pour résoudre ces problèmes, on applique une transformation de Foldy-Wouthuysen aux sections efficaces relativistes données par l'électrodynamique quantique. Ainsi, un nouveau terme d'interaction lumière-matière émerge, que nous avons appelé "spin-position". Une approche performante a été développée pour calculer la section efficace d'absorption afin d'obtenir le dichroïsme circulaire magnétique de rayons X (XMCD) et le dichroïsme circulaire naturel de rayons X (XNCD). La méthode numérique repose sur la théorie de la fonctionnelle de la densité en ondes planes avec des pseudopotentiels. Nous constatons que le terme couplant l'opérateur dipolaire électrique avec l'opérateur spin-position contribue significativement au XMCD au seuil K du fer, du nickel et du cobalt ferromagnétiques et nous l'expliquons grâce aux règles de somme. Nous avons également appliqué la méthode aux calculs du XMCD dans FeH et  $CrO_2$ . Dans les deux cas, la combinaison de l'expérience et de la théorie conduit à un enrichissement mutuel.