UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1558

Relativistic theory of laser-induced magnetization dynamics

RITWIK MONDAL

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017

ISSN 1651-6214
ISBN 978-91-513-0070-2
urn:nbn:se:uu:diva-315247

Dissertation presented at Uppsala University to be publicly examined in Polhemsalen, Lägerhyddsvägen 1, Uppsala, Friday, 27 October 2017 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Paul-Antoine Hervieux (Université de Strasbourg).

Abstract

Mondal, R. 2017. Relativistic theory of laser-induced magnetization dynamics. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1558. 115 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0070-2.

Ultrafast dynamical processes in magnetic systems have become the subject of intense research during the last two decades, initiated by the pioneering discovery of femtosecond laser-induced demagnetization in nickel. In this thesis, we develop theory for fast and ultrafast magnetization dynamics. In particular, we build relativistic theory to explain the magnetization dynamics observed at short timescales in pump-probe magneto-optical experiments and compute from first-principles the coherent laser-induced magnetization.

In the developed relativistic theory, we start from the fundamental Dirac-Kohn-Sham equation that includes *all* relativistic effects related to spin and orbital magnetism as well as the magnetic exchange interaction and any external electromagnetic field. As it describes both particle and antiparticle, a separation between them is sought because we focus on low-energy excitations within the particle system. Doing so, we derive the extended Pauli Hamiltonian that captures all relativistic contributions in first order; the most significant one is the full spin-orbit interaction (gauge invariant and Hermitian). Noteworthy, we find that this relativistic framework explains a wide range of dynamical magnetic phenomena. To mention, (i) we show that the phenomenological Landau-Lifshitz-Gilbert equation of spin dynamics can be rigorously obtained from the Dirac-Kohn-Sham equation and we derive an exact expression for the tensorial Gilbert damping. (ii) We derive, from the gauge-invariant part of the spin-orbit interaction, the existence of a relativistic interaction that linearly couples the angular momentum of the electromagnetic field and the electron spin. We show this spin-photon interaction to provide the previously unknown origin of the angular magneto-electric coupling, to explain *coherent ultrafast magnetism*, and to lead to a new torque, the optical spin-orbit torque. (iii) We derive a definite description of magnetic inertia (spin nutation) in ultrafast magnetization dynamics and show that it is a higher-order spin-orbit effect. (iv) We develop a unified theory of magnetization dynamics that includes spin currents and show that the nonrelativistic spin currents naturally lead to the current-induced spin-transfer torques, whereas the relativistic spin currents lead to spin-orbit torques. (v) Using the relativistic framework together with *ab initio* magneto-optical calculations we show that relativistic laser-induced spin-flip transitions do not explain the measured large laser-induced demagnetization.

Employing the *ab initio* relativistic framework, we calculate the amount of magnetization that can be imparted in a material by means of circularly polarized light – the so-called inverse Faraday effect. We show the existence of both spin and orbital induced magnetizations, which surprisingly reveal a different behavior. We establish that the laser-induced magnetization is *antisymmetric* in the light's helicity for nonmagnets, antiferromagnets and paramagnets; however, it is only *asymmetric* for ferromagnets.

Keywords: Relativistic quantum electrodynamics, magneto-optics, spin-orbit coupling, ultrafast demagnetization, inverse Faraday effect, magnetic inertia, Gilbert damping

Ritwik Mondal, Department of Physics and Astronomy, Materials Theory, Box 516, Uppsala University, SE-751 20 Uppsala, Sweden.

© Ritwik Mondal 2017

ISSN 1651-6214
ISBN 978-91-513-0070-2
urn:nbn:se:uu:diva-315247 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-315247)

Dedicated to my parents

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **Ab initio investigation of light-induced relativistic spin-flip effects in magneto-optics**
Ritwik Mondal, Marco Berritta, Karel Carva, Peter M. Oppeneer
Phys. Rev. B **91**, 174415 (2015)

II **Relativistic interaction Hamiltonian coupling the angular momentum of light and the electron spin**
Ritwik Mondal, Marco Berritta, Charles Paillard, Surendra Singh, Brahim Dkhil, Peter M. Oppeneer, Laurent Bellaiche
Phys. Rev. B **92**, 100402(R) (2015) - *editor's suggestion*

III **New relativistic Hamiltonian: the Angular MagnetoElectric coupling** (invited paper)
Charles Paillard, Ritwik Mondal, Marco Berritta, Brahim Dkhil, Surendra Singh, Peter M. Oppeneer, Laurent Bellaiche
Proc. SPIE **9931**, 99312E (2016)

IV **Relativistic theory of spin relaxation mechanisms in the Landau-Lifshitz-Gilbert equation of spin dynamics**
Ritwik Mondal, Marco Berritta, Peter M. Oppeneer
Phys. Rev. B **94**, 144419 (2016)

V **Ab initio Theory of Coherent Laser-Induced Magnetization in Metals**
Marco Berritta, Ritwik Mondal, Karel Carva, Peter M. Oppeneer
Phys. Rev. Lett. **117**, 137203 (2016)

VI **Signatures of relativistic spin-light coupling in magneto-optical pump-probe experiments**
Ritwik Mondal, Marco Berritta, Peter M. Oppeneer
J. Phys.: Condens. Matter **29**, 194002 (2017) - *JPhys+ News Views*

VII **Relativistic theory of magnetic inertia in ultrafast spin dynamics**
Ritwik Mondal, Marco Berritta, Ashis K. Nandy, Peter M. Oppeneer
Phys. Rev. B **96**, 024425 (2017) - *editor's suggestion*

VIII **Unified relativistic theory of magnetization dynamics with spin-current tensors**
Ritwik Mondal, Marco Berritta, Peter M. Oppeneer
Preprint (2017)

Articles which are co-authored by me but those are not included in this thesis

IX **Magnetization switching of FePt nanoparticle recording medium by femtosecond laser pulses**
R. John, M. Berritta, D. Hinzke, C. Müller, T. Santos, H. Ulrichs, P. Nieves, J.Walowski, R. Mondal, O. Chubykalo-Fesenko, J. McCord, P. M. Oppeneer, U. Nowak, M. Münzenberg
Sci. Rep. **7**, 4114 (2017)

Reprints were made with permission from the publishers.

Comments on my own contribution

In all the papers where I am the first author (Paper I, II, IV, VI, VII, VIII), I had the main responsibility for carrying out the work. I carried out all the analytical calculations presented, discussed with the other co-authors, and wrote the initial manuscript of those papers. Paper III is an extended version of Paper II, and there I was mostly involved in the original derivation and the correction of the manuscript. For Paper V, I was involved in all the discussions and contributed to the analysis of the calculations.

Contents

Part I:
Introduction

1. Introduction

The ever increasing amount of human-made digital information requires novel technological solutions that can be summarized in two main keywords *smaller* and *faster*. *Smaller* refers to the area of a magnetic bit, which needs to be minimal to achieve high-density information storage, while *faster* refers to the speed with which huge data amounts can be processed. Early data storage devices were slow, large in size, expensive and had a small storage capacity. For example, the first computer hard disk drives in 1956 used 50 discs each of 24 inches, however, the total data storage was only up to 5 MB [1]. Over the last few decades data storage technology has been proceeding according to Moore's law prediction which states that the area of a single storage element decreases by a factor of two in every 18 months [2].

In electronic storage devices the information storage is commonly manipulated by an electric current, which, due to unavoidable resistivity losses, causes Joule heating and thus puts a limit on achieving smaller devices. Conversely, magnetic storage devices are non-volatile and based on fundamentally different principles. The latter exploit the magnetic "spin", which can be viewed as a small vector that is either pointing up or down just like in logic bits ("one" and "zero"). In order to record digital data, one has then to reverse the spin from up to down or vice-versa. Now the important, open question is: how fast can the spins be reversed efficiently? One possible way is to use a magnetic field pulse in the opposite direction of the spins alignment [3, 4]. It is known that the fastest reversal can be obtained by the precession of spins, where the angular precession frequency, the Larmor frequency, is proportional to the magnitude of the magnetic field [5–8]. If the magnetic field is infinitely strong, the switching will be quickest, however, restricted to the magnetic field pulse duration. It has been shown that if the magnetic field pulse width is less than 2 picoseconds (ps), the precession leads to unwanted *non-deterministic* spin switching [9] which effectively sets an upper limit to the reachable speed of controllable magnetization reversal.

Another route to achieve fast switching could be to use optical laser pulses instead of magnetic field pulses, because laser pulses are the fastest man-made events that can have pulse durations of femtoseconds (fs) down to attoseconds (10^{-18} s) [10]. Other important properties of contemporary laser sources are: monochromaticity (single frequency), coherency (no phase separation), and collimation (do not diverge over long distances). The very first observation towards ultrafast modification of magnetic information was done in 1996, when Beaurepaire and co-workers used femtosecond laser pulses and demonstrated

laser-induced ultrafast demagnetization within a timescale of less than a ps in ferromagnetic nickel [11]. They used a 60 fs laser pulse on an about 20 nm thick nickel sample and measured the magneto-optical Kerr effect (MOKE) signal in a pump-probe ultrafast spectroscopy set-up. Their observation indicated that it could be possible to manipulate magnetization with femtosecond laser pulses at previously unthought, short timescales. Their discovery initiated moreover the birth of a new branch of solid-state physics, that of *ultrafast magnetism*, which has since then emerged as a burgeoning field of much current interest. Thenceforth many groups have begun to probe the ultrafast magnetization dynamics using ultrashort laser pulse sources. Additionally, the understanding of the fundamental mechanism of ultrafast laser-induced demagnetization has become a topic of intense debates. A few important questions that have been raised in the on-going discussion can be summarized as: magnetism or optics or opto-magnetism? [12–21]. At the same time the microscopic channel that can provide an ultrafast transfer of spin angular momentum has been hotly debated [22–26]. Understandably, the outcome of this debate can have far-reaching consequences for the development of the next-generation ultrafast storage devices.

When the magnetic material is irradiated with a laser pulse, the absorption of laser light by the electron system leads to the rapid increase of the electron temperature, after which the received energy dissipates into the other available degrees of freedom of the system (e.g., the lattice and spins). Therefore the laser-induced magnetization changes can be of *thermal* origin, which limits reaching faster recording because a cooling time is required to reach again the initial values of the system's parameters [27]. Also, heating acts as a non-deterministic process on the magnetization direction. A possible ultrafast, non-thermal and all-optical manipulation of spins was shown by Kimel *et al.* in 2005 [28], who used circularly polarized laser pulses to excite and coherently control the spin dynamics in a magnetic insulator. In particular they deduced that circularly polarized light pulses of 200 fs could act as the equivalent of a magnetic field pulse with strength of the order of 5 Tesla, by means of a special non-linear effect, the inverse Faraday effect (IFE). A basic view on this is that, since the circularly polarized light carries angular momentum, the interaction between the light's angular momentum and spins would induce an effective magnetic field, which is opposite for right- and left-circular polarizations of the laser beam. After this first observation the helicity of light has become an important tool for opto-magnetic recording, as one could employ one helicity to write deterministically a magnetization direction and erase it with the opposite helicity, giving birth to an emerging new research field, that of All-Optical Helicity-Dependent Switching (AO-HDS) [29–33].

Laser-induced demagnetization and switching have become topics of intense discussions, because of their importance for potential applications in future memory devices. Also, in spite of the numerous experimental observations the underlying fundamental processes are not yet understood properly,

leaving many open questions yet to be answered. For example, how does the angular momentum of light interact with the electron spin? What is the magnitude of the effective induced magnetic field by the light? How does the electromagnetic field of the laser light govern the magnetization dynamics on fast and ultrafast timescales? What is the dissipation channel that leads to the removal of spin angular momentum on sub-picosecond timescales? What are the ruling parameters that determine the strength of the dynamical processes? What are the relevant expressions of these parameters? Are these *ab initio* calculable? Are these microscopic parameters related to a quantity that is experimentally measurable? To answer these aforementioned questions is sometimes phenomenologically possible, however, their understanding from a fundamental theory is largely missing. For example, phenomenologically the IFE was thought to induce a magnetic field [28], yet it is not completely understood whether the IFE induces a magnetic field or magnetization. If it is an induced magnetization, is it only the spins that contribute or also laser-induced orbital magnetic effects to be considered, too?

A further, related field of current interest is that of magnetization dynamics [34–36]. Traditionally, the time-evolution of a magnetization distribution in a small volume element is described by the phenomenological Landau-Lifshitz equation [37–40], which forms the foundation of contemporary micromagnetic simulations. The Landau-Lifshitz equation, and its extension, the Landau-Lifshitz-Gilbert equation, which typically are valid on nanosecond timescales, describe the Larmor precession of the local magnetization around an effective magnetic field as well as the slow damping of the magnetization until it is aligned along the magnetic field vector. Although both equations are extensively used for spin dynamics simulations, they are nevertheless purely phenomenological, i.e., without an existing derivation from fundamental principles. Similarly, the important Gilbert damping parameter is often considered to be of relativistic microscopic origin, however, this has not been proven explicitly from relativistic theory and it is not yet unambiguously established what the microscopic expression for the Gilbert damping is. Furthermore, there has recently been interest in magnetic inertial dynamics that might contribute to ultrafast deterministic switching [41]. The effect of magnetic inertia can be phenomenologically introduced as a torque due to a second-order time-derivative of the magnetization [41, 42]. In spite of potential applications of inertia, the fundamental origin of magnetic inertial dynamics and the corresponding expression for the inertia parameter are still unknown.

On a similar note, additional terms have recently been added to the Landau-Lifshitz-Gilbert equation to capture phenomenologically the effects of spin transport [43–45]. In typical magnetic metallic heterostructures that are used in magnetic random-access memories (MRAM) there are nonmagnetic metallic layers and ferromagnetic layers [46–48]. The spins of the electrons in the non-ferromagnet are randomly oriented whereas the spins in the ferromagnet point overall in the same direction. When an electric current is applied to

the metallic heterostructure, electrons from the non-ferromagnet enter into the ferromagnet and can become spin-polarized during the transfer and thus generate a spin-polarized current flow. The interaction of the spin-polarized current with the ferromagnetic magnetization of a further layer provides a spin-transfer torque which can be employed to switch the magnetization of this "free" layer [49–51]. Another spin torque which has recently drawn much attention is the spin-orbit torque, where a relativistic effect, the spin Hall effect, generates a spin accumulation in a nonmagnetic layer that can then influence the orientation of the magnetization of an adjacent ferromagnetic layer [45, 52–54]. The various spin torques play obviously paramount roles in nanoscale spintronic devices [55–58]. It has recently been shown that laser-induced spin currents can be used, too, to manipulate ultrafast the magnetization in metallic heterostructures [59–61]. In spite of these recent successes, the origin of spin torques from a fundamental theoretical framework is missing in the literature.

Along the same line of reasoning, a phenomenologically predicted coupling of the angular momentum density of light and the magnetic moment could explain several interesting phenomena [62]. This coupling, called angular magneto-electric coupling, is based on the understanding that in the presence of a ferro-toroidic order in a multiferroic material, the toroidic dipole moment (moment of magnetization, expressed as $\boldsymbol{T} = \frac{1}{2}\int \boldsymbol{r} \times \boldsymbol{M}\, d\boldsymbol{r}$ with magnetization $\boldsymbol{M}$ and position $\boldsymbol{r}$) can be controlled by crossed electric and magnetic fields [63]. Such a coupling term has been suggested to be important for the anomalous Hall effect, planar Hall effect, and anisotropic magnetoresistance in ferromagnets [64, 65]. However, the mere existence of such coupling term from fundamental principles has not yet been shown.

This thesis

In the present thesis we establish the missing fundamental origins of several above-mentioned effects and equations from the Dirac theory, specifically, from the Dirac-Kohn-Sham framework. The reason for this is that the Dirac equation is both compact and foundational, as it contains all the information needed to describe a relativistic particle (and antiparticle) with spin-$\frac{1}{2}$. As we treat in particular magnetic materials, we also include the effect of the magnetic exchange interaction as an effective Kohn-Sham mean field. Although this might appear as an approximation the underlying Density-Functional Theory (DFT) [66–69] in fact provides a formally exact mapping of the ground state of the correlated many-particle problem onto that of a single particle moving in an effective mean field. To account for the light-induced relativistic effects in a magnetic solid we consider the effect of the laser light (an electromagnetic field) through the change in momentum by means of the so-called minimal coupling. The considered Dirac-Kohn-Sham Hamiltonian is a 4×4

matrix describing both the particles and antiparticles; to examine in detail its low-energy excitations, however, the separation of particles and antiparticles is mandatory for any given momentum [70]. The Foldy-Wouthuysen transformation has been a successful way to achieve this [71], however, it was previously only done *without* accounting for the exchange field. Performing the Foldy-Wouthuysen transformation we derive a Hamiltonian which describes the relativistic particles with spin-$\frac{1}{2}$ and has terms similar to those of the nonrelativistic Pauli Hamiltonian, but includes as well all the relativistic corrections. The most important first order relativistic correction is the spin-orbit coupling, which, in its full form, is shown to be gauge invariant and Hermitian. The other derived relativistic corrections include the Darwin term (independent of spin), the relativistic mass correction and previously unknown relativistic corrections to the exchange interaction. The higher-order relativistic correction terms give higher-order contributions to the spin-orbit coupling, which, as we show, also play an interesting role for spin dynamics on ultrafast timescales.

Once we have performed the Foldy-Wouthuysen transformation and derived all terms, relativistic and nonrelativistic, of the effective single-particle Hamiltonian, we proceed to investigate their effects in, and their contributions to, ultrafast magneto-optics, magnetization dynamics on fast and ultrafast timescale, spin torques, the inverse Faraday effect, and the relativistic spin-photon, or angular magneto-electric coupling. Therefore, the sequence of the topics addressed in this thesis is as follows:

Chapter 2 discusses the most general relativistic quantum formulation, starting from the Dirac equation. The separation of particle and antiparticle is discussed extensively within the Foldy-Wouthuysen transformation. The discrepancies of the present transformation and other methods are pointed out and discussed in detail. To this end, we treat the Dirac-Kohn-Sham Hamiltonian for the magnetic system of our interest and we derive the extended Pauli Hamiltonian with relativistic corrections up to the second order in the small quantity $\frac{1}{c^4}$, where c is the speed of light in vacuum.

Chapter 3 discusses magnetization dynamics in the most general relativistic formulation. Within the Heisenberg picture, the dynamical equation of motion of a local magnetization (or spin) is derived from the Hamiltonian obtained in Chapter 2. In particular, we focus on the origin of the Gilbert damping parameter and derive, within the Dirac-Kohn-Sham theory, an exact new expression for tensorial Gilbert damping, which can easily be calculated in an *ab initio* framework. The Gilbert damping tensor is shown to contain, in general, an isotropic Heisenberg-like contribution, an anisotropic Ising-like contribution, and a chiral, or Dzyaloshinskii-Moriya-like contribution. In addition, we show that the magnetic inertial dynamics originates from a higher order contribution to the spin-orbit coupling, and that it is therefore expected to play a role only at ultrashort timescales. Utilizing the same platform, we derive the exact spin current contributions that arise from the nonrelativistic

and relativistic Hamiltonians as well as the spin torques generated by them in a generalized theory for magnetization dynamics and spin transport.

Chapter 4 deals with the effects that the derived relativistic Hamiltonian terms have on laser-induced ultrafast demagnetization processes. We start with discussing the various proposed mechanisms for ultrafast laser-induced demagnetization, where one of these is that demagnetization is due to relativistic laser-induced spin-flip processes. We implement the derived ultra-relativistic terms in the conductivity tensor in a Kubo linear-response theory formulation and quantify the difference in the optical conductivity and MOKE spectra while treating the nonrelativistic as well as all the relativistic terms. Moreover, we discuss the role of relativistic effects in coherent ultrafast magnetism and show that the newly derived relativistic spin-photon coupling can explain the measured coherent ultrafast magnetism [22].

Chapter 5 discusses the theoretical understanding of the inverse Faraday effect. We start with a classical description, followed by the quantum description. Within the latter description, we calculate the laser-imparted orbital and spin magnetizations for a wide range of materials (nonmagnetic, ferromagnetic, synthetic antiferromagnetic) within an *ab initio* framework. To end with, we discuss the spin-photon or angular magneto-electric coupling. First, we establish a new relativistic Hamiltonian which describes the coupling of the light's angular momentum to the electron's spin, and thereby, to the magnetic moment. We show, furthermore, that the newly derived Hamiltonian can contribute as well to the inverse Faraday effect.

These main chapters will be followed by further chapters as summary and outlook, a summary in Swedish (populärvetenskaplig sammanfattning på svenska) and acknowledgements.

Part II:
Summary of the results

2. Relativistic Hamiltonian Formulation

2.1 Introduction

Quantum theory has been the key to understand the fundamental properties and dynamics of subatomic particles e.g., electrons, protons, neutrons etc. The fundamental description of quantum physics starts with the Schrödinger equation [72, 73] which describes a nonrelativistic quantum particle very efficiently. It is effectively a single particle theory and the corresponding wave function bears all the information of that particle. However, there are few discrepancies in the Schrödinger description: (a) In the quantum nature of a particle, the concept of "spin" was missing in the Schrödinger equation. (b) As the atomic number increases in the periodic table, the nuclear diameter is no longer negligible and the approximation of potential as $1/r$, the Coulomb potential, is not valid anymore. (c) If the particle velocity is comparable to the speed of light, c, relativity comes into play. (d) Schrödinger theory is not Lorentz covariant as it involves different orders of space and time derivatives. (e) Creation and annihilation of particles are not allowed in the Schrödinger theory as the integration of probability density of the particle all over the space is unity.

Thereafter, Gordon [74] and Klein [75] attempted to derive a relativistic wave equation to describe a relativistic particle where the difficulties are eliminated, if not all. The obvious choice was to use the relativistic energy-momentum relation $E^2 = p^2c^2 + m^2c^4$, where $\boldsymbol{p}$ is the particle momentum, m and E define the mass and the energy of the particle, for a Lorentz covariant theory [76–78]. Then the classical energy and momentum operators are replaced by the their corresponding quantum operator as $E \to i\hbar\frac{\partial}{\partial t}$ and $\boldsymbol{p} \to -i\hbar\boldsymbol{\nabla}$ and thus we obtain the Klein-Gordon (KG) equation

$$\left(\partial_\mu \partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\psi = 0\,, \tag{2.1}$$

for a four-component wave function, ψ, and four-component space-time position operator $r^\mu = (r^0 = ct, \boldsymbol{r})$ with the definition $\partial_\mu = \frac{\partial}{\partial r^\mu}$. The free particle solution of the KG equation takes the form of plane waves, $e^{i(Et-\boldsymbol{p}\cdot\boldsymbol{r})/\hbar}$, thus both positive and negative energies are involved as $E = \pm\sqrt{p^2c^2 + m^2c^4}$. This essentially leads to the transition of particles from positive to low-lying negative energy states. As the KG equation contains the second-order time-derivative, the corresponding probability density is not positive definite (see

the book by J. J. Sakurai [79] or by P. Strange [70] for details), rather for a strong potential the KG probability density can become negative which is absurd. Moreover, once again the concept of particle "spin" was not taken into account in the KG relativistic theory. In accordance with the Schrödinger and the KG theory, therefore, we have to look for a relativistic equation with first order time-derivative yet including the spin. In 1928 Dirac [80] derived such an equation and in 1930 he re-explained the positive and negative energy states in the KG theory will be occupied by the particles and antiparticles respectively [81] such that the creation and annihilation of particles are allowed - which is the main theme in many-particle theory.

2.2 Dirac theory

The KG theory was interpreted properly by Dirac. However to include the particle spin, the search for a new relativistic equation led to the Dirac equation [80, 82] which is first order in the time-derivative and according to Lorentz covariance the spatial derivative has to be first-order as well. The main objective of Dirac was to derive a equation that describes electrons i.e., spin-$\frac{1}{2}$ particles.

2.2.1 Towards the Dirac equation

Dirac considered once again the relativistic energy relation and put it in a different form to obtain a first-order time-derivative

$$E\psi = \sqrt{p^2c^2 + m^2c^4}\,\psi\,. \tag{2.2}$$

As the right-hand side should also be first order in space-derivative and the momentum operator is exactly the one that fulfills this condition, Dirac suggested to write the equation as

$$i\hbar\frac{\partial\psi}{\partial t} = \left[c\left(\alpha_1 p_x + \alpha_2 p_y + \alpha_3 p_z\right) + \beta mc^2\right]\psi\,, \tag{2.3}$$

where α and β are the Dirac variables. At this point, the forms of α_i and β are not known, however, they can not be scalar quantities as it would violate KG theory [83]. They can not be functions of time, t or position as all points in the space-time are equivalent [80]. To satisfy the KG equation the character of the Dirac matrices can be revealed as

$$\begin{aligned} \alpha_i^2 = \beta^2 = 1\,,\\ \alpha_i\beta + \beta\alpha_i = \{\alpha_i, \beta\}_+ = 0\,,\\ \alpha_i\alpha_j + \alpha_j\alpha_i = \{\alpha_i, \alpha_j\}_+ = 2\delta_{ij}\,. \end{aligned} \tag{2.4}$$

The condition that the Hamiltonian has to be Hermitian leads to the fact that α_i and β must be Hermitian. To construct the 4×4 matrix representation of the Dirac equation, Dirac introduced Pauli spin matrices as

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}. \tag{2.5}$$

Here the vectors are represented by the **bold** characters. $\boldsymbol{\sigma}$ are the 2×2 Pauli spin matrices, I_2 represents the 2×2 unit matrix. Finally, the Dirac equation is written in the appropriate form

$$i\hbar\frac{\partial\psi}{\partial t} = \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2\right)\psi, \tag{2.6}$$

with ψ as a four-component Dirac bi-spinor. Now we look at the success of the Dirac equation by solving the discrepancies as we stated in the beginning of the chapter. The effect of quantum spin has been included through Pauli matrices. It describes the relativistic particle. It is Lorentz covariant as is demanded from special relativity. As the Dirac equation is similar to the Schrödinger equation, the probability density is always positive definite. In the presence of an electromagnetic (EM) field, the momentum of the Dirac particle (electron) will be changed by the *minimal coupling* (gauge-invariant substitution) as $\boldsymbol{p}\to\boldsymbol{p}-e\boldsymbol{A}$, where $\boldsymbol{A}(\boldsymbol{r},t)$ is the vector potential and $e<0$ is the electronic charge.

2.2.2 Nonrelativistic limit

Even though Dirac equation describes a relativistic particle with spin efficiently, in the nonrelativistic limit it should converge to the Schrödinger equation. For a spin-$\frac{1}{2}$ Dirac particle, the four-component wave function can be written in the two-component wave function

$$\psi(\boldsymbol{r},t) = \begin{pmatrix} \Xi(\boldsymbol{r},t) \\ \xi(\boldsymbol{r},t) \end{pmatrix}. \tag{2.7}$$

The upper two-components Ξ define the positive energy solutions (spin-up and spin-down) and the lower two-components ξ are for negative energy solutions. Therefore, the 4×4 Dirac equation essentially becomes two coupled 2×2 equations as

$$i\hbar\frac{\partial\Xi(\boldsymbol{r},t)}{\partial t} = c\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\xi(\boldsymbol{r},t) + mc^2\Xi(\boldsymbol{r},t), \tag{2.8}$$

$$i\hbar\frac{\partial\xi(\boldsymbol{r},t)}{\partial t} = c\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\Xi(\boldsymbol{r},t) - mc^2\xi(\boldsymbol{r},t). \tag{2.9}$$

For a free particle at *rest*, we have two decoupled first order differential equations, which can be solved easily. However, when the particle is in motion (momentum), the positive and negative energy solutions are coupled. In the nonrelativistic limit ($c \to \infty$) the upper components, Ξ are considered as "large" and the lower components, ξ as "small". In this case as the rest energy, mc^2 is the largest energy, the free particle Dirac solutions are given as

$$\psi(\boldsymbol{r},t) \approx e^{-\frac{imc^2}{\hbar}t}\psi(\boldsymbol{r}). \tag{2.10}$$

This follows for the small components as

$$i\hbar\frac{\partial\xi(\boldsymbol{r},t)}{\partial t} \approx mc^2\xi(\boldsymbol{r},t). \tag{2.11}$$

Insert this equation back in Eq. (2.9), the lower components are approximated as

$$\xi(\boldsymbol{r},t) = \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{2mc^2}\Xi(\boldsymbol{r},t). \tag{2.12}$$

Putting this back in Eq. (2.8) to get the equation for the upper components gives

$$i\hbar\frac{\partial\Xi(\boldsymbol{r},t)}{\partial t} = \frac{1}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\Xi(\boldsymbol{r},t) + mc^2\Xi(\boldsymbol{r},t). \tag{2.13}$$

Here we use the Dirac identity: $(\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{\sigma}\cdot\boldsymbol{b}) = \boldsymbol{a}\cdot\boldsymbol{b} + i\,\boldsymbol{\sigma}\cdot(\boldsymbol{a}\times\boldsymbol{b})$, for any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. At the end we obtain

$$i\hbar\frac{\partial\Xi(\boldsymbol{r},t)}{\partial t} = \frac{p^2}{2m}\Xi(\boldsymbol{r},t) + mc^2\Xi(\boldsymbol{r},t). \tag{2.14}$$

This is an interesting result as in the nonrelativistic limit, the Dirac equation takes the form of the Schrödinger equation which is expected. However, a new rest-mass energy, mc^2 is introduced. Thus, to describe electrons, we have to redefine the energy and now in the Dirac Hamiltonian, we will use the shift in rest-mass energy as $(\beta - \mathbb{1})mc^2$ [84]. By doing so we do not lose any information about electrons, just that the energy levels are shifted.

Until now, we have not discussed the effect of a potential on a relativistic Dirac Particle. A significant difference is noticed between the Schrödinger and the Dirac theory when a particle is not *free*, e.g., the particle moves under the influence of an EM field. As we have pointed out before, the momentum will go through minimal coupling, thus Eq. (2.13) will be re-written as

$$i\hbar\frac{\partial\Xi(\boldsymbol{r},t)}{\partial t} = \frac{1}{2m}\boldsymbol{\sigma}\cdot(\boldsymbol{p}-e\boldsymbol{A})\,\boldsymbol{\sigma}\cdot(\boldsymbol{p}-e\boldsymbol{A})\,\Xi(\boldsymbol{r},t). \tag{2.15}$$

Note that the rest-mass energy term is dropped for obvious reasons. Using a similar formalism, and the exact form of the Dirac equation in the presence of

a EM field, the nonrelativistic limit can be expressed as

$$i\hbar\frac{\partial\Xi(\boldsymbol{r},t)}{\partial t}=\left[\frac{(\boldsymbol{p}-e\boldsymbol{A})^2}{2m}-\frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B}\right]\Xi(\boldsymbol{r},t)\,. \tag{2.16}$$

Here, the associated magnetic field $\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{A}$. This equation is called the nonrelativistic Schrödinger-Pauli equation while in the Schrödinger description only the first term of Eq. (2.16) is obtained. The second term in Eq. (2.16) describes the interaction of the electron's spin with the external field, which is known as the Zeeman coupling. We will come back to it later in a detailed discussion in section (2.5.1).

To summarize, in the nonrelativistic limit, the momentum of a particle becomes small compared to the rest-mass energy, then, the upper two-components of the Dirac bi-spinor describe the Pauli theory in the lowest order. However, this is not valid anymore for any given momentum of the particle. Moreover, when we go beyond the lowest order approximation, the Hamiltonian terms (e.g., spin-orbit coupling) associated with the upper components become non-Hermitian [71, 85]. Furthermore, in the Dirac theory the momentum operator is $mc\boldsymbol{\alpha}$ while in the Pauli theory the momentum operator is $\boldsymbol{p}$ in the lowest order. As the Pauli spin matrices do not commute with each other, different components of the momentum operator in the Dirac theory can not be measured simultaneously. In contrast, different components of momentum do commute in the Pauli theory and they can be measured at the same time [71]. These discrepancies create a question mark of converting four-components Dirac theory to two-components Pauli theory. In what follows, Foldy and Wouthuysen in 1950 provided a way out in separating two-components from four-components [71]. In this context, we also mention that an alternative route towards achieving the nonrelativistic limit, using exact operations, was employed by Kraft *et al.* [86]

2.3 Foldy-Wouthuysen transformation

Separating the particles (upper two-components) from antiparticles (lower two-components) is not assured in the Dirac equation because they are coupled by the off-diagonal components of the 4×4 Dirac Hamiltonian. However, the Foldy-Wouthuysen (FW) transformation provides us with a sufficient tool to separate the particles from antiparticles for any given momentum. Note that for any given momentum we can not use the notation of "large" and "small" components as we had done in the nonrelativistic limit. The essential idea of FW transformation is to find a new representation of the Dirac theory where the off-diagonal i.e., the odd operators become negligible so that the upper and lower two-components can be separated. The FW transformation has extensively been used in condensed matter physics [87–90], optics [91–93], quan-

tum field theory [94], and many others including electrodynamics [95] and gravity [96].

2.3.1 Characteristic features

The general features of the FW transformation are discussed in the following [70, 71, 97, 98]:

- It is a unitary and canonical transformation of the Dirac Hamiltonian.
- It transforms the four-component Dirac equation into two decoupled two-component equations. First two-components (upper) describe positive energy states and the other two-components (lower) describe negative energy states.
- For a free Dirac particle, the FW transformation is exact. Otherwise, the transformation must be achieved by an infinite sequence of transformations.
- For a Dirac particle in the presence of any external field, the transformation is not exact. The associated field has to be sufficiently weak to have finite numbers of terms in the sequences of transformations. For strong fields, the FW transformation is of doubtful value as the series will be poorly convergent to incorporate higher order terms.
- In the nonrelativistic limit, the first part of the FW transformed Hamiltonian resembles Pauli Hamiltonian and the rest can be identified as the relativistic corrections of the order $1/c^2$ or more.

2.3.2 Original FW transformation

The Dirac Hamiltonian for a relativistic particle is given as [80]

$$\mathscr{H} \quad = \quad (\beta - \mathbb{1})mc^2 + \mathscr{O} + \mathscr{E}\,. \tag{2.17}$$

The second term in the Hamiltonian is odd as it contains only the off-diagonal components in the matrix representation and $\mathscr{E}$ represents all the even terms i.e., the diagonal components. Although, the Hamiltonian described in Eq. (2.6) does not have even terms because of being a free particle, but as soon as the Dirac particle experiences an external field, the even terms will be important. We will discuss this further in Sec. 2.5. We have deliberately used the shifted energy in order to describe electrons within the Pauli theory [99].

Investigating the commutation relation between the operators β and $\mathscr{O}$, it is observed that they anticommute with each other i.e., $\beta\mathscr{O} = -\mathscr{O}\beta$. However, β and $\mathscr{E}$ commute with each other i.e., $\beta\mathscr{E} = \mathscr{E}\beta$.

Now, we seek for such a transformation that makes the odd parts smaller and smaller as we move up higher orders. In a consequence, the transformation decouples the upper and lower components of the Dirac bispinor, ψ. Following a unitary and canonical transformation, the transformation of the bispinor

takes the form $\psi'(\boldsymbol{r},t) = e^{iU_{\mathrm{FW}}}\psi(\boldsymbol{r},t)$ with the operator U_{FW} which has to be Hermitian [71]. By doing so, we note that the probability density remains the same $|\psi(\boldsymbol{r},t)|^2 = |\psi'(\boldsymbol{r},t)|^2$. U_{FW} represents the unitary operator which has to be chosen suitably. In general, U_{FW} can be time-dependent as is the case for our formalism through the presence of the magnetic vector potential $\boldsymbol{A}(\boldsymbol{r},t)$. Following the Dirac equation in Eq. (2.6) we can write down

$$i\hbar\frac{\partial\psi}{\partial t} = \mathscr{H}\psi\,. \tag{2.18}$$

The left-hand and right-hand sides can be expanded using the transformed bispinor as

$$\begin{aligned} i\hbar e^{-iU_{\mathrm{FW}}}\frac{\partial\psi'}{\partial t} + i\hbar\frac{\partial e^{-iU_{\mathrm{FW}}}}{\partial t}\psi' &= \mathscr{H}e^{-iU_{\mathrm{FW}}}\psi' \\ \Rightarrow i\hbar e^{-iU_{\mathrm{FW}}}\frac{\partial\psi'}{\partial t} &= \left(\mathscr{H}e^{-iU_{\mathrm{FW}}} - i\hbar\frac{\partial e^{-iU_{\mathrm{FW}}}}{\partial t}\right)\psi'\,. \end{aligned} \tag{2.19}$$

Multiplying both sides by $e^{iU_{\mathrm{FW}}}$ from the left-hand side and using the property of unitary operators, we arrive at

$$\begin{aligned} i\hbar\frac{\partial\psi'}{\partial t} &= e^{iU_{\mathrm{FW}}}\left(\mathscr{H}e^{-iU_{\mathrm{FW}}} - i\hbar\frac{\partial e^{-iU_{\mathrm{FW}}}}{\partial t}\right)\psi' \\ &= \left[e^{iU_{\mathrm{FW}}}\left(\mathscr{H} - i\hbar\frac{\partial}{\partial t}\right)e^{-iU_{\mathrm{FW}}} + i\hbar\frac{\partial}{\partial t}\right]\psi' \\ &\equiv \mathscr{H}_{\mathrm{FW}}\psi'\,. \end{aligned} \tag{2.20}$$

Thus, the FW transformed Hamiltonian takes the form [71, 95, 100]:

$$\mathscr{H}_{\mathrm{FW}} = e^{iU_{\mathrm{FW}}}\left(\mathscr{H} - i\hbar\frac{\partial}{\partial t}\right)e^{-iU_{\mathrm{FW}}} + i\hbar\frac{\partial}{\partial t}\,, \tag{2.21}$$

or

$$\mathscr{H}_{\mathrm{FW}} = e^{iU_{\mathrm{FW}}}\mathscr{H}e^{-iU_{\mathrm{FW}}} - i\hbar\,e^{iU_{\mathrm{FW}}}\frac{\partial e^{-iU_{\mathrm{FW}}}}{\partial t}\,. \tag{2.22}$$

This can be expanded in a series which will involve the commutators of U_{FW} and the terms present in the considered Dirac Hamiltonian.

2.3.3 Time-independent FW transformation

As seen in the previous section, there are time-derivatives involved in the finally transformed Hamiltonian in Eq. (2.21) or (2.22). Hence, it is obvious that for a time-independent FW (TIFW) transformation, one only works with the expansion of the first term in Hamiltonian (2.22). This is the case for a particle where the particle does not experience any time-dependent fields (e.g.,

see Refs. [70, 98] for more details). In fact, for a free particle, it is possible to find an exact transformation [98], otherwise one has to work with the series expansion in powers of $1/m$. However, it is possible to find an exact FW transformation for new classes of external fields those are static [101–103].

For TIFW transformation, the unitary operator has to be chosen in a way such that the following condition is satisfied [71, 101, 104]

$$\mathscr{H}_{\rm FW} = e^{iU_{\rm FW}}\mathscr{H}_{\rm D}\,e^{-iU_{\rm FW}} = \beta\sqrt{m^2c^4+p^2c^2}\,, \tag{2.23}$$

where $\mathscr{H}_{\rm D} = \beta mc^2 + c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}$ is the free particle Dirac Hamiltonian. A few simple algebraic steps will result in the obvious choice of the operator [101]

$$U_{\rm FW} = -\frac{i}{2|p|}\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}\tan^{-1}\left(\frac{|p|}{mc}\right). \tag{2.24}$$

It should be noted that the derived $U_{\rm FW}$ is odd as it contains $\boldsymbol{\alpha}$ matrices which are Hermitian and most importantly time-independent. Consequently, we get

$$e^{\pm iU_{\rm FW}} = e^{\pm\frac{1}{2|p|}\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}\tan^{-1}\left(\frac{|p|}{mc}\right)} = \cos\phi \pm \frac{\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\sin\phi\,, \tag{2.25}$$

where $\phi = \frac{1}{2}\tan^{-1}\left(\frac{|p|}{mc}\right)$. It is easy to prove that this operation is, indeed, unitary by evaluating $e^{iU_{\rm FW}}e^{-iU_{\rm FW}} = 1$. The last equation (2.25) involves the expansion of $e^{\beta\frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\phi}$ and the use of following properties,

$$\left(\frac{\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\right)^n = -1 \;\forall\, n\in\text{even} \tag{2.26}$$

$$= -\frac{\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|} \;\forall\, n\in\text{odd}\,. \tag{2.27}$$

One can easily prove that the Eriksen condition is satisfied [101] for an exact TIFW transformation for spin 1/2 particles,

$$\begin{aligned}\beta e^{iU_{\rm FW}} &= \beta\left(\cos\phi + \frac{\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\sin\phi\right)\\ &= \left(\cos\phi - \frac{\beta\,\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\sin\phi\right)\beta = e^{-iU_{\rm FW}}\beta\,.\end{aligned} \tag{2.28}$$

Using the FW transformation in Eq. (2.22) if one evaluates the Hamiltonian, the subsequent steps will be as follows,

$$\begin{aligned}\mathscr{H}_{\mathrm{FW}} &= e^{iU_{\mathrm{FW}}}\left(\beta mc^2 + c\boldsymbol{\alpha}\cdot\boldsymbol{p}\right)\left(\cos\phi - \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\sin\phi\right)\\ &= e^{iU_{\mathrm{FW}}}\left(\cos\phi + \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\sin\phi\right)\left(\beta mc^2 + c\boldsymbol{\alpha}\cdot\boldsymbol{p}\right)\\ &= \left(\cos 2\phi + \frac{\beta\boldsymbol{\alpha}\cdot\boldsymbol{p}}{|p|}\sin 2\phi\right)\left(\beta mc^2 + c\boldsymbol{\alpha}\cdot\boldsymbol{p}\right)\\ &= \beta mc^2\left(\cos 2\phi + \frac{|p|}{mc}\sin 2\phi\right) + c\boldsymbol{\alpha}\cdot\boldsymbol{p}\left(\cos 2\phi - \frac{mc}{|p|}\sin 2\phi\right).\end{aligned} \tag{2.29}$$

Note that, the second term is odd here. Since the original idea was to remove the odd terms in the transformation, the required condition to eliminate second term is

$$\tan 2\phi = \frac{|p|}{mc}. \tag{2.30}$$

This follows to find the corresponding $\sin 2\phi$ and $\cos 2\phi$ which take the value

$$\sin 2\phi = \frac{|p|c}{\sqrt{p^2c^2 + m^2c^4}} \text{ and } \cos 2\phi = \frac{mc^2}{\sqrt{p^2c^2 + m^2c^4}}, \tag{2.31}$$

and thus the transformed Hamiltonian will take the form

$$\mathscr{H}_{\mathrm{FW}} = \beta mc^2\left(\frac{mc^2 + \frac{p^2}{m}}{\sqrt{p^2c^2 + m^2c^4}}\right) = \beta\sqrt{p^2c^2 + m^2c^4}. \tag{2.32}$$

Finally, the Hamiltonian is diagonalized with the energy of the KG equation. Due to the presence of Dirac matrices β in the final energy, the Hamiltonian is four-component - two components for the positive energy of $+\sqrt{p^2c^2 + m^2c^4}$ and the other two for the negative energy of $-\sqrt{p^2c^2 + m^2c^4}$.

2.3.4 Time-dependent FW transformation

When time-varying fields are involved in the description a Dirac particle, the time-dependent FW (TDFW) transformation is introduced. The fields are usually taken as weak fields such that the kinetic and potential energies are smaller than $2mc^2$ and the fields act as a perturbation. However, if the fields are strong, the positive and negative energies can be equivalent to $2mc^2$. Therefore, for strong fields the separation of positive and negative energy states are not guaranteed [71] and we encounter the so called "Klein Paradox" [105].

Such a transformation is not exact and thus, we have to rely on the series expansion in powers of $1/m$. For a TDFW transformation one has to keep the second term in Eq. (2.21) for obvious reasons. Expanding the time-dependent part and using the Baker-Campbell-Hausdorff (BCH) formula [106–108], one obtains a series of commutators as

$$e^{iU_{\rm FW}}\left(\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right)e^{-iU_{\rm FW}}=\mathscr{H}-i\hbar\frac{\partial}{\partial t}+i\left[U_{\rm FW},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]$$
$$+\frac{i^2}{2!}\left[U_{\rm FW},\left[U_{\rm FW},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]+\frac{i^3}{3!}\left[U_{\rm FW},\left[U_{\rm FW},\left[U_{\rm FW},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]\right]$$
$$+\ldots..\tag{2.33}$$

From Eq. (2.21), the transformed Hamiltonian can be expressed as a series of commutators as following,

$$\mathscr{H}_{\rm FW}=\mathscr{H}+i\left[U_{\rm FW},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]+\frac{i^2}{2!}\left[U_{\rm FW},\left[U_{\rm FW},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]$$
$$+\frac{i^3}{3!}\left[U_{\rm FW},\left[U_{\rm FW},\left[U_{\rm FW},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]\right]+\ldots..\tag{2.34}$$

Essentially, the FW transformation leads the Dirac equation towards the non-relativistic Pauli Hamiltonian plus all the higher order relativistic corrections. Following the operator given in Eq. (2.24), $\tan^{-1}\left(\frac{|p|}{mc}\right)$ is expanded in a Taylor series and only the first term is retained which is $\frac{|p|}{mc}$. Thus, the unitary operator will be expressed as

$$U_{\rm FW}=-\frac{i}{2mc^2}\beta\left(c\boldsymbol{\alpha}\cdot\boldsymbol{p}\right)\equiv-\frac{i}{2mc^2}\beta\mathscr{O}\,.\tag{2.35}$$

This choice of operator is justified because it is Hermitian and also the Eriksen condition is satisfied as β and $\mathscr{O}$ anticommute [101].

Now, we have to evaluate commutators in Eq. (2.34) by using the operator from Eq. (2.35) and the appropriate form of the Dirac Hamiltonian in Eq. (2.17). These commutators will generate higher order terms in $1/m$. We have already seen before that the kinetic and the Zeeman terms are of the order $1/m$ within Pauli theory (see Eq. (2.16)). In what follows, we will restrict ourselves up to the second order of relativistic correction i.e., we evaluate all the terms up to the order $1/m^3$. We mention that more higher order terms will only be important for stronger fields [71, 95, 109–113].

♣ First transformation

We consider the Dirac Hamiltonian in Eq. (2.17) and thus it is obvious that both the terms $\mathscr{E}$ and $i\hbar\frac{\partial}{\partial t}$ transform in a similar way. In the following, we consider the definition $\mathscr{F}=\mathscr{E}-i\hbar\frac{\partial}{\partial t}$ [100]. With these considerations, we

evaluate the appearing commutators in the series expansion of Eq. (2.34) as follows

$$i\left[U_{\mathrm{FW}},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]=-\mathscr{O}+\frac{1}{mc^2}\beta\mathscr{O}^2+\frac{\beta}{2mc^2}\left[\mathscr{O},\mathscr{F}\right]. \tag{2.36}$$

At this point, already, it is important to notice that the first term $-\mathscr{O}$ in the right-hand side already cancels with the odd terms in the Hamiltonian $\mathscr{H}$. Our goal to eliminate the odd terms is hence achieved, although new odd terms are generated of higher orders. Calculations of next commutators go as follows

$$\begin{aligned}\frac{i^2}{2!}\left[U_{\mathrm{FW}},\left[U_{\mathrm{FW}},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]=&-\frac{1}{2mc^2}\beta\mathscr{O}^2-\frac{1}{2m^2c^4}\mathscr{O}^3\\&-\frac{1}{8m^2c^4}\left[\mathscr{O},[\mathscr{O},\mathscr{F}]\right],\end{aligned} \tag{2.37}$$

$$\begin{aligned}\frac{i^3}{3!}\left[U_{\mathrm{FW}},\left[U_{\mathrm{FW}},\left[U_{\mathrm{FW}},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]\right]=&\frac{1}{6m^2c^4}\mathscr{O}^3-\frac{1}{6m^3c^6}\beta\mathscr{O}^4\\&-\frac{\beta}{48m^3c^6}\left[\mathscr{O},[\mathscr{O},[\mathscr{O},\mathscr{F}]]\right],\end{aligned} \tag{2.38}$$

$$\frac{i^4}{4!}\left[U_{\mathrm{FW}},\left[U_{\mathrm{FW}},\left[U_{\mathrm{FW}},\left[U_{\mathrm{FW}},\mathscr{H}-i\hbar\frac{\partial}{\partial t}\right]\right]\right]\right]=\frac{1}{24m^3c^6}\beta\mathscr{O}^4. \tag{2.39}$$

We have only kept up to the terms $1/m^3$ and therefore, only one term is retained for the last commutator. Having all these commutators derived, we write down the new Hamiltonian after first transformation as

$$\begin{aligned}\mathscr{H}_{\mathrm{FW}}&=(\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}\left[\mathscr{O},[\mathscr{O},\mathscr{F}]\right]\\&\quad+\frac{\beta}{2mc^2}\left[\mathscr{O},\mathscr{F}\right]-\frac{\mathscr{O}^3}{3m^2c^4}-\frac{\beta}{48m^3c^6}\left[\mathscr{O},[\mathscr{O},[\mathscr{O},\mathscr{F}]]\right]\\&=(\beta-\mathbb{1})mc^2+\mathscr{E}'+\mathscr{O}'.\end{aligned} \tag{2.40}$$

This is the expression of the Hamiltonian after first FW transformation up to the order $1/m^3$. Looking through the new Hamiltonian in Eq. (2.40), one notices that the transformation has already eliminated the previous odd terms of $1/m^0$. However, new even and odd terms are generated, which we denote as $\mathscr{E}'$ and $\mathscr{O}'$ respectively and they have the order $1/m$ or higher. Thus, one has to perform another second transformation to eliminate the higher order odd terms.

♣ Second transformation

Now, the new odd operator, $\mathscr{O}'$ will be used to form the new unitary operator

$$U'_{\mathrm{FW}}=-\frac{i}{2mc^2}\beta\mathscr{O}', \tag{2.41}$$

where the odd operators are collected from Eq. (2.40) as

$$\mathscr{O}' = \frac{\beta}{2mc^2}\left[\mathscr{O},\mathscr{F}\right] - \frac{\mathscr{O}^3}{3m^2c^4} - \frac{\beta}{48m^3c^6}\left[\mathscr{O},\left[\mathscr{O},\left[\mathscr{O},\mathscr{F}\right]\right]\right]. \tag{2.42}$$

Following the similar prescription, after the second transformation the Hamiltonian becomes

$$\begin{aligned}\mathscr{H}'_{\rm FW} = \mathscr{H}_{\rm FW} + i\left[U'_{\rm FW}, \mathscr{H}_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right] + \frac{i^2}{2!}\left[U'_{\rm FW}, \left[U'_{\rm FW}, \mathscr{H}_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right]\right] \\ + \frac{i^3}{3!}\left[U'_{\rm FW}, \left[U'_{\rm FW}, \left[U'_{\rm FW}, \mathscr{H}_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right]\right]\right] + ... \end{aligned} \tag{2.43}$$

With the help of the new unitary operator in Eq. (2.41) and the first transformed Hamiltonian in Eq. (2.40), we have to evaluate the further commutators that are involved in the second transformation i.e., Eq. (2.43). We keep the terms only up to the order $1/m^3$, while the higher order terms have been dropped. Considering that, the commutators can be evaluated as

$$\begin{aligned} & i\left[U'_{\rm FW}, \mathscr{H}_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right] \\ & = -\frac{\beta}{2mc^2}\left[\mathscr{O},\mathscr{F}\right] + \frac{1}{3m^2c^4}\mathscr{O}^3 + \frac{\beta}{48m^3c^6}\left[\mathscr{O},\left[\mathscr{O},\left[\mathscr{O},\mathscr{F}\right]\right]\right] \\ & - \frac{\beta}{8m^3c^6}\left\{\left[\mathscr{O},\mathscr{F}\right],\mathscr{O}^2\right\} - \frac{i}{6m^4c^8}\mathscr{O}^5 - \frac{\beta}{4m^3c^6}\left[\mathscr{O},\mathscr{F}\right]^2 \\ & + \frac{1}{4m^2c^4}\left[\left[\mathscr{O},\mathscr{F}\right],\mathscr{F}\right] - \frac{\beta}{6m^3c^6}\left[\mathscr{O}^3,\mathscr{F}\right], \end{aligned} \tag{2.44}$$

$$\frac{i^2}{2!}\left[U'_{\rm FW}, \left[U'_{\rm FW}, \mathscr{H}_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right]\right] = \frac{\beta}{8m^3c^6}\left[\mathscr{O},\mathscr{F}\right]^2. \tag{2.45}$$

The higher order commutators have the order of $1/m^4$ or more and therefore, those are not taken into consideration further. After the second transformation, the new transformed Hamiltonian is written as

$$\begin{aligned}\mathscr{H}'_{\rm FW} &= (\beta - \mathbb{1})mc^2 + \beta\left(\frac{\mathscr{O}^2}{2mc^2} - \frac{\mathscr{O}^4}{8m^3c^6}\right) + \mathscr{E} - \frac{1}{8m^2c^4}\left[\mathscr{O},\left[\mathscr{O},\mathscr{F}\right]\right] \\ & - \frac{\beta}{8m^3c^6}\left[\mathscr{O},\mathscr{F}\right]^2 - \frac{\beta}{8m^3c^6}\left\{\left[\mathscr{O},\mathscr{F}\right],\mathscr{O}^2\right\} + \frac{1}{4m^2c^4}\left[\left[\mathscr{O},\mathscr{F}\right],\mathscr{F}\right] \\ & - \frac{\beta}{6m^3c^6}\left[\mathscr{O}^3,\mathscr{F}\right] \\ & = (\beta - \mathbb{1})mc^2 + \mathscr{E}'' + \mathscr{O}''. \end{aligned} \tag{2.46}$$

Note that, the odd terms have been eliminated up to the order $1/m^2$ in the above Hamiltonian. However, by doing so, the new odd terms have been generated and those are of higher orders i.e., $1/m^3$ or more. The newly generated even and odd terms are denoted as $\mathscr{E}'$ and $\mathscr{O}'$.

♣ Third transformation

Following the same procedure, a new odd operator $\mathscr{O}''$ is formed so that the corresponding unitary operator is given by

$$U''_{\mathrm{FW}} = -\frac{i}{2mc^2}\beta\mathscr{O}'', \tag{2.47}$$

where the odd operator can be expressed as,

$$\mathscr{O}'' = -\frac{\beta}{8m^3c^6}\left\{[\mathscr{O},\mathscr{F}],\mathscr{O}^2\right\} + \frac{1}{4m^2c^4}\left[[\mathscr{O},\mathscr{F}],\mathscr{F}\right] - \frac{\beta}{6m^3c^6}\left[\mathscr{O}^3,\mathscr{F}\right]. \tag{2.48}$$

Notice that the newly formed odd operator has already the order $1/m^3$ or more. This indicates that less commutators have to be evaluated as we restrict the series only up to $1/m^3$. In fact, only the following commutator will contribute as,

$$\begin{aligned} & i\left[U''_{\mathrm{FW}}, \mathscr{H}'_{\mathrm{FW}} - i\hbar\frac{\partial}{\partial t}\right] \\ &= \frac{\beta}{8m^3c^6}\left\{[\mathscr{O},\mathscr{F}],\mathscr{O}^2\right\} - \frac{1}{4m^2c^4}\left[[\mathscr{O},\mathscr{F}],\mathscr{F}\right] + \frac{\beta}{6m^3c^6}\left[\mathscr{O}^3,\mathscr{F}\right] \\ &+ \frac{\beta}{8m^3c^6}\left[[[\mathscr{O},\mathscr{F}],\mathscr{F}],\mathscr{F}\right]. \end{aligned} \tag{2.49}$$

The evaluation of the other commutators results in the terms of order $1/m^4$ or more and those are dropped. The newly transformed Hamiltonian is given as

$$\begin{aligned} \mathscr{H}''_{\mathrm{FW}} &= (\beta - \mathbb{1})mc^2 + \beta\left(\frac{\mathscr{O}^2}{2mc^2} - \frac{\mathscr{O}^4}{8m^3c^6}\right) + \mathscr{E} - \frac{1}{8m^2c^4}\left[\mathscr{O},[\mathscr{O},\mathscr{F}]\right] \\ &\quad - \frac{\beta}{8m^3c^6}\left[\mathscr{O},\mathscr{F}\right]^2 + \frac{\beta}{8m^3c^6}\left[[[\mathscr{O},\mathscr{F}],\mathscr{F}],\mathscr{F}\right] \\ &= (\beta - \mathbb{1})mc^2 + \mathscr{E}'''' + \mathscr{O}''''. \end{aligned} \tag{2.50}$$

At this point, one can notice that only the last term in the Hamiltonian of Eq. (2.50) is odd, all the rest have already been transformed to be even.

♣ Fourth transformation

A further transformation is needed as we see that the new odd operator, $\mathscr{O}''''$ is formed. The new unitary operator is given as

$$U''''_{\mathrm{FW}} = -\frac{i}{2mc^2}\beta\mathscr{O}'''', \tag{2.51}$$

with

$$\mathscr{O}'''' = \frac{\beta}{8m^3c^6}\left[[[\mathscr{O},\mathscr{F}],\mathscr{F}],\mathscr{F}\right]. \tag{2.52}$$

The commutator that contributes, is evaluated as,

$$i\left[U_{\mathrm{FW}}^{\prime\prime\prime},\mathscr{H}_{\mathrm{FW}}^{\prime\prime}-i\hbar\frac{\partial}{\partial t}\right]=-\frac{\beta}{8m^3c^6}\left[\left[\left[\mathscr{O},\mathscr{F}\right],\mathscr{F}\right],\mathscr{F}\right], \tag{2.53}$$

that is odd. Notice that, this term exactly cancels with the existing odd term in Eq. (2.50). Hence, the final transformed Hamiltonian will be given as

$$\begin{aligned}\mathscr{H}_{\mathrm{FW}}^{\prime\prime\prime}=(\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}\left[\mathscr{O},\left[\mathscr{O},\mathscr{F}\right]\right]\\ -\frac{\beta}{8m^3c^6}\left[\mathscr{O},\mathscr{F}\right]^2.\end{aligned} \tag{2.54}$$

This is the correct form of the FW transformed Hamiltonian up to the order $1/m^3$ and we see that all the terms are even. The odd operators up to the same order are eliminated, in addition, the higher orders are neglected. This gives a *semirelativistic* expression where the zeroth order Hamiltonian reveals exactly the Schrödinger Hamiltonian. Comparing with the Dirac Hamiltonian for a crystal, $\mathscr{O}=c\boldsymbol{\alpha}\cdot\boldsymbol{p}$ and $\mathscr{E}=V$, the crystal potential, the second and fourth term of Eq. (2.54) lead to the kinetic and potential energies of the Schrödinger equation.

This method of the FW transformed Hamiltonian was proposed by using "step-by-step" transformations or an iterative method and calculating the commutators. Although Foldy and Wouthuysen took the weak fields into account, the higher orders are of doubtful value. In fact, they derived only the first order relativistic correction terms i.e., up to the terms $1/m^2$ for spin-$\frac{1}{2}$ particle strictly [71]. The calculations of second and higher order terms do not derive all the necessary relativistic correction terms. Therefore, a correction in FW transformation is needed. We mention that, in 1954, Case generalized the idea of FW transformation and extended it to include particles with integral spin as well as arbitrary half-integral spin [114]. Furthermore, the FW operators have been found for any arbitrary spin and it has been a research topic of current interest [115–117].

In the next section, we provide other methods to recover the correction terms in higher order expansion of FW transformation.

2.4 Corrections to the FW transformation

The fact that the original FW transformation is based on a series expansion towards a semirelativistic Hamiltonian, renders the higher order terms of doubtful value. Furthermore, the original method fails when the relativistic corrections i.e., the higher orders diverge. Note that the exponential operator $e^{iU_{\mathrm{FW}}}$ in each step of the FW transformation should have the property that it is *odd* and *Hermitian* [101, 118].

2.4.1 Eriksen's method

In 1958, Eriksen developed a general form of the unitary operator, U_{FW} and subsequently the FW transformation which is single step, not "step-by-step". Within Eriksen's method, the FW transformation can uniquely be defined if the exponential operator $e^{iU_{\mathrm{FW}}}$ is odd and Hermitian [101, 118]. These conditions are equivalent to

$$\beta e^{iU_{\mathrm{FW}}} = e^{-iU_{\mathrm{FW}}}\beta\,. \tag{2.55}$$

Eriksen found an operator which is non-exponential in contrast to the original FW transformation. The search for such an operator which commutes with $\mathscr{H}$ and it has eigenvalues either $+1$ or -1, results in a sign operator λ in the following way $\lambda = \frac{\mathscr{H}}{\sqrt{\mathscr{H}^2}}$. Thereby, the operator $\beta\lambda$ or $\lambda\beta$ are unitary given the fact that $\lambda^2 = 1$. The operator $1+\beta\lambda$ has the property that for β matrix elements equal to 1 or -1, the operator cancels either the lower or upper components of the Dirac spinor. Instead of an exponential operator, Eriksen proposed the following operator for an exact FW transformation [100, 101]

$$e^{iU_{\mathrm{FW}}} = S_{\mathrm{FW}} = \frac{1+\beta\lambda}{\sqrt{2+\beta\lambda+\lambda\beta}}\,. \tag{2.56}$$

It is appealing to test the convergence of the operator. Let us consider an eigenfunction, u_δ of $\beta\lambda$ such that $\beta\lambda u_\delta = e^{i\delta}u_\delta$. From the left multiplying with $e^{-i\delta}\lambda\beta$ in both the sides results in $e^{-i\delta}u_\delta = \lambda\beta u_\delta$. Taking the sum of both we get

$$(\beta\lambda+\lambda\beta)\,u_\delta = \left(e^{i\delta}+e^{-i\delta}\right)u_\delta = 2\cos\delta\,u_\delta\,. \tag{2.57}$$

The denominator in the Eriksen operator in Eq. (2.56) becomes

$$(2+\beta\lambda+\lambda\beta)\,u_\delta = 2\,(1+\cos\delta)\,u_\delta\,. \tag{2.58}$$

The convergence condition will be given by $2+2\cos\delta > 0$ or $\cos\delta > -1$. A few other properties of the operator has to be noted: $S_{\mathrm{FW}}^2 = \beta\lambda$, $S_{\mathrm{FW}}^\dagger = \lambda\beta S_{\mathrm{FW}}$ and then $S_{\mathrm{FW}}S_{\mathrm{FW}}^\dagger = 1$. Once the Eriksen operator is constructed, it is interesting to see whether the original FW transformation operator is achieved (see Appendix A for details).

The Eriksen operator for the FW transformation can be used for a particle with any spin [100]. The FW transformation within Eriksen's method will be given as following for the time-independent case [101],

$$\mathscr{H}_{\mathrm{FW}}^{\mathrm{Erik}} = S_{\mathrm{FW}}\mathscr{H}S_{\mathrm{FW}}^\dagger\,. \tag{2.59}$$

Note that this transformation is not iterative because it does not involve the exponential operators, unlike the original FW transformation.

♣ Free Dirac particle

For a free Dirac particle, the following Eriksen operator is derived

$$\lambda = \frac{\beta mc^2 + \mathscr{O}}{\sqrt{m^2c^4 + \mathscr{O}^2}}, \tag{2.60}$$

with the fact that $\mathscr{E} = 0$ and $\mathscr{O} = c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}$. Since, β and $\lambda\mathscr{H}$ commute with each other and $\lambda\mathscr{H} = \mathscr{H}\lambda$, we get $(1+\beta\lambda)\mathscr{H} = \mathscr{H}(1+\lambda\beta)$. This implies that $\mathscr{H}S_{\rm FW}^{\dagger} = S_{\rm FW}\mathscr{H}$ and the Eriksen condition is satisfied. The transformation in this case is rather easy and will be given as

$$\mathscr{H}_{\rm FW}^{\rm Erik} = S_{\rm FW}\mathscr{H}S_{\rm FW}^{\dagger} = S_{\rm FW}^2\mathscr{H} = \beta\lambda\mathscr{H} = \beta\sqrt{m^2c^4 + \mathscr{O}^2}. \tag{2.61}$$

This result is exactly the same as the result presented in Sec. 2.3.3. Again, it is important to note that Eriksen's method is not "step-by-step", rather it is a single transformation.

♣ Dirac particle in a potential

When a Dirac Particle experiences a potential then $\mathscr{E} \neq 0$. Thus, we have to calculate the λ operator carefully, and that involves the determination of $\sqrt{\mathscr{H}^2}$ in the denominator. The latter will be written as

$$\begin{aligned}\sqrt{\mathscr{H}^2} &= \sqrt{m^2c^4 + \mathscr{O}^2 + \mathscr{E}^2 + 2\beta mc^2\mathscr{E} + \{\mathscr{O},\mathscr{E}\}} \\ &= \beta mc^2\sqrt{1 + \frac{\mathscr{O}^2 + \mathscr{E}^2 + 2\beta mc^2\mathscr{E} + \{\mathscr{O},\mathscr{E}\}}{m^2c^4}}.\end{aligned} \tag{2.62}$$

Now this has to be expanded in a Taylor series and the transformation will have the relativistic correction terms in the powers of $\mathscr{O}/m$ and $\mathscr{E}/m$. The analytic expansion is non-trivial and rather cumbersome. However, by using an analytic computer program E. de Vries and J. E. Jonker derived all the correct relativistic terms up to the order $1/c^8$ [85]. To include the time dependency, one works with $\mathscr{F} = \mathscr{E} - i\hbar\frac{\partial}{\partial t}$ instead of $\mathscr{E}$ and finally the results up to the order $1/m^3$ will be given in a compact form as [85, 100, 110]

$$\begin{aligned}\mathscr{H}_{\rm FW}^{\rm Erik} = (\beta - \mathbb{1})mc^2 + \beta\left(\frac{\mathscr{O}^2}{2mc^2} - \frac{\mathscr{O}^4}{8m^3c^6}\right) + \mathscr{E} - \frac{1}{8m^2c^4}[\mathscr{O},[\mathscr{O},\mathscr{F}]] \\ + \frac{\beta}{16m^3c^6}\{\mathscr{O},[[\mathscr{O},\mathscr{F}],\mathscr{F}]\}.\end{aligned} \tag{2.63}$$

Again, this approach does not need any subsequent iterations. Let us verify the two Hamiltonians - one obtained by the original FW transformation in Eq. (2.54) and the other obtained by Eriksen's method in Eq. (2.63). All the Hamiltonian terms are the same in both cases except the last term which has the order $1/m^3c^6$. This proves that the original FW transformation does not

produce the correct terms in the higher order, as already discussed earlier. We can write down the last term in Eq. (2.63) in the following way

$$\frac{\beta}{16m^3c^6}\{\mathscr{O},[[\mathscr{O},\mathscr{F}],\mathscr{F}]\} = \frac{\beta}{16m^3c^6}\left[\left[\mathscr{O}^2,\mathscr{F}\right],\mathscr{F}\right] - \frac{\beta}{8m^3c^6}\left[\mathscr{O},\mathscr{F}\right]^2 . \tag{2.64}$$

Thus, we see that the original FW transformation could capture only one term e.g., only the last term in Eq. (2.64). However, the first term in Eq. (2.64) has the same order of the coefficient and even term that was not present in the original FW transformation. The reason is discussed in the next Section.

2.4.2 Other methods

In the original FW transformation, the first unitary operator $U_{\rm FW}$ is of the order of $1/c^2$ and during the iterative process, any subsequent operator is $1/c^2$ times smaller than the previous one. The expansion with the BCH formula thus generates errors which was pointed out by Eriksen and Kolsrud [118]. The correction of errors can be performed in the following ways [100].

Let us take two preceding exponential operators as $e^{iU_{\rm FW}}$ and $e^{iU'_{\rm FW}}$ in two successive transformations. The main problem underlies the fact that, the commutator $[U_{\rm FW},U'_{\rm FW}] \neq 0$. Neglecting the higher order commutators, we arrive at[1]

$$e^{iU'_{\rm FW}}e^{iU_{\rm FW}} = e^{i(U'_{\rm FW}+U_{\rm FW})}\,e^{\frac{1}{2}\left[U_{\rm FW},U'_{\rm FW}\right]} . \tag{2.65}$$

Notice here that the first exponential term in the right side, is odd and Hermitian. However, the correction term, the second one is even and of the order $1/c^6$. This is the reason that the original FW transformation is only able to produce correct terms up to the order $1/c^4$ (see the Eqs. (2.54) and (2.63)). Therefore, the correction term does not add any odd terms of higher order, however, adds even "missing" terms. Calculating the latter commutator we find,

$$\left[U_{\rm FW},U'_{\rm FW}\right] = -\frac{\beta}{8m^3c^6}\left[\mathscr{O}^2,\mathscr{F}\right] . \tag{2.66}$$

As this is of the order $1/c^6$, the correction term will add the leading even "missing" terms of the same order or higher to the original FW transformation. Following the original FW transformation the correction term will be used as

[1] For two exponential operators, the product is defined as

$$e^A e^B = e^{(A+B)}\,e^{\frac{1}{2}[A,B]}\,e^{\frac{1}{12}([A,[A,B]]-[B,[A,B]])}\ldots$$

following:

$$\mathscr{H}_{\rm FW}^{\rm corr} = e^{\frac{1}{2}\left[U_{\rm FW},U'_{\rm FW}\right]}\left(\mathscr{H}'''_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right)e^{-\frac{1}{2}\left[U_{\rm FW},U'_{\rm FW}\right]} + i\hbar\frac{\partial}{\partial t}$$
$$= \mathscr{H}'''_{\rm FW} - \left[\frac{1}{2}\left[U_{\rm FW},U'_{\rm FW}\right],\left(\mathscr{H}'''_{\rm FW} - i\hbar\frac{\partial}{\partial t}\right)\right]. \tag{2.67}$$

Thus, the leading order correction will be the given by the commutator,

$$\left[\frac{1}{2}\left[U_{\rm FW},U'_{\rm FW}\right],\mathscr{E} - i\hbar\frac{\partial}{\partial t}\right] = -\frac{\beta}{16m^3c^6}\left[\left[\mathscr{O}^2,\mathscr{F}\right],\mathscr{F}\right]. \tag{2.68}$$

This is exactly the "missing" even term, that was not obtained by the original FW transformation (see Eq. 2.64). Hence, we have shown that up to an order $1/c^6$, the Eriksen method and the corrected FW method result in the same transformed Hamiltonian i.e.,

$$\mathscr{H}_{\rm FW}^{\rm corr} = \mathscr{H}_{\rm FW}^{\rm Erik}. \tag{2.69}$$

For the higher orders, one needs to take into account the other commutators for corrections e.g., $[U_{\rm FW},(U'_{\rm FW}+U''_{\rm FW})]$ (see Ref. [100] for details).

To summarize, the original iterative method proposed by Foldy and Wouthuysen [71] does not produce all the correct even terms in higher order and therefore it is not trustworthy. The fact that the different unitary operators do not commute with each other leads to the "missing" even terms in the original FW transformation. The correct FW transformation can be achieved by two methods that are described above. One of them is to consider the corrected even operator in each iterative process that will produce the missing even terms in the first place. On the other hand, Eriksen's method is not iterative rather a single step towards the direct FW transformation and can capture all the higher order even relativistic terms. However, the Eriksen method cannot be used in practical purposes because it involves the square root of different Dirac matrices (see Eqs. (2.56), (2.62)). Thus, the most applied methods are based on the correction of the FW transformation and the use of exponential operators. The intermediate operators in the successive transformations are odd and Hermitian in the FW transformation. However, the correction operators are even and Hermitian, and compensate with the "missing" terms in the original FW transformation. Let us also mention that the series obtained by semirelativistic methods diverge when $p > mc$ [100, 111, 117].

Once we have all the relativistic correction terms to the nonrelativistic Pauli-Schrödinger Hamiltonian, we apply those terms in more physical (e.g., magnetic) systems in the next Section.

♣ Summary of FW transformation

The unitary operators and the corresponding Hamiltonians in each steps of the original TDFW transformation are given below (these are in agreement with Ref. [100]):

$$
\begin{aligned}
U_{\mathrm{FW}} &= -\frac{i}{2mc^2}\beta\mathscr{O} \\
\mathscr{H}_{\mathrm{FW}} &= (\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}[\mathscr{O},[\mathscr{O},\mathscr{F}]] \\
&\quad+\frac{\beta}{2mc^2}[\mathscr{O},\mathscr{F}]-\frac{\mathscr{O}^3}{3m^2c^4}-\frac{\beta}{48m^3c^6}[\mathscr{O},[\mathscr{O},[\mathscr{O},\mathscr{F}]]] \\
U'_{\mathrm{FW}} &= -\frac{i}{4m^2c^4}[\mathscr{O},\mathscr{F}]+\frac{i\beta}{6m^3c^6}\mathscr{O}^3+\frac{i}{96m^4c^8}[\mathscr{O},[\mathscr{O},[\mathscr{O},\mathscr{F}]]] \\
\mathscr{H}'_{\mathrm{FW}} &= (\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}[\mathscr{O},[\mathscr{O},\mathscr{F}]] \\
&\quad-\frac{\beta}{8m^3c^6}[\mathscr{O},\mathscr{F}]^2-\frac{\beta}{8m^3c^6}\left\{[\mathscr{O},\mathscr{F}],\mathscr{O}^2\right\}+\frac{1}{4m^2c^4}[[\mathscr{O},\mathscr{F}],\mathscr{F}] \\
&\quad-\frac{\beta}{6m^3c^6}\left[\mathscr{O}^3,\mathscr{F}\right] \\
U''_{\mathrm{FW}} &= \frac{i}{16m^4c^8}\left\{[\mathscr{O},\mathscr{F}],\mathscr{O}^2\right\}-\frac{i\beta}{8m^3c^6}[[\mathscr{O},\mathscr{F}],\mathscr{F}]+\frac{i}{12m^4c^8}\left[\mathscr{O}^3,\mathscr{F}\right] \\
\mathscr{H}''_{\mathrm{FW}} &= (\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}[\mathscr{O},[\mathscr{O},\mathscr{F}]] \\
&\quad-\frac{\beta}{8m^3c^6}[\mathscr{O},\mathscr{F}]^2+\frac{\beta}{8m^3c^6}[[[\mathscr{O},\mathscr{F}],\mathscr{F}],\mathscr{F}] \\
U'''_{\mathrm{FW}} &= -\frac{i}{16m^4c^8}[[[\mathscr{O},\mathscr{F}],\mathscr{F}],\mathscr{F}] \\
\mathscr{H}'''_{\mathrm{FW}} &= (\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}[\mathscr{O},[\mathscr{O},\mathscr{F}]] \\
&\quad-\frac{\beta}{8m^3c^6}[\mathscr{O},\mathscr{F}]^2 \\
\mathscr{H}^{\mathrm{Erik}}_{\mathrm{FW}} &= (\beta-\mathbb{1})mc^2+\beta\left(\frac{\mathscr{O}^2}{2mc^2}-\frac{\mathscr{O}^4}{8m^3c^6}\right)+\mathscr{E}-\frac{1}{8m^2c^4}[\mathscr{O},[\mathscr{O},\mathscr{F}]] \\
&\quad+\frac{\beta}{16m^3c^6}\{\mathscr{O},[[\mathscr{O},\mathscr{F}],\mathscr{F}]\} \\
&\quad=\mathscr{H}^{\mathrm{corr}}_{\mathrm{FW}}
\end{aligned}
$$

2.5 FW transformation for a magnetic solid in an electromagnetic field

We have already pointed out that the free particle Dirac Hamiltonian is given by the Eq. (2.6). Such a theory has been used to describe the ground state properties of any system by using the concept of particle density, $n(\boldsymbol{r})$ in DFT [66, 67]. To describe a relativistic particle within relativistic DFT, the Dirac-Kohn-Sham (DKS) Hamiltonian is written as [70, 119, 120]

$$\mathscr{H}_{\mathrm{D}}^{\mathrm{KS}} = c\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}_{\mathrm{eff}})+\beta mc^2+V_{\mathrm{eff}}\,, \tag{2.70}$$

where the effective scalar and vector potentials are the functions of external potentials, exchange-correlation energies [121]. To account for the magnetic solids, the magnetic exchange interaction is of particular interest as the exchange interaction can give rise to an exchange field of the order of 10^3 T. The magnetic exchange interaction arises due to the Pauli exclusion principle which says that two electrons with same spin cannot have the same 'position' in a specific orbital. The exchange interaction is responsible for the different types of spontaneous ordering of atomic magnetic moments occurring in magnetic solids e.g., ferromagnetism, antiferromagnetism and ferrimagnetism.

The detailed discussion of different exchange interactions are beyond the scope of this thesis, however, we stress the point that the corresponding exchange field, $\boldsymbol{B}_{\mathrm{xc}}$ is different from the one of usual Maxwell's fields, or in other words, $\boldsymbol{B}_{\mathrm{xc}}$ does not obey the Maxwell's equations [122]. The reason is because $\boldsymbol{B}_{\mathrm{xc}}$ couples *only* to the spin degrees of freedom, not to the other degrees of freedom [122]. Thus, this field is not a proper magnetic field and cannot be included as a vector potential, $\boldsymbol{A}_{\mathrm{xc}}$. Instead, we have to treat the effect of exchange field in a separate term within the DKS Hamiltonian as follows [123, 124]

$$\mathscr{H}_{\mathrm{DKS}} = c\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A})+(\beta-\mathbb{1})mc^2+V+e\Phi-\mu_{\mathrm{B}}\beta\boldsymbol{\Sigma}\cdot\boldsymbol{B}_{\mathrm{xc}}\,. \tag{2.71}$$

The electromagnetic field is taken care by the vector potential as minimal coupling and scalar potential energy $e\Phi$, V defines the unpolarized crystal potential and the exchange field is separately accounted. Now, following the description of the FW transformation, the Hamiltonian in Eq. (2.71) contains odd terms and even terms. The Dirac matrices $\boldsymbol{\alpha}$ have off-diagonal elements in the matrix formalism and therefore, they form odd operators and β contains diagonal elements constructing even operators. Following Eq. (2.71), the odd and even operators can be written as below

$$\mathscr{O} = c\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A})\,, \tag{2.72}$$

$$\mathscr{E} = V+e\Phi-\mu_{\mathrm{B}}\beta\boldsymbol{\Sigma}\cdot\boldsymbol{B}_{\mathrm{xc}}\,. \tag{2.73}$$

Therefore, the Hamiltonian can be expressed as in the form of Eq. (2.17). As the vector potential, $\boldsymbol{A}(\boldsymbol{r},t)$ depends on time, a TDFW transformation is

performed for the separation of particles from antiparticles. The next step will be to substitute these odd and even operators in Eq. (2.63) and calculate the Hamiltonian. The corrected FW transformation gives the Hamiltonian terms discussed further below, which describe spin-$\frac{1}{2}$ particles.

Putting together all the derived Hamiltonian terms, we write down the extended Pauli Hamiltonian including the relativistic effects up to an order of $1/c^4$:

$$
\begin{aligned}
&\mathscr{H}_{\text{Pauli}}^{\text{extn.}} = \\
&= \underbrace{\frac{(\boldsymbol{p}-e\boldsymbol{A})^2}{2m} + V + e\Phi - \mu_{\text{B}}\,\boldsymbol{\sigma}\cdot\left(\boldsymbol{B}+\boldsymbol{B}_{\text{xc}}\right)}_{\text{Pauli Hamiltonian}} - \underbrace{\frac{(\boldsymbol{p}-e\boldsymbol{A})^4}{8m^3c^2}}_{\text{mass correction}} \\
&+ \underbrace{\frac{e\hbar}{8m^3c^2}\left\{(\boldsymbol{p}-e\boldsymbol{A})^2, \boldsymbol{\sigma}\cdot\boldsymbol{B}\right\}}_{\text{indirect field}-\text{spin coupling}} - \underbrace{\frac{e\hbar^2}{8m^2c^2}\boldsymbol{\nabla}\cdot\boldsymbol{E}_{\text{tot}}}_{\text{Darwin term}} \\
&- \underbrace{\frac{e\hbar}{8m^2c^2}\boldsymbol{\sigma}\cdot\left[\boldsymbol{E}_{\text{tot}}\times(\boldsymbol{p}-e\boldsymbol{A}) - (\boldsymbol{p}-e\boldsymbol{A})\times\boldsymbol{E}_{\text{tot}}\right]}_{\text{spin}-\text{orbit and spin}-\text{photon coupling}} \\
&+ \underbrace{\frac{\mu_{\text{B}}}{8m^2c^2}\boldsymbol{\sigma}\cdot\boldsymbol{B}_{\text{corr}}^{\text{xc}} + \frac{i\mu_B}{4m^2c^2}[(\boldsymbol{p}\times\boldsymbol{B}_{\text{xc}})\cdot(\boldsymbol{p}-e\boldsymbol{A})]}_{\text{relativistic corrections to the exchange field}} \\
&- \underbrace{\frac{ie\hbar^2}{16m^3c^4}\boldsymbol{\sigma}\cdot[\partial_t\boldsymbol{E}_{\text{tot}}\times(\boldsymbol{p}-e\boldsymbol{A}) + (\boldsymbol{p}-e\boldsymbol{A})\times\partial_t\boldsymbol{E}_{\text{tot}}]}_{\text{higher}-\text{order spin}-\text{orbit coupling}} \qquad (2.74)
\end{aligned}
$$

These relativistic Hamiltonian terms seems to be complicated, however, their physical meanings are immediately explained.

2.5.1 Pauli Hamiltonian

The second and fourth terms in Eq. (2.63) constitute the nonrelativistic Pauli Hamiltonian, however, yet including the magnetic exchange field. The Pauli Hamiltonian is written as,

$$
\mathscr{H}_{\text{P}} = \frac{(\boldsymbol{p}-e\boldsymbol{A})^2}{2m} + V + e\Phi - \mu_{\text{B}}\,\boldsymbol{\sigma}\cdot\boldsymbol{B} - \mu_{\text{B}}\,\boldsymbol{\sigma}\cdot\boldsymbol{B}_{\text{xc}}, \qquad (2.75)
$$

where the external magnetic field is given by $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$ and $\mu_{\text{B}} = \frac{e\hbar}{2m}$, defines the Bohr magneton. The last two terms of the Pauli Hamiltonian explains the Zeeman coupling with the external magnetic field and exchange field respec-

tively. We choose for simplicity a gauge such that

$$\boldsymbol{A} = \frac{\boldsymbol{B} \times \boldsymbol{r}}{2}, \tag{2.76}$$

which fulfills the Coulomb gauge ($\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$) for the *uniform* magnetic field[2]. It is important to note that with this Pauli Hamiltonian one can show how the different magnetic contributions arise. Incorporating the gauge choice, the Pauli Hamiltonian can be rewritten as

$$\mathscr{H}_{\mathrm{P}} = \left(\frac{p^2}{2m} + V \right) + e\Phi - \frac{e}{m} \boldsymbol{B} \cdot (\boldsymbol{L} + g\boldsymbol{S}) + \frac{e^2}{8m} (\boldsymbol{B} \times \boldsymbol{r})^2 - \frac{e}{m} \boldsymbol{S} \cdot \boldsymbol{B}_{\mathrm{xc}}, \tag{2.77}$$

with g the Landé g-factor, which is approximately 2 for spin degrees of freedom. The spin angular momentum is represented by $\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$ and the orbital angular momentum is defined by $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$. The first term of Eq. (2.77) is obviously the unperturbed Hamiltonian - the Schrödinger terms, the third term is the dominant perturbation - the paramagnetic contribution and the fourth term is recognized as the diamagnetic contribution [125].

Note that, the external magnetic field couples to both the spin and the orbital angular momentum operators, as it should be. However, the exchange field only couples to the spin degrees of freedom as discussed earlier.

2.5.2 Relativistic mass correction

The very first relativistic correction appears in the form of relativistic mass correction that is derived from the third term of Eq. (2.63) as

$$\mathscr{H}_{\mathrm{RMC}} = -\frac{(\boldsymbol{p} - e\boldsymbol{A})^4}{8m^3c^2}. \tag{2.78}$$

In the same context, we note that in special relativity, the mass correction arises from the relativistic momentum-energy relation in Eq. (2.2) such as

$$E = mc^2 \left(1 + \frac{p^2}{m^2c^2} \right)^{\frac{1}{2}} \approx mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2}. \tag{2.79}$$

As the relativistic mass correction does not concern us about the electron spin, this will not be considered in our discussions in this thesis which will mostly focus on the magnetic systems. The expansion of third term of Eq. (2.63) includes two more terms but those belong to the indirect field-spin coupling [90]. The spin dependent indirect field-spin coupling term is derived as

$$\mathscr{H}_{\mathrm{indirect}} = \frac{e\hbar}{8m^3c^2} \left\{ (\boldsymbol{p} - e\boldsymbol{A})^2, \boldsymbol{\sigma} \cdot \boldsymbol{B} \right\}. \tag{2.80}$$

[2] $\boldsymbol{\nabla} \times \boldsymbol{A} = \frac{1}{2}\boldsymbol{\nabla} \times (\boldsymbol{B} \times \boldsymbol{r}) = \frac{1}{2}\left[\boldsymbol{B}(\boldsymbol{\nabla} \cdot \boldsymbol{r}) - \boldsymbol{r}(\boldsymbol{\nabla} \cdot \boldsymbol{B}) + (\boldsymbol{r} \cdot \boldsymbol{\nabla})\boldsymbol{B} - (\boldsymbol{B} \cdot \boldsymbol{\nabla})\boldsymbol{r}\right] = \frac{1}{2}\left[3\boldsymbol{B} - \boldsymbol{B}\right] = \boldsymbol{B}$
$\boldsymbol{\nabla} \cdot \boldsymbol{A} = \frac{1}{2}\boldsymbol{\nabla} \cdot (\boldsymbol{B} \times \boldsymbol{r}) = (\boldsymbol{\nabla} \times \boldsymbol{B}) \cdot \boldsymbol{r} - \boldsymbol{B} \cdot (\boldsymbol{\nabla} \times \boldsymbol{r}) = 0$

2.5.3 Darwin term

The expansion of the fifth term in Eq. (2.63) gives the so called relativistic Darwin terms as

$$\mathscr{H}_{\text{Darwin}} = -\frac{e\hbar^2}{8m^2c^2}\boldsymbol{\nabla}\cdot\boldsymbol{E}_{\text{tot}}\,. \tag{2.81}$$

The corresponding total electric field has two components, the intrinsic electric field arises from the crystal field and is denoted as $\boldsymbol{E}_{\text{int}} = -\frac{1}{e}\boldsymbol{\nabla}V$, however, the external electric field is derived as $\boldsymbol{E}_{\text{ext}} = -\frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{\nabla}\Phi$. This is a relativistic contribution to the energy which does not have any nonrelativistic analogue. Within the Darwin theory, the electron cannot be considered as a particle, rather its position is spread out of the order Compton wavelength $\hbar/mc$, thus producing a peculiar motion of electrons - *Zitterbewegung* [70, 126]. As we are mostly focusing on the magnetic systems, the Darwin terms will not be discussed furthermore in the following thesis.

2.5.4 Spin-orbit and spin-photon coupling

The other term which also comes from the expansion of the fifth term in Eq. (2.63), is so-called spin-orbit coupling and has the form

$$\mathscr{H}_{\text{SOC}} = -\frac{e\hbar}{8m^2c^2}\boldsymbol{\sigma}\cdot\Big[\boldsymbol{E}_{\text{tot}}\times(\boldsymbol{p}-e\boldsymbol{A}) - (\boldsymbol{p}-e\boldsymbol{A})\times\boldsymbol{E}_{\text{tot}}\Big]\,. \tag{2.82}$$

We notice that there exist two different types of spin-orbit coupling: One couples to the intrinsic crystal electric field which is usually taken as time independent - intrinsic spin-orbit. The other one couples to the external electric field which is time-dependent - extrinsic spin-orbit.

In the same context, we note that in special relativity, when a particle moves in an electric field, $\boldsymbol{E}$, with velocity $\boldsymbol{v}$, it experiences a magnetic field in its reference frame, which has the expression $\boldsymbol{B} = -\frac{1}{c^2}\boldsymbol{v}\times\boldsymbol{E}$. If the particle is an electron and it moves in a spherically symmetric potential, this magnetic field will be proportional to the orbital angular momentum, $\boldsymbol{l}$, as

$$\boldsymbol{B} = \frac{1}{ec^2}\frac{1}{r}\frac{dV}{dr}\boldsymbol{v}\times\boldsymbol{r} = -\frac{1}{mec^2}\frac{1}{r}\frac{dV}{dr}\boldsymbol{r}\times(m\boldsymbol{v}) = -\frac{1}{mec^2}\frac{1}{r}\frac{dV}{dr}\boldsymbol{l}\,. \tag{2.83}$$

When this magnetic field couples to the electron spin, the interaction energy is then recognized as the spin-orbit coupling, $\sim \boldsymbol{\sigma}\cdot\boldsymbol{l}$ [78]. Thus, it is clear that the traditional spin-orbit coupling has the following form $\boldsymbol{\sigma}\cdot(\boldsymbol{E}\times\boldsymbol{p})$ and it is neither gauge invariant nor Hermitian [127].

However, our derived spin-orbit coupling Hamiltonian in Eq. (2.82) is gauge invariant by the introduction of minimal coupling and notably Hermitian. Following the Hamiltonian in Eq. (2.82), we denote the traditional spin-orbit term

as $\boldsymbol{\sigma}\cdot(\boldsymbol{E}_{\rm tot}\times\boldsymbol{p})$, however, the spin-photon coupling as $\boldsymbol{\sigma}\cdot(\boldsymbol{E}_{\rm tot}\times\boldsymbol{A})$. Alternatively, we mention that such spin-orbit coupling terms can also be derived using Møller's idea in special relativity together with the Schrödinger Hamiltonian [128].

Throughout this thesis, we will show that the relativistic spin-orbit and spin-photon coupling is very much important in explaining many complex physical phenomena in condensed matter magnetic systems.

2.5.5 Relativistic corrections to the exchange field

The fifth term in Eq. (2.63) also provides the derivation of relativistic corrections to the exchange field while considering the exchange field as the even terms in the FW transformation. These terms can be written as[3]

$$\begin{aligned}\mathscr{H}_{\rm RCXC} = \frac{\mu_{\rm B}}{8m^2c^2}\boldsymbol{\sigma}\cdot\Big\{\big[p^2\boldsymbol{B}_{\rm xc}\big] + 2(\boldsymbol{p}\boldsymbol{B}_{\rm xc})\cdot(\boldsymbol{p}-e\boldsymbol{A}) + 2(\boldsymbol{p}\cdot\boldsymbol{B}_{\rm xc})(\boldsymbol{p}-e\boldsymbol{A}) \\ + 4[\boldsymbol{B}_{\rm xc}\cdot(\boldsymbol{p}-e\boldsymbol{A})](\boldsymbol{p}-e\boldsymbol{A})\Big\} + \frac{i\mu_B}{4m^2c^2}[(\boldsymbol{p}\times\boldsymbol{B}_{\rm xc})\cdot(\boldsymbol{p}-e\boldsymbol{A})]\,. \quad (2.84)\end{aligned}$$

The terms within $\{\}$ can be recognized as the effective Zeeman-like field due to the exchange field which are relativistic corrections, they can be written as

$$\begin{aligned}\boldsymbol{B}^{\rm xc}_{\rm corr} = \big[p^2\boldsymbol{B}_{\rm xc}\big] + 2(\boldsymbol{p}\boldsymbol{B}_{\rm xc})\cdot(\boldsymbol{p}-e\boldsymbol{A}) + 2(\boldsymbol{p}\cdot\boldsymbol{B}_{\rm xc})(\boldsymbol{p}-e\boldsymbol{A}) \\ + 4[\boldsymbol{B}_{\rm xc}\cdot(\boldsymbol{p}-e\boldsymbol{A})](\boldsymbol{p}-e\boldsymbol{A})\,. \quad (2.85)\end{aligned}$$

These terms can be explained as the spin-orbit coupling due to the exchange field. This is more apparent with the substitution of momentum operators in the Eq. (2.84) as radial and angular parts: $\boldsymbol{p} = \boldsymbol{e}_r\hat{p}_r - \frac{\boldsymbol{r}\times\boldsymbol{L}}{r}$, where $\hat{p}_r = -i\hbar\partial/\partial r$ is the radial part of the momentum operator and $\boldsymbol{e}_r$ is the unit radial vector. As the exchange fields are usually very strong for magnetic systems (ferromagnets), their relativistic corrections cannot be ignored at all.

Apart from the last term in Eq. (2.84), as pointed out, the rest can be written together to form an effective exchange correction field as $\boldsymbol{B}^{\rm xc}_{\rm corr}$ and the same can be re-written as

$$\mathscr{H}_{\rm RCXC} = \frac{\mu_{\rm B}}{8m^2c^2}\boldsymbol{\sigma}\cdot\boldsymbol{B}^{\rm xc}_{\rm corr} + \frac{i\mu_B}{4m^2c^2}[(\boldsymbol{p}\times\boldsymbol{B}_{\rm xc})\cdot(\boldsymbol{p}-e\boldsymbol{A})]\,. \quad (2.86)$$

Notice that, these terms contain all the informations of the interaction among the spins, the electromagnetic fields and the exchange fields. We also state that the inclusion of $\boldsymbol{B}$xc in the Dirac equation is not the same as the inclusion in the

[3]The following vector identity is used for any three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$

$$(\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{\sigma}\cdot\boldsymbol{b})(\boldsymbol{\sigma}\cdot\boldsymbol{c}) = (\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{b}\cdot\boldsymbol{c}) + (\boldsymbol{a}\cdot\boldsymbol{b})(\boldsymbol{\sigma}\cdot\boldsymbol{c}) - \boldsymbol{a}(\boldsymbol{\sigma}\cdot\boldsymbol{b})\cdot\boldsymbol{c} + i(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}$$

Pauli equation. There have been attempts to derive the relativistic correction terms of exchange fields, however, only a component of the exchange field was used [86]. Here, in our derivation, we use a general exchange field and use the corrected FW transformation to obtain the relativistic correction terms - this has not been done before.

2.5.6 Higher-order spin-orbit coupling

The last term of Eq. (2.63) is a higher order in the transformation as it is of the order $1/c^6$. The expansion of the last term gives the Hamiltonian which is recognized as the higher order spin-orbit coupling that takes the form,

$$\mathscr{H}_{\mathrm{SOC}}^{\mathrm{higher}} = -\frac{ie\hbar^2}{16m^3c^4}\boldsymbol{\sigma}\cdot\left[\partial_t \boldsymbol{E}_{\mathrm{tot}} \times (\boldsymbol{p}-e\boldsymbol{A}) + (\boldsymbol{p}-e\boldsymbol{A})\times\partial_t \boldsymbol{E}_{\mathrm{tot}}\right]. \qquad (2.87)$$

Note that, the difference between Eq. (2.82) and Eq. (2.87) are: (i) the latter is $\frac{1}{c^2}$ times smaller than the usual spin-orbit coupling Hamiltonian and (ii) a first-order time-derivative of the electric field is present in the latter one while no time-derivative is involved in the usual spin-orbit coupling. But it is interesting that both of them have similar formulation and they are Hermitian and gauge invariant. The intrinsic electric field is time independent, thus, their contributions will be zero in the higher-order spin-orbit coupling, restricting to the external electric field which *only* contributes. These higher order spin-orbit coupling have not often been discussed in the literature but they can have much importance as well [90]. We note that, the expansion of the last term in Eq. (2.63) further gives another term which is spin independent and thus we do not intend to it include here [90].

To summarize, we have used the DKS Hamiltonian in an electromagnetic environment, however, treated the exchange field separately. Using the corrected FW transformation, we have derived the extended Pauli Hamiltonian that includes all the relativistic correction terms up to an order $1/c^4$. In the next chapters, we will describe how these derived terms contribute to many physical phenomena in magnetic systems.

3. Magnetization Dynamics: Precession, Damping, Nutation, Torque

3.1 Introduction

Magnetism was known since ancient ages; it was probably the ancient Greeks who first reflected upon the unusual properties of magnetite (iron ore) [129]. It was used in devices, for example the most common use were the compasses in navigation. Despite of being used in the ancient times, the fundamental origin of magnetism in a material was mysterious and concealed. In the middle of 19^{th} century James Clerk Maxwell compiled the theory of electromagnetism, which was the greatest step in explaining the electricity and magnetism [78]. In particular, the Maxwell equations establish a relationship between the electric systems (charge density and electric field) and the magnetic systems (current density and magnetic field) and they describe how the change of electric field produces a magnetic field and vice-versa. These equations provide the cornerstone for understanding electromagnetic phenomena and are even heavily used today [130].

Although magnetism and electricity can be expressed through the Maxwell equations, the origin of spontaneous magnetism (presence of magnetic field without any application of an electric field) in some materials cannot be understood from the Maxwell equations. To this end, the analysis of magnetism from a quantum theory is required. Here, in the following, first the classical theory of magnetic moment is given, followed by the quantum theory and theory of magnetic ordering.

3.2 Magnetic moment

In the classical theory of electromagnetism, when a charged particle moves in a current loop, it produces a magnetic dipole moment. The magnitude of the moment is proportional to the amount of current and the area of the loop. Similarly, when the electron orbits around a nucleus (atom), it gives rise to an angular momentum. The Einstein-de Haas effect[1] could explain that magnetism is related to angular momentum, specifically, the magnetic moment is proportional to the angular momentum of the orbiting electron as

$$\boldsymbol{\mu}_{\mathrm{e}} = \frac{e}{2m}\boldsymbol{L}. \tag{3.1}$$

[1] A freely unmagnetized suspended rod starts to rotate to compensate the angular momentum when a magnetic field is induced in it.

The orbital angular momentum of the electron is $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ and due to the negative electronic charge, $e = -|e|$, the magnetic moment and the angular momentum are antiparallel to each other [125].

However, the classical explanation of arising magnetic moment is not sufficient, because of the following reasoning. The partition function (describes the statistical properties e.g., energy, entropy etc. of a system under thermodynamic equilibrium) of N classical particles each of charge q can be written as

$$Z = \int\int ... \int \mathrm{Exp}\Big[-\frac{E\{\boldsymbol{r}_i, \boldsymbol{p}_i\}}{k_\mathrm{B} T} \Big] d\boldsymbol{r}_1 d\boldsymbol{r}_2 ... d\boldsymbol{r}_N \, d\boldsymbol{p}_1 d\boldsymbol{p}_1 ... d\boldsymbol{p}_N \,, \tag{3.2}$$

where $E\{\boldsymbol{r}_i, \boldsymbol{p}_i\}$ is the energy associated to the N charged particles with positions and momentum $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$ respectively. When a magnetic field (defined by the vector potential $\boldsymbol{A}(\boldsymbol{r},t)$) is applied to the system, the momentum gets changed by $\boldsymbol{p}_i \rightarrow (\boldsymbol{p}_i - e\boldsymbol{A})$ (see the end of Sec. 2.2.1). Therefore the free energy of the system, $F = -k_\mathrm{B} T \log Z$ is not a function of magnetic field. The corresponding magnetization is $\boldsymbol{M} = -\partial F / \partial \boldsymbol{H}$ is then expected to be zero, which indicates that considering the ensemble of classical charged particles, one cannot explain the spontaneous magnetization of materials. The aforementioned statement is known as Bohr van-Leeuwen theorem. The theorem justifies that the magnetic properties of materials can be properly accounted only when we go beyond classical physics, into quantum mechanics.

The *intrinsic* property of an electron is spin, a purely quantum mechanical concept. The electrons not only move in an orbit around the nucleus, but also precess around their own axes, leading to the origin of spin angular momentum. For electrons, the spin quantum number, s takes the value $\frac{1}{2}$ and the spin algebra is connected to the Pauli spin matrices $\boldsymbol{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ in spinor space. The spin angular momentum is then denoted by $\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$. Similar to Eq. (3.1), the spin moment is expressed as

$$\boldsymbol{\mu}_\mathrm{s} = g \frac{e}{2m} \boldsymbol{S}, \tag{3.3}$$

where $g \approx 2$, is the Landé g-factor for electrons.

The values of any component of spin angular momentum are discrete and can take only one of the $2s+1$ values, provided that the normal component is determined by the quantum number $m_s = \pm\frac{1}{2}$ (corresponding to spin up and spin down). This is in accordance with the orbital angular momentum which can also take any of $2l+1$ values, where the orbital quantum number is l and the z-components are determined by $m_l = \pm l, \pm(l-1)$ and so on. The magnitude of the total spin angular momentum for atoms with several electrons is found as $S = \sum m_s$, while the orbital angular momentum is $L = \sum m_l$. The total magnetic moment is then the sum of the orbital contribution and the spin contribution.

For determining the total moments of electrons on atoms the filling of electron shells is important. These values of the moment and angular momentum

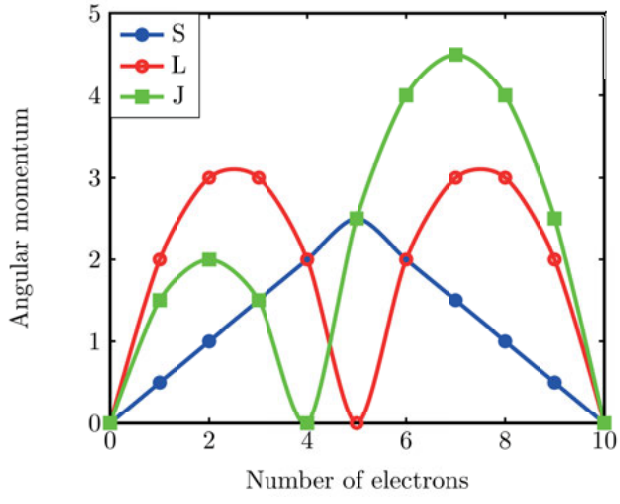

Figure 3.1. Hund's rules for the electrons in a d orbital. S, L and J represent the spin, orbital and total angular momentum, respectively. A d orbital can contain a maximum of 10 electrons. According to the Hund's rule, the filling of electrons in the available shells dictates the angular momentum in the ground state configuration.

can be calculated from the ground of the electronic configuration of an atom. The ground state can be obtained from the Hund's rules that have three main statements:
(a) *Maximize S*: this means the spin up electrons will be arranged first in the available shells and then the spin down, or the other way around. For example, if there are two electrons, the total spin angular momentum will be $S = 1/2 + 1/2 = 1$. The blue plot in Fig. 3.1 refers to such an arrangement for d orbital ($l = 2$).
(b) *Maximize L*: this means once S is maximized, the choice of l has to be such that L is maximized. For example, a d orbital has the z-component of angular momentum as $m_l = \pm 2, \pm 1, 0$ and therefore for two electrons the values of m_l is chosen as $+2$ and $+1$ so that the total orbital angular momentum is $L = 2 + 1 = 3$. The red plot in Fig. 3.1 refers to such an arrangement for electrons filling the d orbital.
(c) *Maximize J*: this means if the shell is more than half-filled, the total angular momentum is taken as $J = L + S$, otherwise, the total angular momentum can be calculated as $J = L - S$. For example, the d orbital can have maximum of 10 electrons, therefore the availability of two electrons follows the less than half filled arrangement. The total angular momentum will be $J = 3 - 1 = 2$. The green plot in Fig. 3.1 refers to such an arrangement for the d orbital.

Therefore, using the Hund's rules, one can conclude that the partially filled shells give rise to nonzero magnetic moment in the ground state and thus can explain the origin of magnetism.

♣ Magnetic ordering

Depending on the orientation of the magnetic moments, the magnetic systems can be classified into mainly three classes: *ferromagnet*, *antiferromagnet* and *ferrimagnet*. The ferromagnetic order is obtained when the magnetic moments are aligned in parallel to each other, producing a finite (nonzero) value of net magnetization. The antiparallel alignment of the magnetic moments gives rise to the antiferromagnetic orders resulting a net zero magnetization. The ferrimagnetic order arises when the magnetic moments are antiparallel to each other but the magnitude of the magnetic moments are not equal. As a consequence, the net magnetization is nonzero. Apart from these three usual magnetic ordering, there exist many complex orders such as the noncollinear magnetic moments.

Microscopically, the ordered magnetic structure is the result of correlation among the direction of electrons (spins). This correlation is due to the fact that the space symmetry of the wave function depends on the magnitude of the resultant spin of the system of electrons. Thus, different values of the energy of the system in general correspond to wave functions with different space symmetry. This effect is called exchange effect and the dependence of energy on the magnitude of spins i.e., the dependence of energy on the symmetry properties of the wave functions, is referred to as being due to the exchange interactions. The exchange interaction is mainly responsible for the long-range order in a ordered magnetic structure. The exchange interaction can be modeled by the Hamiltonian following the simple Heisenberg prescription [125, 131].

3.3 Magnetization

The magnetization is defined by the magnetic moment per unit volume. Therefore, one must take the expectation value of the total magnetic moment $\boldsymbol{\mu}$ as [131]

$$\boldsymbol{M} = \int \psi^{\star} \boldsymbol{\mu} \psi \prod_i d\boldsymbol{r}_i \,, \tag{3.4}$$

where ψ is a many-body wave function. But, the main problem lies with the fact that the exact many-body wave function is never known for a particular system. To approximate it, one introduces a set of eigenfunctions of the system so that the wave function can be written in a superposition of the eigen functions (ϕ_k) of the system as $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N, t) = \sum_k a_k(t)\, \phi_k(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N)$. The magnetization can then be recast as

$$\boldsymbol{M}(\boldsymbol{r}, t) = \sum_{k,k'} a_k^{\star}(t) a_{k'}(t)\, \boldsymbol{\mu}_{kk'} \,, \tag{3.5}$$

where $\boldsymbol{\mu}_{kk'}$ defines the magnetic moment matrix elements. Finally, taking a time-average and realizing the fact that the statistical density matrix $\rho_{kk'} =$

$\overline{a_k^\star(t)a_{k'}(t)}$, the magnetization of the entire system becomes

$$\boldsymbol{M} = \mathrm{Tr}\left[\rho\boldsymbol{\mu}\right] . \tag{3.6}$$

However, if one is interested in defining a magnetization at the point $\boldsymbol{r}$, we have to introduce the magnetic moment density as follows

$$\boldsymbol{\mu}(\boldsymbol{r}) = \frac{1}{2}\sum_j \Big[\boldsymbol{\mu}_j\,\delta(\boldsymbol{r}-\boldsymbol{r}_j) + \delta(\boldsymbol{r}-\boldsymbol{r}_j)\,\boldsymbol{\mu}_j\Big] . \tag{3.7}$$

For the orbital moment, as it is a function of position and momentum operator, the symmetrized expression has been used. But, for the spin moment, the symmetric product is not mandatory because the spin moment and the density operator commute with each other. Note that, the δ-function defines the density operator which has the dimension of inverse volume. For a strong ferromagnet, the orbital moment is usually quenched because of the cubic crystal field and the magnetization is largely given by the spin moment only. Thus, throughout the dynamics, we consider the spin moments, the orbital moments are ignored. However, it is important to note that a full magnetization description, including orbital and spin contributions, is necessary to describe the dynamics for other systems (e.g., lanthanides). We write the magnetization as

$$\boldsymbol{M}(\boldsymbol{r},t) = \frac{g\mu_\mathrm{B}}{\hbar}\sum_j \mathrm{Tr}\left[\rho\,\boldsymbol{S}_j\,\delta(\boldsymbol{r}-\boldsymbol{r}_j)\right] , \tag{3.8}$$

where μ_B is the Bohr magneton which serves as a unit of the magnetic moment. Once the magnetization is defined for a system, the magnetization dynamics can be analyzed appropriately.

3.4 Magnetization dynamics

The time evolution of a magnetization element has been successfully described within the phenomenological Landau-Lifshitz (LL) equation of motion [37]. The corresponding dynamics involves a precession of the magnetization vector around an effective field and a subsequent transverse relaxation of magnetization that is called damping. Because of the form of the transverse damping, the magnitude of the magnetization does not change during the LL dynamics. The proposed phenomenological LL equation for a magnetization element, $\boldsymbol{M}(\boldsymbol{r},t)$ (magnetic moment per unit volume at each point) is defined as

$$\frac{\partial\boldsymbol{M}}{\partial t} = -\gamma\boldsymbol{M}\times\boldsymbol{H}_\mathrm{eff} - \lambda\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{H}_\mathrm{eff}) , \tag{3.9}$$

where γ is the gyromagnetic ratio, $\boldsymbol{H}_\mathrm{eff}$ is the effective magnetic field and λ is the *scalar* isotropic damping parameter that dictates the rate at which

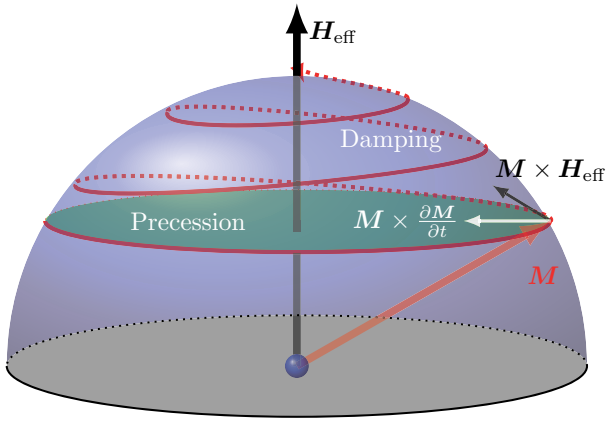

Figure 3.2. Magnetization dynamics has been schematically depicted. The red arrow shows the magnetization vector. It precesses around the effective field and the damping curves are shown in red color as well.

the z-magnetization direction is dissipated. The effective field includes the externally applied field, demagnetizing field, magnetocrystalline anisotropy, spin-spin exchange interaction etc. [132, 133].

However, to include the effect of large (anomalous) damping for real ferromagnets, Gilbert modified the preexisting LL equation, accounting for the Gilbert damping parameter α in the Landau-Lifshitz-Gilbert (LLG) dynamics [38–40]

$$\frac{\partial \boldsymbol{M}}{\partial t} = -\gamma' \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} + \alpha \boldsymbol{M} \times \frac{\partial \boldsymbol{M}}{\partial t}\,. \tag{3.10}$$

This LLG equation of motion can be re-written in the form of LL dynamics and both the LL and LLG equations are shown to be mathematically equivalent to each other with the fact that $\gamma = \gamma'/(1+\alpha^2 M^2)$ and $\lambda = \alpha\gamma'/(1+\alpha^2 M^2)$. The corresponding two dynamics (precession and transverse damping) are schematically depicted in the Fig. 3.2. Originally the damping parameters were attributed to the relativistic origins and predicted to be scalar quantities [37].

For practical purposes, the extension to the LLG equation has been useful to describe several other dynamical effects. For example, to include the effect of magnetic inertia (spin nutation), the LLG equations have been extended with an inertia term. Nutation means a continuous change of the precession angle of the magnetization vector while it precesses around an effective field. This effect has been predicted to be observed in ultrashort timescales i.e., sub-ps [134, 135]. While the Gilbert damping effect enters into the LLG equation as a first order time-derivative of the magnetization, the nutation dynamics has been related to the *second* order time-derivative of the magnetization [42, 135].

Another extension of the LLG equation is to include the effect of spin currents, which naturally appear in the spin continuity equation [43, 44]. These spin currents can be of nonrelativistic and of relativistic origin. By including the effect of spin currents in the extended LLG equation, it is possible to describe phenomenologically the current induced phenomena such as spin-transfer torque, spin-orbit torque etc.

The derivations of the LLG equations and their extension have been purely phenomenological and lack the footing obtained from a fundamental equation. Here, in the below, we provide the origin of all the dynamical motions from the DKS theory and the extended Pauli Hamiltonian that has been derived in Chapter 2. Before proceeding to the details of magnetization dynamics, we point out that using the the Hamiltonian in Eq. (2.74), it is possible to derive the single Dirac particle spin dynamics (see **paper VII** for details).

3.4.1 Our approach towards magnetization dynamics

To find the magnetization dynamics, it is important to define the magnetization properly. In this regard, we note that the magnetization is defined for a many-body Hamiltonian in Eq. (3.8). However, as our starting point is the DKS Hamiltonian which is a single particle theory, one cannot use the definition in Eq. (3.8). Instead, the definition of magnetization for a single particle theory can be written as the expectation value of the spin density

$$\boldsymbol{M}(\boldsymbol{r},t) = \frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left[\rho\hat{\boldsymbol{S}}(t)\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}(t))\right]. \tag{3.11}$$

Next, we use the dynamical evolution of magnetization within the Heisenberg picture where the operators are time-dependent, however, the states do not evolve in time. Thus, the magnetization dynamics will be given by taking the time-derivative in the both sides of Eq. (3.11), we arrive to

$$\frac{\partial\boldsymbol{M}(\boldsymbol{r},t)}{\partial t} = \frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left[\rho\,\frac{\partial\hat{\boldsymbol{S}}}{\partial t}\,\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}) + \rho\,\hat{\boldsymbol{S}}\,\frac{\partial}{\partial t}\delta(\boldsymbol{r}-\hat{\boldsymbol{r}})\right]. \tag{3.12}$$

In a next step, we employ the Heisenberg equation[2] of motion for the time evolutions of the operators. At the same time, we use the charge continuity equation such that $\frac{\partial}{\partial t}\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}) = -\frac{1}{e}\boldsymbol{\nabla}\cdot\hat{\boldsymbol{j}}_{\mathrm{e}}$, where $\hat{\boldsymbol{j}}_{\mathrm{e}}$ being the charge current density operator. Therefore, the same dynamics as in Eq. (3.12) can further be

[2]The Heisenberg equation of motion is defined for any operator $\hat{A}$ as

$$\frac{\partial\hat{A}}{\partial t} = \frac{1}{i\hbar}\left[\hat{A},\mathscr{H}\right]$$

written as

$$\frac{\partial \boldsymbol{M}}{\partial t}+\frac{1}{e}\frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left[\rho\,\hat{\boldsymbol{S}}(\boldsymbol{\nabla}\cdot\hat{\boldsymbol{j}}_{\mathrm{e}})\right]=\frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left[\frac{1}{i\hbar}\rho\,\left[\hat{\boldsymbol{S}},\mathscr{H}_{\mathrm{spin}}\right]\,\delta(\boldsymbol{r}-\hat{\boldsymbol{r}})\right]. \quad (3.13)$$

Notice that, the associated Hamiltonian on the right-hand side of the last equation can, in principle, be the full extended Pauli Hamiltonian that has been derived in Eq. (2.74). However, in practice, only the spin Hamiltonian contributes, otherwise the commutator vanishes. In addition, we consider that the space derivative of the spin angular momentum is zero. This enables one to write the Eq. (3.13) in a compact way as [3]:

$$\frac{\partial \boldsymbol{M}}{\partial t}+\frac{1}{e}\frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left[\rho\,\boldsymbol{\nabla}\cdot(\hat{\boldsymbol{j}}_{\mathrm{e}}\otimes\hat{\boldsymbol{S}})\right]=\frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left[\frac{1}{i\hbar}\rho\,\left[\hat{\boldsymbol{S}},\mathscr{H}_{\mathrm{spin}}\right]\,\delta(\boldsymbol{r}-\hat{\boldsymbol{r}})\right]. \quad (3.14)$$

Now, we define the spin current tensor as $J_{\mathrm{S}}=\frac{1}{e}\frac{g\mu_{\mathrm{B}}}{\hbar}\mathrm{Tr}\left(\rho(\hat{\boldsymbol{j}}_{\mathrm{e}}\otimes\hat{\boldsymbol{S}})\right)$, where $\otimes$ is the tensor product. In order to calculate the trace, we further employ the separation of fast and slow variables. The orbital degrees of freedom (current densities) are usually the fast variables, where as the spin degrees of freedom (magnetization) are the slow variables. Now, due to the spin-orbit coupling present in a system, one cannot separate the fast and slow variables because neither the orbital nor the spin degrees of freedom are conserved. However, in this case, magnetic systems are being treated where the exchange energies are much larger than the available relativistic spin-orbit coupling. Therefore, one can approximate the density matrix to write it in the form as: $\rho=\rho_{\mathrm{o}}\otimes\rho_{\mathrm{s}}$, where ρ_{o} and ρ_{s} define the reduced density matrix for orbital and spin parts, respectively. In particular, if $\hat{O}$ define the orbital degrees of freedom and $\hat{S}$ are for the spins, the trace in Eq. (3.14) can be approximated as following $\mathrm{Tr}\big(\rho\hat{O}\hat{S}\big)=\mathrm{Tr}\big(\rho_{\mathrm{o}}\hat{O}\big)\,\mathrm{Tr}\big(\rho_{\mathrm{s}}\hat{S}\big)$. One has also to remember that, in the out-of-equilibrium dynamics, the orbital degrees of freedom (electric motion) are much faster than the dynamics of spin (magnetization). With this assumption, one can write the spin current tensor as $J_{\mathrm{S}}=\frac{1}{e}\,(\boldsymbol{j}_{\mathrm{e}}\otimes\boldsymbol{M})$ [136].

The associated current density has to be calculated from the Hamiltonian in Eq. (2.74). One can notice that, as the Hamiltonian contains nonrelativistic and relativistic terms, the separation of nonrelativistic and relativistic current densities can be written as $\boldsymbol{j}_{\mathrm{e}}=\boldsymbol{j}_{\mathrm{e}}^{\mathrm{NR}}+\boldsymbol{j}_{\mathrm{e}}^{\mathrm{R}}$. Similarly, the involved spin Hamiltonian in Eq. (3.14) can be written as a sum of relativistic and nonrelativistic

[3] For any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ the divergence of their tensor product is expressed as

$$\boldsymbol{\nabla}\cdot(\boldsymbol{a}\otimes\boldsymbol{b})=(\boldsymbol{\nabla}\cdot\boldsymbol{a})\boldsymbol{b}+(\boldsymbol{a}\cdot\boldsymbol{\nabla})\boldsymbol{b}$$

terms. These contributions can be recovered from the Eq. (2.74) as:

$$\mathscr{H}_{\text{spin}}^{\text{NR}} = -\mu_B \boldsymbol{\sigma} \cdot (\boldsymbol{B} + \boldsymbol{B}_{\text{xc}}) \,, \tag{3.15}$$

$$\begin{aligned}\mathscr{H}_{\text{spin}}^{\text{R}} = &\frac{\mu_B}{8m^2c^2} \boldsymbol{\sigma} \cdot \boldsymbol{B}_{\text{corr}}^{\text{xc}} + \frac{e\hbar}{8m^3c^2} \left\{ (\boldsymbol{p} - e\boldsymbol{A})^2 , \boldsymbol{\sigma} \cdot \boldsymbol{B} \right\} \\ &- \frac{e\hbar}{8m^2c^2} \boldsymbol{\sigma} \cdot [\boldsymbol{E}_{\text{tot}} \times (\boldsymbol{p} - e\boldsymbol{A}) - (\boldsymbol{p} - e\boldsymbol{A}) \times \boldsymbol{E}_{\text{tot}}] \\ &- \frac{ie\hbar^2}{16m^3c^4} \boldsymbol{\sigma} \cdot [\partial_t \boldsymbol{E}_{\text{tot}} \times (\boldsymbol{p} - e\boldsymbol{A}) + (\boldsymbol{p} - e\boldsymbol{A}) \times \partial_t \boldsymbol{E}_{\text{tot}}] \,. \end{aligned} \tag{3.16}$$

Therefore, it is possible to write down the magnetization dynamics due to non-relativistic and relativistic contributions. In the following, we describe the effects of the spin Hamiltonian terms and current densities in the magnetization dynamics i.e., Eq. (3.14).

3.5 Precession

We have described the origin of magnetic moments in Sec. 3.2. It is also interesting to note that when a magnetic moment, $\boldsymbol{\mu}$ is kept in a magnetic field $\boldsymbol{B}$, the energy of the magnetic moment is given by $E = \boldsymbol{\mu} \cdot \boldsymbol{B}$. This means that the energy will be minimized when the angle between the magnetic moment and magnetic field is zero. As a result, if the angle is nonzero, there will be a torque acting on the magnetic moment and the torque is the rate of change of angular momentum, given by

$$\frac{d\boldsymbol{\mu}}{dt} = -\gamma \boldsymbol{\mu} \times \boldsymbol{B} \,. \tag{3.17}$$

Thus, the existence of precession of a magnetic moment around a magnetic field is realized. Note that, this scenario is exactly analogous to the case of a classical symmetric spinning top [137].

In our approach of the magnetization dynamics, the precession of magnetization derives from the Zeeman-like coupling terms in Eq. (3.15) and their commutators with the spin momentum in Eq. (3.14). The derivations are explicitly carried out in the **paper IV**. With an effective Zeeman field $\boldsymbol{B}_{\text{eff}}$ the precession dynamics becomes [124]

$$\frac{\partial \boldsymbol{M}}{\partial t} = -\gamma \boldsymbol{M} \times \boldsymbol{B}_{\text{eff}} \,. \tag{3.18}$$

Using the linear relationship of magnetic induction and magnetization as $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M})$, the precession dynamics can be written in the conventional form as in Eq. (3.9). In addition, the relativistic counterpart of the precession is derived from the first two terms in the Hamiltonian Eq. (3.16).

It is also important to note that the intrinsic spin-orbit coupling [third term in Eq. (3.16)] also contributes to the relativistic precession for the symmetric crystal potential (see **paper IV** for details).

3.6 Damping

The dissipation of energy during the precession of magnetization, is known as damping. The traditional LLG damping is in the transverse direction as shown in Fig. 3.2. However, there are other types of damping effects reported in the literature e.g., exchange relaxation of Bar'yakhtar type [132, 138, 139], longitudinal relaxation of Bloch type [140, 141]. Here, in the below, we focus on the Gilbert damping and its derivation from the DKS Hamiltonian.

3.6.1 Our theory

Already in the original proposition of Landau and Lifshitz, they attributed the damping processes to relativistic effects. In accordance, within our approach, the Gilbert damping derives from the relativistic external spin-orbit coupling and their commutator with the spin momentum in the right-hand side of Eq. (3.14). The explicit derivation is not provided here as it has been nicely carried out in **paper IV**. Importantly, the Gilbert damping parameter that we derive is not a scalar, rather a *tensor*. We mention that the original Gilbert parameter has been proposed as a scalar. For a system driven by a harmonic (ac) field, our derived Gilbert damping dynamics and the damping parameter read as

$$\frac{\partial \boldsymbol{M}}{\partial t} = \boldsymbol{M} \times \left[A \cdot \frac{\partial \boldsymbol{M}}{\partial t} \right], \tag{3.19}$$

$$A_{ij} = -\frac{e\mu_0}{8m^2c^2} \sum_{l,k} \Big[\langle r_i p_k + p_k r_i \rangle - \langle r_l p_l + p_l r_l \rangle \delta_{ik} \Big] \times (\mathbb{1} + \chi^{-1})_{kj} \,. \tag{3.20}$$

The details of the involved parameters are immediately described below. We found that the damping parameter has two main contributions:

♣ **Electronic**: This involves the expectation value of the matrix elements of the product of different components of position and momentum operators. This contribution is present in the first part of Eq. (3.20) within [...]. The expectation value can be calculated with the states from e.g., the DKS Hamiltonian in Eq. (2.71). As provided in Appendix C of **paper IV**, the interband matrix elements can be calculated as

$$\langle r_i p_j \rangle = -\frac{i\hbar}{2m} \sum_{n,m,\boldsymbol{k}} \frac{f(E_{n\boldsymbol{k}}) - f(E_{m\boldsymbol{k}})}{E_{n\boldsymbol{k}} - E_{m\boldsymbol{k}}} p^i_{nm} p^j_{mn} \,, \tag{3.21}$$

where $f(E_{n\boldsymbol{k}})$ defines the Fermi-Dirac distribution function and $E_{n\boldsymbol{k}}$ represents the band energy. For an intraband contribution of the electronic contributions, the gradient of Fermi-Dirac distributions with respect to the band energy, has to be taken into consideration. Importantly, these contributions can be calculated *ab initio* using Eq. (3.21). The interband and intraband contributions of the damping parameter has been found from first principles calculations within the torque-torque correlation model as well [142–144]. These contributions reflect as a tensor, however, the off-diagonal matrix elements have

been found to be symmetric. Therefore, the electronic contribution provides a Gilbert damping parameter as a tensor which is symmetric.

♣ **Magnetic**: This involves the magnetic susceptibility tensor, χ, for a harmonic field. Here is the reason: for a ferromagnet with uniform magnetization, $\boldsymbol{M}$, the following relation holds $\boldsymbol{M} = \boldsymbol{M}_0 + \boldsymbol{m}(t)$, where the ac magnetization, $\boldsymbol{m}(t)$, is driven by an ac field as $\boldsymbol{m}(t) = \chi \cdot \boldsymbol{h}(t)$ and $\boldsymbol{M}_0$ is the spontaneous magnetization that would be present without any applied field. This linear relationship of field and magnetization through the susceptibility, is only limited to the case where $\boldsymbol{h}(t) \sim e^{i\omega t}$; so that the susceptibility depends on the frequency of the driving field $\chi = \chi(\omega)$. Therefore, the presence of susceptibility is only valid for the excitation field $\boldsymbol{h}(t)$ which is a single harmonic. For a general time-dependent field, the LLG equation will have additional torques. We will come back to it in Sec. 3.6.4. Nonetheless, the susceptibility tensor can be calculated from e.g., the spin-spin response function. Different expressions of χ have been discussed in the context of Gilbert damping tensor [145–147]. These expressions would be suitable for the *ab initio* calculations of χ from a DFT framework. From Eq. (3.21), we notice that the electronic contribution is imaginary (the complex number i present in the expression). Therefore, only the imaginary part of the susceptibility will contribute to the energy dissipations. These results are in agreement with other findings as well [148–151].

Note that, the tensorial form of the damping parameter has been found in agreement with other recent findings [145, 149, 152–154]. However, we provide the microscopic derivation while starting from the fundamental DKS Hamiltonian [124].

Now, any tensor can be decomposed in a symmetric and an antisymmetric part. These parts will take the form as $A_{ij}^{\text{sym}} = \frac{1}{2}(A_{ij} + A_{ji})$ (symmetric) and $A_{ij}^{\text{antisym}} = \frac{1}{2}(A_{ij} - A_{ji})$ (antisymmetric) respectively. The latter can further be written as a product of a Levi-Civita tensor, ε_{ijk}, and a vector, $\boldsymbol{D}$ as: $A_{ij}^{\text{antisym}} = \varepsilon_{ijk} D_k$. This contribution to the damping parameter, we recognize as the Dzyaloshinskii-Moriya (DM) terms. However, the symmetric part can be decomposed in two different contributions as $A_{ij}^{\text{sym}} = \alpha\delta_{ij} + \mathbb{I}_{ij}$, where α contains the diagonal components i.e., isotropic Heisenberg-like and $\mathbb{I}_{ij}$ defines the Ising-like contributions. We mention here that if the Heisenberg contributions are such that $\alpha = \frac{1}{3}\text{Tr}(A_{ij}^{\text{sym}})$, the trace of Ising-like contributions will become zero. With the aforementioned descriptions, the derived Gilbert damping dynamics can be split into three parts as

$$\frac{\partial \boldsymbol{M}}{\partial t} = \alpha \boldsymbol{M} \times \frac{\partial \boldsymbol{M}}{\partial t} + \boldsymbol{M} \times \left[\mathbb{I} \cdot \frac{\partial \boldsymbol{M}}{\partial t} \right] + \boldsymbol{M} \times \left[\boldsymbol{D} \times \frac{\partial \boldsymbol{M}}{\partial t} \right]. \tag{3.22}$$

It is important here to point out that the DM terms can be expanded using the usual vector identities. Since the local magnetization length is conserved i.e.,

$\boldsymbol{M} \cdot \partial \boldsymbol{M}/\partial t = 0$, the dynamics in Eq. (3.22) can be recast as

$$(1 + \boldsymbol{M} \cdot \boldsymbol{D}) \frac{\partial \boldsymbol{M}}{\partial t} = \alpha \boldsymbol{M} \times \frac{\partial \boldsymbol{M}}{\partial t} + \boldsymbol{M} \times \left[\mathbb{I} \cdot \frac{\partial \boldsymbol{M}}{\partial t} \right]. \tag{3.23}$$

Therefore, one can see that the DM-like damping parameter only contributes to the renormalization of the LLG equation. Nonetheless, the presence of DM-like damping can become important for magnetic textures (e.g., domain wall motion, skyrmion dynamics etc.). Overall, we notice that Eq. (3.23) provides more insight to the the LLG equation and particularly the Gilbert damping. Phenomenologically it was thought that the latter is only a scalar quantity. However, we derive that apart from the scalar and isotropic Gilbert damping (the first term in the right side of Eq. (3.23)), there exists a tensorial Gilbert damping which is anisotropic (the second term in right side of Eq. (3.23)). We also point out that often for simulations of LLG equations, the LL form of the spin dynamics is sought. Thus, a transformation of LLG to LL is needed. For a scalar damping parameter, this transformation is rather simple and straightforward. However, the transformation is much more involved for the case of anisotropic and tensorial Gilbert damping (see Appendix B of **paper IV**). In the context of Gilbert damping, in the following, we present one of the important models that has been extensively used to evaluate the *ab initio* damping parameter and to this end a comparison is made between our derived microscopic theory and the existing theory - so called torque-torque correlation model.

3.6.2 Torque-torque correlation model

In 1976, Kamberský [155] computed the damping parameter from the ac susceptibility, his theory was initially called the breathing Fermi surface model and later refined and termed torque-torque correlation model. It was suggested that the source of damping is a combined effect of the spin-orbit interaction and the scattering of electrons by the lattice or defects [156, 157]. The theory describes that the damping originates due to two processes [142]:
(i) A magnon decays into an electron-hole pair.
(ii) The scattering of electron and hole with the lattice.
The physical arguments are described in the following: the spin-orbit torque (variation of spin-orbit energies of the states as the magnetization changes direction) produces a "breathing Fermi surface", which in turn annihilates an uniform mode of spin wave (magnon) and creates an electron-hole pair. The electron-hole pair then travels through the lattice with scattering and eventually collapses. The electron and hole go through lattice and are thought to be a single quasiparticle with a lifetime given by the electron-lattice scattering time, τ. If the electron and hole occupy the same band, one names it *intraband* transition. However, the transition between two different bands is called *interband*.

The torque-torque correlation model predicts that the damping rate (the damping parameter divided by the scattering time) can be written as [142, 143],

$$\lambda = \pi\hbar\frac{\gamma'^2}{\mu_0}\sum_{n,m}\int\frac{dk^3}{(2\pi)^3}\,|\Gamma^-_{nm}(k)|^2\,W_{nm}(k)\,, \tag{3.24}$$

with the redefined gyromagnetic ratio $\gamma' = \frac{g\mu_0\mu_B}{\hbar}$. The matrix elements $\Gamma^-_{nm}(k)$ are given as $\Gamma^-_{nm}(k) = \langle n,k|\,[\sigma^-,\mathscr{H}_{\rm soc}]\,|m,k\rangle$. This measures the transition matrix elements between band indices m and n by the spin-orbit torque, where $\sigma^- = \sigma_x - i\sigma_y$. The nature of the scattering events are measured by the weight of the spectral overlap integral

$$W_{nm}(k) = \int d\varepsilon\,\eta(\varepsilon)A_{nk}(\varepsilon)A_{mk}(\varepsilon)\,. \tag{3.25}$$

Obviously, $n \neq m$ ($n = m$) determines the *interband* (*intraband*) contributions. $\eta(\varepsilon)$ is the first order derivative of Fermi function with respect to energy - $df/d\varepsilon$. The electron spectral distribution function is represented as $A_{nk}(\varepsilon)$ and they follow Lorentzians profile in energy space.

The *ab initio* damping rates by torque-torque correlation model produces qualitatively the experimental behavior for ferromagnets e.g., iron, cobalt, nickel, however, quantitatively the calculated damping rates are off from the experimental values [142, 143, 145, 158–161].

♣ Physical understanding of the scenario

As the magnetization precesses around an effective field, the energies of the states change through variations in the spin-orbit coupling energy and transitions between the states. These two effects, (a) changing energies of the states, (b) transitions between the states, contribute to the total effective field [143]. If the energy of a single electronic state is denoted by ε_{nk} and the population of the state is ρ_{nk}, the total electronic energy of the system will be given as $E = \sum_{nk}\rho_{nk}\varepsilon_{nk}$. Therefore, the effective field is calculated as the variation of the energy with respect to the the magnetization direction (the magnitude of magnetization does not change within LLG formulation, only the magnetization direction changes) as [143]

$$\boldsymbol{H}_{\rm eff} = -\frac{1}{\mu_0}\frac{\partial E}{\partial \boldsymbol{M}} = -\frac{1}{\mu_0}\sum_{n,k}\left[\rho_{nk}\,\frac{\partial\varepsilon_{nk}}{\partial\boldsymbol{M}} + \frac{\partial\rho_{nk}}{\partial\boldsymbol{M}}\,\varepsilon_{nk}\right]. \tag{3.26}$$

The first term determines the changes of spin-orbit torque energy. Spin-orbit coupling mainly causes the Fermi surface to swell and contract as the magnetization precesses around. This effect is extensively called “breathing” of the Fermi surface. It has been also shown that the spin-orbit torque is the reason for *intraband* contributions in the Gilbert damping. On the other hand,

the second term refers to the transition between states when the magnetization precesses and accounts for changes in energy due to the transitions. This latter process does not change the energy of the states, but create electron-hole pair by exciting an electron from lower to upper bands. This produces a "bubbling" of individual electrons on the Fermi surface. It has also been shown that these bubbling contributes to the *interband* terms of the effective field [143].

3.6.3 Comparison of our theory and torque-torque correlation model

Essentially, our derivation of Gilbert damping from the DKS Hamiltonian and the torque-torque correlation model, both are derived from the spin-orbit coupling, a relativistic contribution. In fact, the matrix elements of the commutators between Pauli spin matrices and spin-orbit coupling Hamiltonian is considered in both cases, where *interband* and *intraband* contributions are present. However, a fundamental difference is that the torque-torque correlation model is phenomenologically derived but we derive the Gilbert damping parameter from the fundamental Dirac equation. In the torque-torque model, the Gilbert dynamics is assumed and the corresponding parameter is calculated by the commutator of spin angular momentum with the spin-orbit interaction Hamiltonian, however, our theory does not assume the damping dynamics, it has been derived from the aforementioned commutator [124]. Moreover, the torque-torque correlation model relies on the Fermi surface based calculation, thus it is only applicable to the metals. The semiconductor or insulator systems cannot be treated within this model. In contrast, our derivation does not take into consideration of the Fermi surface, so in principle, our derivation should equally be valid for any systems which are not metallic. It is important to mention that our derivation does not include the spin relaxation effects due to the the interaction with the lattice or magnons or scattering with the defects. Therefore, we do not anticipate the scattering time (in the torque-torque correlation model) within our derived theory. However, we point out that the influence of the electron interaction with quasiparticles can be treated by introducing finite relaxation time through δ, which can be used to evaluate the electronic contribution in Eq. (3.21).

3.6.4 Field-derivative torque

So far the discussion of Gilbert damping and its derivation has been considered in the case when the system is driven by an ac harmonic field which is single harmonic. In many experiments, the system is driven by a pulse (e.g., THz pulse, pulsed laser excitation) which can be nonharmonic. In those cases, the above-mentioned Gilbert damping dynamics will not be adequate to capture the involved physics. To be more specific, as argued before, the definition and

introduction of susceptibility is valid only for the ac field. Once the ac field is switched off, the relation between the magnetization and the magnetic field does not hold anymore (see **paper IV** for details). For a general (nonharmonic) time-dependent field, the Gilbert damping dynamics and the damping tensor will thus be given as follows

$$\frac{\partial \boldsymbol{M}}{\partial t} = \boldsymbol{M} \times \left[\bar{A} \cdot \left(\frac{\partial \boldsymbol{M}}{\partial t} + \frac{\partial \boldsymbol{H}}{\partial t} \right) \right], \tag{3.27}$$

$$\bar{A}_{ij} = -\frac{e\mu_0}{8m^2c^2} \sum_k \left[\langle r_i p_j + p_j r_i \rangle - \langle r_k p_k + p_k r_k \rangle \delta_{ij} \right]. \tag{3.28}$$

Notice that the time-dependent field generates a torque which we name - *field-derivative torque* (FDT). If we consider the corresponding parameter as a scalar, the FDT becomes: $\boldsymbol{M} \times \frac{\partial \boldsymbol{H}}{\partial t}$. The existence of this torque, therefore, leads to the modification of traditional LLG equation for a nonharmonic and time-dependent field. In particular, considering a field pulse which is initially

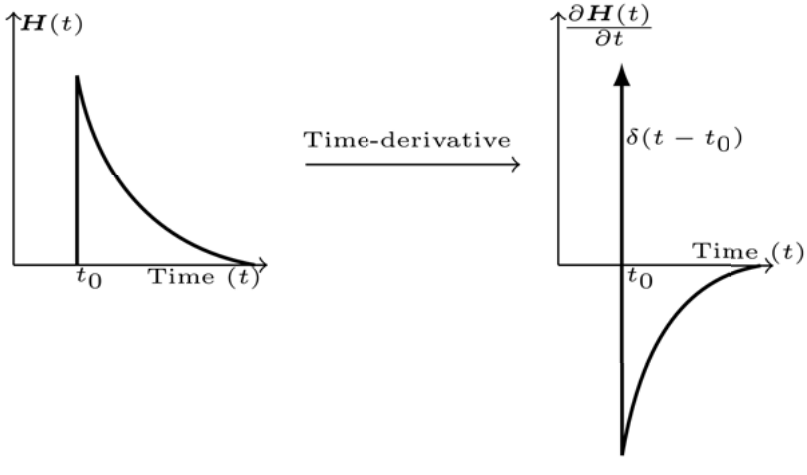

Figure 3.3. The possible shape of the field-derivative term ($\partial \boldsymbol{H}/\partial t$) for a nonharmonic field pulse is shown.

very steep and relaxes slowly back to the initial value or a step-like field pulse for a ultrashort time (see Fig. 3.3), the time-derivative of such a field will initially be a δ-function. Therefore, this very short but large time-derivative will exert a torque which might initiate the switching. We are the first to report the existence of this torque in the literature and it has to be studied in more details to understand the underlying physics of what it might offer. For example, atomistic simulations of the LLG equation including the FDT would be very interesting. In this regard, we mention that, for a scalar damping parameter, α, the LL form of magnetization dynamics including FDT adopts the form (see

Appendix A of **paper VIII**):

$$\left(1+\alpha^2 M^2\right)\frac{\partial \boldsymbol{M}}{\partial t} = -\gamma_0 \boldsymbol{M}\times\left(\boldsymbol{H}_{\text{eff}} - \frac{\alpha}{\gamma_0}\frac{\partial \boldsymbol{H}}{\partial t}\right) - \gamma_0\,\alpha \boldsymbol{M}\times\left[\boldsymbol{M}\times\left(\boldsymbol{H}_{\text{eff}} - \frac{\alpha}{\gamma_0}\frac{\partial \boldsymbol{H}}{\partial t}\right)\right]. \tag{3.29}$$

Finally, we point out that any nonharmonic field can be written as the linear combination of several harmonic fields, where the harmonic fields can have several frequencies. In this case, the appearance of FDT is still valid because the susceptibility can only be introduced for a single harmonic field; when two or more harmonics are present in a field, one has to treat it as a nonharmonic field.

3.6.5 Optical spin-orbit torque

We have already discussed that the external spin-orbit coupling is responsible for the Gilbert damping and field-derivative torque. However, the gauge invariant part of the spin-orbit coupling in Eq. (2.82) is also important. This part of the Hamiltonian can be written in the form as $\boldsymbol{\sigma}\cdot(\boldsymbol{E}\times\boldsymbol{A})$. Now, within the paraxial approximation, the spin angular momentum of an EM field is equal to $\boldsymbol{j}_{\text{s}} = \varepsilon_0(\boldsymbol{E}\times\boldsymbol{A})$ [113, 162]. In consequence, the corresponding interaction

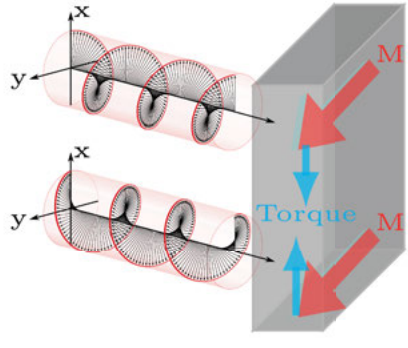

Figure 3.4. Schematic illustration of the possible effect due to optical spin-orbit torque. The material magnetization is shown by the red arrow, the right and left circular polarized light exerts opposite torque on the existing magnetization.

Hamiltonian can be written as $\boldsymbol{\sigma}\cdot\boldsymbol{j}_{\text{s}}$. This interaction is quite similar to the one of the usual spin-orbit coupling. However, the angular momentum of light interacts with the spin here [163, 164]. This implies to a new route of manipulating the spins with light. In the magnetization dynamics this term can be written within the LLG equation as (see **paper IV**)

$$\frac{\partial \boldsymbol{M}}{\partial t} = -\frac{e^2}{2m^2c^2\varepsilon_0}\boldsymbol{M}\times\boldsymbol{j}_{\text{s}}\,. \tag{3.30}$$

This torque is called optical spin-orbit torque. Eq. (3.30) explains the torque by the optical angular momentum acting on the electrons' spin moment, which has been depicted in Fig. 3.4. The manipulation of magnetization using the optical angular momentum has been attempted recently in experiments [32, 165, 166].

3.7 Nutation

In classical mechanics, when a symmetric top is spinning, it shows three different dynamics: (a) precession, (b) nutation and (c) spinning. On top of that if the body is not driven by a constant force, eventually the dynamics will be damped and will come to rest after a certain time. Nutation describes the continuous change of the precession angle as the dynamics progresses. The nutation arises because of the rigid-body dynamics i.e., the inertia of the body. In the same connection, when an electron spin has precession around an effective field, it shows nutation [167]. This effect has been schematically shown in Fig. 3.5. The reason of this motion is known as *magnetic inertia*.

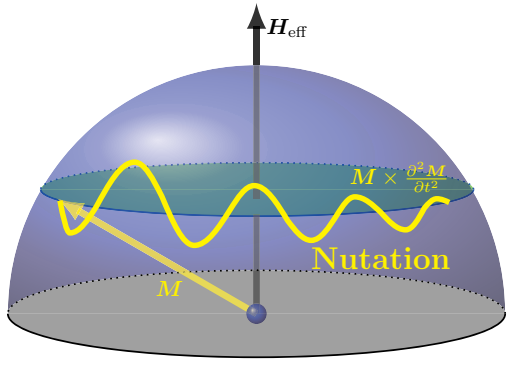

Figure 3.5. The nutation effect has been schematically depicted. The yellow arrow shows the magnetization vector. It precesses around the effective field and the yellow curve shows the nutation on top of the precession plane.

As mentioned before, to include the nutation dynamics, one has to extend the LLG equation to incorporate a torque due to the second order time-derivative of magnetization as [42]

$$\frac{\partial \boldsymbol{M}}{\partial t} = -\gamma \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} + \alpha \boldsymbol{M} \times \frac{\partial \boldsymbol{M}}{\partial t} + \iota \boldsymbol{M} \times \frac{\partial^2 \boldsymbol{M}}{\partial t^2}, \tag{3.31}$$

where ι is the magnetic inertia parameter which is scalar. The origin of magnetic inertial dynamics have not been understood from a fundamental point of view. Nonetheless, there have been efforts to simulate such a dynamics which indicate that the effect is much prominent in shorter timescales e.g., sub-ps

[134, 135]. It was shown that the extension of breathing Fermi surface model could lead to the inertial dynamics as well [151, 168, 169]. Alongside, the magnetic inertial dynamics has been discussed in the case of slow manifold and Hannay angle that the inertial dynamics is rather fast on top of the slow precession dynamics and the nature of nutation angle depends on the slowness variable [170–173]. Experimental observations of inertial dynamics have been reported recently in Fe-Ni permalloy and Co films, confirming its importance in ultrashort timescale [174]. The possible future applications of inertial dynamics have been illustrated by Kimel *et al.* [41] that when a particle does not have inertia, a continuous driving force is needed in order to pull it over a potential barrier and eventually to switch. In contrast, the particle with inertia does not need any continuous deriving force, just an ultrashort pulse is enough to acquire adequate linear momentum for overcoming the potential barrier. Therefore, the magnetic inertia could be of great interest for switching.

Despite of so much interest and possible potential applications, the fundamental understanding of magnetic inertial dynamics was missing until recently. In **paper VII**, we rigorously derive the corresponding dynamics starting from the DKS Hamiltonian and thus giving a strong fundamental origin. In particular, we find that the inertial dynamics derives from the higher-order spin-orbit coupling (the last term in Eq. (3.16)) and its commutator with the spin angular momentum in Eq. (3.14). Similar to the Gilbert damping, for a harmonic field, the inertial dynamics can be expressed as

$$\frac{\partial \boldsymbol{M}}{\partial t} = \boldsymbol{M} \times \left[\mathscr{I} \cdot \frac{\partial^2 \boldsymbol{M}}{\partial t^2} \right], \tag{3.32}$$

where the inertia parameter, $\mathscr{I}$, is a *tensor* given by the following expression

$$\mathscr{I}_{ij} = \frac{\gamma \mu_0 \hbar^2}{8 m^2 c^4} \left(\mathbb{1} + \chi_m^{-1} \right)_{ij}. \tag{3.33}$$

Notice that the phenomenological parameter ι and our derived parameter $\mathscr{I}_{ij}$ are fundamentally different because our derived inertia parameter is a tensor while the original proposition considered it as a scalar quantity. The tensorial behavior of this parameter was also predicted in other studies as well [150, 151]. Notably, the inertia parameter will be given by the real part of the susceptibility because no complex number is involved in the expression of the inertia parameter. As it is shown that the magnetic inertia is a higher order relativistic effect, it is expected to be significant only at the ultrashort timescale. At the same time, the magnitude of the magnetic inertia is much smaller than the Gilbert damping parameter because the former is a second-order relativistic correction ($\sim 1/c^4$) while the latter is a first-order relativistic correction ($\sim 1/c^2$). For a harmonic field, the inertial dynamics can be described by the Eq. (3.32), however, for a nonharmonic field, the inertial dynamics will be complicated because it will involve a second-order field-derivative torque (see

paper VII for details), in contrast to the first-order field-derivative torque that appeared in the derivation of Gilbert damping parameter.

At this point, it is worth to make a comparison table of the Gilbert damping parameter and the inertia parameter as we have derived them on the same footing i.e., from a relativistic formulation.

3.7.1 Comparison of Gilbert damping parameter and inertia parameter

Our derivations reveal the respective expressions of the Gilbert damping and inertia parameter for a harmonic field as [175],

$$A_{ij} = -\zeta \sum_{n,k} \left[\frac{\langle r_i p_k + p_k r_i \rangle - \langle r_n p_n + p_n r_n \rangle \delta_{ik}}{i\hbar} \right] \times \mathtt{Im}\big(\chi_m^{-1}\big)_{kj}, \tag{3.34}$$

$$\mathscr{I}_{ij} = \frac{\zeta \hbar}{2mc^2} \Big[\mathbb{1} + \mathtt{Re}\big(\chi_m^{-1}\big)_{ij} \Big], \tag{3.35}$$

with $\zeta \equiv \frac{\mu_0 \gamma \hbar}{4mc^2}$. The real and imaginary parts are related to each other by the Kramers-Kronig relation, which means that both the parameters are interrelated. However, the Gilbert damping parameter is $\sim c^2$ times larger than the inertia parameter as revealed in our derivation. Their strengths can nonetheless be comparable depending on the real and imaginary parts of the susceptibility. In fact, if the real part of the susceptibility is much higher than the imaginary part, the existence of 1 in the inertia parameter will play a major role. A comparison table is made between Gilbert and inertia parameters in the following Table 3.1.

Subjects	Gilbert damping (A)	magnetic inertia ($\mathscr{I}$)
Contributions	✓ magnetic ✓ electronic	✓ magnetic
Susceptibility	✓ imaginary part	✓ real part
Timescale	✓ fast (ns, ps)	✓ ultrafast (ps, fs)
Dimension	✓ dimensionless	✓ time
Relation	$\mathscr{I} \propto -A\bar{\tau}$ with $\bar{\tau} = \frac{\hbar}{mc^2}$	

Table 3.1. *A comparison table of Gilbert damping parameter vs. inertia parameter.*

As discussed in **paper VII**, while deriving the inertial dynamics, a Gilbert damping parameter is also found of order $1/c^4$. However, the dominant contribution to the Gilbert damping will be of order $1/c^2$. In this connection, it

is worth mentioning that the Gilbert damping parameter and inertia parameter can be seen as a series of higher orders of $1/c^2$. This will be visible if one derives the higher order terms in the FW transformation. The corresponding terms will have higher order time derivatives of electric and magnetic field which will then follow our derivations.

We note that the Gilbert and inertia parameters that we derive are intrinsic and applicable to a pure and isolated system. There are several other extrinsic factors that can contribute to these parameters as well. For example, the environmental effects, radiation, interactions of spins with other quasiparticles (phonons), interfacial effects, spin pumping in magnetic multi-layers can modify these parameters [161, 176–179].

3.8 Spin currents

So far, in the discussions, the spin currents have not been considered. To explain many dynamical phenomena (e.g., domain wall motion, skyrmion dynamics etc.) in magnetic systems, one has to incorporate the spin current terms [52]. These spin-polarized currents can induce a local torque on the magnetization that has been proposed by Slonczewski [43] and Berger [44]. This effect is called spin-transfer torque (STT) which only requires the flow of an electrical current. This is realized in a multilayer system where the two ferromagnetic layers are separated by a nonmagnetic layer. The possible effect of STT is pictorially described in Fig. 3.6. To account for the current-induced

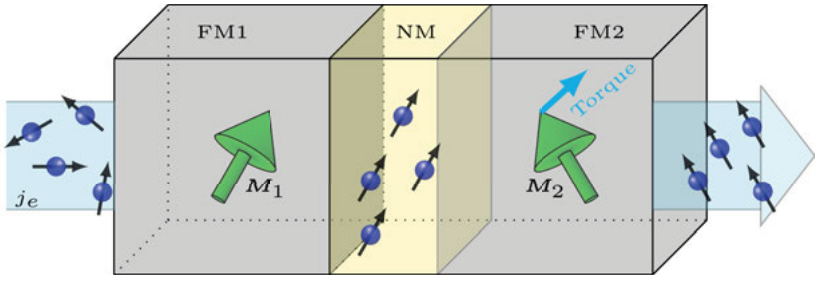

Figure 3.6. Schematic diagram for STT. First, an electrical current j_e enters into a ferromagnetic layer, the electrons get spin-polarized when they leave the first layer. This generated spin-polarized current then travels to the next ferromagnetic layer causing a torque to the pre-existing magnetization.

STT, the LLG equation has to be extended and these torques has been written as [44, 45, 136]

$$\begin{aligned}\frac{\partial \boldsymbol{M}}{\partial t} &= -\gamma \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} + \alpha \boldsymbol{M} \times \frac{\partial \boldsymbol{M}}{\partial t} \\ &\quad - b_{\text{J}} \boldsymbol{M} \times [\boldsymbol{M} \times (\boldsymbol{j}_{\text{e}} \cdot \boldsymbol{\nabla}) \boldsymbol{M}] - c_{\text{J}} \boldsymbol{M} \times (\boldsymbol{j}_{\text{e}} \cdot \boldsymbol{\nabla}) \boldsymbol{M}, \end{aligned} \tag{3.36}$$

where the last two terms explain the STT. b_{J} is termed as adiabatic STT parameter, while c_{J} defines the nonadiabatic STT. These torques account for the action of a spin-polarized current on the magnetization gradient.

In the theory of magnetization dynamics, these torques derive from the nonequilibrium spin density (nonequilibrium conduction electrons) which exerts a torque on the pre-existing local magnetization [136]. These current induced torques can also be derived by using a symmetry considerations [180]. However, in our formalism, the corresponding continuity equation of Eq. (3.14) provides deep insight into the torques. Following the discussion of Sec. 3.4.1, one can write down the nonrelativistic magnetization dynamics as

$$\frac{\partial \boldsymbol{M}}{\partial t}\Big|_{\mathrm{NR}} + \frac{1}{e}\left[\boldsymbol{M}\left(\boldsymbol{\nabla}\cdot\boldsymbol{j}_{\mathrm{e}}^{\mathrm{NR}}\right) + \left(\boldsymbol{j}_{\mathrm{e}}^{\mathrm{NR}}\cdot\boldsymbol{\nabla}\right)\boldsymbol{M}\right] = -\gamma_0 \boldsymbol{M}\times\boldsymbol{H}_{\mathrm{eff}}\,. \tag{3.37}$$

The nonrelativistic dynamics precisely shows the precession of magnetization and the effect of nonrelativistic current on the dynamics. Now, as there are no source or sink available in the system, the divergence of the current density must go to zero, implying the second term in the left-hand side of Eq. (3.37) vanishes. The other current term in the left-hand side of Eq. (3.37) takes into account the inhomogeneous magnetization and the corresponding torque can be written as the current induced adiabatic torque that takes the form $\sim \left(\boldsymbol{j}_{\mathrm{e}}^{\mathrm{NR}}\cdot\boldsymbol{\nabla}\right)\boldsymbol{M}$ [45, 136, 181, 182]. Therefore, the adiabatic STT can be recovered within the nonrelativistic magnetization dynamics.

However, the relativistic magnetization dynamics shows completely different dynamical behavior. We have extensively discussed that the Gilbert damping originates from the relativistic spin-orbit coupling. Alongside, there is a relativistic correction to the precession term that is coming from the intrinsic spin-orbit coupling as well. The relativistic magnetization dynamics can be written as

$$\frac{\partial \boldsymbol{M}}{\partial t}\Big|_{\mathrm{R}} + \frac{1}{e}\left[\boldsymbol{M}\left(\boldsymbol{\nabla}\cdot\boldsymbol{j}_{\mathrm{e}}^{\mathrm{R}}\right) + \left(\boldsymbol{j}_{\mathrm{e}}^{\mathrm{R}}\cdot\boldsymbol{\nabla}\right)\boldsymbol{M}\right] = -\gamma_0' \boldsymbol{M}\times\boldsymbol{H}_{\mathrm{eff}} + \boldsymbol{M}\times\left(A\cdot\frac{\partial \boldsymbol{M}}{\partial t}\right), \tag{3.38}$$

where γ_0' parameter corresponds to the relativistic precession, A_{ij} is the Gilbert damping tensor and $\boldsymbol{j}_{\mathrm{e}}^{\mathrm{R}}$ is the relativistic current density that has to be derived from the relativistic Hamiltonian. The latter contains the effect of spin-orbit coupling plus the additional contributions from the relativistic corrections to the exchange field (see Sec. 2.5.5) as $\boldsymbol{j}_{\mathrm{e}}^{\mathrm{R}} = \boldsymbol{j}_{\mathrm{soc}} + \boldsymbol{j}_{\mathrm{other}}$. The current density due to the spin-orbit coupling can be written as $\boldsymbol{j}_{\mathrm{soc}} = \kappa \boldsymbol{M}\times\boldsymbol{E}_{\mathrm{tot}}$ which depends on the magnetization; the parameter $\kappa = -e/(2m^2c^2g\mu_{\mathrm{B}})$. Using this relativistic current density one can see that the corresponding conductivity is proportional to the magnetization which is at the heart of the anomalous Hall effect. Moreover, if the electric field is directed in a certain direction, the current points along the plane to that is perpendicular to the direction of the

field and magnetization. Therefore, the effect of current density due to spin-orbit coupling will be more significant in the interfaces where the spin-orbit coupling is strong (see Fig. 3.6). Nonetheless, the effect can also be equally visible for a single heavy metal where the spin-orbit coupling is strong and if it has magnetization inhomogeneity. As discussed in **paper VIII**, accounting that the curl of the magnetization is zero (magnetization profile is irrotational), the relativistic magnetization dynamics in Eq. (3.38) results in

$$\frac{\partial \boldsymbol{M}}{\partial t}\Big|_{\mathrm{R}} + \frac{\kappa}{e}\boldsymbol{M} \times [(\boldsymbol{E}_{\mathrm{tot}} \cdot \boldsymbol{\nabla})\boldsymbol{M}] = -\gamma_0' \boldsymbol{M} \times \boldsymbol{H}_{\mathrm{eff}} + \boldsymbol{M} \times \left(A \cdot \frac{\partial \boldsymbol{M}}{\partial t}\right). \quad (3.39)$$

We have not incorporated the effect of the current density due to other relativistic effects than spin-orbit coupling. We know that, the electric field and the current are proportional to each other through the conductivity tensor, therefore, the torque (second term) in the left side of Eq. (3.39) can be shown to adopt the form $\sim \boldsymbol{M} \times [(\boldsymbol{j}_{\mathrm{e}} \cdot \boldsymbol{\nabla})\boldsymbol{M}]$. This exactly resembles to the torque of nonadiabatic STT that has been discussed earlier. Here, the nonadiabaticity parameter comprises the strength of spin-orbit coupling and the conductivity tensor. Now, due to spin-orbit coupling, the spin-flip processes are possible, those spin-flip relaxation times can be accounted through the calculation of conductivity tensor. Hence, we derive the current-induced STT within the spin continuity equation, while starting from the DKS Hamiltonian.

To summarize, the LLG and extended LLG terms have been derived from the fundamental DKS Hamiltonian. Within the magnetization dynamics, the precession of magnetization comes from the Zeeman-like couplings between spins and effective field. The intrinsic spin-orbit coupling also contributes to the relativistic precession when a spherically symmetric potential is assumed. On the other hand, the Gilbert damping derives from the extrinsic spin-orbit coupling. The corresponding Gilbert damping parameter has been shown to have two-fold origin: electronic and magnetic which can be calculated from first principles. More importantly, the damping parameter is a *tensorial* quantity while it has been known previously to be scalar. The damping tensor has also been shown to contain an isotropic Heisenberg-like damping, anisotropic Ising-like damping and a chiral DM-like damping. Furthermore, for a system driven by nonharmonic fields, an additional torque is generated - field derivative torque. This torque can become important for pulsed shaped external fields which will exert an instantaneous but huge torque on the magnetization that might lead to the switching of magnetization. On the same path, we have derived the optical spin-orbit torque that accounts for the manipulation of spins using the angular momentum of light. We have also derived the magnetic inertial dynamics within an extended LLG equation and the origin of magnetic inertia relates to the higher-order spin-orbit coupling. Finally, deriving the spin-currents in the spin continuity equation, we have derived the origin of adiabatic and nonadiabatic spin transfer torques within extended LLG equation.

4. Relativistic Effects in Ultrafast Magnetism

4.1 Introduction

The study of magnetization dynamics is an exceedingly vast research area with the involvement of profound physics as discussed in the previous chapter. However, realizing these dynamics in technological devices have been proven to be challenging. Importantly, the unanswered question remains: how fast can the magnetization be manipulated? For example, it has been shown that the magnetization reversal by an applied magnetic field pulse occurs in a nondeterministic way if the pulse duration is less than 2 ps [9]. Therefore, the possibility of manipulating the spins in an ultrashort timescale appears unachievable with the application of a magnetic field pulse. The use of an optical field pulse (laser) is another option. The main advantage of the latter is that the contemporary laser pulses are available in the laboratory with a wide range of pulse durations, where the pulse width can be as low as up to a few attoseconds [10].

Recent advancements of new techniques using optical laser pulses in pump-probe experiments have made it possible to access the sub-picosecond dynamics. In a typical pump-probe set up, an ultrashort laser beam is split into two separate beams, among which, one acts as pump (higher intensity) and the other acts as probe (lower intensity). First, the pump pulse is directed towards the sample to excite and the probe consequently detects the modification (that has been caused by the pump) within the sample. The pump stimulates the optical excitations which are very short-lived, therefore, an ultrashort probe pulse is needed to measure the modification of the sample. The pulse durations of the pump and probe are usually kept in about a few tens of fs. Thus, the set up allows one to detect a temporal as well as a spatial resolution of the dynamics, given the fact that the probe pulse is delayed with respect to the pump pulse.

Ultrafast demagnetization was discovered more than two decades ago by E. Beaurepaire and co-workers in such a pump-probe magneto-optical Kerr effect (MOKE) experiment on a nickel (Ni) thin-film of $\sim$ 20 nm. In particular, a sudden loss of magnetization was observed within an ultrashort timescale less than a ps after the laser pulse hit the sample. The magnetization response of Ni was measured in pump-probe ultrafast spectroscopy as a function of pump-probe time delay (Δt) as shown in Fig. 4.1 [11].

The above-described effects have been roughly accounted for using the phenomenological extension of the so-called two temperature model (electrons

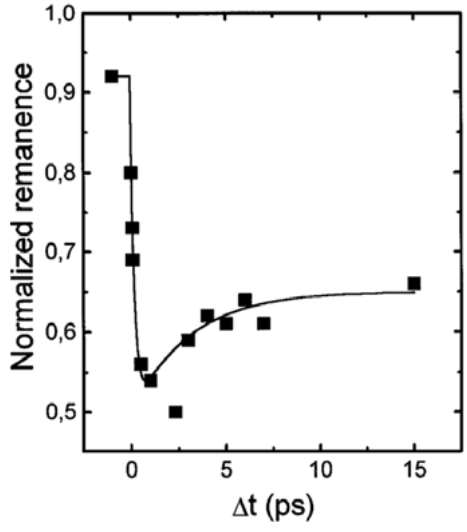

Figure 4.1. Ultrafast demagnetization is shown. The black dots represent the experimental observations and the smooth line shows the follow up of the experimental data points. Taken with permission from Beaurepaire *et al.*, Phys. Rev. Lett. **76**, 4250 (1996). Copyright © 1996, The American Physical Society.

and lattice) including the spins. Thus the system has three different reservoirs and they evolve according to three coupled differential equations (taking into account their specific heat and the transfer of energy from one to another). The solution of the equations resembles the experimentally observed electron temperature evolution within the ultrashort timescale [11].

The experiment was followed by several others leading to a similar trend of ultrafast demagnetization [15, 17, 23, 183–185]. Among the metals, mostly the transition metals (Fe, Co, Ni) have been extensively studied [18, 186–188]. Other systems like ferrimagnets, antiferromagnets and ferromagnetic insulators have shown much more complicated dynamics [189–192]. Many theoretical investigations have been made till the date, however, the underlying mechanisms are not properly understood in the ultrashort timescales as yet. In the following, we provide a few of the important proposed mechanisms during ultrafast demagnetization processes.

4.2 Mechanisms of ultrafast demagnetization

4.2.1 Direct spin-photon coupling

The prime aspect during ultrafast demagnetization is the transfer of angular momentum within different available degrees of freedom within a material. It has been predicted that the direct coupling of laser photons with the spins could lead to the transfer of angular momentum from light to the magnetic

moments and thus to the demagnetization [16, 193]. However, the estimations on Ni reveal that the direct photo-quenching can only change the magnetic moments by 10^{-4} μ_B per Ni atom [194]. Therefore, this negligible amount of photo-quenching cannot be the reason behind the experimentally observed ultrafast demagnetization [195].

4.2.2 Electron-electron interaction

The Coulomb interaction among the electrons might also be another mechanism of ultrafast demagnetization as the electron-electron interaction is a very fast process [23]. Krauß *et al.* investigated the effect of electron-electron interaction within Elliott-Yafet (EY) like mechanisms that was originally proposed for spin-lattice relaxation [196]. The ultrafast demagnetization is possible in EY theory because the spin-orbit interaction does not conserve spin momentum and any momentum dependent scattering mechanism transfers the momentum from the spins to another reservoir. It has been found that the electron-electron EY-type scattering mechanisms are very important in ultrafast demagnetization [197]. Model calculations have also been performed including the electron-electron and electron-phonon scatterings within the spin-resolved Boltzmann equation. The study suggests that the electron-electron scattering is required for demagnetization, while the electron-phonon scattering is responsible for remagnetization [198].

4.2.3 Electron-phonon interaction

Spin-flip due to EY electron-phonon scattering mechanism has been discussed as the key issue for demagnetization [199, 200]. The main idea is the transfer of angular momentum from spins to the lattice. Carva and co-workers have developed an *ab initio* theory to calculate the effect of spin-flip due to electron-phonon scattering where they have concluded that the effect is too small to explain the amount of observed demagnetization [24, 25]. However, the question of angular momentum has not been addressed properly as the calculations with electron-phonon scattering considers the linearly polarized phonon states (they do not carry angular momentum). This consideration assumes the lattice to be a perfect sink of angular momentum. In the following, we provide explanations of the debate over the angular momentum conservation.

Angular momentum conservation

After the fs laser pulse hits the ferromagnetic sample, the microscopic mechanisms of the angular momentum transfer from spins to other degrees of freedom are important and have often been debated. One way to recognize the conservation of angular momentum is that the magnetic sample is not free in the space, but it is often mounted on a sample holder which then connects to a

table on the floor. Thus, the total system is isotropic as a whole and the angular momentum is conserved, but not the angular momentum of the sample (part of the total system) [201]. One also has to note that, during demagnetization there are emissions of electromagnetic waves which are linearly polarized and they carry the energy of THz regime [202, 203]. These THz waves are able to interact with the phonon and they can be re-absorbed in the sample. Therefore the strict conservation of angular momentum during fs demagnetization is arguable [201].

4.2.4 Superdiffusive spin transport

A conceptually different approach than the usual sub-atomic interactions has been employed by Battiato *et al.*, that is called superdiffusion theory. Within the superdiffusive spin transport, the majority spins travel faster into the substrate as compared to the minority spins leading to the demagnetization in the sample [26, 204]. Experimental measurements have been performed in magnetic multi-layered system of Au/Ni, where the Au layer was hit by the fs laser pulse generating a non-equilibrium hot electrons. Those hot electrons travel through the Au/Ni interface towards the Ni region and efficiently explain the demagnetization in ferromagnet [187]. Similar experiments reveal the importance of ultrafast spin transfer current that directly indicates that the superdiffusive spin-transport as the key mechanism for ultrafast demagnetization [205–207]. The superdiffusion theory, however, depends critically on the composition of the sample [208] and it has not been the dominant effect in single ferromagnetic thin films on insulator substrates [209].

4.2.5 Spin-orbit coupling and excitation

Within this framework, a full *ab initio* time-dependent DFT formalism has been used to propose an explanation of the observed ultrafast demagnetization [210]. The ultrashort laser-induced demagnetization has been proposed as a two step process: (1) the laser excites a fraction of electrons, and (2) the spin-orbit coupling of the remaining localized electrons helps in spin-flip transitions. Thus, the spin-orbit coupling is shown to take a fundamental role. The latter has been included in the Pauli Hamiltonian taking into account the vector potential from the laser pulse to explain the amount of observed demagnetization [211]. However, in the calculations a very short pulse has been used with a duration of $\sim$ 2.2 fs and a very high laser intensity (2-3 orders of magnitude higher than experiments) has been employed. Moreover, the demagnetization time has been found to be very short about $\sim$ 20 fs in bulk Ni in comparison to the experimentally observed time [210]. In a recent study, a comparison is made between bulk and thin films demagnetization using *ab initio* TDDFT [212].

4.2.6 Relativistic spin-flip processes

To understand the effect of relativistic spin-flip processes in ultrafast demagnetization, a single 50 fs duration of laser pulse has been used in a pump-probe ultrafast spectroscopy experiment [22]. From the experimental observations, it has been proposed that the laser field couples to magnetism in the ferromagnetic Ni film efficiently during its propagation [22]. The mechanism is described in the following: the fs laser pulse interacts coherently with the charges and spins of the magnetic Ni thin film, inducing a demagnetization. This coherent step is followed by the incoherent step due to thermalization of charges and spins and their interactions with the lattice. A linear decrease of complex Kerr rotation and ellipticity has been found in Ni, which was suggested to have its origin in relativistic quantum electrodynamics involving the spin-orbit coupling, beyond the ionic potential [22].

To verify the above-described processes, we need to use the spin density functional theory with the charge and spin correlations, yet including the effective exchange field terms with lowest order relativistic corrections up to $1/c^2$ that includes the proper description of spin-orbit coupling (which was missing in Ref. [22]). In other words, we have to use simply the Hamiltonian in Eq. (2.74) that has been derived throughout the Chapter 2. This will be discussed in more details in Sec. 4.3.

In connection to the previous discussion, it is worth mentioning that Vonesch and Bigot used a similar Hamiltonian, however, they considered a strong static homogeneous magnetic field, $\boldsymbol{B}_{\mathrm{M}}$, which was given by the vector potential $\boldsymbol{A}_{\mathrm{M}} = 1/2\, \boldsymbol{B}_{\mathrm{M}} \times \boldsymbol{R}$, with $\boldsymbol{R}$ as the position of electrons. They also separately used an electric field, coming from laser $\boldsymbol{E}_{\mathrm{L}}$ and the corresponding vector potential $\boldsymbol{A}_{\mathrm{L}}$ [213]. Once again, we emphasize that the consideration of homogeneous magnetic field to account for the exchange fields cannot be included as a vector potential [122]. Therefore, the minimal coupling can *only* be used for the vector potential from laser light as $(\boldsymbol{p} - e\boldsymbol{A}_{\mathrm{L}})$. In their work, they found a significant contribution in the transition matrix elements of the interaction $\boldsymbol{A}_{\mathrm{L}} \cdot \boldsymbol{A}_{\mathrm{M}}$, while this interaction does not appear in our formulation derived from the DKS theory (see Eq. (2.74) for details).

To summarize, the fundamental processes during the ultrafast demagnetization in shorter timescale are still debatable and do not have any accurate description [21, 184, 214, 215]. It is mainly because of the fact that at the ultrashort timescale, the electrons immediately become non-thermal. Thereafter, the electrons gets heavily excited in or out of the Fermi sea. In this situation, the description of electron or spin temperatures is not strictly valid as the temperatures are usually defined for thermalized reservoirs. Moreover, at the ultrashort timescale, the classical descriptions become invalid and the quantum effects become prominent. Along the same line, the transfer of angular momentum has been a long standing question during the process of ultrafast demagnetization.

Despite of the debate on underlying mechanisms, one can conclude that there is not a single mechanism that could explain the huge magnetization loss. It is probably the combination of several mechanisms that could explain the experimental observation. Nonetheless, in the following, we describe the influence of relativistic spin-flip effects in the demagnetization processes.

4.3 Relativistic magneto-optics

MOKE is itself a relativistic effect due to its direct relation with the spin-orbit coupling [216, 217]. To be specific, when MOKE is calculated in *ab initio* electronic structure theory, the wave functions correspond to the relativistic Hamiltonian that includes the spin-orbit coupling. MOKE spectra have been calculated for uranium [86, 218] and transition metal compounds [219–223] and these have been found to be fully satisfactory. The Kerr spectra can be directly obtained through the diagonal and off-diagonal parts of the optical conductivity tensor ($\sigma_{\alpha\beta}$) by the following relation [224]

$$\Phi_{\mathrm{K}}(\omega) = \theta_{\mathrm{K}}(\omega) + i\varepsilon_{\mathrm{K}}(\omega) = \frac{-\sigma_{xy}(\omega)}{\sigma_{xx}(\omega)\sqrt{1+\frac{4\pi i}{\omega}\sigma_{xx}(\omega)}}, \tag{4.1}$$

where θ_{K} and ε_{K} are the Kerr rotation angle and ellipticity respectively. Note that the above expression is particularly valid for polar MOKE where the material magnetization is perpendicular to the reflection surface and parallel to the plane of incidence. These optical conductivity tensors are calculated within Kubo linear response formalism that can be expressed as a current-current or momentum-momentum response function [224, 225]. To derive the response function, the first order interaction Hamiltonian is taken into account (see Appendix B of **paper I**). We know that the effect of an intensive laser pulse is considered through the vector potential, $\boldsymbol{A}(\boldsymbol{r},t)$. Now, within an uniform magnetic field (see Eq. 2.76) and Coulomb gauge one finds $\boldsymbol{r}\cdot\boldsymbol{A} = 0$. A simple formulation of its time derivative gives:

$$\begin{aligned} \frac{d}{dt}(\boldsymbol{r}\cdot\boldsymbol{A}) = 0 \Rightarrow \frac{d\boldsymbol{r}}{dt}\cdot\boldsymbol{A} &= \boldsymbol{r}\cdot\boldsymbol{E} \\ \Rightarrow \frac{e}{m}\boldsymbol{\Pi}\cdot\boldsymbol{A} &= e\,\boldsymbol{r}\cdot\boldsymbol{E}\,. \end{aligned} \tag{4.2}$$

In this case, we only retain the transverse part of the electric field, $\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t}$ for the choice of Coulomb gauge. The left-hand side in Eq. (4.2) defines the interaction Hamiltonian, which is the product of momentum operator $\boldsymbol{\Pi}$ and the vector potential [226]. Thus, we show that the first-order interaction Hamiltonian can equally be written as the product of positions of electrons ($\boldsymbol{r}_i$) and the electric field acting upon them. This formulation is true for the nonrelativistic, semirelativistic or fully relativistic case [123]. The conductivity tensor is

thereafter derived from the commutators of current density and the first-order interaction Hamiltonian. The conductivity tensor adopts the expression of a momentum-momentum response function (see Appendix B of **paper I**) as

$$\sigma_{\alpha\beta}(\omega) = -\frac{ie^2}{m^2\hbar V}\sum_{nn'\boldsymbol{k}}\left[\frac{f(\varepsilon_{n\boldsymbol{k}})-f(\varepsilon_{n'\boldsymbol{k}})}{\omega_{nn'}(\boldsymbol{k})}\frac{\Pi^{\alpha}_{n'n}(\boldsymbol{k})\Pi^{\beta}_{nn'}(\boldsymbol{k})}{\omega-\omega_{nn'}(\boldsymbol{k})+i/\tau}\right], \tag{4.3}$$

with the optical energy $\hbar\omega$, Fermi-Dirac distribution functions $f(\varepsilon_{n\boldsymbol{k}})$ and the difference between the band energies $\hbar\omega_{nn'}(\boldsymbol{k}) = \varepsilon_{n\boldsymbol{k}} - \varepsilon_{n'\boldsymbol{k}}$. The involved momentum operator can be taken as $\boldsymbol{\Pi} = \boldsymbol{p}$ for the nonrelativistic case, $\boldsymbol{\Pi} = mc\boldsymbol{\alpha}$ for the fully relativistic case. In our case, we derive the corresponding momentum operator from the Hamiltonian in Eq. (2.74) that involves the nonrelativistic momentum operator plus the additional contributions coming from the relativistic correction and can be as expressed by

$$\begin{aligned}\boldsymbol{\Pi} = \boldsymbol{p} + \frac{1}{4mc^2}\Big[&\frac{2p^2}{m}\boldsymbol{p} + i\boldsymbol{\sigma}\times(\boldsymbol{p}V) + \mu_B\Big\{\boldsymbol{\sigma}\cdot(\boldsymbol{p}\boldsymbol{B}^{\mathrm{xc}}) + (\boldsymbol{p}\cdot\boldsymbol{B}^{\mathrm{xc}})\boldsymbol{\sigma} \\ &+ 2\boldsymbol{B}^{\mathrm{xc}}(\boldsymbol{\sigma}\cdot\boldsymbol{p}) + 2\boldsymbol{\sigma}(\boldsymbol{B}^{\mathrm{xc}}\cdot\boldsymbol{p}) + i(\boldsymbol{p}\times\boldsymbol{B}^{\mathrm{xc}})\Big\}\Big].\end{aligned} \tag{4.4}$$

Within relativistic magneto-optics, one calculates the diagonal and off-diagonal conductivity tensors and the complex MOKE spectra of Kerr rotation and Kerr ellipticity [224].

4.3.1 Relativistic spin-flip effects in ultrafast demagnetization

In this section, we discuss the contribution of relativistic spin-flip effects during ultrafast demagnetization. We have calculated *ab initio* the real and imaginary part of the conductivity tensors (diagonal and off-diagonal components) using the relativistic four component extension of the augmented spherical wave (ASW) method [227], within local spin density approximation in DFT. To do so, we have solved numerically the DKS Hamiltonian in order to find out the electron spinor wave function which includes the effect relativistic corrections e.g., spin-orbit interaction. The momentum matrix elements have been calculated using the relativistic wave functions. The momentum operator in Eq. (4.4) includes the spin-orbit coupling and thus accounts for the spin-flip effects because the spin-orbit interaction and other relativistic effects do not conserve the spin angular momentum. To quantitatively account for the spin-flip effects, we calculate the conductivity and complex MOKE spectra of Ni and Fe in the following taking into consideration the relativistic and nonrelativistic momentum operator. The contributions due to the relativistic spin-flip effects can be captured by taking their difference.

As described in **paper I**, the calculations of conductivity spectra and complex MOKE overlap with each other while taking the nonrelativistic momentum matrix elements in contrast to the relativistic momentum matrix elements

for Ni. In particular, we have also calculated the difference between the contributions of relativistic momentum and nonrelativistic momentum and the difference has been found to be of the order of 10^{-3} for Ni, showing the effect of relativistic spin-flip contributions are of $\sim 0.1\%$ [123]. Henceforth, we have concluded that the relativistic photon-spin interactions are present during an ultrafast pump-probe measurement, however, they do not provide a significant channel of laser-induced magnetization loss.

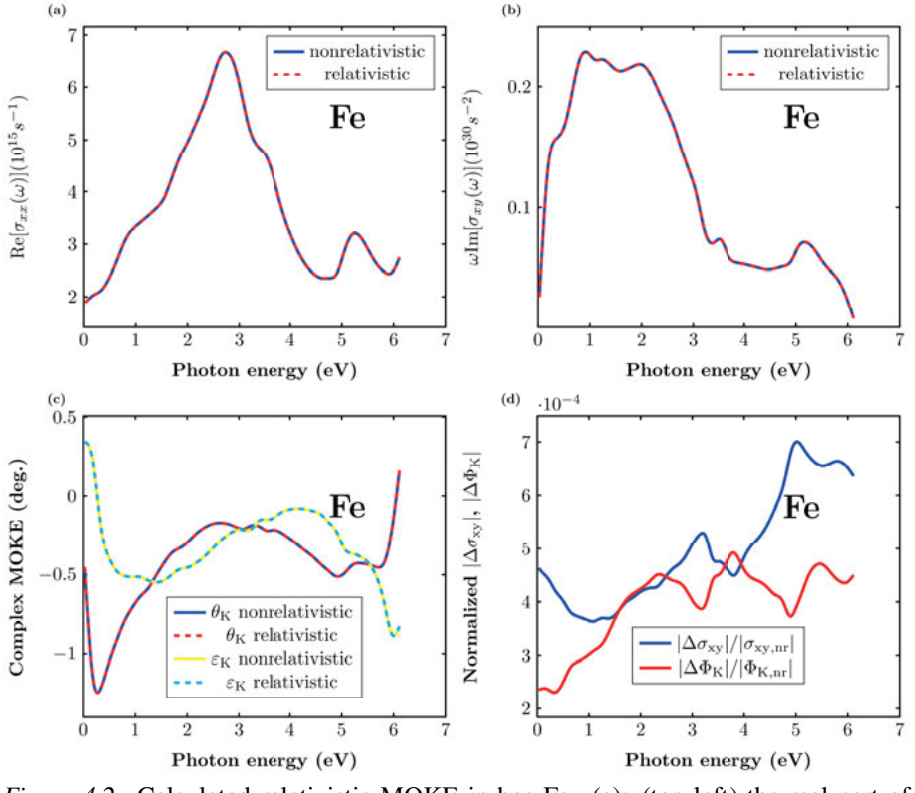

Figure 4.2. Calculated relativistic MOKE in bcc Fe. (a): (top left) the real part of diagonal conductivity matrix elements are shown for relativistic and nonrelativistic momentum. (b): (top right) The imaginary part of off-diagonal conductivity matrix elements are shown for relativistic and nonrelativistic momentum. (c): (bottom left) The Kerr angle and Kerr ellipticity have been calculated for relativistic and nonrelativistic momentum operator. (d): (bottom right) The absolute value of the difference in σ_{xy} (Φ_K) normalized to the quantities $|\sigma_{\mathrm{xy,nr}}|$ ($|\Phi_{\mathrm{K,nr}}|$) calculated by the nonrelativistic momentum operator.

It is important to point out that, we have used the linear response theory here, while neglecting the nonlinear effects. The reason is that, when the pump laser intensity is of the order of 10^{11} W/cm^2 (corresponding fluence is of the order 1 mJ/cm^2 with a pulse duration $\sim$ 50 fs), the number of electrons that are removed from below the Fermi surface is of the order of 0.01 [15]. We do

not expect the role of additional many-body effects for such a small number of electrons in the exited states, thus the normal linear response theory within DFT will be reasonably enough to capture the experimental evidences. However, we note that for a higher intensity ($\sim 10^{14} - 10^{15}$ W/cm^2), alternatively higher fluence (intensity per second), one has to consider nonlinear effects, beyond linear response theory [210].

In connection to the magneto-optics in Ni, we furthermore show the calculated relativistic magneto-optical effects for bcc Fe in Fig. 4.2.

We find the similar conclusion for Fe as in the case of fcc Ni. It is seen from the plots in Fig. 4.2 that the relativistic and nonrelativistic contributions of momentum matrix elements to the MOKE spectra almost overlap with each other. Additionally, we find that their normalized differences are very small and the order is $5 \times 10^{-4} (\times 10^{15}\, s^{-1})$ for both the conductivity response and MOKE spectra. Therefore, we conclude that the relativistic photon-spin interactions are present (and also the relativistic spin-flips) during the magneto-optical phenomena, however, their contributions are negligibly small to explain the large loss of magnetization within ultrashort timescales.

4.4 Coherent ultrafast magnetism

The ultrafast demagnetization is experimentally determined by using a pump-probe spectroscopy technique for the MOKE spectra. In this experiment, the magnetic sample is being exposed to the pump pulse and consequently the magnetization is probed by measuring the Kerr rotation angle of the outgoing probe pulse. The intensity of the pump pulse is much higher than the probe. The differential Kerr angle in Ni is plotted against the pump-probe time-delay, τ, in the left panel of Fig. 4.3, nicely showing the behavior of ultrafast demagnetization within ~ 250 fs range. The differential Kerr angle is defined using the following relation:

$$\frac{\Delta\theta(\tau)}{\theta} = \frac{\theta^{\mathrm{pu}}_{\mathrm{pr}}(\tau) - \theta^{\mathrm{pr}}_{\mathrm{none}}}{\theta^{\mathrm{pr}}_{\mathrm{none}}}. \tag{4.5}$$

We mention that $\theta^{\mathrm{pu}}_{\mathrm{pr}}(\tau)$ is measured as the Kerr rotation angle of the reflected probe beam, once the pump has hit the sample and the angle is expressed as a function of pump-probe time delay. $\theta^{\mathrm{pr}}_{\mathrm{none}}$ obviously determines *only* the reflected probe Kerr rotation without the pump beam and thus it is not a function of time-delay.

As shown in Fig. 4.3(a), the MOKE measurements in Ref. [22] have been carried out with two different configuration set ups of pump and probe: (i) the pump and probe polarizations are parallel to each other i.e., $\theta_{\mathrm{PP}} = 0^{\mathrm{o}}$ that has been marked with red solid dots and (ii) the pump and probe polarizations are perpendicular to each other i.e., $\theta_{\mathrm{PP}} = 90^{\mathrm{o}}$ that has been marked with blue hol-

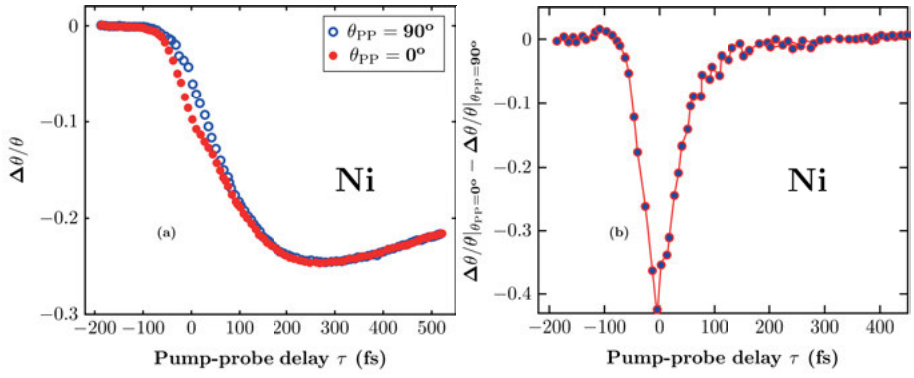

Figure 4.3. Coherent ultrafast magnetism in Ni. (a) (left panel) The normalized differential Kerr angle is plotted against the pump-probe delay in fs for two set ups. The red data is for when the pump polarization is parallel to the probe polarization and the blue corresponds to pump polarization being perpendicular to the probe. (b) (right panel) The difference between the two plots in the left panel is presented, which clearly shows the coherent pump-probe contribution only. [Reproduced with permission from Bigot *et al.*, Nat. Phys. **5**, 515 (2009). Copyright © 2009, Macmillan Publishers Limited. All rights reserved.]

low dots. One can observe that the two configurations produce the same measurements (overlap with each other) in longer timescales showing the ultrafast demagnetization and remagnetization, however, a non-negligible difference is observed within first 100 fs. This is more clear from the plot in Fig. 4.3(b), showing the difference between two Kerr rotations, $\theta_{\rm pr}^{\rm pu}\big|_{\rm PP=0^o} - \theta_{\rm pr}^{\rm pu}\big|_{\rm PP=90^o}$ as a function of time-delay, where a dip within 100 fs is observed. Now it is obvious that the latter plot completely depends on the polarization of pump and probe. This observation is known as the *coherent ultrafast magnetism* [22].

Already in their work, Bigot *et al.* [22] proposed the microscopic origin to be the relativistic quantum electrodynamics, beyond spin-orbit interaction involving ionic potential. To account for the relativistic Hamiltonians, efforts have been made to derive the FW transformation of the Dirac Hamiltonian and finding the higher-order relativistic correction terms, that can be useful to describe the experimental observation [90]. Under the extreme light-matter conditions, the modification of electron-electron interactions has been studied with the Dirac-Breit Hamiltonian and the corresponding FW transformation, where it has been shown that the electric field from the light can coherently modify the two electron interactions [228]. There have also been attempts to classically model the observed coherent ultrafast experiments by using the self-consistent mean-field approximations, which includes the nonlinear effects [229–233]. There, the consideration is to use the nonlinear Drude-Voigt model and include the effective magnetic field inside the material, together with up to the third order polarization created by the bound and free charges. The total mean-field is considered as the combination of the static molecular

mean field (that exists without the light perturbation) and the light-induced mean field of Weiss type. The latter includes the spin-orbit interaction of electric field from light with the material spins. However, the set of electric field and vector potential are considered within the material as ($\boldsymbol{E}_{\rm mat}$, $\boldsymbol{A}_{\rm mat}$) and from the laser as ($\boldsymbol{E}$,$\boldsymbol{A}$) and different possible combination of interactions e.g., $\boldsymbol{E}_{\rm mat}\times\boldsymbol{A}$, $\boldsymbol{E}\times\boldsymbol{A}_{\rm mat}$ etc. have been treated [233].

It should be noted that, our approach is different as we argue that the material's magnetic field is very large (order of 10^3 Tesla), which comes from the exchange field. It has been concluded that the material's magnetic field is about 510 Tesla in Ref. [233]. Therefore, it is clear that $\boldsymbol{A}_{\rm mat}$ is nothing but a measure of the exchange field. We have already pointed out earlier (in Sec. 2.4) that the consideration of vector potential within the material ($\boldsymbol{A}_{\rm mat}$) as a minimal coupling, is not strictly valid for exchange field and the corresponding magnetic magnetic field can not be written as $\boldsymbol{B}_{\rm mat}=\boldsymbol{\nabla}\times\boldsymbol{A}_{\rm mat}$. Instead, we have found in Eq. (2.74), a quadratic interaction term of fields from light, which is recognized as the gauge invariant part of the spin-orbit coupling. The corresponding Hamiltonian which we call spin-photon coupling Hamiltonian, is found to have the expression [163, 164]

$$\mathscr{H}_{\rm spin-photon}=\frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\boldsymbol{E}\times\boldsymbol{A})\,, \tag{4.6}$$

where the electric field is defined as $\boldsymbol{E}=-\partial\boldsymbol{A}/\partial t$. Using the plane wave solutions to the Maxwell's equations, the corresponding Hamiltonian can be written as

$$\mathscr{H}_{\rm spin-photon}=\frac{e^2\hbar}{8m^2c^2\omega}\boldsymbol{\sigma}\cdot\mathtt{Re}\Big[-i\left(\boldsymbol{E}^0\times\boldsymbol{E}^{0\star}\right)\Big]\,, \tag{4.7}$$

where '`Re`' defines the real part and general elliptically polarized light fields are assumed as $\boldsymbol{E}^0=E_0/\sqrt{2}(\boldsymbol{e}_x+e^{i\eta}\boldsymbol{e}_y)$ with the ellipticity parameter η. Now in a typical pump-probe experimental set up, the above-mentioned electric fields can be due to pump or probe or both. The energies of both the pump and probe have been accounted to have the same ω in Eq. (4.7), however, the same expression is not valid if the photon energies of pump and probe are different. A detailed calculation of the Hamiltonian with different frequencies of pump and probe has been presented in **paper VI**. In the following, we discuss the effect of such a Hamiltonian in explaining the coherent ultrafast magnetism that has been observed in the experiment.

4.4.1 Induced optomagnetic field

As discussed in earlier Sec. 3.6.5, the Hamiltonian in Eq. (4.7) can be understood as the coupling of light's angular momentum to the spin. Therefore, the Hamiltonian provides new insight to optically manipulate the spins. In

the same context, considering a Zeeman-like coupling as $-g\mu_{\mathrm{B}}\boldsymbol{\sigma}\cdot\boldsymbol{B}_{\mathrm{opt}}$, the induced optomagnetic can be expressed as

$$\boldsymbol{B}_{\mathrm{opt}} = \frac{e^2\hbar}{8m^2c^2\omega g\mu_{\mathrm{B}}}\mathrm{Re}\Big[i\left(\boldsymbol{E}^0\times\boldsymbol{E}^{0\star}\right)\Big]. \tag{4.8}$$

We should note here that the electric fields involved are purely from the laser light. To explain the origin of coherent ultrafast signals, we argue that during the first 100 fs, pump and probe beams interact with each other, given the fact that their pulse duration is of the order of 50 fs. Therefore, we consider all the possible interactions between pump and probe as described by the following terms: $\boldsymbol{E}^0_{\mathrm{pu}}\times\boldsymbol{E}^{0\star}_{\mathrm{pu}}$ (pump-pump), $\boldsymbol{E}^0_{\mathrm{pr}}\times\boldsymbol{E}^{0\star}_{\mathrm{pr}}$ (probe-probe), $\boldsymbol{E}^0_{\mathrm{pu}}\times\boldsymbol{E}^{0\star}_{\mathrm{pr}}$ (pump-probe) and $\boldsymbol{E}^0_{\mathrm{pr}}\times\boldsymbol{E}^{0\star}_{\mathrm{pu}}$ (pump-probe) [164]. It is worth to mention here that, the pump-pump and the probe-probe interactions will only contribute if they have nonzero ellipticity, for linearly polarized pump and probe, their contribution will become zero. However, the pump-probe interaction will always contribute even for linearly polarized pulses. Considering the linearly polarized pump and probe beams, traveling towards z-direction, if the polarizations of pump and probe are parallel to each other ($\theta_{\mathrm{PP}} = 0^{\mathrm{o}}$); the electric fields can be taken as, $\boldsymbol{E}^0_{\mathrm{pu}} = (E^0_{\mathrm{pu}}\boldsymbol{e}_x, 0)$ and $\boldsymbol{E}^0_{\mathrm{pr}} = (E^0_{\mathrm{pr}}\boldsymbol{e}_x, 0)$. In this case, the induced optomagnetic field will be zero as all the above-mentioned contributions will be zero. On the other hand, if the pump and probe polarizations are perpendicular to each other ($\theta_{\mathrm{PP}} = 90^{\mathrm{o}}$); the electric fields can be taken as, $\boldsymbol{E}^0_{\mathrm{pu}} = (E^0_{\mathrm{pu}}\boldsymbol{e}_x, 0)$ and $\boldsymbol{E}^0_{\mathrm{pr}} = (0, E^0_{\mathrm{pr}}\boldsymbol{e}_y)$. In this case, the induced optomagnetic field will not be zero as all the pump-probe interaction contributions will be non-zero. In particular, the magnitude of the optomagnetic field that we have derived has the expression [164]:

$$B_{\mathrm{opt}} = \frac{e\hbar}{4mc^2}\frac{|E^0_{\mathrm{pu}}||E^0_{\mathrm{pr}}|}{\hbar\omega}. \tag{4.9}$$

Therefore, the strength of the optomagnetic field depends on the amplitudes of the pump and probe beams. Considering the experimental values of pump and probe electric field amplitudes, the optomagnetic field has been found to be as large as few μT upto mT.

Along with the magnitude, we also mention that the induced optomagnetic field depends on the pump-probe delay; if the delay time is longer, the pump and probe do not interact with each other and the corresponding optomagnetic field will be zero. A full picture involving different interactions to the induced optomagnetic field, is summarized in the following Table 4.1 for linearly polarized pump and probe beams.

It is important to mention that, in pump-probe experiments, the pump can be elliptically polarized to capture the effect of light's helicity. This is employed for helicity dependent investigations e.g., AO-HDS. For example, we can mention the study of light-induced spin oscillations in $DyFeO_3$ by Kimel

Contributions to induced optomagnetic field					
Pump beam	Probe beam	$\boldsymbol{E}^0_{\rm pu}\times\boldsymbol{E}^{0\star}_{\rm pu}$	$\boldsymbol{E}^0_{\rm pr}\times\boldsymbol{E}^{0\star}_{\rm pr}$	$\boldsymbol{E}^0_{\rm pu}\times\boldsymbol{E}^{0\star}_{\rm pr}$	$\boldsymbol{E}^0_{\rm pr}\times\boldsymbol{E}^{0\star}_{\rm pu}$
$(E^0_{\rm pu}\boldsymbol{e}_x,0)$	$(E^0_{\rm pr}\boldsymbol{e}_x,0)$	0	0	0	0
$(0,E^0_{\rm pu}\boldsymbol{e}_y)$	$(0,E^0_{\rm pr}\boldsymbol{e}_y)$	0	0	0	0
$(E^0_{\rm pu}\boldsymbol{e}_x,0)$	$(0,E^0_{\rm pr}\boldsymbol{e}_y)$	0	0	$\neq 0$	$\neq 0$
$(0,E^0_{\rm pu}\boldsymbol{e}_y)$	$(E^0_{\rm pr}\boldsymbol{e}_x,0)$	0	0	$\neq 0$	$\neq 0$
$E^0_{\rm pu}(\boldsymbol{e}_x,\boldsymbol{e}_y)$	$E^0_{\rm pr}(\boldsymbol{e}_x,\boldsymbol{e}_y)$	0	0	$\neq 0$	$\neq 0$

Table 4.1. *All possible configurations of pump and probe for linearly polarized light.*

et al. [28], where circularly polarized pump pulses have been used and probed by linear polarized light. As discussed in the last paragraph, the impact of pump-pump interaction will become important in this case and will have larger impact on the induced optomagnetic field because of the fact that the intensity of the pump beam is about 10^3 times higher than that of the probe. Such contributions have been listed for possible configurations of elliptically polarized pump and linear polarized probe in Table 4.2. Considering the experimental values of pump electric field amplitude, the pump-pump interaction produces an optomagnetic field of the order of few mT [164].

For elliptically polarized light, η defines the ellipticity parameter which takes the value $\eta = 0$ for linear polarization, while $\eta = +\pi/2$ and $\eta = -\pi/2$ for right and left circular polarized light, respectively. Needless to say that, when both the pump and probe are elliptically or circularly polarized, all the above four interactions will become non-zero and thus they all will contribute to the induced optomagnetic field.

Contributions to induced optomagnetic field					
Pump beam	Probe beam	$\boldsymbol{E}^0_{\rm pu}\times\boldsymbol{E}^{0\star}_{\rm pu}$	$\boldsymbol{E}^0_{\rm pr}\times\boldsymbol{E}^{0\star}_{\rm pr}$	$\boldsymbol{E}^0_{\rm pu}\times\boldsymbol{E}^{0\star}_{\rm pr}$	$\boldsymbol{E}^0_{\rm pr}\times\boldsymbol{E}^{0\star}_{\rm pu}$
$E^0_{\rm pu}(\boldsymbol{e}_x,e^{i\eta}\boldsymbol{e}_y)$	$(E^0_{\rm pr}\boldsymbol{e}_x,0)$	$\neq 0$	0	$\neq 0$	$\neq 0$
$E^0_{\rm pu}(\boldsymbol{e}_x,e^{i\eta}\boldsymbol{e}_y)$	$(0,E^0_{\rm pr}\boldsymbol{e}_y)$	$\neq 0$	0	$\neq 0$	$\neq 0$

Table 4.2. *All possible configurations of elliptically polarized pump and linearly polarized probe light.*

Using the experimental parameters of the pump-probe intensity, wavelength, photon energies, pulse durations, we can estimate the induced optomagnetic field numerically. We have found that the optomagnetic fields are small and range from the order of few mT to μT depending on the pump and probe intensities that have been used in the experiments. Such small fields can not be neglected at all, in fact they can be very much important for thin films. For example, here we mention that a magnetic field of about 10-20 mT of magnetic field is able to switch the magnetization of a thin film [234].

To summarize, in the first part of this chapter, we have quantitatively studied the effect of relativistic spin-flip processes in laser-induced ultrafast demagnetization. We have carried out *ab initio* calculations of optical conductivity and MOKE spectra with relativistic and nonrelativistic momentum matrix elements. Due to the spin-orbit coupling and other relativistic effects, the relativistic momentum should capture all the possible relativistic spin-flip processes. We have found that the modification of the optical conductivity and MOKE spectra due to the relativistic momentum matrix elements are very small ($\sim 0.1\%$) in Ni and Fe. Therefore, we conclude that although the relativistic spin-flip processes do exist during ultrafast demagnetization, they do not provide a notable channel for the huge magnetization loss that has been observed in the demagnetization experiments.

In the later part of the chapter, we have explained the origin of coherent ultrafast magnetism with the help of a novel relativistic Hamiltonian which efficiently couples the electron's spin to the angular momentum of light. The corresponding Hamiltonian stems from the gauge invariant part of the spin-orbit coupling. We have rewritten this Hamiltonian in order to explain a Zeeman-like coupling between the electron's spin and an optically induced magnetic field which is proportional to the ellipticity and the intensity of light. Using the relativistic optomagnetic field expression, we have evaluated its strength during the interaction of pump and probe within the first 100 fs of the dynamics. We have found that the optomagnetic field strongly depends on the polarization of pump and probe; for linearly polarized pump and probe, if the polarizations are parallel to each other, the optomagnetic field becomes zero. However, if the polarizations are perpendicular to each other, the induced optomagnetic field becomes non-zero and it is directly proportional to the amplitude of pump, amplitude of probe, however, inversely proportional to the frequencies of both. For the latter case, the optomagnetic field depends on the pump-probe time delay as well. It is however worth to note that, if the time delay is large then the interactions between pump and probe cannot be accounted. Finally, we remark that similar optomagnetic fields can be anticipated for the elliptically polarized laser pulses.

5. Inverse Faraday Effect

5.1 Introduction

The Faraday effect (FE) is a magneto-optical phenomenon, which describes the rotation of the plane of polarization of linearly polarized light, while it passes through a magnetic medium [235]. If the medium does not possess any intrinsic magnetic moment, the medium should be kept in a magnetic field. In contrast, the inverse Faraday effect (IFE) occurs when circularly polarized light propagates through a nonabsorbing or absorbing crystal, where it induces a magnetic moment [236]. The induced magnetization due to the IFE changes sign when the circular polarization is changed from right to left circular polarization. Now, as linearly polarized light can be considered as the superposition of equal amounts of right and left circularly polarized rays, it was thought that linearly polarized light cannot induce IFE. However, using the orbital angular momentum of light, it has been shown that linearly polarized light can induce a magnetization as well [237]. The IFE is observed in paramagnets [238, 239] and plasmas [240–243] with circularly polarized light. In the same context, various models for the IFE have been proposed in the last decades [244–247]. However, there are many open questions. For example, it is not properly understood whether the IFE induces a magnetic field or a magnetization within a material.

Here, we describe the IFE in a way that it induces a magnetic moment (magnetization), $\Delta \boldsymbol{M}$ which is proportional to the cross product of electric fields, $\sim \boldsymbol{E} \times \boldsymbol{E}^{\star}$ i.e., the intensity of the incident light. This induced magnetization behaves differently for nonmagnetic systems and for magnetic systems as illustrated in Fig. 5.1. The induced magnetization can exert a torque on the preexisting magnetization and thus can assist to the magnetization switching in a magnetic material [248]. Therefore, the IFE has been instrumental to achieve all-optical switching in the ultrafast regime [28, 29]. This shows that the microscopic understanding of the IFE is potentially very rich. In the following, we describe the classical theory of the IFE followed by its quantum description, however, including the effect of light absorption.

5.2 Classical description

The classical description of the IFE starts with the assumption of a free electron gas, neglecting their quantum effects [249]. In the free electron model,

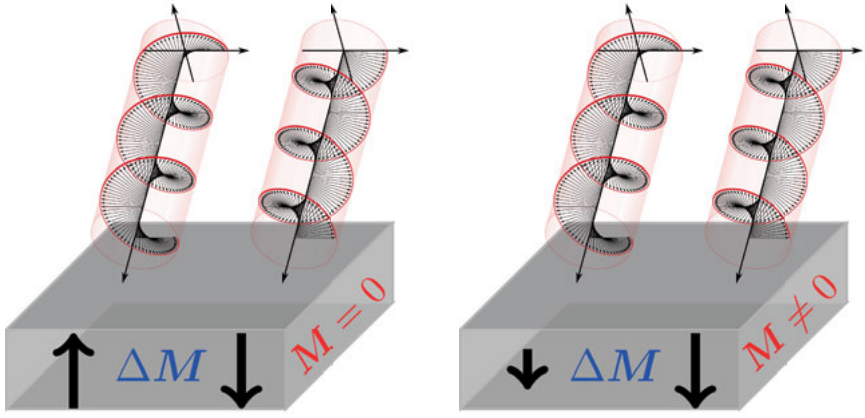

Figure 5.1. A schematic view of the inverse Faraday effect. (Left) For nonmagnetic and antiferromagnetic systems, where the net magnetization is zero, the opposite light's helicity induce same but opposite magnetizations ($\Delta \boldsymbol{M}$). (Right) For magnetic systems, where there is a net magnetization, $\boldsymbol{M}$ present, the opposite light's helicity induce asymmetric magnetizations.

the collisions of the electrons with the other electrons and ions are neglected. Moreover, in the absence of any external electromagnetic field, the electrons move freely on a straight line unless they collide with the others. While moving inside the metal, the electrons can collide with each other. These collisions are instantaneous and as a consequence, the electrons abruptly change their velocity and momentum. The probability of an electron having a collision within a time interval dt is γdt and that does not depend on the position of the electrons (γ is the collision rate, which takes into account the absorption of light). The electrons are assumed to have achieved thermal equilibrium with the surroundings after the collisions. With these assumptions, and according to Newton's second law, the dynamics of the electrons can be modeled by the Drude-Lorentz theory [245, 250]

$$m\frac{d^2\boldsymbol{r}(t)}{dt^2} + m\gamma\frac{d\boldsymbol{r}(t)}{dt} + m\omega_0^2\boldsymbol{r}(t) = \boldsymbol{F}(t)\,. \tag{5.1}$$

$\boldsymbol{r}(t)$ denotes the instantaneous position of the electron and the electromagnetic force $\boldsymbol{F}(t) = e\boldsymbol{E}$, acting on the electron. γ accounts for energy loss processes i.e., dissipations ($\gamma = 0$ for the nonabsorbing materials) and ω_0 is the angular frequency of the electron, which represents the electrons are bound by potential $\omega_0^2\boldsymbol{r}(t)$. To obtain the IFE induced effects, the electric field is taken as circular polarization $\boldsymbol{E} = E_0\,(\sin\omega t\,\boldsymbol{e}_x + \cos\omega t\,\boldsymbol{e}_y)$, with ω, the frequency of the incident light while the light is propagating towards the z-direction. The next step is to solve Eq. (5.1) for the position, $\boldsymbol{r}(t)$ and the velocity $\boldsymbol{v}(t) = \frac{d\boldsymbol{r}(t)}{dt}$. Now, we know that for a classical magnetization in materials, the magnetic moment is proportional to the function $\boldsymbol{r} \times \boldsymbol{j}$, where $\boldsymbol{j} = ne\boldsymbol{v}$ defines the cur-

rent density [251]. The induced orbital magnetization is then computed as $\boldsymbol{r}(t) \times \frac{d\boldsymbol{r}(t)}{dt}$ [250]. For the asymptotic solution the induced magnetization can be explicitly calculated and will be given as

$$\boldsymbol{M}_{\mathrm{IFE}} = -\frac{Ne^3}{2m^2} \frac{\omega}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2} E_0^2 \, \boldsymbol{e}_z \,, \tag{5.2}$$

where N is the electron density. Note that the magnetization is induced in the same direction of the light's propagation. We notice that the induced magnetization is proportional to the intensity (square of the amplitude of electric field) of the light, as originally proposed by Pitaevskii [236]. However, the effect of absorption was not considered in the thermodynamic description by Pitaevskii. Therefore, the consideration of the parameter γ underlines the main difference between the above-given classical Drude-Lorentz description and the approach of Pitaevskii [250]. This is realized by the following relation between the parameters ruling FE (ν_{FE}) and IFE (ν_{IFE}) [250, 252]:

$$\frac{\nu_{\mathrm{IFE}} - \nu_{\mathrm{FE}}}{\nu_{\mathrm{IFE}} + \nu_{\mathrm{FE}}} = -\frac{i\gamma\omega}{\omega^2 - \omega_0^2} \,. \tag{5.3}$$

These parameters are called Verdet constants for FE and IFE. As the derivation by Pitaevskii [236] was based on the dissipationless media ($\gamma = 0$), the equivalence of two parameters ruling FE and IFE was concluded. In the classical theory, however, only if one considers $\gamma = 0$, this equivalence is realized. The classical treatment thus provides a first step to understand the IFE while it accounts for the light's absorption. However, for realistic and more material specific calculations of the IFE, the quantum effects have to be included.

In the following, we present the quantum theory of the IFE and the implementation of the theory to calculate the IFE induced parameters.

5.3 Quantum description

The microscopic theory of IFE was described in the context of the possibility of achieving ultrafast optical control of magnetism [253]. The proposal was that IFE is an effect stemming from a Raman-like coherent scattering mechanism which generates an ultrafast and efficient magnetic field. When the laser pulse hits the sample, an electron is excited to a higher quantum state. This leads to the transition of the electron back to an another quantum state. However, if the transition is driven by an electric dipole, the spin state cannot be changed because within the electric dipole approximation spin does not interact with the electromagnetic field of the laser light. On the other hand, a magnetic dipole transition allows the spin-flip due to the interaction between the electron spin and the magnetic field. Therefore, first the electric dipole

transition followed by a magnetic transition would leave the electron in a final magnetic state that is different than the initial one. The whole process is schematically described in Fig. 5.2.

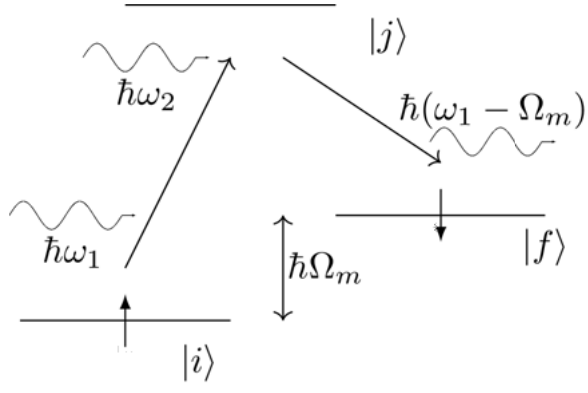

Figure 5.2. The Λ-like transitions for three level system. The laser light causes the transition from initial state, $|i\rangle$ to a intermediate state $|j\rangle$ and then finally back to the final state $|f\rangle$ which has different magnetic state from the one in initial. Adapted with permission from Hansteen *et al.*, Phys. Rev. B **73**, 014421 (2006). Copyright © 2006, The American Physical Society.

First, an optical excitation of $\hbar\omega_1$ initiates the transition of an electron from an initial to a higher virtual state (intermediate), with a large spin-orbit coupling. As the spin is not conserved, the spin-flip becomes allowed. Radiation at the energy $\hbar\omega_2$ then stimulates the relaxation of the electron back to the final ground state. Note that, due to a spin-flip in the intermediate state, the electron will be in a different magnetic state. During the final transition a radiation of $\hbar(\omega_1 - \Omega_m)$ energy will be released and a magnon of energy $\hbar\Omega_m$ will be created to compensate the energy. To account for the Raman-like processes and population of the different states during the IFE, an effective Hamiltonian approach has been considered, including the electron-electron interactions and spin-orbit coupling [246, 254].

In a further study by Battiato, Barbalinardo and co-workers [204, 250, 252], they employed the second order density response of the states to the circularly polarized light. In their work, the quantum Liouville-von Neumann equation was solved along with the perturbation from the circularly polarized light. The perturbation was treated as the interaction Hamiltonian $V(t) = e\sum_i \boldsymbol{r}_i \cdot \boldsymbol{E} = \frac{e}{m}\sum_i \boldsymbol{p}_i \cdot \boldsymbol{A}$, where $\boldsymbol{r}$ and $\boldsymbol{p}$ are the electron position and momentum operators respectively, $\boldsymbol{E}$ and $\boldsymbol{A}$ define the time-varying electric field and corresponding vector potential from the light. As discussed in Ref. [204], the static part of the second order density matrix response, $\hat{\rho}^{[2]}$, has been shown to contribute to the inverse Faraday effect. The induced magnetization is calculated as the contributions from the orbital and the spin angular momentum and they are

evaluated as

$$\boldsymbol{M}_{\mathrm{IFE}} = \mu_{\mathrm{B}} \mathrm{Tr}\left[(\hat{\boldsymbol{L}} + g\hat{\boldsymbol{S}})\hat{\rho}^{[2]}\right], \tag{5.4}$$

where $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{S}}$ denote the orbital and spin angular momentum operators respectively. The static part of the second order response has three terms and their complex conjugates. This means that the optically induced magnetization has three contributions which can be written as

$$M_{\mathrm{IFE}} = (\mathscr{K}_{\mathrm{o}} + \mathscr{K}_{\mathrm{dA}} + \mathscr{K}_{\mathrm{dB}} + \mathrm{c.c.})\, E_0^2, \tag{5.5}$$

with E_0 the amplitude of the light's electric field. These three contributions can be expressed in terms of the momentum and magnetization matrix elements as follows [255]

$$\begin{aligned}
\mathscr{K}_{\mathrm{o}} &= \frac{e^2}{m^2\omega^2} \sum_{n\neq m;l} M_{mn} \frac{\frac{p^+_{nl}p^-_{lm}(f_m - f_l)}{E_l - E_m + i\hbar\Gamma_{lm} - \hbar\omega} - \frac{p^-_{nl}p^+_{lm}(f_l - f_n)}{E_n - E_l + i\hbar\Gamma_{nl} - \hbar\omega}}{E_n - E_m + i\hbar\Gamma_{nm}}, \\
\mathscr{K}_{\mathrm{dA}} &= \frac{e^2}{m^2\omega^2} \sum_{n,l} M_{nn} \left(\frac{p^+_{nl}p^-_{ln}(f_l - f_n)}{(E_l - E_n + i\hbar\Gamma_{ln} - \hbar\omega)^2} - \frac{p^-_{nl}p^+_{ln}(f_n - f_l)}{(E_n - E_l + i\hbar\Gamma_{nl} - \hbar\omega)^2} \right), \\
\mathscr{K}_{\mathrm{dB}} &= \frac{e^2}{m^2\omega^2} \sum_{n,l} \frac{M_{nn}}{\hbar\omega} \frac{p^+_{nl}p^+_{ln}(f_n - f_l)(i\hbar\Gamma_{ln} - \hbar\omega)}{(E_l - E_n)^2 + (\hbar\Gamma_{ln} + i\hbar\omega)^2}.
\end{aligned} \tag{5.6}$$

Here, p_{nl} are the matrix elements of the momentum operator. These are calculated with the relativistic electronic states $|n\boldsymbol{k}\rangle$ corresponding to the energies $E_{n\boldsymbol{k}}$. For shake of brevity the momentum $\boldsymbol{k}$ has been suppressed in Eq. (5.6). $f_n = f(E_{n\boldsymbol{k}})$ are the Fermi-Dirac distribution functions. To account for the right and left circular polarization of the light, we have introduced the raising and lowering momentum operator as $\hat{\boldsymbol{p}}^{\pm} = (\hat{\boldsymbol{p}}_x \pm i\hat{\boldsymbol{p}}_y)$, remembering the fact that the light propagates towards the z-direction. The laser photon energy is $\hbar\omega$, and the line-widths are taken into account through Γ_{nl} matrix elements and those are taken as positive parameters, usually chosen to be state independent.

The above-mentioned expressions of the laser-induced magnetization contributions seems mathematically complicated, however, their physical meanings are immediately explained. The first term, $\mathscr{K}_{\mathrm{o}}$, arises from the off-diagonal components in the second order perturbed density matrix. Therefore, it includes all the possible optical transitions among different states - electronic Raman-like scattering. First the circular laser light is absorbed by the system under study, bringing the system into a virtual excited state. At this stage, the excited state can couple to the other available electronic or vibrational states of the system and therefore finally the system reaches to a state that is different from the initial one. These transitions involve the absorption and the emission of light but not by the equal amount. The second and third terms, $\mathscr{K}_{\mathrm{dA}}$

and $\mathscr{K}_{\mathrm{dB}}$, arise from the diagonal components in the second order perturbed density matrix - electronic Rayleigh-like scattering. In these cases, upon the absorption of light, the system reaches to an excited state. But at this stage, the excited states do not couple to the other available states of the system and consequently, the initial and final states are the same. These transitions, thus, involve the absorption and emission of light by same amount.

Notice that, in particular, all the terms in Eq. (5.6) diverge when $\omega \to 0$ because of its appearance in the the denominator. Hence, it is expected that the induced magnetization may also diverge at $\omega \to 0$. To understand this unusual result, it is important to keep in mind that the treated field is that inside the material. For metals, the reflectivity approaches 100% when $\omega \to 0$, implying that the field in the material is strongly reduced. While the laser-induced magnetization is calculated, eventually it induces a corresponding optomagnetic field that can be obtained from the susceptibility of the material as $\boldsymbol{H} = \boldsymbol{M}/\chi$. We note that, this approach would be valid if both induced spin and orbital magnetizations behave in the same way, that is, behave as a response to an applied Zeeman field.

5.3.1 *Ab initio* calculations

We have implemented these contributions to the induced magnetization in an *ab initio* relativistic DFT framework. The effect of absorption which was ignored in previous theories [236, 256], is included in our calculation. The important point here that we address is that, there are contributions from the orbital and the spin parts to the induced magnetization. Our configuration set up is such that the laser light travels towards the z-direction, thus for circularly polarized light the electric field components will be in the x- and y-directions. This means that the induced magnetization will only have the z-component. Therefore, following the definition of laser-imparted magnetization in Eq. (5.4), we get

$$\begin{aligned} M^{z}_{\mathrm{IFE}} &= \mu_{\mathrm{B}} \mathrm{Tr}\left[\hat{L}^{z}\hat{\rho}^{[2]}\right] + 2\mu_{\mathrm{B}} \mathrm{Tr}\left[\hat{S}^{z}\hat{\rho}^{[2]}\right], \\ &= \left[K^{L}_{\mathrm{IFE}}(\omega) + K^{S}_{\mathrm{IFE}}(\omega)\right] \frac{I}{c}. \end{aligned} \tag{5.7}$$

The laser intensity is denoted by $I = c\varepsilon_0 E_0^2/2$ and the frequency dependent orbital and spin IFE parameters are $K^{L}_{\mathrm{IFE}}(\omega)$ and $K^{S}_{\mathrm{IFE}}(\omega)$ (see **paper V** for details). In this way, the parameters $K_{\mathrm{IFE}}(\omega)$ have the dimension of inverse Tesla.

In practice, when the circularly polarized light hits on a magnetic or non-magnetic system, depending on their magnetic ordering, the absorption of the light and all the possible excitations (transitions) happen differently for right and left circularly polarized light. Therefore, it is not possible to predict

clearly the behavior of the light-induced magnetizations for two circular polarizations, unless they are calculated *ab initio*. In the following, we present *ab initio* calculations of the IFE which consist of the contributions of the orbital, spin and total induced magnetization in a photon energy span of 0 - 4 eV.

♣ Detail of the calculations

In particular, we calculate the given three expressions in Eq. (5.6) starting from *ab initio* calculations. To calculate those, we need informations of the several parameters involved in the expressions. The momentum matrix elements (p_{nl}) are calculated with the relativistic electronic states which are the solutions of the DKS equation, solving the electron spinor equations for the large and small components numerically. The corresponding relativistic band energies are given by E_n. The magnetization matrix elements are also calculated similarly with the relativistic electronic states. The calculations are performed with the Fermi-Dirac distributions (f_l) taken as a heavy-side step functions because the current calculations do not account for temperature effects. The parameter for the consideration of broadening or line width is taken as, $\hbar\Gamma = 0.03$ R$_y$. This value has been known to give a very good description of *ab initio* calculations of linear-order optical response of metals [224]. The polarization of the electric field is considered for the right circularly polarized light (we denote it as σ^+) as $\boldsymbol{E} = E_0\,(\sin\omega t\,\boldsymbol{e}_x + \cos\omega t\,\boldsymbol{e}_y)$. Such polarization for a ray, the wave propagating along $-z$-direction corresponds to a left circular polarization (we denote it as σ^-). In general, the space and time dependent electric field has the following form

$$\boldsymbol{E}(\boldsymbol{r},t) = E_0\,\mathrm{Re}\left[i|\Psi_m\rangle e^{i(kz-\omega t)}\right]. \tag{5.8}$$

Here $|\Psi_m\rangle$ are the vectors in the $x-y$ plane. For the right and left circular polarization, we define those vectors as

$$\begin{aligned}
|\Psi_{\mathrm{R}}\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}}\,(\boldsymbol{e}_x - i\boldsymbol{e}_y)\,, \\
|\Psi_{\mathrm{L}}\rangle &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}}\,(\boldsymbol{e}_x + i\boldsymbol{e}_y)\,.
\end{aligned} \tag{5.9}$$

The momentum matrix elements of the contributions in Eq. (5.6) are calculated within the augmented spherical wave (ASW) method with the relativistic band energies and wave functions. The calculations of momentum matrix elements use the Bessel and Hankel functions as nicely described in the Ref. [220].

♣ Results

The calculated induced magnetization is shown for σ^+ and σ^- and separately for the contributions from orbital, spin and total magnetization.

Nonmagnetic metals:

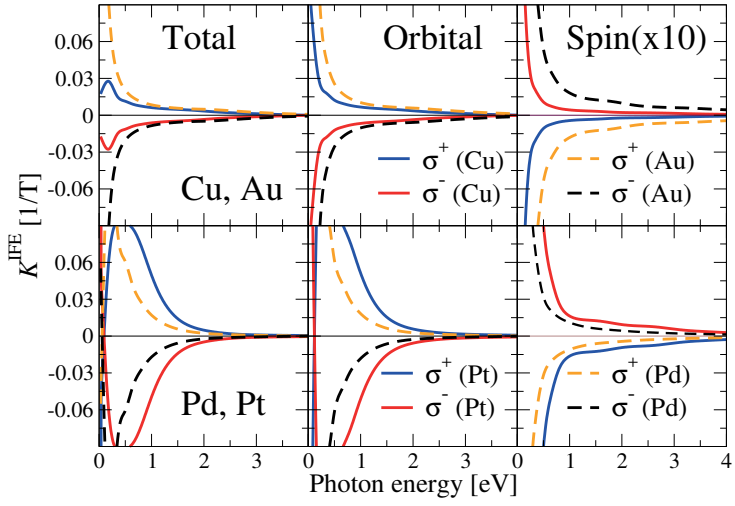

Figure 5.3. Calculated different contributions to the IFE for nonmagnetic metals as a function of photon energy. (Top panels): The orbital, spin and total IFE effects are shown for Au (black and yellow dashed lines) and Cu (red and blue solid lines). (Bottom panels): The orbital, spin and total IFE effects are shown for Pd (black and yellow dashed lines) and Pt (red and blue solid lines). Adapted with permission from Berritta *et al.*, Phys. Rev. Lett. **117**, 137203 (2016). Copyright © 2016, The American Physical Society.

In Fig. 5.3, we show the IFE induced magnetization for nonmagnetic metals (net magnetization is zero) in Cu, Au, Pd and Pt. The calculations show the existence of finite contributions to not only the spin IFE, but also the orbital IFE. Note that, the orbital IFE was unknown previously and also was ignored in the original work by Pitaevskii [236]. In addition, we note that the magnitude of the orbital IFE is 10 times larger than the spin IFE. However, these contributions are in the opposite directions for a particular helicity, and therefore they compete with each other to the total IFE.

The IFE contributions are visibly materials specific. We notice that the induced magnetization is larger in Au than the one in Cu in both cases of orbital and spin IFE. A similar behavior can also be concluded for the case of Pt and Pd, where the induced magnetization in Pt is much higher than Pd. All together, the total induced magnetization in Pd and Pt is larger than the one in Cu and Au. We can anticipate the fact that the IFE induced magnetization diverges at $\omega \to 0$ as there are $1/\omega^2$ in the prefactor of the set of Eqs. (5.6). More importantly, the IFE induced effects are antisymmetric in the light's helicity which means right circularly polarized light induces exactly the same and opposite orbital (spin) IFE as left circular polarized light [255].

Ferromagnetic metals: Notably, the ferromagnetic metals show a completely different behavior for the induced IFE than the nonmagnetic metals. The cal-

culated induced magnetizations are shown for the 3d transition metals, namely, bcc Fe, hcp Co and fcc Ni in Fig. 5.4.

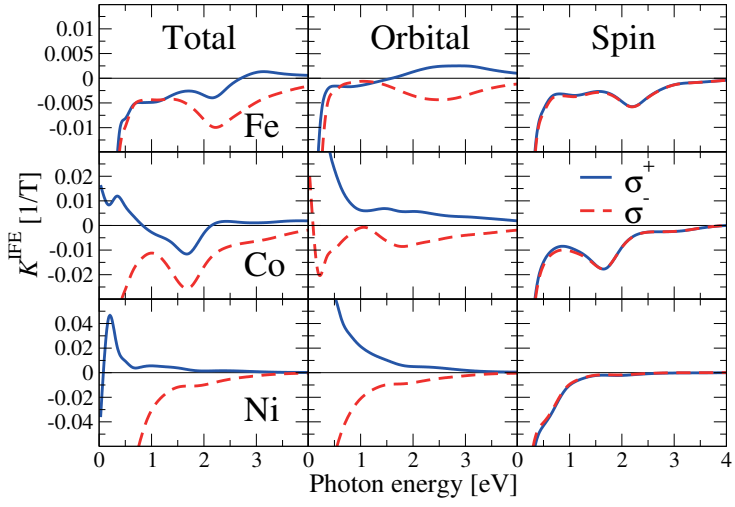

Figure 5.4. Calculated different contributions to the IFE for ferromagnetic metals. The red dashed and blue solid lines show the contributions induced by left and right circularly polarized light respectively. (Top panels): The orbital, spin and total IFE effects are shown for bcc Fe. (Middle panels): the orbital, spin and total IFE effects are shown for hcp Co. (Bottom panels): The orbital, spin and total IFE effects are shown for fcc Ni. Adapted with permission from Berritta *et al.*, Phys. Rev. Lett. **117**, 137203 (2016). Copyright © 2016, The American Physical Society.

In the following, we analyze the obtained results for each material [255].

Bcc Fe: The contribution to the spin IFE does not depend on the helicity of the light, for right and left circularly polarized light, the spin IFE curves fall on top of each other. However, the orbital IFE depends on the helicity of the circularly polarized light. The right and left circular polarization induce an orbital magnetization in the same direction, except in a span of photon energies 2 - 4 eV. Therefore, the induced total magnetization also points in the same direction for both helicities. Unlike the nonmagnetic metals, the spin and orbital IFE have the similar order of magnitude and they contribute to the total IFE.

Hcp Co: The behavior of induced spin IFE is the same as for bcc Fe, they almost overlap with each other for opposite light helicities. The orbital IFE shows a different scenario than bcc Fe. Except for the low energies, the right circularly polarized light induces orbital IFE in the opposite direction of that induced by the left circularly polarized light and they are *asymmetric* in the light's helicity. Unlike bcc Fe, the right circularly polarized light induces orbital and spin IFE in the opposite direction, hence, the total induced magnetization is less. On the other hand, the left circular polarized light induces spin

and orbital IFE in the same direction, increasing the total induced magnetization.

Fcc Ni: The effect of spin IFE is effectively the same as Fe or Co. However, the orbital IFE are approximately *antisymmetric* in the light helicity. As seen in hcp Co, a similar behavior of spin and orbital IFE for the right and left circularly polarized light are anticipated. A reason of the asymmetric behavior is that, for ferromagnets where the net magnetization is nonzero, the time reversal symmetry is notably broken already and thus a full antisymmetric behavior is not expected.

Antiferromagnetic materials:

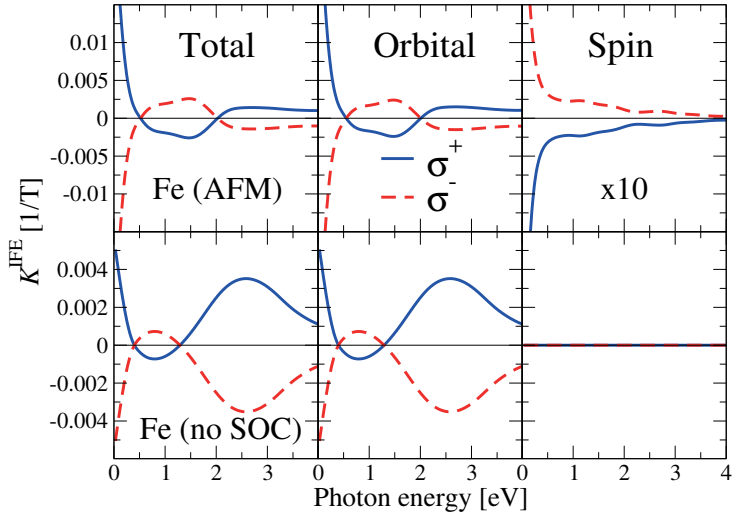

Figure 5.5. Calculated different contributions to the IFE for antiferromagnetic Fe and ferromagnetic Fe without spin-orbit coupling. The red dashed and blue solid lines show the contributions induced by left and right circularly polarized light respectively. (Top panel): The orbital, spin and total IFE effects are shown for synthetic antiferromagnetic Fe. (Bottom panel): The orbital, spin and total IFE effects are shown for bcc Fe computed without spin-orbit coupling. Adapted with permission from Berritta *et al.*, Phys. Rev. Lett. **117**, 137203 (2016). Copyright © 2016, The American Physical Society.

For antiferromagnetic materials, the net magnetization is zero and once again the IFE is expected to have antisymmetric nature with respect to the light's helicity. The calculations are presented in Fig. 5.5, for synthetic antiferromagnetic Fe. Note that, the spin IFE is about 10 times smaller than the orbital IFE, as was also concluded in the nonmagnetic case. Similar to the nonmagnetic materials, the orbital and spin IFE is induced in opposite direction for a particular helicity. Thereby, the total induced magnetization is also antisymmetric in the light's helicity. In addition, antiferromagnetic Fe shows

an photon energy dependence that is different from that of ferromagnetic Fe [255].

To summarize, the *ab initio* calculations reveal that the IFE is more complex at the fundamental level than it was thought. It is found that the induced magnetization is antisymmetric in the helicity for the magnetic system with zero net magnetization. The same is expected for paramagnetic materials as well. However, for ferromagnetic and ferrimagnetic systems, where a net magnetization exists, the induced magnetization is shown to be asymmetric in the helicity of the light. Furthermore, in Fig. 5.5, the calculated IFE effects are shown without spin-orbit coupling for ferromagnetic Fe. We see that the without spin-orbit coupling, the spin IFE is zero irrespective of the photon energy of the light. This reveals that the spin-orbit coupling is purely responsible for the laser-imparted magnetization in the spin components.

5.4 Angular magneto-electric coupling

There exist mainly four classes of ferroic orders in a multi-ferroic material. These are: ferroelastic, ferroelectric, ferromagnetic and ferrotoroidic orders. These can be classified in terms of the space inversion and time reversal symmetry considerations [257]. For example, the space inversion and the time reversal symmetry is conserved for the case of ferroelastic materials because only mechanical processes (stress, strain) are involved. On the other hand, ferroelectric materials are electrically polarizable means they have spontaneous electric polarization which comes from the charge density. Now, as the charge density depends on the spatial coordinates, the space inversion symmetry is broken for ferroelectric materials, however, the time reversal symmetry is conserved. The opposite picture is realized for ferromagnetic materials. In this case, the the magnetism arises due to the current density which changes sign under the time reversal, however, stays invariant under space inversion. Both the symmetries are broken for the ferrotoroidic materials, where the alignment of the magnetic vortices gives rise to a toroidal moment. The observation of ferrotoroidic order has been reported recently [258]. These different ferroic orders are pictorially classified in Fig. 5.6.

In this part of the chapter, we are mostly interested in the ferrotoroidic systems and the eventual manipulation of the toroidal moments using electric and magnetic fields. In the multipole expansion of the magnetic vector potential, the toroidal dipole moment arises from the second order term in the expansion [257]. The toroidal moment can be seen as the cross product between position and magnetization, $\boldsymbol{M}$. According to the definition, the toroidal moment has the following expression [257]

$$\boldsymbol{T} = \frac{1}{2}\int \boldsymbol{r}\times\boldsymbol{M}\, d^3r\,. \tag{5.10}$$

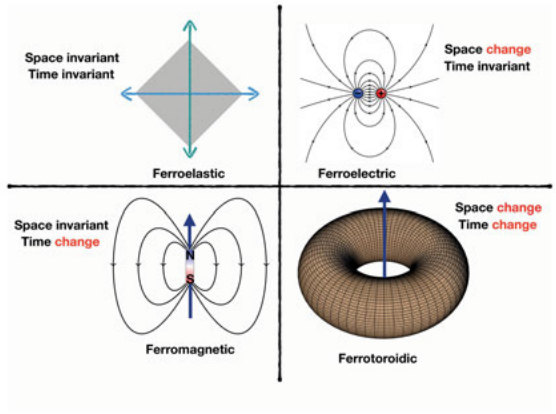

Figure 5.6. The classification of the various ferroic orders according to space and time symmetry considerations.

In a toroid, the current flowing perpendicular to the toroid creates the moment (see bottom-right part of Fig. 5.6). These current loops will be shifted by the application of a magnetic field, inducing a polarization - a magnetoelectric coupling [258]. This manipulation of the toroidal dipole moment is potentially important for technological purposes e.g., in data storage [257, 259]. The manipulation has been studied by using a static inhomogeneous field [260], however, the crossed electric and magnetic fields ($\boldsymbol{E} \times \boldsymbol{H}$) can also be used [63]. The reason is that the crossed electric and magnetic field breaks both the time reversal and space inversion symmetry. Therefore, both fields are needed for manipulation of the toroidal moment. The coupling between the toroidal moment and the crossed electric and magnetic fields leads to the introduction of a physical energy [62]

$$\mathscr{E} = -\frac{1}{2}\int \xi[\boldsymbol{r} \times (\boldsymbol{E} \times \boldsymbol{H})] \cdot \boldsymbol{M}\, d^3 r. \tag{5.11}$$

The electric and magnetic fields are considered as $\boldsymbol{E}$ and $\boldsymbol{H}$ and ξ is the material dependent strength of the coupling parameter which is unknown. Notice that, $\boldsymbol{r} \times (\boldsymbol{E} \times \boldsymbol{H})$ exactly measures the total angular momentum density of an EM field. Therefore, essentially, the above energy describes the coupling of angular momentum of an EM field with the electron spin, however, is only put forward from a phenomenological basis and simple symmetry considerations. Nonetheless, the energy can be useful in describing many complex phenomena e.g., the anomalous Hall effect [64], spin current model in multiferroics [62], the planar Hall effect and anisotropic magnetoresistance [65].

As of now, the physical energy in Eq. (5.11) is very useful, however, it is phenomenologically described. On the other hand, the energy expression vaguely resembles to the one of well-known spin-orbit coupling, but with only the difference that the angular momentum of light ($\propto \boldsymbol{r} \times (\boldsymbol{E} \times \boldsymbol{H})$) couples to

the spin. This energy has been called angular magneto-electric (AME) coupling energy [62].

Here, in the below we provide the microscopic derivation of the AME energy in the optical regime while starting from the relativistic DKS Hamiltonian.

5.4.1 Microscopic derivation

In **paper II** and **paper III**, we rigorously derive the physical energy of Eq. (5.11) from the gauge invariant part of the spin-orbit interaction in Eq. (2.82). Thus, we name the corresponding Hamiltonian as the spin-photon coupling that has the expression [163]

$$\mathscr{H}_{\text{spin}-\text{photon}} = \frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\boldsymbol{E}\times\boldsymbol{A}) \tag{5.12}$$

Note that, here in contrast to the AME consideration in Eq. (5.11) with static applied fields, we consider general time-dependent photon fields. As already discussed in Chapter 2, there are two electric fields present, the internal and external. Therefore, the spin-photon coupling Hamiltonian can be split in an intrinsic part, caused by the microscopic field generated by the crystal potential ($\boldsymbol{E}_{\text{int}} = -\frac{1}{e}\boldsymbol{\nabla}V$), and an extrinsic part that is linked to the external electric field ($\boldsymbol{E}_{\text{ext}} = -\partial\boldsymbol{A}/\partial t$) such that $\boldsymbol{E} = \boldsymbol{E}_{\text{int}} + \boldsymbol{E}_{\text{ext}}$. Now, within a material, the internal electric field will be changed by the perturbation from the external field. Within the linear response framework, we write

$$\boldsymbol{E}_{\text{int}} = \boldsymbol{E}^0_{\text{int}} + \gamma\cdot\boldsymbol{E}_{\text{ext}}\,, \tag{5.13}$$

with $\boldsymbol{E}^0_{\text{int}}$, the internal electric field which exists even without applying any field and γ is a material linear response to the applied field. The parameter γ is usually a tensor, however, for a cubic system can be considered as a scalar quantity. With this simplification, the Hamiltonian in Eq. (5.12) can be expressed in two terms:

$$\mathscr{H}_{\text{spin}-\text{photon}} = \frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\boldsymbol{E}^0_{\text{int}}\times\boldsymbol{A}) + \frac{e^2\hbar(1+\gamma)}{4m^2c^2}\boldsymbol{\sigma}\cdot(\boldsymbol{E}_{\text{ext}}\times\boldsymbol{A})\,. \tag{5.14}$$

The first term, here, can be dropped because the microscopic field does not exist in the bulk of a solid when it is averaged over an unit cell. Therefore, we proceed while taking the second term into account.

We point out that, the derivation of the Eq. (5.11) from the second term of Eq. (5.14) is based on several strong assumptions. First, the magnetic field is assumed to be uniform that is taken into account through the choice of gauge as $\boldsymbol{A} = \boldsymbol{B}\times\boldsymbol{r}/2$. This is because we assume that the skin depth of the applied EM field is much larger than the thickness of the system being studied. Second,

the results are derived within the Coulomb gauge that means $\boldsymbol{\nabla}\cdot\boldsymbol{A}=0$, in accordance with the Maxwell equations. With the choice of gauge, we insert the vector potential in second term of Eq. (5.14) and taking its expectation value results in

$$\mathscr{E}_{\mathrm{AME}}=\frac{\xi}{2}\int \boldsymbol{M}\cdot[\boldsymbol{E}_{\mathrm{ext}}\times(\boldsymbol{B}\times\boldsymbol{r})]\,d\boldsymbol{r}\,, \tag{5.15}$$

where ξ is a material dependent parameter which can be realized due to the fact that $\xi\propto(1+\gamma)$. Third, we have used a linear proportional relationship between the magnetization and magnetic field. As we stressed in Chapter 3, this is typically not true in ferromagnets, where only a small component of the magnetization, the induced magnetization, is proportional to the applied magnetic field, in contrast to the spontaneous magnetization existing even at zero field. Nonetheless, we expand the expression in Eq. (5.15) and use the proportionality relationship in the second term of the expansion to interchange $\boldsymbol{B}$ and $\boldsymbol{M}$ so that we get [163]

$$\begin{aligned}\mathscr{E}_{\mathrm{AME}}&=\frac{\xi}{2}\int\Big[(\boldsymbol{M}\cdot\boldsymbol{B})(\boldsymbol{r}\cdot\boldsymbol{E}_{\mathrm{ext}})-(\boldsymbol{M}\cdot\boldsymbol{E}_{\mathrm{ext}})(\boldsymbol{B}\cdot\boldsymbol{r})\Big]\,d\boldsymbol{r}\\&=-\frac{\xi}{2}\int\boldsymbol{M}\cdot[\boldsymbol{r}\times(\boldsymbol{E}_{\mathrm{ext}}\times\boldsymbol{B})]\,d\boldsymbol{r}\,.\end{aligned} \tag{5.16}$$

Noticeably, this is exactly the physical energy of the phenomenologically proposed AME coupling (see Eq. (5.11)) [62]. We thus provide the microscopic derivation of manipulation of magnetization by the angular momentum of light; it straightforwardly derives from the DKS Hamiltonian, in particular, from the gauge invariant part of the spin-orbit coupling - the spin-photon coupling.

As there are assumptions involved in the derived AME energy in Eq. (5.16), this expression is not general. Instead, we use, in the following, the general spin-photon coupling Hamiltonian of Eq. (5.12) and discuss how it derives the relativistic correction to the IFE.

5.5 Relativistic contributions to the inverse Faraday effect

In the discussions of earlier sections (Sec. 5.3), the IFE has been attributed to the second order density response to the circularly polarized light. The interaction Hamiltonian of light and electrons that derives IFE is taken as non-relativistic as $\boldsymbol{p}\cdot\boldsymbol{A}$. In this way, the IFE induces a magnetization in a material. On a different note, the essence of IFE was thought to reside on the induced optomagnetic field [28, 30, 261]. The latter defines the IFE in a magnetic material as an effective Zeeman field - optomagnetic field as $\boldsymbol{B}_{\mathrm{opt}}\cdot\boldsymbol{M}_i$, acting on

the atomic spin moment $\boldsymbol{M}_i$ with conserved length (see the discussion of optomagnetic field in the Sec. 4.4.1). As the IFE is a quadratic response of the applied electric field, the corresponding Hamiltonian has the following form [30]

$$\mathscr{H}_{\text{IFE}} = \zeta \, \text{Re}\left[i(\boldsymbol{E} \times \boldsymbol{E}^{\star})\right] \cdot \boldsymbol{\sigma}\,, \tag{5.17}$$

where ζ is a material dependent parameter. If the electric fields are taken as the plane wave solutions of the Maxwell equations i.e., $\boldsymbol{E} = \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$, only the amplitude of the electric field contributes to the IFE induced optomagnetic field. Comparing with the Zeeman field, the optomagnetic field ($\boldsymbol{B}_{\text{opt}}$) will be proportional to the real part of the crossed electric field amplitudes such that the Eq. (5.17) can be written in the form: $\mathscr{H}_{\text{IFE}} = -\frac{\zeta}{g\mu_{\text{B}}}\boldsymbol{B}_{\text{opt}} \cdot \boldsymbol{\sigma}$.

Here, in the following, we show that the spin-photon coupling Hamiltonian naturally derives the relativistic optomagnetic field of IFE. We use the transverse part of the electric field $\boldsymbol{E} = -\partial \boldsymbol{A}/\partial t$, such that the vector potential can be calculated with

$$\boldsymbol{A} = -\text{Re}\left[\frac{i}{\omega}\boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right]. \tag{5.18}$$

Using these considerations, the second term of the spin-photon coupling Hamiltonian in Eq. (5.14) takes the form

$$\mathscr{H}_{\text{spin-photon}} = -\frac{e^2\hbar(1+\gamma)}{4m^2c^2\omega}\boldsymbol{\sigma} \cdot \text{Re}\left[i(\boldsymbol{E}_0 \times \boldsymbol{E}_0^{\star})\right]. \tag{5.19}$$

The derived Hamiltonian thus provides precisely the energy related to the IFE (see Eq. 5.17). Considering that the IFE induces a relativistic optomagnetic field, the latter will be given by [163]

$$\boldsymbol{B}_{\text{opt}} = \frac{e^2\hbar(1+\gamma)}{4m^2c^2\omega g\mu_{\text{B}}} \, \text{Re}\left[i(\boldsymbol{E}_0 \times \boldsymbol{E}_0^{\star})\right]. \tag{5.20}$$

We see that the relativistic optomagnetic field is material dependent and strongly depends (inversely proportional) on the angular frequency of the incident light, however, linearly proportional to the intensity of the light. Note that, a similar optomagnetic field has been used to describe the coherent ultrafast magnetism in pump-probe spectroscopy in ultrashort timescales (see Sec. 4.4.1). In particular, if we consider an elliptically polarized light of amplitude E_0 and ellipticity parameter η so that

$$\boldsymbol{E}_0 = \frac{E_0}{\sqrt{2}}\left(\boldsymbol{e}_x + e^{i\eta}\boldsymbol{e}_y\right), \tag{5.21}$$

the magnitude of the induced optomagnetic field in Eq. (5.20) takes the simpler form in terms of the intensity of light $I = c\varepsilon_0 E_0^2/2$ and ellipticity as

$$B_{\text{opt}} = \frac{e^2\hbar(1+\gamma)}{4m^2c^3\varepsilon_0\omega g\mu_{\text{B}}} \sin\eta \; I\,. \tag{5.22}$$

Again, we point out that as the incoming light has the electric field components in the x- and y-directions, the optomagnetic field will be induced in the z-direction. This way, the induced field is zero for linearly polarized wave because $\eta = 0$ or π leads to $\sin\eta = 0$. For circularly polarized light ($\eta = \pm\pi/2$), the induced optomagnetic field is maximal ($\sin\eta = \pm 1$). The magnitude of the induced field is equal but opposite for the case of right ($\eta = +\pi/2$) and left ($\eta = -\pi/2$) circularly polarized light, for the same intensity and same angular frequency. Now, the linearly polarized light is a combination of equal amounts of right and left circularly polarized light and thus the effective induced field will be zero for the linear polarization.

In order to estimate the strength of the induced relativistic optomagnetic fields, we use the experimentally used values of the intensity and wavelength of the contemporary laser source which are largely available. We consider circularly polarized light of intensity $I = 10\ \mathrm{GW/cm^2}$ and wavelength centered about 800 nm (angular frequency, $\omega = 2.35 \times 10^{15}$ Hz). The magnitude of the IFE induced field in this case will be $B_{\mathrm{opt}} = 8 \times (1+\gamma)\sin\eta\ \mu\mathrm{T}$. For circularly polarized light and in the limit $\gamma \to 0$ (no screening response of the system to the applied field) the relativistic optomagnetic field is of the order of 8 μT. However, this value is the lowest determined in the zero response limit, and it will be increased once we include the response of the system. As discussed towards the end of **paper II** and **paper III**, the value of γ would possibly be high leading to an induced optomagnetic field of mT up to Tesla. Lastly, we stress that the induced field strongly depends on the intensity of the incoming light and more intense light would induce higher optomagnetic fields.

To summarize, the gauge invariant part of the spin-orbit coupling i.e., the spin-photon coupling Hamiltonian has been shown to explain the microscopic origin of the phenomenologically predicted the angular magneto-electric coupling. The latter coupling defines the interaction of angular momentum of a EM field and the electron spin. The spin-photon coupling Hamiltonian has been overlooked for many years, however, we show that this relativistic coupling could explain many important physical phenomena. Although, the derivation involves mainly two assumptions, the true origin of the AME coupling energy is revealed. Nonetheless, the spin-photon coupling derives from the DKS Hamiltonian and it is exact. Later, considering elliptically polarized light, we show that this spin-photon coupling leads to a Zeeman-like field, which in turn defines the physical energy behind the explanation of the relativistic correction to the IFE. As we have discussed in the first part of this chapter that main IFE is a different effect, this relativistic spin-photon coupling only serves as a relativistic contribution to the IFE. Although, the main IFE is described as the induced magnetization, whereas in this way, the IFE correction has been described as an induced optomagnetic field. Finally, we also comment that the strength of the optomagnetic field strongly depends on the intensity and frequency of the incident light and the material's response to the light.

6. Summary and Outlook

Relativistic quantum mechanics has been the main cornerstone of physics that can describe massive particles propagating at any velocities, limited only by to the velocity of light. Relativistic theories have been proven to be very much useful in describing many complex phenomena in high-energy physics, particle physics, and atomic physics. Throughout this thesis, it has been shown that relativistic theories are equally important in condensed matter physics as well. In particular, we have investigated the relativistic effects within a general magnetic system excited by an electromagnetic field pulse.

To elucidate the relativistic effects in magnetic systems, we have considered the Dirac-Kohn-Sham Hamiltonian which efficiently describes the particle and antiparticle simultaneously. We have included a spin-polarized Kohn-Sham potential to account for the magnetic exchange interactions. In the low energy excitation, the separation between particle and antiparticle is necessary because only the particles can be described within the extended Pauli Hamiltonian. This separation is realized by the so called Foldy-Wouthuysen transformation that is exact for a free Dirac particle. However, if the relativistic particle experiences a time-dependent field, this separation is not exact and the transformed Hamiltonian needs further corrections. Nonetheless, using a corrected FW transformation, the true form of the extended Pauli Hamiltonian is revealed up to an order of $1/c^4$. The latter Hamiltonian includes the nonrelativistic Schrödinger-Pauli Hamiltonian and other relativistic corrections (e.g., relativistic mass correction, Darwin terms, spin-orbit coupling, relativistic corrections to exchange field, higher-order spin-orbit coupling). These relativistic correction terms bear plenty of enlightenment in the area of laser-induced relativistic magnetization dynamics.

We have extensively derived the Landau-Lifshitz and Landau-Lifshitz-Gilbert magnetization dynamics from the extended Pauli Hamiltonian. As we focus on the spin dynamics, the spin Hamiltonian is taken into account in this case. Our results reveal: the Zeeman-like coupling with the external field and exchange field gives the magnetization Larmor precession. The intrinsic spin-orbit coupling contributes to the relativistic correction to the precession. Furthermore, for a system driven by a harmonic external field, the extrinsic spin-orbit coupling exactly provides the Gilbert damping but here the damping parameter is a tensor. Along with the isotropic Gilbert damping, we have shown the existence of anisotropic Ising-like and chiral Dzyaloshinskii-Moriya-like damping. On the other hand, for a system driven by nonharmonic fields, the Gilbert damping alone is not adequate to describe the dynamics; we have derived the existence

of the field-derivative torque that might lead to the switching of magnetization when a sharp ultrashort laser pulse is employed. Moreover, we have derived an additional torque - the optical spin-orbit torque which suggests new ways to manipulate spins using the angular momentum of light. From the higher-order spin-orbit coupling, we derive the magnetic inertial dynamics that could be significant at the ultrashort timescales. We have made a comparison between the derived Gilbert damping parameter and inertia parameter; they scale with the imaginary and real part of the magnetic susceptibility, respectively. However, the inertia parameter is in general much smaller than the Gilbert damping (by $1/c^2$). We have derived the spin-current tensors within our description and we have established the adiabatic and nonadiabatic spin-transfer torques in our extended magnetization dynamics. In particular, we show that the adiabatic spin-transfer torque of Berger form arises within the nonrelativistic magnetization dynamics, however, the nonadiabatic spin-transfer torque derives from the relativistic magnetization dynamics.

Using the relativistic momentum operator, we have calculated *ab initio* the relativistic magneto-optical Kerr rotation and ellipticity of Ni and Fe, and the results underline that the relativistic spin-flip effects can modify the complex Kerr spectra by only $\sim 0.1\%$. This small amount of modification does not explain towards the huge loss of magnetization in an ultrafast demagnetization experiment. However, using the relativistic spin-photon coupling, we have explained the origin of the experimentally observed *coherent ultrafast magnetism* that depends on the polarization of light.

In the last part of the thesis, we have presented the coherent laser-imparted magnetization (inverse Faraday effect) in metals using a second order density-matrix response. The latter includes the Raman-like and Rayleigh-like scattering matrix elements. Our *ab initio* calculations show that the inverse Faraday effect is much more complex than it was earlier described in the literature. For nonmagnetic metals, we find that the induced magnetization is exactly antisymmetric in the light's helicity. However, the imparted magnetization in spins is very much smaller than that in orbitals. The total induced magnetization thus resembles to the orbital inverse Faraday effect. On the other hand, we show that for ferromagnetic metals, the spin inverse Faraday effect does not depend on the helicity of light. However, the induced orbital magnetization is asymmetric in the light's helicity, therefore, the total induced magnetization is asymmetric. Also, antiferromagnetic materials show a similar behavior to that of nonmagnetic case.

In a future work, the magnetization dynamics including the field-derivative torque and optical spin-orbit torque has to be explored in more detail in order obtain a better understanding of its physical importance. As in the experiments of ultrafast magnetization dynamics, often an ultrashort laser pulse is used, this field-derivative torque might be interesting for switching applications. On the other hand, there have recently been attempts to experimentally manipulate the spins in a material using the light's angular momentum

that leads to first experimental evidences for the optical spin-orbit torque. At the same time, the tensorial Gilbert damping has to be implemented within the atomistic spin dynamics simulation to elucidate the underlying physics of anisotropic damping and chiral, Dzyaloshinskii-Moriya damping. The latter can become useful in the case of magnetic texture dynamics e.g., domain-wall motion. The anisotropic damping can be realized in a ferromagnetic resonance experiment because when the magnetization vector rotates around a field, the direction of magnetization continuously changes which serves as anisotropy to the damping. The anisotropic damping can have importance for the skyrmion dynamics as well on the size and velocity of skyrmions [262]. It has to be noted that as we have done, it is possible to show that the Gilbert damping parameter can be expressed as a series of higher-order relativistic terms. To realize that, one has to evaluate the higher-order relativistic corrections and derive the magnetization dynamics taking into account those terms. As the Gilbert damping parameter is notably relativistic, the higher-order terms will be smaller and they will have less impact. Nonetheless, it is important to know the exact expression for the Gilbert damping parameter.

In our formulations of the Landau-Lifshitz-Gilbert magnetization dynamics, we have not considered the orbital contributions, only the spin dynamics has been accounted. However, the orbital angular momentum too contributes to the magnetization of the material. Therefore, another aspect is to include the orbital contributions in the magnetization dynamics; this is very important for the materials where the spin-orbit coupling is strong. In particular, one can obtain the spin and orbital angular momentum dynamics with the Hamiltonian which we have derived, in the form of Heisenberg operator dynamics. In the components form, one then has to solve six first-order differential equations which are coupled to each other. Then the transfer of angular momentum from spins to orbitals, and to lattice will become clearer from such relativistic theory. In this way one achieves the complete spin and orbital angular momentum dynamicswithin the Dirac-Kohn-Sham theory, which would be sufficient to describe the involved many-particle interactions.

7. Relativistisk teori om ljus-spinnsamverkan och inverterad Faradayeffekt (Svensk sammanfattning)

Ljuset (solen) är den huvudsakliga källan av energi i vårt dagliga liv. I termer av klassisk- och kvantfysik, har ljus våg-partikeldualistisk natur som betyder att ljuset består av många partiklar (fotoner) och att flödet av fotoner bildar en elektromagnetisk våg. Varje elektron bär med sig ett enskilt energikvanta och därför har ljuset en energi som är proportionell till inversen av våglängden av ljusvågen. Fastän fotonen är en partikel utan massa, är det också känt att ljuset bär på en rörelsemängd i form av spinn och bandelar. Å andra sidan, är materia uppbyggt av atomer (kärnor) och subatomära partiklar (elektroner). Enligt Bohrmodellen, rör sig elektronerna runt atomerna i bestämda omloppsbanor och de har energi och bär på ett moment i form av spinn och bankvanttal också.

När ljuset samverkar med materian, kan energin hos ljuset absorberas av elektronerna och de exciteras till en högre energinivå, "till slut" kommer elektronerna avge energikvanta (strålning) och hoppa ner till en lägre energinivå. Det här hör samman med många olika fenomen som beror på intensiteten hos det infallande ljuset, om intensiteten är hög (t.ex. röntgenstrålning) kommer de innersta elektronerna bli exciterade och relativistiska effekter gör sig gällande. De relativistiska effekterna blir mer avgörande när spinnet hos elektronerna kopplar med frihetsgraderna från bankvanttalen - spinn-bankoppling.

I den här avhandlingens härleder vi möjliga relativistiska effekter under ljusets samverkan med materian och hur de påverkar de dynamiska processerna. För att undersöka den relativistiska ljus-materiesamverkan startar vi från en Dirac-Kohn-Sham-Hamiltonian. Eftersom den beskriver både partikel- och antipartikelteorin samt deras samverkan inom energiexcitationer med låg energi, vill vi gärna sortera ut partiklarna från antipartiklarna. Vi använder oss av en transformation, en så kallad Foldy-Wouthuysentranformation, för att koppla loss partikelvågfunktionen från antipartikelversionen. Partiklarna kan beskrivas med den härledda utökade Pauli-Hamiltonianen, vilken innehåller alla möjliga relativistiska korrigeringar av första ordningen och den mest betydelsefulla är den totala spinn-bankopplingen (gaugeinvariant och Hermitsk). De relativistiska effekterna är: masskorrigering, Darwinterm, spinn-bankoppling (gaugeinvariant och Hermitsk), högre ordningens spinn-bankoppling och relativistiska korrigeringstermer för utbytessamverkan (de här är nyheter och härledda för första gången). I stort sätt är spinn-bankopplingen av tvåtyper:

den ena är den vanliga spinn-bankopplingen av elektroner som beror på den interna potentialen, den andra uppstår när rörelsemängden hos ljuset kopplar med det elektroniska spinnet. Vi såg, intressant nog, att de här relativistiska termerna täcker upp en stor del av fysikaliska fenomen.

Den relativistiska spinn-bankopplingen (Hermitsk) ger oss en inblick i ursprunget till Gilbertdämpning i snabb Landau-Lifshitz-Gilbertmagnetiseringsdynamik; vi härleder den exakta Gilbertdämpningsparametern från Dirac-Kohn-Shamekvationen. Vi kommer fram till att Gilbertdämpningsparametern är en tensor, vilken förklarar uppkomsten av topografisk dämpning. Vi härleder också förekomsten av det nya och tidigare förbisedda fältderiverade momentet, vilket kanske kan spela en stor roll för magnetisk omkopplingsmekanism driven av en optisk puls.

Den gaugeinvarianta delen av spinn-bankopplingen förklarar det okända ursprunget till den relativistiska kopplingen mellan spinn och ljus. Den tillhörande relativistiska Hamiltonianen har visat sig vara samverkan mellan rörelsemängden hos ljuset och elektronspinnet. Dessutom visar vi också att den sistnämnda samverkan kan ses som det verkande magnetiska fältet från laserpulsen. De verkande fälten har undersökts i olika konfigureringar av pump-probeundersökningsutrustningar i ultrasnabba avmagnetiseringsexperiment. Genom att göra så, har vi visat att den här samverkan förklarar den koherenta ultrasnabba magnetismen i den femtosekundslånga avmagnetiseringen. Förutom det kan det verkande fältet från ljuset ses utöva ett moment på spinnen - det optiska spinn-banmomentet.

Högre ordningens spinn-bansamverkan härleds och de förklarar uppkomsten av spinnutation (magnetisk tröghetsdämpning) i ultrasnabb magnetismdynamik. Den tillhörande Gilbertdämpningsparametern och tröghetsparametern är tensorer och de proportionella med de respektive imaginära och reella delarna av susceptibilitetstensorn i ett harmoniskt fält.

I fortsatta studier behandlas spinnströmmar inom utökad magnetismdynamik och vi ser att ickerelativistiska spinnströmmar naturligt leder till strömin-ducerade spinnförflyttningsmomenten, medan de relativistiska spinnströmmarna (p.g.a. spinn-bansamverkan) leder till de fältinducerade momenten - spinn-banmomenten. De här spinn-banmomenten verkar väldigt lovande för att driva switchningen hos magnetismdynamiken.

Med ab-initio-beräkningar har vi beräknat storleken på laserinducerad magnetisering i flera klasser av magnetiska material med hjälp av cirkulärpolariserat ljus - den så kallade inverterade Faradayeffekten (inverse Faraday effect, IFE). Vi beräknar den inducerade magnetiseringen med frihetsgrader från både bankvanttal och spinnkvanttal, vilket konstigt nog visar på olika resultat. Vi ser att magnetiseringen inducerad av IFE är antisymmetrisk i ljusets helicitet för ickemagneter, antiferromagneter och paramagneter, men helt osymmetrisk för ferromagneter.

8. Acknowledgements

The whole thesis work would not have been possible if I had not come across remarkable and motivating persons like you!

First and foremost, I am really fortunate to have my supervisor Prof. Peter M. Oppeneer. Thank you for giving me the opportunity to work in the field where you are the expert. The help that you have provided during my 4.5 years of stay in Uppsala is invaluable and cannot be described in words. I learned a lot from you but still an apprentice. You had (have) always time for any discussions, whenever I knocked at your office even if it was late in the evening. Thank you for your valuable comments on improving this thesis. Thank you Maria for providing me the shelter for the first days in Uppsala, the super delicious food and all those amazing Christmas dinners! The first days were quite difficult for me and you both have been very helpful, whenever I needed. A mere "thank you" is not enough for you.

Marco Berritta, you are an awesome collaborator and friend, I learned a lot of physics and mathematics from you. I always appreciate your long and "yes and no" explanations about everything - thank you very much. Thank you Karel Carva for helping me with the initial code learning and fruitful collaborations. To all our "ultrafast magnetism and superconductivity" group members - Peter Oppeneer, Pablo Maldonado, Alex Aperis, Marco Berritta, Ashis K. Nandy, Ulrike Ritzmann, thank you for the usual Thursday group meetings that we have had and also the Friday afternoon meetings in the beginning of my PhD, when we spent almost 3-4 hours discussing physics. Those meetings and discussions helped me to grow my knowledge towards a better researcher. Thank you Prof. Laurent Bellaiche and Dr. Charles Paillard for bringing in fruitful collaborations. Thank you Dr. Jan Rusz for a helpful discussion during our meetings. This thesis is much improved only because of the thesis checking committee members: Tanumoy, Sudip, Ulrike, Peter thank you very much for your time and effort to make sure that this thesis is readable. Thank you Soumyajyoti for your help with the design of the cover page of this thesis.

To my teaching colleagues Jonathan, Kostas, Anna, Samara and Vivek thank you very much for making my life easier during the teaching sessions and the corrections of the reports/presentations. I enjoyed teaching with you and special thanks to Kostas (my long time teaching partner) for all the funny discussions during the spring school, the pedagogic course and all the lunches. Thank you John Wärnå for your extensive effort and keenness for the Swedish translation. To my nicest office-mates Marco and Robert, it is nice that we shared the "best office" in the world, even though we never discovered what

was inside the “safe”. To Robert, Kostas, Marco, Alekos, John, Jonathan thank you for all the bowling nights and the coffee times. Each and every members of materials theory division have been amazing, thank you for providing such a friendly environment. Thank you to our running group, Soumyajyoti-da (the inspiration), Somnath-da (the motivator and Indian elite in Sweden), Vivek (the energy booster at the final), Pralay-da (the believer), that I could think of running a half-marathon and actually crossed the finish line. To Sudip-da, a friend and philosopher in Uppsala, thank you for your motivation in science and the never-ending discussions during lunch and dinner sessions. To the Indian Bengali community - Biplab-da, Suparna-di, Sumanta-da, Soumyajyoti-da, Sudip-da, Shreemoyee-di, Swarup-da, Mala, Eheshan-da, Anjum-di, Somnath-da, Soma-di, Khusi, Pralay-da, Anik-da, Antara, Ashis-da, Tanumoy-da, Madhumita-di, Rudra-da, Kathakali, Souvik-da, Amitava, Vivek, Nishant, Rahuly (surely I am missing some names here) thank you very much for creating such a wonderful homely environment in Uppsala through all the gatherings, celebrations and of course mouth-watering authentic Indian food (how can I forget that?). Special thanks to Sumanta-da, Soumyajyoti-da, Sudip-da, Amitava, Vivek for the weekdays lunch, weekend parties, spontaneous dinner plans and the trips to amazing places.

Time to acknowledge the persons, who are not in Uppsala, but still motivated me and have been with me along my journey. Arnab Rudra, a big thank you for the great friendship we have, and hope to continue forever, our late night skype chat helped me a lot during tough times. Arnab, Sudipto, Sanjoy-da, Mouli, Ankit, Raj, at times your skype presence and the laugh made my life easier in Uppsala. To my dearest RKMV-Physics friends, I will cherish all the reunions that we had and hopefully we will continue doing it (at least once in a year), just to have the fun, laugh and leg-pulling each other.

Nothing was possible for me without the constant support and motivation of my parents (baba, maa) and my brother (dada). I am standing here, at this stage, is only because of your encouragement, care and love throughout my studies. I do not dare to thank you, I always love you.

This thesis work has been supported by the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 281043, FEMTOSPIN, the Swedish Research Council (VR), the Knut and Alice Wallenberg Foundation (Contract No. 2015.0060), the European Union’s Horizon2020 Research and Innovation Programme under grant agreement No. 737709 (FEMTOTERABYTE, http://www.physics.gu.se/femtoterabyte), and the Swedish National Infrastructure for Computing (SNIC).

References

[1] "IBM 350 disk storage unit," `https://www-03.ibm.com/ibm/history/exhibits/storage/storage_350.html`.

[2] G. E. Moore, Electronics **38**, 8 (1965).

[3] C. H. Back, D. Weller, J. Heidmann, D. Mauri, D. Guarisco, E. L. Garwin, and H. C. Siegmann, Phys. Rev. Lett. **81**, 3251 (1998).

[4] H. C. Siegmann, E. L. Garwin, C. Y. Prescott, J. Heidmann, D. Mauri, D. Weller, R. Allenspach, and W. Weber, J. Magn. Magn. Mater. **151**, L8 (1995).

[5] T. Gerrits, H. A. M. van den Berg, J. Hohlfeld, L. Bar, and Th. Rasing, Nature **418**, 509 (2002).

[6] T. Gerrits, H. A. M. van den Berg, J. Hohlfeld, O. Gielkens, K. J. Veenstra, L. Bar, and Th. Rasing, IEEE Trans. Magn. **38**, 2484 (2002).

[7] B. Hillebrands and K. Ounadjela, *Spin Dynamics in Confined Magnetic Structures II*, Physics and Astronomy Online Library (Springer, 2003).

[8] H. W. Schumacher, C. Chappert, P. Crozat, R. C. Sousa, P. P. Freitas, J. Miltat, J. Fassbender, and B. Hillebrands, Phys. Rev. Lett. **90**, 017201 (2003).

[9] I. Tudosa, C. Stamm, A. B. Kashuba, F. King, H. C. Siegmann, J. Stöhr, G. Ju, B. Lu, and D. Weller, Nature **428**, 831 (2004).

[10] G. Ndabashimiye, S. Ghimire, M. Wu, D. A. Browne, K. J. Schafer, M. B. Gaarde, and D. A. Reis, Nature **534**, 520 (2016).

[11] E. Beaurepaire, J.-C. Merle, A. Daunois, and J.-Y. Bigot, Phys. Rev. Lett. **76**, 4250 (1996).

[12] J. Hohlfeld, E. Matthias, R. Knorren, and K. H. Bennemann, Phys. Rev. Lett. **78**, 4861 (1997).

[13] A. Scholl, L. Baumgarten, R. Jacquemin, and W. Eberhardt, Phys. Rev. Lett. **79**, 5146 (1997).

[14] G. Ju, A. V. Nurmikko, R. F. C. Farrow, R. F. Marks, M. J. Carey, and B. A. Gurney, Phys. Rev. Lett. **82**, 3705 (1999).

[15] B. Koopmans, M. van Kampen, J. T. Kohlhepp, and W. J. M. de Jonge, Phys. Rev. Lett. **85**, 844 (2000).

[16] G. P. Zhang and W. Hübner, Phys. Rev. Lett. **85**, 3025 (2000).

[17] H.-S. Rhie, H. A. Dürr, and W. Eberhardt, Phys. Rev. Lett. **90**, 247201 (2003).

[18] P. M. Oppeneer and A. Liebsch, J. Phys.: Condens. Matter **16**, 5519 (2004).

[19] T. Ogasawara, K. Ohgushi, Y. Tomioka, K. S. Takahashi, H. Okamoto, M. Kawasaki, and Y. Tokura, Phys. Rev. Lett. **94**, 087202 (2005).

[20] G. Zhang, W. Hübner, E. Beaurepaire, and J.-Y. Bigot, "Laser-induced ultrafast demagnetization: Femtomagnetism, a new frontier?" in *Spin Dynamics in Confined Magnetic Structures I*, edited by B. Hillebrands and K. Ounadjela (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002) pp. 245–289.

[21] J. Stöhr and H. C. Siegmann, "Ultrafast magnetization dynamics," in *Magnetism: From Fundamentals to Nanoscale Dynamics* (Springer-Verlag Berlin Heidelberg, 2006).
[22] J.-Y. Bigot, M. Vomir, and E. Beaurepaire, Nat. Phys. **5**, 515 (2009).
[23] B. Koopmans, G. Malinowski, F. Dalla Longa, D. Steiauf, M. Fähnle, T. Roth, M. Cinchetti, and M. Aeschlimann, Nat. Mater. **9**, 259 (2010).
[24] K. Carva, M. Battiato, and P. M. Oppeneer, Phys. Rev. Lett. **107**, 207201 (2011).
[25] K. Carva, M. Battiato, D. Legut, and P. M. Oppeneer, Phys. Rev. B **87**, 184425 (2013).
[26] M. Battiato, K. Carva, and P. M. Oppeneer, Phys. Rev. Lett. **105**, 027203 (2010).
[27] J. Hohlfeld, T. Gerrits, M. Bilderbeek, Th. Rasing, H. Awano, and N. Ohta, Phys. Rev. B **65**, 012413 (2001).
[28] A. V. Kimel, A. Kirilyuk, P. A. Usachev, R. V. Pisarev, A. M. Balbashov, and Th. Rasing, Nature **435**, 655 (2005).
[29] C. D. Stanciu, F. Hansteen, A. V. Kimel, A. Kirilyuk, A. Tsukamoto, A. Itoh, and Th. Rasing, Phys. Rev. Lett. **99**, 047601 (2007).
[30] K. Vahaplar, A. M. Kalashnikova, A. V. Kimel, S. Gerlach, D. Hinzke, U. Nowak, R. Chantrell, A. Tsukamoto, A. Itoh, A. Kirilyuk, and Th. Rasing, Phys. Rev. B **85**, 104402 (2012).
[31] S. Mangin, M. Gottwald, C.-H. Lambert, D. Steil, V. Uhlíř, L. Pang, M. Hehn, S. Alebrand, M. Cinchetti, G. Malinowski, Y. Fainman, M. Aeschlimann, and E. E. Fullerton, Nat. Mater. **13**, 286 (2014).
[32] C.-H. Lambert, S. Mangin, B. S. D. C. S. Varaprasad, Y. K. Takahashi, M. Hehn, M. Cinchetti, G. Malinowski, K. Hono, Y. Fainman, M. Aeschlimann, and E. E. Fullerton, Science **345**, 1403 (2014).
[33] A. Kirilyuk, A. V. Kimel, and Th. Rasing, Rev. Mod. Phys. **82**, 2731 (2010).
[34] W. Brown, *Micromagnetics*, Interscience tracts on physics and astronomy (Interscience Publishers, 1963).
[35] O. Chubykalo-Fesenko, U. Nowak, R. W. Chantrell, and D. Garanin, Phys. Rev. B **74**, 094436 (2006).
[36] U. Atxitia, O. Chubykalo-Fesenko, R. W. Chantrell, U. Nowak, and A. Rebei, Phys. Rev. Lett. **102**, 057203 (2009).
[37] L. D. Landau and E. M. Lifshitz, Phys. Z. Sowjetunion **8**, 101 (1935).
[38] T. L. Gilbert, Ph.D. thesis, Illinois Institute of Technology, Chicago, 1956.
[39] T. L. Gilbert and J. M. Kelly, in *American Institute of Electrical Engineers* (New York, October 1955) pp. 253–263.
[40] T. L. Gilbert, IEEE Trans. Magn. **40**, 3443 (2004).
[41] A. V. Kimel, B. A. Ivanov, R. V. Pisarev, P. A. Usachev, A. Kirilyuk, and Th. Rasing, Nat. Phys. **5**, 727 (2009).
[42] M.-C. Ciornei, J. M. Rubí, and J.-E. Wegrowe, Phys. Rev. B **83**, 020410 (2011).
[43] J. C. Slonczewski, J. Magn. Magn. Mater. **159**, L1 (1996).
[44] L. Berger, Phys. Rev. B **54**, 9353 (1996).
[45] D. C. Ralph and M. D. Stiles, J. Magn. Magn. Mater. **320**, 1190 (2008).
[46] C. Chappert, A. Fert, and F. N. Van Dau, Nat. Mater. **6**, 813 (2007).

[47] J. Åkerman, Science **308**, 508 (2005).
[48] S. S. P. Parkin, C. Kaiser, A. Panchula, P. M. Rice, B. Hughes, M. Samant, and S.-H. Yang, Nat. Mater. **3**, 862 (2004).
[49] Z. Diao, Z. Li, S. Wang, Y. Ding, A. Panchula, E. Chen, L.-C. Wang, and Y. Huai, J. Phys.: Condens. Matter **19**, 165209 (2007).
[50] L. Liu, C.-F. Pai, Y. Li, H. W. Tseng, D. C. Ralph, and R. A. Buhrman, Science **336**, 555 (2012).
[51] M. Gajek, J. J. Nowak, J. Z. Sun, P. L. Trouilloud, E. J. O'Sullivan, D. W. Abraham, M. C. Gaidis, G. Hu, S. Brown, Y. Zhu, R. P. Robertazzi, W. J. Gallagher, and D. C. Worledge, Appl. Phys. Lett. **100**, 132408 (2012).
[52] P. Gambardella and I. M. Miron, Phil. Trans. Roy. Soc. London A **369**, 3175 (2011).
[53] C. Ciccarelli, L. Anderson, V. Tshitoyan, A. J. Ferguson, F. Gerhard, C. Gould, L. W. Molenkamp, J. Gayles, J. Železný, L. Šmejkal, Z. Yuan, J. Sinova, F. Freimuth, and T. Jungwirth, Nat. Phys. **12**, 855 (2016).
[54] H. Kurebayashi, J. Sinova, D. Fang, A. C. Irvine, T. D. Skinner, J. Wunderlich, V. Novák, R. P. Campion, B. L. Gallagher, E. K. Vehstedt, L. P. Zârbo, K. Výborný, A. J. Ferguson, and T. Jungwirth, Nat. Nano. **9**, 211 (2014).
[55] Z. Li and S. Zhang, Phys. Rev. B **68**, 024404 (2003).
[56] K. Xia, P. J. Kelly, G. E. W. Bauer, A. Brataas, and I. Turek, Phys. Rev. B **65**, 220401 (2002).
[57] N. Locatelli, V. Cros, and J. Grollier, Nat. Mater. **13**, 11 (2014).
[58] A. N. Slavin and I. N. Krivorotov, *Spin-torque devices*, US patent (2010).
[59] A. J. Schellekens and B. Koopmans, Phys. Rev. Lett. **110**, 217204 (2013).
[60] R. R. J. C. de Wit, Ph.D. thesis, Technische Universiteit Eindhoven (TUE), Eindhoven, 2013.
[61] Y. Fan, P. Upadhyaya, X. Kou, M. Lang, S. Takei, Z. Wang, J. Tang, L. He, L.-T. Chang, M. Montazeri, G. Yu, W. Jiang, T. Nie, R. N. Schwartz, Y. Tserkovnyak, and K. L. Wang, Nat. Mater. **13**, 699 (2014).
[62] A. Raeliarijaona, S. Singh, H. Fu, and L. Bellaiche, Phys. Rev. Lett. **110**, 137205 (2013).
[63] M. Baum, K. Schmalzl, P. Steffens, A. Hiess, L. P. Regnault, M. Meven, P. Becker, L. Bohatý, and M. Braden, Phys. Rev. B **88**, 024414 (2013).
[64] L. Bellaiche, W. Ren, and S. Singh, Phys. Rev. B **88**, 161102 (2013).
[65] R. Walter, M. Viret, S. Singh, and L. Bellaiche, J. Phys.: Condens. Matter **26**, 432201 (2014).
[66] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[67] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[68] G. Vignale and M. Rasolt, Phys. Rev. Lett. **59**, 2360 (1987).
[69] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[70] P. Strange, *Relativistic Quantum Mechanics: With Applications in Condensed Matter and Atomic Physics* (Cambridge University Press, 1998).
[71] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[72] E. Schrödinger, Ann. Physik **384**, 489 (1926).
[73] E. Schrödinger, Ann. Physik **81**, 109 (1926).
[74] W. Gordon, Z. Physik **40**, 117 (1926).

[75] O. Klein, Z. Physik **41**, 407 (1927).
[76] C. Møller, *The Theory of Relativity*, 2nd ed. (Oxford University Press, 1972).
[77] S. Weinberg, *Gravitation and cosmology: principles and applications of the general theory of relativity* (Wiley, 1972).
[78] J. D. Jackson, *Classical electrodynamics* (John Wiley & Sons, Inc., 1975).
[79] J. J. Sakurai, *Advanced Quantum Mechanics*, Addison-Wesley series in advanced physics (Pearson Education, 1967).
[80] P. A. M. Dirac, Proc. Roy. Soc. (London) **A117**, 610 (1928).
[81] P. A. M. Dirac, Proc. Roy. Soc. (London) **A126**, 360 (1930).
[82] P. A. M. Dirac, Proc. Roy. Soc. (London) **A118**, 351 (1928).
[83] J. D. Bjorken and S. D. Drell, *Relativistic quantum mechanics*, International series in pure and applied physics (McGraw-Hill, 1964).
[84] M. Reiher and A. Wolf, *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science* (Wiley, 2009).
[85] E. de Vries and J. Jonker, Nucl. Phys. B **6**, 213 (1968).
[86] T. Kraft, P. M. Oppeneer, V. N. Antonov, and H. Eschrig, Phys. Rev. B **52**, 3561 (1995).
[87] K. Y. Bliokh, Europhys. Lett. **72**, 7 (2005).
[88] P. Gosselin, A. Bérard, and H. Mohrbach, Euro. Phys. J. B **58**, 137 (2007).
[89] P. Gosselin, A. Bérard, H. Mohrbach, and S. Ghosh, Euro. Phys. J. C **59**, 883 (2009).
[90] Y. Hinschberger and P.-A. Hervieux, Phys. Lett. A **376**, 813 (2012).
[91] R. Jagannathan, Phys. Rev. A **42**, 6674 (1990).
[92] S. A. Khan, Optik - International Journal for Light and Electron Optics **117**, 481 (2006).
[93] F. A. Reuse, *Electrodynamique et Optique Quantiques*, Ph.D. thesis, Presses Polytechniques et Universitaires Romandes, Lausanne (2007).
[94] V. P. Neznamov, Phys. Part. Nucl. **37**, 86 (2006).
[95] A. J. Silenko, J. Math. Phys. **44**, 2952 (2003).
[96] A. J. Silenko and O. V. Teryaev, Phys. Rev. D **71**, 064016 (2005).
[97] W. Greiner, *Relativistic quantum mechanics. Wave equations* (Springer, Berlin, 2000).
[98] F. Schwabl, R. Hilton, and A. Lahee, *Advanced Quantum Mechanics*, Advanced Texts in Physics Series (Springer Berlin Heidelberg, 2008).
[99] M.-F. Li, T. Ariizumi, and S. Suzuki, J. Phys. Soc. Jpn. **76**, 054702 (2007).
[100] A. J. Silenko, Phys. Rev. A **93**, 022108 (2016).
[101] E. Eriksen, Phys. Rev. **111**, 1011 (1958).
[102] A. G. Nikitin, J. Phys. A: Math. Gen. **31**, 3297 (1998).
[103] G. Murguía and A. Raya, J. Phys. A: Math. Theor. **43**, 402005 (2010).
[104] L. L. Foldy, Phys. Rev. **87**, 688 (1952).
[105] O. Klein, Z. Physik **53**, 157 (1929).
[106] F. Hausdorff, Ber. Verh. Kgl. Sächs. Ges. Wiss. Leipzig. Math.-phys. Kl. **58**, 106 (1906).
[107] J. Campbell, Proc. Lond. Math. Soc. **28**, 381 (1897).
[108] H. Baker, Proc. Lond. Math. Soc. **34**, 347 (1902).
[109] A. J. Silenko, Phys. Rev. A **77**, 012116 (2008).
[110] A. Y. Silenko, Theor. Math. Phys. **176**, 987 (2013).

[111] A. J. Silenko, Phys. Rev. A **91**, 022103 (2015).
[112] H. Bauke, S. Ahrens, and R. Grobe, Phys. Rev. A **90**, 052101 (2014).
[113] H. Bauke, S. Ahrens, C. H. Keitel, and R. Grobe, New J. Phys. **16**, 103028 (2014).
[114] K. M. Case, Phys. Rev. **95**, 1323 (1954).
[115] D. L. Pursey, Nucl. Phys. **53**, 174 (1964).
[116] R. F. Guertin, Ann. Phys. **91**, 386 (1975).
[117] A. J. Silenko, Phys. Rev. A **94**, 032104 (2016).
[118] E. Eriksen and M. Korlsrud, Nuovo Cimento Suppl. **18**, 1 (1960).
[119] A. H. MacDonald and S. H. Vosko, J. Phys. C: Solid State Phys. **12**, 2977 (1979).
[120] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer-Verlag, Berlin and Heidelberg, 1990).
[121] A. Crépieux and P. Bruno, Phys. Rev. B **64**, 094434 (2001).
[122] H. Eschrig and V. D. P. Servedio, J. Comput. Chem. **20**, 23 (1999).
[123] R. Mondal, M. Berritta, K. Carva, and P. M. Oppeneer, Phys. Rev. B **91**, 174415 (2015).
[124] R. Mondal, M. Berritta, and P. M. Oppeneer, Phys. Rev. B **94**, 144419 (2016).
[125] S. Blundell, *Magnetism in Condensed Matter* (Oxford University Press, Oxford, 2001).
[126] C. G. Darwin, Proc. Roy. Soc. (London) **A118**, 654 (1928).
[127] C. Paillard, R. Mondal, M. Berritta, B. Dkhil, S. Singh, P. M. Oppeneer, and L. Bellaiche, Proc. SPIE **9931**, 99312E (2016).
[128] R. F. O'Connell, J. Phys. A: Math. Theor. **50**, 085306 (2017).
[129] D. C. Mattis, "History of magnetism," in *The Theory of Magnetism I: Statics and Dynamics* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1981) pp. 1–38.
[130] B. Mahon, Nat. Photon. **9**, 2 (2015).
[131] R. M. White, *Quantum Theory of Magnetism: Magnetic Properties of Materials* (Springer, Heidelberg, 2007).
[132] V. G. Bar'yakhtar, Sov. Phys. JETP **60**, 863 (1984).
[133] M. Lakshmanan, Phil. Trans. Roy. Soc. London A **369**, 1280 (2011).
[134] D. Böttcher and J. Henk, Phys. Rev. B **86**, 020404 (2012).
[135] M.-C. Ciornei, Ph.D. thesis, Ecole Polytechnique, Universidad de Barcelona, 2010.
[136] S. Zhang and Z. Li, Phys. Rev. Lett. **93**, 127204 (2004).
[137] L. D. Landau and E. M. Lifshitz, *Mechanics*, Volume 1 of A Course of Theoretical Physics (Pergamon Press, 1969).
[138] V. G. Bar'yakhtar, in *Frontiers in Magnetism of Reduced Dimension Systems* (Springer, 1998) pp. 63–94.
[139] V. G. Bar'yakhtar and A. G. Danilevich, Low Temp. Phys. **39**, 993 (2013).
[140] F. Bloch, Phys. Rev. **70**, 460 (1946).
[141] D. A. Garanin, Phys. Rev. B **55**, 3050 (1997).
[142] K. Gilmore, Y. U. Idzerda, and M. D. Stiles, Phys. Rev. Lett. **99**, 027204 (2007).
[143] K. Gilmore, Ph.D. thesis, Montana State University, 2007.

[144] K. Gilmore, Y. U. Idzerda, and M. D. Stiles, J. Appl. Phys. **103**, 07D303 (2008).
[145] I. Garate and A. H. MacDonald, Phys. Rev. B **79**, 064403 (2009).
[146] Y. Tserkovnyak, G. A. Fiete, and B. I. Halperin, Appl. Phys. Lett. **84**, 5234 (2004).
[147] E. M. Hankiewicz, G. Vignale, and Y. Tserkovnyak, Phys. Rev. B **78**, 020404 (2008).
[148] A. Sakuma, J. Appl. Phys. **117**, 013912 (2015).
[149] M. Fähnle and C. Illg, J. Phys.: Condens. Matter **23**, 493201 (2011).
[150] S. Bhattacharjee, L. Nordström, and J. Fransson, Phys. Rev. Lett. **108**, 057204 (2012).
[151] D. Thonig, O. Eriksson, and M. Pereiro, Sci. Rep. **7**, 931 (2017).
[152] D. Steiauf and M. Fähnle, Phys. Rev. B **72**, 064450 (2005).
[153] K. Gilmore, M. D. Stiles, J. Seib, D. Steiauf, and M. Fähnle, Phys. Rev. B **81**, 174414 (2010).
[154] H. Ebert, S. Mankovsky, D. Ködderitzsch, and P. J. Kelly, Phys. Rev. Lett. **107**, 066603 (2011).
[155] V. Kamberský, Czech. J. Phys. B **26**, 1366 (1976).
[156] V. Kamberský, Can. J. Phys. **48**, 2906 (1970).
[157] V. Korenman and R. E. Prange, Phys. Rev. B **6**, 2769 (1972).
[158] J. Kuneš and V. Kamberský, Phys. Rev. B **65**, 212411 (2002).
[159] V. Kamberský, Phys. Rev. B **76**, 134416 (2007).
[160] I. Garate and A. H. MacDonald, Phys. Rev. B **79**, 064404 (2009).
[161] M. A. W. Schoen, D. Thonig, M. L. Schneider, T. J. Silva, H. T. Nembach, O. Eriksson, O. Karis, and J. M. Shaw, Nat. Phys. **12**, 839 (2016).
[162] S. M. Barnett, J. Mod. Opt. **57**, 1339 (2010).
[163] R. Mondal, M. Berritta, C. Paillard, S. Singh, B. Dkhil, P. M. Oppeneer, and L. Bellaiche, Phys. Rev. B **92**, 100402(R) (2015).
[164] R. Mondal, M. Berritta, and P. M. Oppeneer, J. Phys.: Condens. Matter **29**, 194002 (2017).
[165] N. Tesarová, P. Nemec, E. Rozkotová, J. Zemen, T. Janda, D. Butkovicová, F. Trojánek, K. Olejnik, V. Novák, P. Maly, and T. Jungwirth, Nat. Photon. **7**, 492 (2013).
[166] G.-M. Choi, A. Schleife, and D. G. Cahill, Nat. Comm. **8**, 15085 (2017).
[167] K. Wago, D. Botkin, C. S. Yannoni, and D. Rugar, Phys. Rev. B **57**, 1108 (1998).
[168] M. Fähnle, D. Steiauf, and C. Illg, Phys. Rev. B **84**, 172403 (2011).
[169] M. Fähnle, D. Steiauf, and C. Illg, Phys. Rev. B **88**, 219905(E) (2013).
[170] E. Olive, Y. Lansac, M. Meyer, M. Hayoun, and J.-E. Wegrowe, J. Appl. Phys. **117**, 213904 (2015).
[171] J.-E. Wegrowe and M.-C. Ciornei, Am. J. Phys. **80**, 607 (2012).
[172] J.-E. Wegrowe and E. Olive, J. Phys.: Condens. Matter **28**, 106001 (2016).
[173] M. V. Berry and P. Shukla, Eur. J. Phys. **32**, 115 (2011).
[174] Y. Li, A.-L. Barra, S. Auffret, U. Ebels, and W. E. Bailey, Phys. Rev. B **92**, 140413 (2015).
[175] R. Mondal, M. Berritta, A. K. Nandy, and P. M. Oppeneer, Phys. Rev. B **96**, 024425 (2017).

[176] Y. Tserkovnyak, A. Brataas, and G. E. W. Bauer, Phys. Rev. Lett. **88**, 117601 (2002).
[177] A. Kapelrud and A. Brataas, Phys. Rev. Lett. **111**, 097602 (2013).
[178] T. Taniguchi and H. Imamura, Phys. Rev. B **76**, 092402 (2007).
[179] E. Barati, M. Cinal, D. M. Edwards, and A. Umerski, Phys. Rev. B **90**, 014420 (2014).
[180] E. van der Bijl and R. A. Duine, Phys. Rev. B **86**, 094406 (2012).
[181] K. Nomura and D. Kurebayashi, Phys. Rev. Lett. **115**, 127201 (2015).
[182] K.-W. Kim, K.-J. Lee, H.-W. Lee, and M. D. Stiles, Phys. Rev. B **92**, 224426 (2015).
[183] C. Stamm, T. Kachel, N. Pontius, R. Mitzner, T. Quast, K. Holldack, S. Khan, C. Lupulescu, E. F. Aziz, M. Wietstruk, H. A. Dürr, and W. Eberhardt, Nat. Mater. **6**, 740 (2007).
[184] A. V. Kimel, A. Kirilyuk, and Th. Rasing, Laser Photon. Rev. **1**, 275 (2007).
[185] E. Carpene, E. Mancini, C. Dallera, M. Brenna, E. Puppin, and S. De Silvestri, Phys. Rev. B **78** (2008).
[186] A. Weber, F. Pressacco, S. Günther, E. Mancini, P. M. Oppeneer, and C. H. Back, Phys. Rev. B **84**, 132412 (2011).
[187] A. Eschenlohr, M. Battiato, P. Maldonado, N. Pontius, T. Kachel, K. Holldack, R. Mitzner, A. Föhlisch, P. M. Oppeneer, and C. Stamm, Nat. Mater. **12**, 332 (2013).
[188] M. Cinchetti, M. Sánchez Albaneda, D. Hoffmann, T. Roth, J.-P. Wüstenberg, M. Krauß, O. Andreyev, H. C. Schneider, M. Bauer, and M. Aeschlimann, Phys. Rev. Lett. **97**, 177201 (2006).
[189] T. A. Ostler, J. Barker, R. F. L. Evans, R. W. Chantrell, U. Atxitia, O. Chubykalo-Fesenko, S. El Moussaoui, L. Le Guyader, E. Mengotti, L. J. Heyderman, F. Nolting, A. Tsukamoto, A. Itoh, D. Afanasiev, B. A. Ivanov, A. M. Kalashnikova, K. Vahaplar, J. Mentink, A. Kirilyuk, Th. Rasing, and A. V. Kimel, Nat. Comm. **3**, 666 (2012).
[190] C. E. Graves, A. H. Reid, T. Wang, B. Wu, S. de Jong, K. Vahaplar, I. Radu, D. P. Bernstein, M. Messerschmidt, L. Müller, R. Coffee, M. Bionta, S. W. Epp, R. Hartmann, N. Kimmel, G. Hauser, A. Hartmann, P. Holl, H. Gorke, J. H. Mentink, A. Tsukamoto, A. Fognini, J. J. Turner, W. F. Schlotter, D. Rolles, H. Soltau, L. Strüder, Y. Acremann, A. V. Kimel, A. Kirilyuk, Th. Rasing, J. Stöhr, A. O. Scherz, and H. A. Dürr, Nat. Mater. **12**, 293 (2013).
[191] M. Fiebig, N. P. Duong, T. Satoh, B. B. V. Aken, K. Miyano, Y. Tomioka, and Y. Tokura, J. Phys. D: Appl. Phys. **41**, 164005 (2008).
[192] F. Boschini, M. Mansurova, G. Mussler, J. Kampmeier, D. Grützmacher, L. Braun, F. Katmis, J. S. Moodera, C. Dallera, E. Carpene, C. Franz, M. Czerner, C. Heiliger, T. Kampfrath, and M. Münzenberg, Sci. Rep. **5**, 15304 (2015).
[193] G. P. Zhang, W. Hubner, G. Lefkidis, Y. Bai, and T. F. George, Nat. Phys. **5**, 499 (2009).
[194] B. Koopmans, M. van Kampen, and W. J. M. de Jonge, J. Phys.: Condens. Matter **15**, S723 (2003).
[195] C. Illg, M. Haag, and M. Fähnle, Phys. Rev. B **88**, 214404 (2013).

[196] M. Krauß, T. Roth, S. Alebrand, D. Steil, M. Cinchetti, M. Aeschlimann, and H. C. Schneider, Phys. Rev. B **80**, 180407(R) (2009).
[197] D. Steil, S. Alebrand, T. Roth, M. Krauß, T. Kubota, M. Oogane, Y. Ando, H. C. Schneider, M. Aeschlimann, and M. Cinchetti, Phys. Rev. Lett. **105**, 217202 (2010).
[198] B. Y. Mueller, T. Roth, M. Cinchetti, M. Aeschlimann, and B. Rethfeld, New J. Phys. **13**, 123010 (2011).
[199] R. J. Elliott, Phys. Rev. **96**, 266 (1954).
[200] Y. Yafet, in *Solid State Physics*, Vol. 14, edited by F. Seitz and D. Turnbull (Academic, New York, 1963) pp. 1–98.
[201] M. Fähnle, M. Haag, and C. Illg, J. Magn. Magn. Mater. **347**, 45 (2013).
[202] E. Beaurepaire, G. M. Turner, S. M. Harrel, M. C. Beard, J.-Y. Bigot, and C. A. Schmuttenmaer, Appl. Phys. Lett. **84**, 3465 (2004).
[203] T. J. Huisman, R. V. Mikhaylovskiy, A. Tsukamoto, Th. Rasing, and A. V. Kimel, Phys. Rev. B **92**, 104419 (2015).
[204] M. Battiato, K. Carva, and P. M. Oppeneer, Phys. Rev. B **86**, 024404 (2012).
[205] T. Kampfrath, M. Battiato, P. Maldonado, G. Eilers, J. Notzold, S. Mahrlein, V. Zbarsky, F. Freimuth, Y. Mokrousov, S. Blügel, M. Wolf, I. Radu, P. M. Oppeneer, and M. Münzenberg, Nat. Nano. **8**, 256 (2013).
[206] D. Rudolf, C. La-O-Vorakiat, M. Battiato, R. Adam, J. M. Shaw, E. Turgut, P. Maldonado, S. Mathias, P. Grychtol, H. T. Nembach, T. J. Silva, M. Aeschlimann, H. C. Kapteyn, M. M. Murnane, C. M. Schneider, and P. M. Oppeneer, Nat. Comm. **3**, 1037 (2012).
[207] A. Melnikov, I. Razdolski, T. O. Wehling, E. T. Papaioannou, V. Roddatis, P. Fumagalli, O. Aktsipetrov, A. I. Lichtenstein, and U. Bovensiepen, Phys. Rev. Lett. **107**, 076601 (2011).
[208] E. Turgut, C. La-O-Vorakiat, J. M. Shaw, P. Grychtol, H. T. Nembach, D. Rudolf, R. Adam, M. Aeschlimann, C. M. Schneider, T. J. Silva, M. M. Murnane, H. C. Kapteyn, and S. Mathias, Phys. Rev. Lett. **110**, 197201 (2013).
[209] A. J. Schellekens, W. Verhoeven, T. N. Vader, and B. Koopmans, Appl. Phys. Lett. **102**, 252408 (2013).
[210] K. Krieger, J. K. Dewhurst, P. Elliott, S. Sharma, and E. K. U. Gross, J. Chem. Theory Comput. **11**, 4870 (2015).
[211] P. Elliott, K. Krieger, J. K. Dewhurst, S. Sharma, and E. K. U. Gross, New J. Phys. **18**, 013014 (2016).
[212] K. Krieger, P. Elliott, T. Müller, N. Singh, J. K. Dewhurst, E. K. U. Gross, and S. Sharma, J. Phys.: Condens. Matter **29**, 224001 (2017).
[213] H. Vonesch and J.-Y. Bigot, Phys. Rev. B **85**, 180407 (2012).
[214] J.-Y. Bigot and M. Vomir, Ann. Phys. **525**, 2 (2013).
[215] J.-Y. Bigot, W. Hübner, Th. Rasing, and R. Chantrell, *Ultrafast Magnetism I: Proceedings of the International Conference UMC 2013 Strasbourg, France, October 28th - November 1st, 2013*, Springer Proceedings in Physics (Springer International Publishing, 2014).
[216] P. N. Argyres, Phys. Rev. **97**, 334 (1955).
[217] P. M. Oppeneer, J. Sticht, T. Maurer, and J. Kübler, Z. Physik B **88**, 309 (1992).

[218] V. N. Antonov, B. N. Harmon, A. N. Yaresko, and A. Y. Perlov, Phys. Rev. B **59**, 14571 (1999).
[219] H. Ebert, Physica B: Condens. Matter **161**, 175 (1990).
[220] P. M. Oppeneer, T. Maurer, J. Sticht, and J. Kübler, Phys. Rev. B **45**, 10924 (1992).
[221] P. M. Oppeneer, V. N. Antonov, T. Kraft, H. Eschrig, A. N. Yaresko, and A. Y. Perlov, Solid State Commun. **94**, 255 (1995).
[222] I. Osterloh, P. M. Oppeneer, J. Sticht, and J. Kubler, J. Phys.: Condens. Matter **6**, 285 (1994).
[223] M.-F. Li, T. Ariizumi, and S. Suzuki, J. Phys. Soc. Jpn. **76**, 054702 (2007).
[224] P. M. Oppeneer, in *Handbook of Magnetic Materials*, Vol. 13, edited by K. H. J. Buschow (Elsevier, Amsterdam, 2001) pp. 229–422.
[225] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
[226] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed., Vol. 3 (Butterworth-Heinemann, Oxford, 1981).
[227] A. R. Williams, J. Kübler, and C. D. Gelatt, Phys. Rev. B **19**, 6094 (1979).
[228] Y. Hinschberger and P.-A. Hervieux, Phys. Lett. A **379**, 2261 (2015).
[229] A. Dixit, Y. Hinschberger, J. Zamanian, G. Manfredi, and P.-A. Hervieux, Phys. Rev. A **88**, 032117 (2013).
[230] Y. Hinschberger and P.-A. Hervieux, Phys. Rev. B **88**, 134413 (2013).
[231] Y. Hinschberger, A. Dixit, G. Manfredi, and P.-A. Hervieux, Phys. Rev. A **91**, 012101 (2015).
[232] Y. Hinschberger, G. Manfredi, and P.-A. Hervieux, Phys. Rev. A **93**, 042117 (2016).
[233] Y. Hinschberger and P.-A. Hervieux, J. Appl. Phys. **118**, 243902 (2015).
[234] G. Gubbiotti, L. Albini, G. Carlotti, M. G. Pini, P. Politi, A. Rettori, P. Vavassori, M. Ciria, K. Ha, and R. C. O'Handley, J. Appl. Phys. **89**, 7386 (2001).
[235] M. Faraday, T. Martin, and Roy. Inst. of Great Britain, *Faraday's diary: being the various philosophical notes of experimental investigation made by Michael Faraday, during the years 1820-1862 and bequeathed by him to the Royal Institution of Great Britain* (G. Bell and Sons. Ltd., 1936).
[236] L. P. Pitaevskii, Sov. Phys. JETP **12**, 1008 (1961).
[237] S. Ali, J. R. Davies, and J. T. Mendonca, Phys. Rev. Lett. **105**, 035001 (2010).
[238] J. P. van der Ziel, P. S. Pershan, and L. D. Malmstrom, Phys. Rev. Lett. **15**, 190 (1965).
[239] P. S. Pershan, J. P. van der Ziel, and L. D. Malmstrom, Phys. Rev. **143**, 574 (1966).
[240] J. Deschamps, M. Fitaire, and M. Lagoutte, Phys. Rev. Lett. **25**, 1330 (1970).
[241] Y. Horovitz, S. Eliezer, A. Ludmirsky, Z. Henis, E. Moshe, R. Shpitalnik, and B. Arad, Phys. Rev. Lett. **78**, 1707 (1997).
[242] Y. Horovitz, S. Eliezer, Z. Henis, Y. Paiss, E. Moshe, A. Ludmirsky, M. Werdiger, B. Arad, and A. Zigler, Phys. Lett. A **246**, 329 (1998).
[243] Z. Najmudin, M. Tatarakis, A. Pukhov, E. L. Clark, R. J. Clarke, A. E. Dangor, J. Faure, V. Malka, D. Neely, M. I. K. Santala, and K. Krushelnick, Phys. Rev. Lett. **87**, 215004 (2001).
[244] R. Hertel, J. Mag. Magn. Mater. **303**, L1 (2006).

[245] H.-L. Zhang, Y.-Z. Wang, and X.-J. Chen, J. Mag. Magn. Mater. **321**, L73 (2009).
[246] D. Popova, A. Bringer, and S. Blügel, Phys. Rev. B **84**, 214421 (2011).
[247] S. R. Woodford, Phys. Rev. B **79**, 212412 (2009).
[248] R. John, M. Berritta, D. Hinzke, C. Müller, T. Santos, H. Ulrichs, P. Nieves, J. Walowski, R. Mondal, O. Chubykalo-Fesenko, J. McCord, P. M. Oppeneer, U. Nowak, and M. Münzenberg, Sci. Rep. **7**, 4114 (2017).
[249] N. W. Ashcroft and N. D. Mermin, *Solid state physics*, Science: Physics (Saunders College, 1976).
[250] M. Battiato, G. Barbalinardo, and P. M. Oppeneer, Phys. Rev. B **89**, 014413 (2014).
[251] L. D. Landau, L. P. Pitaevskii, and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed. (Elsevier, Oxford, 1984).
[252] G. Barbalinardo, *Quantum theory of the inverse Faraday effect for ultrafast magneto-optics*, Masters thesis, Uppsala University, Uppsala (2011).
[253] F. Hansteen, A. V. Kimel, A. Kirilyuk, and Th. Rasing, Phys. Rev. B **73**, 014421 (2006).
[254] D. Popova, A. Bringer, and S. Blügel, Phys. Rev. B **85**, 094419 (2012).
[255] M. Berritta, R. Mondal, K. Carva, and P. M. Oppeneer, Phys. Rev. Lett. **117**, 137203 (2016).
[256] A. R. Khorsand, M. Savoini, A. Kirilyuk, A. V. Kimel, A. Tsukamoto, A. Itoh, and Th. Rasing, Phys. Rev. Lett. **108**, 127205 (2012).
[257] N. A. Spaldin, M. Fiebig, and M. Mostovoy, J. Phys.: Condens. Matter **20**, 434203 (2008).
[258] B. B. Van Aken, J.-P. Rivera, H. Schmid, and M. Fiebig, Nature **449**, 702 (2007).
[259] N. Papasimakis, V. A. Fedotov, V. Savinov, T. A. Raybould, and N. I. Zheludev, Nat. Mater. **15**, 263 (2016).
[260] S. Prosandeev, I. Ponomareva, I. Kornev, I. Naumov, and L. Bellaiche, Phys. Rev. Lett. **96**, 237601 (2006).
[261] A. Kirilyuk, A. V. Kimel, and Th. Rasing, Rep. Prog. Phys. **76**, 026501 (2013).
[262] K. M. D. Hals and A. Brataas, Phys. Rev. B **89**, 064426 (2014).

Appendix A.
FW operator from the Eriksen operator

Here, we provide derivation steps that lead to obtaining the original FW operator while starting from the Eriksen operator. From Eqs. (2.55) and (2.56) one has

$$e^{iU_{FW}} = \cos U_{\rm FW} + i \sin U_{\rm FW} = \frac{1+\beta\lambda}{\sqrt{2+\beta\lambda+\lambda\beta}}\,, \tag{8.1}$$

$$e^{-iU_{FW}} = \cos U_{\rm FW} - i \sin U_{\rm FW} = \frac{1+\lambda\beta}{\sqrt{2+\beta\lambda+\lambda\beta}}\,. \tag{8.2}$$

Thus, by adding and subtracting the last two equations, and the corresponding tangent function can be expressed by

$$2\cos U_{\rm FW} = \sqrt{2+\beta\lambda+\lambda\beta}\,, \quad 2i\sin U_{\rm FW} = \frac{\beta\lambda-\lambda\beta}{\sqrt{2+\beta\lambda+\lambda\beta}}\,, \tag{8.3}$$

$$\tan U_{\rm FW} = -\frac{i(\beta\lambda-\lambda\beta)}{2+\beta\lambda+\lambda\beta}\,. \tag{8.4}$$

Now, with the Hamiltonian expressed in Eq. (2.17), one derives

$$\beta\lambda-\lambda\beta = \frac{2\beta\mathscr{O}}{\sqrt{\mathscr{H}^2}}\,, \quad \beta\lambda+\lambda\beta = \frac{2\left(mc^2+\beta\mathscr{E}\right)}{\sqrt{\mathscr{H}^2}}\,. \tag{8.5}$$

With these two expressions we write Eq. (8.4) in the following way

$$\tan U_{\rm FW} = -\frac{i\beta\mathscr{O}}{\sqrt{\mathscr{H}^2}+mc^2+\beta\mathscr{E}}\,. \tag{8.6}$$

In the nonrelativistic limit, mc^2 is the highest energy and within the approximation $\sqrt{\mathscr{H}^2} \approx mc^2$, we arrive at

$$\tan U_{\rm FW} = -\frac{i\beta\mathscr{O}}{2mc^2}\,. \tag{8.7}$$

Thus, within the approximation and expansion of $\tan^{-1}$ series keeping only the lowest order term, one obtains exactly the same original FW operator

$$U_{\rm FW} = -\frac{i\beta\mathscr{O}}{2mc^2}\,. \tag{8.8}$$

We observe that the Eriksen operator is more general and robust for FW transformation and within the approximation, the original FW operator is obtained.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1558

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-315247

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017