

# Ultrafast Magnetization Dynamics

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften  
der Rheinisch-Westfälischen Technischen Hochschule Aachen  
zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften  
genehmigte Dissertation

vorgelegt von

Simon Woodford, M.Sc.

aus Kapstadt, Südafrika

Berichter: Universitätsprofessor Dr. Stefan Blügel  
Universitätsprofessor Dr. Peter H. Dederichs

Tag der mündlichen Prüfung: 30. Juni 2008

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

This thesis was written at the institute  
Quanten-Theorie der Materialien, Institut für Festkörperforschung (IFF),  
Forschungszentrum Jülich GmbH, Jülich, Deutschland.

It is available online at the internet pages of the library of  
RWTH Aachen university.

# Abstract

This thesis addresses ultrafast magnetization dynamics from a theoretical perspective. The manipulation of magnetization using the inverse Faraday effect has been studied, as well as magnetic relaxation processes in quantum dots.

The inverse Faraday effect — the generation of a magnetic field by nonresonant, circularly polarized light — offers the possibility to control and reverse magnetization on a timescale of a few hundred femtoseconds. This is important both for the technological advantages and for the deeper fundamental understanding of magnetization dynamics that can be gained. However, several aspects of the inverse Faraday effect have remained poorly understood. The question of whether light can manipulate magnetization alone or whether an additional angular momentum reservoir is needed, in particular, remains unanswered.

This question is answered here: the light beam that causes the inverse Faraday effect provides the angular momentum required for the magnetization to precess. No other reservoir is needed. This implies that manipulation of the magnetization occurs on the timescale of a laser pulse, which can be made extremely short. Even magnetization *reversal* on this timescale could be possible, provided a material with a sufficiently strong magneto-optical response can be found. This is a technical challenge, not a fundamental obstacle.

The Faraday effect in the presence of optical birefringence has also been analyzed. This effect has been used for imaging magnetization dynamics in transparent media on an ultrafast timescale, but transparent magnetic materials usually have a complex crystal structure and complicated optical properties, which render the relationship between Faraday rotation and magnetization unclear. We have shown that the Faraday effect *can* be used to measure the instantaneous magnetization, even in the presence of birefringence, provided certain experimental conditions are met. Suggestions concerning these experimental conditions are made.

The relaxation of magnetization, particularly the relaxation of a spin in a quantum dot, has been studied. This problem is relevant to the fields of quantum computing and highly-multiplexed optical memory, both of which are of great current interest. We have investigated the interaction of the spin with a metallic electrode and calculated the dephasing and dissipation rates. We found that under current experimental conditions, this relaxation pathway is negligible compared to spin-phonon scattering, but as systems are miniaturized, interactions with electrodes become more important.

The methods developed to study the relaxation of a spin were also applied to the relaxation of a charge in a double quantum dot, another important problem in quantum computing. Again, we found that the interaction with a gate electrode is generally much weaker than the interaction with phonons but may be important for smaller systems.

# Kurzfassung

Diese Dissertation behandelt die ultraschnelle Dynamik der Magnetisierung aus einer theoretischen Perspektive. Die Manipulation der Magnetisierung mit Hilfe des inversen Faraday-Effektes wird untersucht, wie auch magnetische Relaxationsprozesse in Quantenpunkten.

Der inverse Faraday-Effekt — die Erzeugung eines magnetischen Feldes durch nicht-resonantes, zirkular-polarisiertes Licht — bietet die Möglichkeit der Kontrolle und Schaltung der Magnetisierung auf einer Zeitskala von wenigen hundert Femtosekunden. Dies ist wichtig sowohl aus technologischen Erwägungen als auch für die Gewinnung eines tieferen Verständnisses der grundlegenden Dynamik der Magnetisierung. Es gibt bis heute einige Aspekte des inversen Faraday-Effektes, die nicht gut verstanden sind. Insbesondere bleibt die Frage unbeantwortet, ob Licht allein eine Magnetisierung manipulieren kann oder ob eine zusätzliche Drehimpulsquelle notwendig ist.

Diese Frage wird hier beantwortet: Der Lichtstrahl, der den inversen Faraday-Effekt bewirkt, stellt auch den Drehimpuls zur Verfügung, der für Präzession erforderlich ist. Keine andere Drehimpulsquelle ist notwendig. Dies bedeutet, dass die Kontrolle der Magnetisierung auf der Zeitskala eines Laser-Pulses, der sehr kurz sein kann, stattfindet. Selbst die *Umkehr* der Magnetisierung könnte in diesem Zeitrahmen möglich sein, vorausgesetzt, dass ein Material mit einer ausreichend starken magnetooptischen Antwort gefunden werden kann. Dies ist eine technische Herausforderung, kein grundlegendes Hindernis.

Der Faraday-Effekt unter dem Einfluss optischer Doppelbrechung wird auch analysiert. Dieser Effekt wird benutzt um die Magnetisierungsdynamik in transparenten Medien auf einer ultraschnellen Zeitskala zu visualisieren. Transparente magnetische Materialien haben jedoch in der Regel eine komplexe Kristallstruktur und komplizierte optische Eigenschaften, die die Beziehung zwischen Faraday-Rotation und Magnetisierung verschleiern. Wir zeigen, dass der Faraday-Effekt verwendet werden kann, um die momentane Magnetisierung zu messen. Dies funktioniert auch unter dem Einfluss von Doppelbrechung, sofern bestimmte experimentelle Bedingungen erfüllt sind. In diesem Zusammenhang werden entsprechende experimentelle Methoden vorgeschlagen.

Die Relaxation der Magnetisierung, vor allem die Relaxation eines Spins in einem Quantenpunkt, wird untersucht. Dieses Problem ist von großer Bedeutung im Gebiet des Quanten-Computing und “highly-multiplexed” optischer Speicher. Wir untersuchen die Wechselwirkung des Spins mit einer metallischen Elektrode und berechnen die Dephasings- und Dissipations-Raten. Wir stellen fest, dass die Relaxation unter den gegenwärtigen experimentellen Bedingungen gegenüber der Spin-Phonon-Streuung vernachlässigt werden kann. Mit fortschreitender Miniaturisierung wird die Wechselwirkung mit den Elektroden jedoch an Bedeutung gewinnen.

Die Methoden, die zur Beschreibung der Spin-Relaxation entwickelt werden, sind auch auf die Relaxation der Ladung in einem doppelten Quantenpunkt angewendet. Letzteres ist ein weiteres wichtiges Problem im Gebiet des Quanten-Computing. Auch hier stellen wir fest, dass die Wechselwirkung mit einer Elektrode bei heute realisierbaren Systemgrößen viel schwächer ist als die Wechselwirkung mit Phononen, jedoch in kleineren Systemen wichtig werden kann.

# Contents

<b>Abstract</b>	<b>1</b>
<b>Kurzfassung</b>	<b>2</b>
<b>1 Introduction</b>	<b>6</b>
1.1 Magnetization dynamics . . . . .	6
1.1.1 Magnetism in technology . . . . .	6
1.1.2 Magnetic switching . . . . .	7
1.1.3 New methods . . . . .	8
1.2 Magneto-optics . . . . .	10
1.2.1 Experiments . . . . .	10
1.2.2 Angular momentum . . . . .	12
1.2.3 Nonthermal switching . . . . .	12
1.3 Quantum Computing . . . . .	14
1.3.1 What is quantum computing? . . . . .	14
1.3.2 Qubits . . . . .	15
1.3.3 Manipulating qubits . . . . .	17
<b>2 The Inverse Faraday Effect</b>	<b>22</b>
2.1 The Faraday effect . . . . .	22
2.2 Energy considerations . . . . .	23
2.3 The effective Hamiltonian formulation . . . . .	26
2.4 The ultrafast inverse Faraday effect . . . . .	31
<b>3 Conservation of Angular Momentum: Small Systems</b>	<b>33</b>
3.1 Introduction . . . . .	33
3.2 Angular momentum of light . . . . .	34
3.3 Effective magnetic field . . . . .	36
3.4 Demonstration using hydrogen atom . . . . .	37
<b>4 Conservation of Angular Momentum: Extended Systems</b>	<b>41</b>
4.1 Description of the problem . . . . .	42
4.2 Orbital angular momentum of light . . . . .	43
4.3 Propagation of light . . . . .	46
4.4 Change of polarization . . . . .	52
4.5 The second interface . . . . .	53

---

4.5.1	First reflection . . . . .	53
4.5.2	Loss of angular momentum . . . . .	55
4.5.3	Multiple reflection . . . . .	56
<b>5</b>	<b>The Difficulty of Measurement</b>	<b>60</b>
5.1	Measuring ultrafast magnetization dynamics . . . . .	60
5.2	Orthoferrites . . . . .	61
5.3	Faraday rotation in a birefringent weak ferromagnet . . . . .	62
5.4	Limiting cases . . . . .	65
5.4.1	Zero birefringence . . . . .	65
5.4.2	Large birefringence, small $\theta$ . . . . .	66
5.5	Numerical exploration of nonperturbative regions . . . . .	69
5.6	Summary . . . . .	70
<b>6</b>	<b>Interaction of a Quantum System with a Heat Bath</b>	<b>73</b>
6.1	Dephasing and dissipation . . . . .	73
6.2	The generalized master equation . . . . .	74
6.3	Solving the master equation . . . . .	77
6.3.1	Numerical integration . . . . .	78
6.3.2	Laplace transformation . . . . .	82
<b>7</b>	<b>Dephasing of Quantum Dots: Coulomb Interaction with a Gate</b>	<b>86</b>
7.1	Overview . . . . .	86
7.2	Dephasing of spin qubits . . . . .	87
7.2.1	Structure of qubit . . . . .	87
7.2.2	Coulomb interaction with the gate . . . . .	90
7.2.3	Response function of the gate . . . . .	93
7.2.4	Laplace transform of $I^>(t)$ . . . . .	95
7.2.5	Results . . . . .	96
7.3	Dephasing of charge qubits . . . . .	99
7.3.1	Difference between charge and spin qubits . . . . .	99
7.3.2	Results . . . . .	101
<b>8</b>	<b>Concluding Remarks</b>	<b>105</b>
<b>A</b>	<b>Propagation and Polarization of Electromagnetic Waves</b>	<b>107</b>
A.1	Monochromatic plane waves . . . . .	107
A.2	Polarization of plane waves . . . . .	108
	<b>Bibliography</b>	<b>120</b>
	<b>List of Publications</b>	<b>121</b>
	<b>Acknowledgments</b>	<b>122</b>
	<b>Curriculum Vitae</b>	<b>124</b>

# Chapter 1

## Introduction

### 1.1 Magnetization dynamics

#### 1.1.1 Magnetism in technology

Magnetism has proven to be extremely useful in the field of data storage. From Poulsen’s early attempts to record telephone conversations on steel wire [62, 85] up to modern hard drives and credit cards, magnetism has offered the possibility of storing information reliably and stably, with ever-increasing densities of data. An excellent review of the impact of magnetism on technology is given in Ref. [62].

One reason for the prevalence of magnetism in data storage is that the binary nature of magnets is compatible with the binary nature of semiconductor computation. Computers consist of an array of transistors that are either “on” or “off”; similarly, magnets consist of a large number of domains that, assuming certain geometrical constraints are met, are directed either “up” or “down”<sup>1</sup> [100]. This allows a very clear mapping between stored data and computations. Consequently, system design is greatly simplified.

The existence of these “up” and “down” states is made possible by geometric- and size-effects [100, 110, 111]. Such states would not be identifiable in a bulk, isotropic ferromagnet, which is a classical example of spontaneous symmetry breaking: it is energetically preferable for a nonzero, homogeneous magnetization to form, but *all directions are equivalent* [119]. To break rotational symmetry and ensure that the magnetization prefers to point either up or down (relative to some fixed axis), either the shape of the magnet must be constrained or a magnetic crystal that is suitably anisotropic must be used [110, 111]. In general, it is more practical to use geometric constraints, e.g. magnetic tapes use needle-like particles of  $\text{Fe}_2\text{O}_3$  [197]. Furthermore, the system must be kept small. As the system passes a certain critical size, weak long-range forces cause the magnetization to break into domains with different orientations of the magnetic vector; by keeping the system below this size, domain effects can be suppressed [3, 23, 141, 181].

The up and down states of a small needle-shaped magnet are very stable against random perturbations. This is related to the fact that the largest energy scale of the system is the exchange energy, which is responsible for the homogeneity of the magnetization. Any perturbation that tries to disorder the magnetism, e.g. thermal fluctuations, is strongly

---

<sup>1</sup>Both of these statements require qualification, but are true at a certain level of abstraction.

suppressed. However, homogeneous changes that occur when writing information to the magnetic medium are not as strongly suppressed [141, 181]. This stability against random fluctuations is a very important property, particularly in archiving applications, where the data may be stored for several years (or even decades).

The dipole interaction is the only long-range interaction between two separate magnets. It is extremely weak, so a region in which all the magnetic needles are in the up-state can be very close to a region of down-state needles without incurring any loss of stability. Each region (usually referred to as a bit) independently holds a piece of information, so if stability of stored information were the only design criterion, high densities would be attained.

However, the density of data storage is not only constrained by stability considerations. It is also important that data can be read and written. Great advances have been made in the field of reading magnetic fields. In particular, the discovery of giant magnetoresistance (GMR) [10, 18] has led to the possibility of accurately measuring extremely small magnetic fields. This has greatly increased the possible storage density and improved the speed at which data can be read, for which reason GMR earned its discoverers the 2007 Nobel prize. The speed at which data can be written, however, remains slower than that of reading and there are numerous investigations underway to increase this speed. This is discussed further in the following sections.

### 1.1.2 Magnetic switching

In many ways, writing is more complicated than reading [131]. Reading is a non-destructive technique, while writing destroys the information that was previously stored. Writing not only involves destabilizing the bit and driving the magnetization in the required direction, it also involves waiting for the system to settle back to equilibrium (the modern equivalent of waiting for ink to dry).

In the context of magnetic data storage, writing is often referred to as magnetization reversal or magnetic switching. I shall use these terms interchangeably in what follows.

The conventional method of writing a magnetic bit involves (approximately) the application of a large magnetic field in the desired direction<sup>2</sup>. This makes the desired state lower in energy and destabilizes the opposite state. If the bit is already in the desired state, the magnetic field does not cause any dynamics. If the bit is in the opposite state, it will start to move towards the lower energy state.

In the case where both the magnetization and applied field are homogeneous, the motion towards the lower energy state is not coherent. The higher energy state is destabilized, but still exists as an unstable equilibrium position. Motion away from it is governed by the growth of perturbations that are caused by quantum and thermal fluctuations. These perturbations are nondeterministic. In the homogeneous case, switching is a stochastic process.

To provide more reliable switching times, an inhomogeneous applied field is used. This causes a magnetic gradient that leads to deterministic and predictable switching. However, even with deterministic switching the motion between the states is not instantaneous. The magnetic field must be applied for long enough to guarantee switching.

---

<sup>2</sup>There are, of course, many technical subtleties, but this is the basic concept. For further details see, e.g., Refs. [131, 197]

Once the magnetization is mostly pointing in the correct direction, the magnetic field can be removed and the “write-head” (the object that provides the magnetic field) can move on to write a different bit. However, the writing process of the original bit is not yet finished. The bit must be in either an up or a down state; it should not be in a “mostly up” state. It requires some time to relax to equilibrium.

Relaxation to equilibrium is a vital part of switching. It allows us to be quite cavalier about the application of the magnetic field. It is not necessary to keep the magnetic field acting until the magnetization is pointing exactly where we want it to; we can remove the magnetic field as soon as we are sure that the magnetization will relax to the state that we want. This is referred to as placing the magnetization in the *basin of attraction* of the final state [201].

However, relaxation has some disadvantages. It reduces the speed of writing by dissipating energy during switching. Furthermore, the heat generated during switching and the subsequent relaxation must be rapidly removed to prevent a loss of stability. Of course, the closer we place the magnetization to the final state, the less time we lose waiting for equilibration and the less heat is produced.

### 1.1.3 New methods

Magnetization reversal using a strong magnetic field in the desired direction is now a mature technology. Advances are still being made to increase the storage density<sup>3</sup>, but hopes for massive increases of storage densities and write speeds seem to lie in other technologies [7, 94, 99, 144]. This is particularly important for the development of magnetic random access memory (MRAM), which is a proposed form of nonvolatile memory that could reduce power consumption by up to 99% [54, 184]. The random access memory of a computer requires very fast read and write speeds, much faster than those needed for long-term data storage devices<sup>4</sup>. The current generation of MRAM is slower than static RAM (SRAM) and has a lower storage density (and is thus more expensive) than dynamic RAM (DRAM). For these reasons, it has not been widely adopted for commercial use. However, with the modern drive towards energy-saving technology, MRAM has a lot to offer and much work is being done on reducing the cost and increasing the storage density and write-speed. This has led to many alternative methods of controlling magnetization being intensely investigated.

One candidate for high-speed writing is *precessional switching* [8, 9, 14, 52, 88, 172]. Precessional switching uses a magnetic field *orthogonal* to the equilibrium direction of magnetization instead of using a field *parallel* to the desired direction and waiting for fluctuations to push the magnetization into the lowest energy state. As with Larmor precession [73], the orthogonally-oriented field exerts a torque on the magnetization, which then starts to rotate. By tuning the magnetic pulse length and shape [52], the rotation can

---

<sup>3</sup>A recent example is perpendicular recording, which was discovered in 1977 [80] and first implemented in commercial disk drives in 2005 by Toshiba; see the press release at [http://www.toshiba.co.jp/about/press/2004\\_12/pr1401.htm](http://www.toshiba.co.jp/about/press/2004_12/pr1401.htm)

<sup>4</sup>The reason for this is that programs must generate data before writing it to long-term storage media, so writing to memory occurs more often than writing to the hard-drive. Furthermore, bottlenecks due to slow hard-drive write-speeds can generally be avoided by using a memory buffer. Writing to memory, on the other hand, occurs frequently and constantly throughout the execution of a program, so it must be extremely fast.

be stopped when the magnetization is very close to the desired final state. Precessional switching is much faster than conventional switching; reliable precessional switching within 200 ps [9, 14, 172] has been demonstrated, while conventional switching requires several nanoseconds [70, 120, 199]. However, the switching speed is determined by the rate of precession, which is proportional to the magnetic field. Fast switching requires strong magnetic fields, and since generation of strong magnetic pulses shorter than around 100 ps is out of the range of modern experiments [88], precessional switching offers limited potential for improvement<sup>5</sup>.

Another method that is attracting much attention is *spin-torque transfer* [17, 176]. This is essentially the inverse of the GMR effect. In this method, a current is passed through a ferromagnetic layer and emerges spin-polarized. This current then passes through the bit that is to be written. If the magnetization of the bit is parallel to the direction of spin-polarization, the current passes through the bit with little effect. If the magnetization is antiparallel, the current scatters strongly. This scattering transfers angular momentum from the current into the magnetization, leading to a torque that acts on the bit and reversing the magnetization.

The discovery of GMR and the spin-torque effect has led to the growing field of spintronics, which extends far beyond simple data storage and memory devices (see, e.g., Refs. [104, 143, 145, 202]). However, spintronics is unlikely to advance into the subpicosecond regime, primarily because ultrashort current pulses broaden rapidly as they propagate through a wire, resulting in poor time resolution.

The above-mentioned methods, precessional switching and spin-torque transfer, rely on the idea that the information is stored in the direction of the magnetization. However, there have been some recent proposals to use vortices to store information instead [6, 28, 72, 162]. In this case, the bit is a thin disc, with the magnetization within the plane. The magnetization can be prepared such that it forms a circular, flux-closure pattern. Such a configuration has a singularity at the centre of the disc where, if it were constrained to the plane, the magnetization would have to point equally in all planar directions and therefore vanish. The magnetization cannot vanish, so it points out of the plane near the centre. With a symmetric sample, the energy is the same if the magnetization points “up” out of the plane or “down” — these are the binary states that store information.

The flux-closure condition allows vortices to be brought closer to each other than conventional magnetic bits. They are also more stable. The biggest problem with using vortex cores to store information is that this high stability makes them difficult to switch. Recent simulations indicate that trying to directly reverse the vortex is not the optimum method of switching [55]. A more efficient and faster method is to nucleate a vortex-antivortex pair with opposite polarity, which results in the antivortex annihilating the existing vortex and leaving a reversed vortex plus some extra energy. (Note that an antivortex is not simply a vortex with reversed vorticity; Fig. 1.1 clarifies the difference<sup>6</sup>.)

---

<sup>5</sup>Recent experiments have shown that even if very strong magnetic pulses could be made without damaging the sample, magnetization dynamics under high fields may be unpredictable [186], which would impose fundamental limits on the maximum precessional switching speed.

<sup>6</sup>Due to the use of APS copyrighted material, this thesis is subject to the following restriction: Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or part, without prior written permission from the American Physical Society.

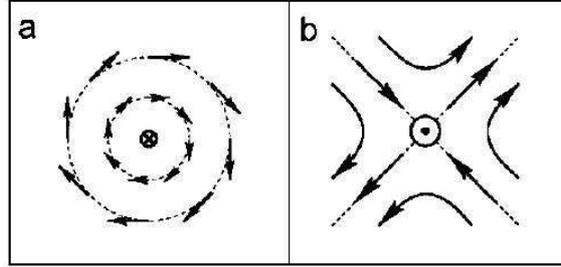


Figure 1.1: Schematic diagram of (a) a vortex and (b) an antivortex, taken from Ref. [72]. Note that an antivortex is not simply a vortex with reversed vorticity. Figure reprinted with permission from R. Hertel and C. M. Schneider, *Physical Review Letters* **97**, 177202 (2006). Copyright (2006) by the American Physical Society.

This method gives switching on a similar timescale to precessional switching. Most of this time is spent waiting for the excess energy to dissipate, since the actual nucleation and annihilation processes require only a few tens of picoseconds [55]. However, because this method relies on either a spin-current or a magnetic field to nucleate the vortex-antivortex pair, the above disadvantages are still present. In particular, this method is very unlikely to advance into the subpicosecond regime.

Indeed, out of all equipment that a solid-state physics researcher may have available, only one device reliably generates pulses of subpicosecond length — the laser. The readily-available Ti:sapphire laser is able to produce light pulses with full-width at half-maximum (FWHM) of less than 100 fs. Therefore the fastest methods of magnetic switching are likely to use the interaction of light and magnetism, usually referred to as *magneto-optics*.

## 1.2 Magneto-optics

### 1.2.1 Experiments

The study of magneto-optics dates back at least as far as Faraday, who discovered that a magnetic field influences the propagation of light through a material. This effect, known as the “Faraday effect”, is discussed in detail in the following chapter. However, the attempt to reverse magnetization using optical pulses is much more recent.

The pioneering experiments in this field were performed in 1996 by the Strasbourg group of Beaurepaire *et al.* [16]. They demonstrated that a 60 fs laser pulse absorbed in a nickel film leads to demagnetization on a subpicosecond timescale. Beaurepaire’s experiments were confirmed by Hohlfeld *et al.* [76], who provided further details on the dynamics within the first 300 fs. These results were very surprising, since Vaterlaus *et al.* had previously shown that spin-lattice relaxation occurs on the timescale of 100 ps [190]; the angular momentum of the spin appears to change much faster than it can flow into the lattice. As discussed below, several mechanisms have been proposed to explain this, most notably that electrons far from equilibrium experience a much stronger spin-lattice interaction [103, 178].

These experiments motivated intense experimental [2, 61, 171] and theoretical [79, 208, 209] investigations. Already by 1999, it appeared that demagnetization could occur within

50 fs of the pulse [61]. This estimate was limited by the resolution of the experiment, so demagnetization could be even faster. This estimate was supported by theoretical work that claimed demagnetization could occur within 20 fs [79, 209]. Of course, demagnetization is not the same as magnetization reversal, but by inducing demagnetization in the presence of an external magnetic field, complete magnetic switching can be achieved [60, 75].

However, in 2000 an alternative interpretation of the experiments was provided by Koopmans *et al.* [103]. They noticed that the instantaneous magnetization was being measured with an optical technique that only admitted reliable interpretation provided the sample was close to equilibrium. In the highly nonequilibrium state that appears when the pulse is absorbed, the correspondence between the optical signal and the magnetization interpretation no longer holds. By measuring both the Kerr rotation and the induced ellipticity, Koopmans *et al.* showed that demagnetization occurs on the timescale of 500 fs; the apparent demagnetization on timescales shorter than this is due to a purely optical effect [103].

The problem with using an optical method to probe magnetization is that light does not interact directly with a magnetization. Instead, the magnetization leads to changes in the electronic properties of the sample and the propagation of light depends on these properties. Any nonmagnetic change of the electronic state also affects the light and must be accurately taken into account. This is exceptionally difficult for a strongly nonequilibrium electron distribution. However, measurement with laser pulses is the only technique available that can provide sufficient temporal resolution to take “snapshots” of the magnetization — the difficulties in interpreting optical signals must be overcome.

In the case of the ultrafast demagnetization experiments, the main nonmagnetic contribution to the optical signal arises from the depletion of electrons from the most optically-active bands [103, 146]. These electrons are excited by the light and, after absorption, have their energy raised by  $\hbar\omega$ , with  $\omega$  the frequency of the light. This means that the sample does not have a hot, thermalized distribution of electrons, but rather a cold thermal distribution with one band depleted and the electrons displaced to higher energies [146]. When a probe pulse of the same frequency as the pump pulse interacts with the sample, its interaction is modified because the electrons with which it would interact most strongly are depleted. This optical effect, known as “bleaching” or “state-filling” [103, 146], is responsible for the apparent instantaneous demagnetization. By using complementary optical measurements, such as the rotation *and* the ellipticity [103], or by using different frequencies for the pump and probe beams, this problem can be avoided.

Many experiments have been performed and there is general consensus that demagnetization occurs within 100-300 fs [101]. Recent experiments using X-rays for the probe beam and optical-frequency light for the pump [178] yielded an estimate of demagnetization within  $120 \pm 70$  fs. However, because of the heat generated and the difficulty of causing magnetization *reversal*, it remains unclear whether this ultrafast demagnetization can be reliably implemented in a commercial device.

The main question that has been asked since the advent of ultrafast demagnetization is, “Where does the angular momentum go?” (see, e.g., the discussion in Ref. [102]). It was precisely this question that led to the discovery of bleaching [103]. The issue of angular momentum is considered in the next section.

### 1.2.2 Angular momentum

The discovery of the Einstein-de Haas effect [42] demonstrated that spin angular momentum, which is responsible for ferromagnetism, can be converted into an equivalent amount of mechanical angular momentum. This important result led to the realization that reversal of magnetization is only possible if there is a reservoir of angular momentum available; it is not enough simply to make the reversed state energetically preferable. The rate at which angular momentum is transferred from the reservoir limits the maximum speed of magnetization reversal. For optically-induced magnetic switching and demagnetization, the question of which reservoir supplies the necessary angular momentum remains controversial.

The magneto-optical switching methods described above are all *thermal* methods, i.e. the laser pulse is absorbed and drives the electrons to a much higher temperature. Several papers discuss the source of angular momentum of these experiments [102, 103, 178]. Some estimates show that demagnetization within 300 fs is too rapid for phonon processes, and it has been suggested that angular momentum is transferred between the spin and orbital components of the electrons [102]. This can lead to a reduced magnetic moment, since the orbital and spin components with the same angular momentum contribute differently to the magnetic moment. However, an alternative explanation is that the nonequilibrium electron distribution experiences a stronger spin-phonon interaction and that the phonons provide the angular momentum [103, 178].

By estimating the number of photons that are absorbed and the dichroism of the substrate, Koopmans *et al.* showed that the photon angular momentum is irrelevant in the nickel experiments [102]. Furthermore, in experiments performed using circularly-polarized light on nickel, changes of the helicity of the light yielded a negligible change in the demagnetization time [126]. On the other hand, some experiments with circularly polarized light on GdFeCo show significant differences in the magnetic response when the sense of polarization is reversed [179]. The difference between these experiments can probably be attributed to the strong spin-orbit coupling of gadolinium, but much remains to be clarified about the role of the photon angular momentum in thermal demagnetization.

Theoretical identification of the relevant angular momentum reservoir is complicated by the highly nonequilibrium dynamics of the electrons. In this thesis, I have concentrated on magneto-optical switching close to equilibrium. This is far more accessible theoretically, but the main motivation for this choice is that I am convinced that *nonthermal switching*, i.e. magnetization reversal without the generation of a hot electron gas, provides more insights into the fundamental speed limits of magnetic switching and will also have more relevance for technological applications, since it does not waste a lot of energy generating heat.

### 1.2.3 Nonthermal switching

The magneto-optical switching processes described above have two clear obstacles preventing their technological implementation. The first is that many of the “ultrafast magnetization reversal” experiments are actually ultrafast demagnetization experiments — the magnetization is driven from its ferromagnetic ground state into a disordered state with a decreased magnetization. The lack of inertia in magnetic dynamics means that it is difficult to ensure that the magnetization reorients itself in the reversed ground state af-

ter dissipation. Experiments have shown that reversal can be attained using an external magnetic field [60, 75], but this introduces the difficulty of having rapidly time-varying magnetic fields if magnetic switching is to be allowed in both directions.

Assuming that a device can be designed in which magnetic switching functions as expected, there is another problem. Absorption of a laser pulse drives the electrons far from equilibrium, which involves a large increase of heat. This heat must be dissipated and the device will suffer a significant cooling time, during which no further writing may occur.

I mentioned above that writing to long-term data storage need not be of comparable speed to writing to memory. As such, good use of memory buffers can ensure that this cooling time is not too problematic for data storage, considered bitwise. However, the cooling of individual bits is not the only thermal process. The most efficient method for the bits to lose heat is to conduct it into the substrate surrounding the bit. If many bits are written in a short period of time, the substrate itself heats up substantially. This heat must be removed, or the stability of the storage device will suffer. Since a hard drive is usually as isolated as possible, this heat must be convected across an air gap, which is not a rapid process. The cooling of the hard drive is likely to be a nontrivial process.

In addition, the problem of cooling time eliminates such devices as candidates for fast random access memories. Because MRAM should allow truly random access, each write cycle must wait until the previous write cycle is finished, including equilibration back to the operational temperature. The question of whether demagnetization occurs within 50 or 500 fs becomes irrelevant compared to the cooling time.

Finally, this excessive heat production is incommensurate with the modern drive towards energy efficiency. Going from a ground state through a highly nonequilibrium situation to another ground state is far from being the lowest energy pathway. Such devices would be very unpopular, both from an environmental perspective and in terms of the cost of power. It seems unlikely that this technology will find large-scale technological application.

What is needed, therefore, is a *nonthermal* optical method of manipulating the magnetization, a method of controlling the electron spins without absorption of the light. In 2005, the experimental group in Nijmegen realized that there is a physical effect that can provide such a method: the *inverse Faraday effect* [99]. The inverse Faraday effect (IFE) is the induction of a magnetization in a sample by nonresonant, circularly polarized light [152, 187]; it is discussed in detail in the following chapter. The Nijmegen group showed that the IFE could trigger magnetic oscillations. Due to the nonresonance condition, the light is not absorbed and there is very little heating of the sample.

These experiments inspired most of the work contained in this thesis. As with the other methods of optically-induced magnetic switching, the source of angular momentum in the IFE was unknown. This problem has now been resolved; as shown in chapters 3 and 4, the angular momentum is provided by the light<sup>7</sup>. This discovery suggests a further technique for estimating the magnetization in the sample by evaluating the angular momentum lost by the light. This is discussed further in chapter 4. Thereafter, in chapter 5, I present our analysis of the Nijmegen experiments. Despite the lack of bleaching due to the proximity to equilibrium, other optical effects hinder the evaluation of the magnetization. By taking

---

<sup>7</sup>This is a nontrivial result, considering the light is not absorbed.

these into account and tailoring the experimental conditions appropriately, experiments become simpler and more reliable.

However, before starting my exposition of the IFE, it should be noted that in the above discussion, I have considered only classical computing methods. I referred to bits, to binary logic and to magnetization reversal. There is an alternative paradigm of computation that is likely to become ever more prominent in the future: *quantum computing*. Since several models of quantum computing use spin to store information, ultrafast magnetization dynamics on a few-electron scale is relevant to quantum computing. In particular, the relaxation rate of a nonequilibrium magnetization must be considered in order to provide a timescale on which quantum computing may be reliable. Chapters 6 and 7 are devoted to physical processes that are relevant to quantum computing. It is therefore appropriate to provide here a brief overview of this rapidly-developing field.

## 1.3 Quantum Computing

### 1.3.1 What is quantum computing?

Simply stated, quantum computing is the use of algorithms that rely on quantum mechanical concepts, particularly superposition and entanglement. Instead of representing data with a sequence of independent zeros and ones, as in classical computing, information is stored in quantum superpositions of zeros and ones. This allows more information to be stored and processed in each computing operation [132]. Indeed, since quantum mechanics contains classical mechanics as a limiting case, quantum computing can only offer greater computing power than classical computing [37]. However, quantum computing requires the manipulation of quantum objects, which is far more complicated than manipulating classical objects. Successful large-scale quantum computers will only be built once the understanding and control of quantum objects reaches a sufficiently advanced level, a goal that has thus far not been achieved.

The idea of using quantum states for computing originated with Feynman [50], who suggested that quantum systems would be more readily simulated on a quantum computer than on a classical one. This idea is closely related to the practice of analog simulation in the early days of computing, where instead of approximating an analog system by a digital one and solving the digital system on a computer, an analog device would be built with identical properties to the system being simulated [185]. The device would be designed so that its parameters were easily adjusted and a simulation would be fast, thus providing results for the original system. Similarly, a complicated quantum system (e.g. a system that is too small to be measured, reacts too quickly, etc.) could be modeled by another quantum device that is easier to prepare and monitor.

Several years later, the first “quantum algorithm” was found by Deutsch [35]. The algorithm itself is not particularly useful (it involves determining whether a certain binary function is constant or not), but it introduced ideas that led to the discovery of far more powerful algorithms. Indeed, the modern popularity of quantum computing can be traced to the 1994 paper of Shor [174], which was inspired by the work of Deutsch. In this paper, Shor showed that large numbers may be factorized into primes exponentially faster on quantum computers than on classical computers. This means that quantum computers render certain public-key cryptography protocols [167] insecure. Fortunately, quantum

computing brings with it the possibility for quantum cryptography [43], recovering the possibility of secure communications. Other well-known applications of quantum computing include finding entries in a database [58], comparing protein sequences [78] and pattern recognition [173]. An excellent overview of the field of quantum information processing is provided in Ref. [142].

The discovery of quantum algorithms led to a search for appropriate hardware on which these algorithms could be implemented. Generally, quantum computing uses quantum bits (qubits), which are two-state quantum systems. During the calculation, an array of qubits will be in an entangled superposition of states. This is in marked contrast to classical computing, where each bit has a definite state and interactions between bits are to be avoided as much as possible. The requirements that a two-state quantum system must fulfill to be a candidate for a qubit have been presented by DiVincenzo [36–38] and I summarize them here.

First, the qubit states must be accurately known and characterized. It is also essential that the system is easily scalable, so that a large number of qubits can be used (and entangled) simultaneously. In addition, it should be possible, and not too difficult, to prepare, manipulate and measure the quantum states. Without the ability to prepare and manipulate states, programming is impossible, while measurement allows results to be obtained in classical terms. Finally, it is important that the system remain “quantum-mechanical” during the computation. That is, it must not interact so strongly with its environment that it decays into a classical state, or the quantum algorithms will fail. An accurate estimate of the decoherence time — the time in which quantum-mechanical coherence is destroyed — is very important; this decoherence even limits the possibility of quantum error correction [25, 39].

Many candidates for quantum computing have been proposed over the years. As yet, none of the proposals satisfy all of the above requirements sufficiently, but work is underway on several of them. Some of the more common proposals are summarized in the following section; for a more detailed discussion see, e.g., Ref. [68]

### 1.3.2 Qubits

The earliest proposals to realize qubits came from the fields of quantum optics [30] and nuclear magnetic resonance [53]. There have been successful demonstrations of quantum algorithms in both types of system [26, 29, 63, 170, 189]. However, it seems that NMR devices are not scalable to large numbers of qubits [97]; quantum computers based on NMR will remain a useful testbed for developing new techniques, but are unlikely to have commercial relevance. Quantum optical devices, such as ion traps, are more scalable, but modern designs are heading in the direction of solid-state implementations, rather than the original all-optical method [196]. Not only is scalability required to increase the power of the computers, but quantum error correction methods also drastically increase the number of qubits needed [25, 39] — solid-state systems appear to be the only viable option.

Solid-state systems are advantageous from the point of view of scalability, but due to the many degrees of freedom available, decoherence presents more of a problem. One of the most important fundamental challenges in designing a solid-state quantum computer is to find suitable qubits that do not interact too strongly with their environments.

There are two types of solid-state systems considered for quantum computing — superconductors and semiconductors. Superconductors have the advantage that they can be macroscopic in size but still behave quantum-mechanically. Several experiments with superconductors have been performed and many of the fundamental quantum computing operations have been implemented [27, 135, 140]. There is even a company, D-Wave Systems<sup>8</sup>, that is aiming to produce 1000-bit quantum computers based on superconductors by the end of 2008. Superconductors are not considered in this thesis and are mentioned here for completeness. More details on superconducting quantum computers can be found in, e.g., Ref. [68] and the references therein.

With semiconductors, there are two basic proposals. The first uses the nuclear spin of an array of donor atoms as the qubit states [91]. An example setup would be  $^{31}\text{P}$  atoms embedded in a Si matrix (using a very pure sample of the spinless isotope  $^{28}\text{Si}$ ). While such a setup can be reliably fabricated, this design has the problem that manipulation of the nuclear spins is difficult.

The second proposal for semiconductor qubits, and the one that we investigated, is to use quantum dots. These are far easier to manipulate and measure since they use electron degrees of freedom. There are two categories of quantum dots: electronically-defined [21, 74, 93] and self-assembled [44, 113]. Electronically-defined quantum dots are formed at the interface between two semiconductors, often GaAs and AlGaAs [69, 77]. Because of the difference of Fermi levels of these materials, a two-dimensional electron gas (2DEG) forms at the interface [100]. Metallic electrodes are patterned onto the surface of one of the semiconductors using electron beam lithography. By setting the potential of these electrodes, regions can be made in the 2DEG where the electrons are depleted. If the potentials of the electrodes are carefully set, the 2DEG is almost completely depleted, with the exception of some small regions. These regions are the quantum dots; electrons can be injected into them, manipulated and then measured.

Self-assembled quantum dots are formed using the material properties of different semiconductors. No external electrodes are needed to form the quantum dot structure, although electrodes are usually present to manipulate and measure the electronic states. Self-assembled quantum dots form during growth of one substance on another, e.g. molecular beam epitaxial growth of a material on a substrate [122]. If the lattice vectors of the substrate and deposited material are not matched, homogeneous growth is unstable and the deposited medium forms small islands; these are the quantum dots. Self-assembled quantum dots are usually smaller than lithographically etched (electronically defined) dots, making them more useful for quantum computing and information storage. However, there is not much control in the final positions or sizes of the quantum dots.

Apart from there being two categories of quantum dots, there are two methods of *encoding information* in quantum dot qubits: spin and charge. In the case of spin, each electron is confined to a single quantum dot and its spin is the two-level system with which quantum computing proceeds. Information encoded in the electron spin is quite stable, since there are not many spin-flip mechanisms available, even in a solid. However, such information is more difficult to read and manipulate. Charge qubits, on the other hand, require a double quantum dot structure for each electron. The position of the electron, i.e. either in the “right” or “left” quantum dot, encodes the information. Because of the strong

---

<sup>8</sup>More information about this company can be found at their website: <http://www.dwavesys.com/>

interaction of charges with their surroundings, such a system will decohere quite rapidly. On the other hand, it is easy to manipulate and measure the position of an electron.

In chapters 6 and 7, I present our work on the dynamics of a spin within a qubit, with particular attention paid to the relaxation of this system. In addition, I present the results of our investigation on the relaxation of a charge qubit. These investigations were motivated by a set of experiments, the details of which are given in the next section.

### 1.3.3 Manipulating qubits

The stability of an electron spin in a quantum dot makes it an ideal candidate for quantum computing and quantum information processing [127]. This stability against various forms of dephasing and dissipation has been studied in great depth and there is a comprehensive literature on the calculation of relaxation rates (see, e.g., Refs. [56, 95, 96, 205] and the references therein). The difficulties involved in initializing, controlling and measuring a spin state are viewed as challenges, not fundamental obstructions, and much work has been done to overcome these [45, 46, 127, 165]. It is somewhat ironic, therefore, that the inspiration for our study of the dephasing of a spin in a quantum dot came from experiments unrelated to quantum computing. These experiments, detailed in Refs. [51, 71, 112, 113] and summarized below, dealt instead with the possibility of constructing spin-based, frequency-selective optical memory devices.

Frequency-selective optical memory using quantum dots was proposed in 1995 by Muto [137]. This proposal relies on the fact that self-assembled quantum dots vary considerably in shape and size and their optical properties change accordingly. More precisely, the inhomogeneous broadening of absorption by an array of self-assembled quantum dots is quite large (around 100 meV), while the homogeneous broadening for each quantum dot is much smaller (several  $\mu\text{eV}$ ). If an array of quantum dots is subjected to a laser field of a specific frequency, only some of the quantum dots absorb the light and produce excitons. The application of a strong electric field prevents recombination of these excitons by causing the hole to tunnel out of the dot. (The electron can be made to remain in the dot, as discussed below.) Thus the quantum dots that absorb at the frequency used will be charged and thereby store information. If this process is repeated with a different frequency, a different set of quantum dots are charged. Because of the large inhomogeneous broadening, very large multiplexing factors are possible: within the same laser spot, it is possible to store several hundred bits of information by using different frequencies.

This proposal was extended in 2004 by Kroutvar *et al.* to include spin selectivity [113], i.e. the quantum dots can be selectively occupied by an electron of a particular spin. This not only offers technological advantages but also provides a useful tool for studying spin dynamics in quantum dots [71]. In the 2004 experiments [113], the relaxation rate of the spins (explained below) was measured and compared to the theoretical result for relaxation via spin-phonon scattering [95, 96, 205]. The fit was reasonably good, but relied on parameters that could not be estimated accurately (e.g. the Dresselhaus velocity was given as  $\beta = (2 \pm 1) \times 10^5 \text{ cm s}^{-1}$ ). As such, we thought other dephasing mechanisms might also be relevant; the results of our investigation are presented in chapter 7.

Spin selectivity is made possible by using circularly polarized light. In the ideal situation, this proceeds as follows. The relevant exciton states consist of electrons ( $L = 0$ ,  $S = \hbar/2$ ,  $J_z = \pm\hbar/2$ ) and heavy holes ( $L = \hbar$ ,  $S = \hbar/2$ ,  $J_z = \pm 3\hbar/2$ ) [15]. Therefore there

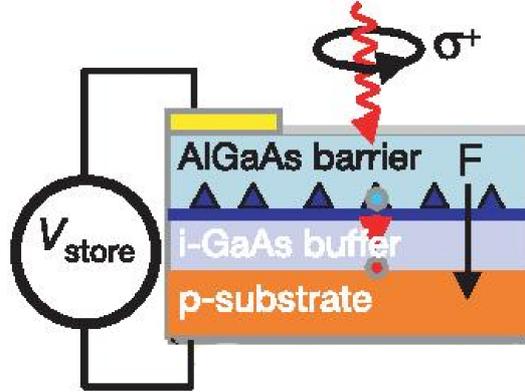


Figure 1.2: The experimental setup for a spin-selective optical memory using self-assembled quantum dots, taken from Ref. [113]. A magnetic field, not depicted, is applied vertically to split the optically active exciton states. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, *Optically programmable electron spin memory using semiconductor quantum dots*, M. Kroutvar, Y. Ducommun, D. Heiss, M. Bichler, D. Schuh, G. Abstreiter and J. J. Finley, **432**, 81 (2004). Copyright (2004).

are four possible excitonic combinations:  $e\uparrow h\downarrow$  ( $J_z = -\hbar$ ),  $e\downarrow h\uparrow$  ( $J_z = \hbar$ ),  $e\uparrow h\uparrow$  ( $J_z = 2\hbar$ ) and  $e\downarrow h\downarrow$  ( $J_z = -2\hbar$ ). The first pair of these are optically active, while the second pair are not. By using circularly polarized light of helicity  $\sigma^+$  or  $\sigma^-$ , one can selectively excite excitons with either  $J_z = +\hbar$  or  $J_z = -\hbar$ , respectively. The hole is removed using a bias field. Thus, depending on the helicity of the light, the dots are either excited with spin up or spin down electrons exclusively.

Of course, real solids are seldom ideal and quantum dots are no exception. The material used in these experiments, GaAs, is not inversion symmetric. In addition, self-assembled quantum dots are generally asymmetrically shaped. These asymmetries mix the two optically active exciton eigenstates, which therefore cannot be independently excited. Spin selectivity is made possible by the application of a large magnetic field (greater than 4 T), which separates the eigenstates.

The experimental setup is depicted in Fig. 1.2, which is taken from Ref. [113]. The quantum dots consist of small islands, roughly 30 nm wide and 5 nm high, of InGaAs grown on a GaAs substrate [71, 113]. A thin layer of GaAs coats the quantum dots, followed by a thick layer of AlGaAs. The structure is finished with a thin (5 nm) layer of titanium; this is thin enough to be transparent to the light.

In a typical experiment, the entire structure is inserted in a homogeneous magnetic field. A negative potential ( $V_{\text{bias}} = -1.2$  V) is applied to the titanium layer and a laser pulse of a particular helicity is sent through the structure. Because of the magnetic field, excitons of a well-defined angular momentum are excited by the light in some of the quantum dots (those with the appropriate absorption energy). The other quantum dots remain in their previous state. The electric field pulls the electron towards the titanium layer and pushes the hole in the opposite direction. However, the material is asymmetric: the hole can easily tunnel out of the dot, but the electrons have a very low probability of

tunneling through the AlGaAs. Thus the dot is charged with an electron of a well-defined spin.

To read the information stored in the structure, the bias field is removed, i.e. the potential of the titanium layer is set to zero. Within a few hundred nanoseconds, holes tunnel into the charged quantum dots, recombination occurs and a photon is emitted. Because each quantum dot has a well-defined energy, the resulting emission spectrum should coincide with the spectrum that was absorbed. The information that was stored can be recovered with the help of a spectrum analyzer. Furthermore, the photons will be left- or right-circularly polarized depending on the spins of the electrons that led to their emission. Information stored as spin can also be recovered.

This description omits the possibility of decay. If the delay between storing and reading the information is long enough, many electrons will have tunneled out of the dot into the titanium electrode. Once this occurs, the information is lost. It is also possible for individual electrons to tunnel into neighbouring dots. This would not only destroy the information stored in the first dot, but also generate a spurious signal from the second dot, i.e. incorrect information would be found in the read-signal. Fortunately, at low temperatures, this effect can be greatly reduced. It was found that below  $T = 50$  K, information is preserved over timescales of  $\tau \gg 1$  ms [71].

The spin is not so fortunate. Because of the strong magnetic field, there is a large Zeeman splitting between up and down electrons. Electrons excited into the lower energy state will keep the same spin, while electrons in the higher state will relax towards the lower. This was verified in the experiments. There is no noticeable change in the degree of polarization of the emitted light for electrons in the lower energy state, even for delay times up to 2 ms. On the other hand, if the higher energy state is excited, the polarization of the emitted light changes as a function of the delay time: for short times, the emitted light has predominantly the same helicity as the light that was absorbed, while for longer times, the emitted light contains an ever-increasing number of photons with the opposite helicity. From these results, the decay time of the spins can be estimated. As an example, the decay time of the higher energy spin state was approximately  $T_1 = 1.1 \pm 0.2$  ms at  $T = 1$  K, in the magnetic field  $B = 8$  T. I should note that the polarization of the emitted photons is never perfectly circular; there is always some admixture of the opposite helicity, even for very short delays.

Clearly, if information was stored in the spin state of the electron and this decays, information has been lost. Furthermore, the mechanism that enables decay to a lower energy state will also cause a loss of phase coherence with other qubits. If the above experimental method can be extended to allow the generation of entangled spins in the quantum dots, decay mechanisms will act to destroy this entanglement. The decay times, both the loss of energy (dissipation) and the loss of phase coherence (dephasing), are important for the application of this process to information storage and quantum computing. It is therefore important to know all the relevant mechanisms through which the spin can decay. This motivated the study presented in chapter 7.

Chapter 7 also addresses the dephasing and dissipation in charge qubits. This investigation was motivated by the experiments presented in Ref. [69]. For these experiments, a double quantum dot structure is defined electronically, i.e. using several gate electrodes at different potentials to deplete regions in the 2DEG formed at the interface between GaAs and AlGaAs. This is shown schematically in Fig. 1.3, which is reproduced with permission

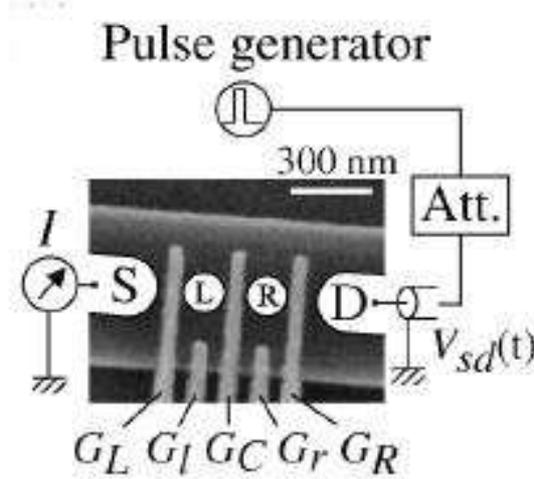


Figure 1.3: Schematic diagram of a double quantum dot, taken from Ref. [69]. The electrodes  $G_L$ ,  $G_l$ ,  $G_C$ ,  $G_r$  and  $G_R$  define the region of the quantum dots  $L$  and  $R$ . Electrons are provided by the source  $S$  and tunnel into the drain  $D$ . Figure reprinted with permission from T. Hayashi, T. Fujisawa, H. D. Cheong, Y. H. Jeong and Y. Hirayama, *Physical Review Letters* **91**, 226804 (2003). Copyright (2003) by the American Physical Society.

from Ref. [69]. The electrodes that define the quantum dot are labeled  $G_L$ ,  $G_l$ ,  $G_C$ ,  $G_r$  and  $G_R$ . The two quantum dots are referred to as the “left” and “right” dots and are labeled  $L$  and  $R$  accordingly. Finally, the source ( $S$ ) and drain ( $D$ ) electrodes are used to initialize and measure the system.

A typical experiment proceeds as follows. First, a large source-drain potential is applied. This leads to a current through the double quantum dot (DQD) structure. Parameters are chosen so that this current is limited to sequential tunneling; electrons hop from  $S$  to  $L$ , then into  $R$  and finally across into  $D$ . This process causes an excess electron to be found in  $L$ . The potential is then removed suddenly, which isolates the DQD system. The potentials of the two quantum dots become equal and coherent tunneling between them occurs. After a predefined delay time  $t_p = 80 - 2000$  ps, the bias field is restored and the current is measured. This current depends on whether the electron is found in  $L$  or  $R$  when the potential is applied. By measuring the dependence of the current on the delay time, oscillations between the quantum dots can be seen. The decay of these oscillations for large delay times yields the decoherence rate.

Because rectangular pulses are used to start the coherent tunneling and to perform the measurement, the time taken to switch from one stage of the experiment to another is negligible. This implies that decoherence and other quantum-mechanical processes are only relevant during the time when the coherent tunneling occurs; simulations need only include this stage of the dynamics.

A brief analysis of decoherence was presented in Ref. [69] along with the experimental results. The authors concluded that cotunneling was one of the main causes of decoherence, but that many other mechanisms were present and individual contributions could not be accurately estimated from the experiments. As for the dephasing of spin, several

theoretical studies of decoherence of charge states can be found in the literature (see, e.g., Refs. [68, 180] and the references therein). However, the mechanism studied in chapter 7, the interaction of the charge with the conduction electrons in the gate electrodes, has not previously been studied.

This concludes the introduction. The work on the relaxation of both spin and charge in quantum dots is contained in chapters 6 and 7, but first I present my work on ultrafast magnetic switching using the inverse Faraday effect.

## Chapter 2

# The Inverse Faraday Effect

The inverse Faraday effect (IFE) has been briefly introduced in the previous chapter. In this chapter, the history and modern understanding of the IFE is reviewed.

### 2.1 The Faraday effect

As the name suggests, the inverse Faraday effect is related to another physical effect: the Faraday effect. The Faraday effect is important, because it allows magnetization dynamics to be measured at a very high time resolution, as discussed further in chapter 5. The Faraday effect is the rotation of the polarization vector of linearly polarized light as it propagates through a transparent sample in a magnetic field [48, 49, 116]. Here, I present a brief derivation of this effect in the simplest case — an isotropic medium.

Consider the propagation of monochromatic light through a sample. As usual for high-frequency fields, the distinction between currents causing a magnetization and a time-dependent polarization density is not well-defined [116, 151], so introducing a magnetization due to the light would be “an unwarrantable refinement” [116]. In this case,  $\mathbf{H} = \mathbf{B}$  and the susceptibility  $\mu = 1$ . Magneto-optical effects are contained in the dielectric tensor.

If the external magnetic field  $\mathbf{H}_{\text{ext}}$  vanishes, then the condition of isotropy ensures that the dielectric tensor is diagonal and can be replaced by a scalar:  $\epsilon_{ij} = \epsilon\delta_{ij}$ . (The subscript “ext” distinguishes the applied magnetic field from the magnetic field of the light  $\mathbf{H}$ .) For a nonzero magnetic field, time-reversal symmetry forces  $\epsilon_{ij}(H_{\text{ext}}) = \epsilon_{ji}(-H_{\text{ext}})$ . Furthermore, to avoid absorption, the dielectric tensor must be Hermitian:  $\epsilon_{ij}(H_{\text{ext}}) = \epsilon_{ji}^*(H_{\text{ext}})$ . These symmetries, together with truncation at linear order in  $H_{\text{ext}}$ , lead to the dielectric tensor having the form

$$\epsilon_{ij} = \epsilon\delta_{ij} + i\gamma \begin{pmatrix} 0 & -H_{\text{ext}}^z & H_{\text{ext}}^y \\ H_{\text{ext}}^z & 0 & -H_{\text{ext}}^x \\ -H_{\text{ext}}^y & H_{\text{ext}}^x & 0 \end{pmatrix} \quad (2.1)$$

which can also be expressed as

$$\mathbf{D} = \epsilon\mathbf{E} - i\gamma\mathbf{E} \times \mathbf{H}_{\text{ext}}. \quad (2.2)$$

In general,  $\gamma H \ll \epsilon$  [210].

Assuming that the light propagates in the  $z$ -direction and that the magnetic field  $\mathbf{H}_{\text{ext}}$  is parallel to this direction,  $\mathbf{H}_{\text{ext}} = H_{\text{ext}}\hat{\mathbf{z}}$ , the normal modes for the propagation of light in the medium are easily found:

$$\begin{aligned}\mathbf{n}_+ &= \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, & k_+ &= k\sqrt{\epsilon + \gamma H_{\text{ext}}} \approx k\sqrt{\epsilon}\left(1 + \frac{\gamma H_{\text{ext}}}{2\epsilon}\right) \\ \mathbf{n}_- &= \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, & k_- &= k\sqrt{\epsilon - \gamma H_{\text{ext}}} \approx k\sqrt{\epsilon}\left(1 - \frac{\gamma H_{\text{ext}}}{2\epsilon}\right).\end{aligned}\quad (2.3)$$

Here and throughout this thesis,  $k = \omega/c$ . Assume that the light enters the sample linearly polarized:

$$\mathbf{E}_{z=0} = \frac{1}{2} \left\{ |E| \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix} e^{i\omega t} + c.c. \right\}, \quad (2.4)$$

where  $\theta$  is the angle between the polarization plane and the  $x$ -axis. Since the  $xy$ -plane is isotropic, this angle can be set to  $\theta = 0$ .

The reflection coefficients are approximately equal for the different normal modes and the electric field in the material is

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{4} T |E| \left( \mathbf{n}_+ e^{i(\omega t - k_+ z)} + \mathbf{n}_- e^{i(\omega t - k_- z)} + c.c. \right), \quad (2.5)$$

with the transmission coefficient  $T$  describing the loss of intensity due to reflection. Eq. (2.5) can be written in the more instructive form

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} T |E| e^{i(\omega t - k\sqrt{\epsilon}z)} \begin{pmatrix} \cos(VH_{\text{ext}}z) \\ \sin(VH_{\text{ext}}z) \\ 0 \end{pmatrix} + c.c. \quad (2.6)$$

where  $V = k\gamma/2\sqrt{\epsilon}$  is the *Verdet constant* [191]. It is clear from Eq. (2.6) that as the light propagates, it remains linearly polarized, but the plane of polarization rotates. The angle through which it has rotated after a distance  $z = d$  is

$$\beta = VH_{\text{ext}}d. \quad (2.7)$$

This rotation, discovered by Faraday in 1845 [48, 49], is referred to as the Faraday effect.

## 2.2 Energy considerations

An experimental verification of the Faraday effect requires only a magnet, a polarizer and a suitable medium. In practice, the third requirement can be ignored since all media will exhibit a Faraday effect provided the magnetic field is strong enough. The ready availability of these simple experimental tools led to the discovery of the Faraday effect in the middle of the nineteenth century. The inverse Faraday effect, on the other hand, required light of a very high intensity before it could be detected. The first experimental demonstration of the IFE in 1965 [187] was made possible by the invention of the laser.

The work that led to this discovery was part of a general search for nonlinear electro-magnetic effects in crystals. The argument motivating this search was that if two fields are simultaneously present in the crystal, with frequencies  $\omega_1$  and  $\omega_2$ , then an effect with the beat frequencies  $\omega_{\pm} = \omega_1 \pm \omega_2$  should be visible. In addition, if fields with frequencies  $\omega_1$  and  $\omega_{\pm}$  are used, an effect with the frequency  $\omega_2$  should be visible and *the magnitude of these two effects should be related*.

The first experimental observation of this type of relationship was the discovery that the linear electro-optic effect is related to the induction of a static polarization by a laser with the amplitude of polarization proportional to the intensity of the light [13]. The linear electro-optic effect is the change of the optical properties of a material when a static electric field is applied. A light beam passing through such a material has  $\omega_1 = \omega$ , the frequency of the light. The other relevant frequency is  $\omega_2 = 0$  because the applied electric field is static. From the above argument, effects can be expected with  $\omega_+ = \omega$ , i.e. at the fundamental frequency of the incident light. Such effects are interpreted as a change of the optical properties and are indeed observed. For the inverse effect, the induction of a static polarization in the presence of light,  $\omega_1 = \omega_+ = \omega$  now describe the input fields (the light, counted twice due to nonlinearity) and  $\omega_2 = 0$  is the frequency of the result. At first glance, these effects may seem to be unrelated. However, by taking nonlinearities into account (i.e. using the above frequency mixing arguments), relationships are found. Indeed, materials exhibiting a strong electro-optic effect also exhibit a strong induced polarization, as predicted by the theory of nonlinear electrodynamics [13].

In 1965, van der Ziel *et al.* considered a similar effect, but with a static *magnetic* field [187]. The propagation of light when such a field is applied is governed by the Faraday effect, as described in the previous section. They found, however, that when the external magnetic field is removed and the laser is circularly polarized, a static magnetization is induced with amplitude proportional to the intensity of the light. As expected by analogy to the electro-optic effect, the strength of this magnetization was found to be related to the strength of the Faraday rotation. Therefore they named the formation of this magnetization the “inverse Faraday effect”.

The possibility of such an effect had already been considered theoretically in 1961 by Pitaevskii [157] and a year later (independently) by Pershan [150], who showed the connection to the Faraday effect. The name “inverse Faraday effect” was first used in Ref. [187], which contained the first experimental demonstration of this effect.

To explain both the Faraday and inverse Faraday effects, the free energy of the material is assumed to include a “potential function” of the form

$$\Delta F = i\chi_{nm}H_{\text{ext}}(E_x E_y^* - E_x^* E_y). \quad (2.8)$$

Here,  $H_{\text{ext}}$  is a static magnetic field oriented along the  $z$ -direction and  $E_{x,y}$  are the components of a laser beam, propagating along the  $z$ -axis:

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2}(E_x \hat{\mathbf{x}} + E_y \hat{\mathbf{y}})e^{i(\omega t - kz)} + c.c. \quad (2.9)$$

The frequency of this beam is taken away from resonance, i.e. no light will be absorbed by the material. Note that for consistency with the rest of this thesis, Eq. (2.9) is different by a factor of 1/2 from the definition used in Refs. [150, 152, 187]; the following equations are changed appropriately.

The potential function (2.8) can be rigorously shown to exist in all materials [150], but the general arguments from which it is derived cannot be used to determine the magnitude of  $\chi_{nm}$  (which could even vanish). Therefore this parameter must be determined from microscopic calculations or by experiment. Note the subscript  $nm$  indicates that the sample is assumed to have no magnetic ordering.

The relationship between the potential function (2.8) and the Faraday effect is demonstrated by noting that the dielectric tensor  $\epsilon_{ij}(\omega)$  can be written in terms of the free energy of the system  $F$  as

$$\epsilon_{ij}(\omega) = 16\pi \frac{\partial^2 F}{\partial E_i \partial E_j^*}. \quad (2.10)$$

Therefore an additional energy of the form (2.8) leads to a change of dielectric function by

$$\Delta\epsilon_{xy}(\omega) = 16i\pi\chi_{nm}H_{\text{ext}}, \quad \Delta\epsilon_{yx}(\omega) = -16i\pi\chi_{nm}H_{\text{ext}} \quad (2.11)$$

Comparison with Eq. (2.1) shows that  $\gamma = -16\pi\chi_{nm}$ . The ‘‘magneto-optical susceptibility’’  $\chi_{nm}$  is related to the Faraday rotation angle. (Of course, this was the reason that a potential function of the form (2.8) was postulated.)

Eq. (2.10) follows from assuming that  $H_{\text{ext}}$  is fixed while  $E_{x,y}$  are variables. Equally, one could consider the laser to be fixed and the magnetic field to vary. Then the potential function causes an additional magnetic induction in the  $z$ -direction:

$$\frac{1}{4\pi}\Delta B = \frac{\partial\Delta F}{\partial H_{\text{ext}}}. \quad (2.12)$$

Assuming that the external field  $H_{\text{ext}}$  vanishes and that the material is nonmagnetic, we find from  $\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}$  that a magnetization is formed:

$$\mathbf{M} = i\chi_{nm}(E_x E_y^* - E_x^* E_y)\hat{\mathbf{z}}. \quad (2.13)$$

This magnetization is proportional to the intensity of the circularly polarized component of the light:  $i(E_x E_y^* - E_x^* E_y)$  is maximized for circularly polarized light and vanishes if the light is linearly polarized. Furthermore, it is linear in  $\chi_{nm}$  and is therefore directly proportional to the strength of the Faraday effect. Materials exhibiting a strong Faraday effect will also exhibit a strong inverse Faraday effect.

If no external magnetic field is present but the material has a static magnetization  $m$ , a potential function similar to (2.8) should be considered [98, 150]:

$$\Delta F = i\chi_f m (E_x E_y^* - E_x^* E_y) \quad (2.14)$$

or more generally

$$\Delta F = i\chi_f^{ijk} m_i E_j E_k^*, \quad (2.15)$$

where  $\chi_f^{ijk}$  is a third-rank tensor. The subscript  $f$  indicates that the material is ferromagnetically ordered. In this case, the inverse Faraday effect is the generation of an effective magnetic field by circularly-polarized light. For the simplest case, where the potential is given by Eq. (2.14), an equation similar to Eq. (2.13) is obtained:

$$\mathbf{H} = i\chi_f (E_x E_y^* - E_x^* E_y)\hat{\mathbf{z}}. \quad (2.16)$$

This equation will be rederived from angular momentum considerations in chapter 4 (compare Eq. (4.31) using  $\chi_f = -\gamma/16\pi$ ). With the potential (2.14), the Faraday effect keeps the same form (2.7), but now the Faraday rotation angle depends on the magnetization instead of an external field. This angle can be measured and is used to estimate the magnetization state of the material. (See chapter 5 for details.)

This derivation of the IFE is correct provided the system is in thermal equilibrium, the light is monochromatic and the light fields  $E_{x,y}$  inside the material are known [98,152]. For thin samples, the fields  $E_{x,y}$  can be taken as the external fields of the laser, as discussed in chapter 4. If the light is not monochromatic, e.g. in the case of ultrashort pulses, it is reasonable to expect that the effective magnetic field has the same time-dependence as the intensity of the light. However, this cannot be proven from thermodynamics. Furthermore, the effective magnetic field will cause precession of the magnetization (if the magnetization is orthogonal to the field), i.e. a deviation from thermal equilibrium. Finally, no clue is given within the thermodynamic theory for determining the parameters  $\chi_{nm}$  or  $\chi_f$ . To overcome these deficiencies, Pershan *et al.* derived the IFE from quantum-mechanical perturbation theory [152]. This work is summarized in the next section.

### 2.3 The effective Hamiltonian formulation

The starting point for the quantum-mechanical derivation of the IFE is to consider the interaction of light with a system with a well-defined “ground-state manifold” [32], i.e. several states lie very close in energy to the ground-state and there is a large gap between these states and higher energy states<sup>1</sup>. The frequency of the light is far from any resonance, so long-lasting transitions to excited states are rare. These excited states can therefore be removed from consideration. However, they are not simply neglected, but rather are used as virtual states to allow transitions between different ground states. This is referred to as constructing an “effective Hamiltonian” [32].

It should be noted that if the driving field (in this case the laser) is perfectly monochromatic, transitions between ground states are not possible unless the ground states have the same energies. Unless this condition is satisfied, the effective Hamiltonian only causes an energy shift of each ground state, although these energy shifts can be different. However, no field can be perfectly monochromatic unless it is present at all times, which is unreasonable. Switch-on effects yield a time dependence that can cause transitions [152,157].

Before presenting the details of the derivation, it is worth spending a little time reviewing virtual processes in quantum mechanics [117]. The state of an arbitrary quantum-mechanical system is given by a vector (which can be infinite-dimensional). The set of possible state vectors spans the underlying Hilbert space and this space admits a basis in which every vector is an eigenvector of the Hamiltonian, i.e. has a well-defined energy. This energy is related to the phase velocity of the vector. If the system starts in an energy eigenstate and is subjected to a perturbation of a given frequency,  $V(t) = V_0 \cos(\omega t)$ , conservation of energy makes it impossible for transitions to occur to states that differ from the original by an amount other than  $\pm\omega$ .

However, quantum mechanics also has uncertainty relations, in particular  $\Delta E \Delta t \geq \hbar$  [117]. This means that a system that should have energy  $E$  can occupy a state with the

<sup>1</sup>Terms like “close” and “large gap” are meant relative to the frequency of the light.

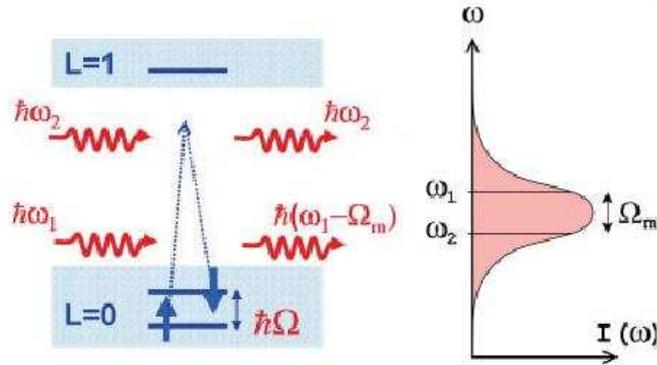


Figure 2.1: Spin-flip via stimulated Raman scattering, taken from Ref. [98]. The excited state  $|L = 1\rangle$  is not occupied for a significant length of time, but contributes as a virtual state. Reproduced with permission from *Femtosecond opto-magnetism: ultrafast laser manipulation of magnetic materials*, A. V. Kimel, A. Kirilyuk and Th. Rasing, *Laser & Photonics Reviews* 1, 275-287 (2007). Copyright Wiley-VCH Verlag GmbH & Co. KGaA.

classically unacceptable energy  $E + \Delta E$  for a length of time  $\Delta t$ , provided  $\Delta t < \hbar/|\Delta E|$ . If nothing occurs during this time, energy conservation asserts itself and the system moves back to the energy  $E$ . On the other hand, it is possible for the perturbation to interact with the system a second time while the system is in an “unacceptable” state. The system then moves into yet another state. If this state is still not energetically allowed, it will rapidly decay, but if it is energetically allowed, the system will remain in this new state until further perturbations act on it.

This is easily visualized using Fig. 2.1, which depicts a three-level system interacting with light and was taken from Ref. [98]. First, consider the case where the light is monochromatic,  $\omega_1 = \omega_2 = \omega$ , and the ground state is degenerate (i.e.  $\Omega_m = 0$  in Fig. 2.1). Assume the energy difference between the  $|L = 1\rangle$  and  $|L = 0\rangle$  states is  $\hbar(\omega + \Delta\omega)$  and define the zero of energy by the  $|L = 0\rangle$  state.

The system starts in the  $|L = 0, M_S = 1/2\rangle$  state. After absorbing a photon, the system should have energy  $\hbar\omega$ . However, there is no available state with this energy, so the system occupies the state  $|L = 1\rangle$  for a time  $\Delta t < 1/\Delta\omega$ . During this time, it can interact with the light again. If it absorbs another photon, then its energy is  $2\hbar\omega$ , which persists for time  $\Delta t < 1/(\omega - \Delta\omega)$ . More interesting is the case where a photon is emitted. If there is a mechanism by which the  $|L = 1\rangle$  spin states mix, the emission of a photon can place the system in the  $|L = 0, M_S = -1/2\rangle$  state. This state has the correct energy and can therefore persist indefinitely. A spin flip has occurred via a virtual transition.

If the ground state is non-degenerate, but the light is still monochromatic, then the spin-flipped state has the incorrect energy and will decay within time  $\Delta t < 1/\Omega_m$ . Finally, if the light has a linewidth greater than  $\Omega_m$ , photons of frequency  $\omega_1$  can excite the  $|L = 0, M_S = 1/2\rangle$  state to  $|L = 1\rangle$  and photons of frequency  $\omega_2 = \omega_1 - \Omega_m$  can stimulate emission down to the  $|L = 0, M_S = -1/2\rangle$  state. In this case, the energy of the final state is correct and this state persists indefinitely. This is stimulated Raman scattering.

Stimulated Raman scattering only occurs if the light is nonresonant, i.e. the frequency

of the light does not match the energy gap. If the light is resonant, the excited state is occupied for an arbitrary length of time. In a perfect solid, this corresponds to lifting electrons into the conduction band [100]. It is possible for these states to undergo stimulated emission back to their ground states, but it is more likely that they will scatter before this occurs. In this case, a nonequilibrium distribution of electrons forms and the dynamics are far more difficult to analyze.

It is appropriate to mention here that virtual processes are also responsible for the formation of the dipole moment of an atom in the presence of a nonresonant electric field. Consider a hydrogen atom. In the ground state, the electron occupies the  $1s$  orbital and the system has no dipole moment. If a nonresonant electric field is applied, one might naively expect (from Fermi's golden rule) that it has no effect as it cannot cause transitions. However, it can virtually excite the electron into the  $2p$  states. This virtual excitation does not last long because of the energy mismatch, but if a time-average is taken, the virtual excitation will be seen. The electron spends a small but well-defined percentage of its time in the  $2p$  state. This causes the dipole moment.

Virtual processes allow energetically inaccessible states to influence the dynamics. However, the length of time spent in such states is usually negligible in terms of the experimental timescale. Particularly for the IFE, these states are only important because they allow transitions between low-energy states that seem disconnected. The removal of these excited states via the introduction of an effective Hamiltonian is now discussed.

Following Ref. [152], the total Hamiltonian is  $H = H_0 + V(t)$ . The perturbation  $V(t)$  is assumed to be a slowly modulated periodic function,  $V(t) = v(t)e^{i\omega t} + c.c.$  that does not cause direct transitions within the ground-state manifold. In the interaction representation, the Schrödinger equation becomes

$$i\hbar \frac{d}{dt} \tilde{\psi}(t) = \tilde{V}(t) \tilde{\psi}(t), \quad (2.17)$$

where

$$\tilde{\psi}(t) = e^{iH_0 t/\hbar} \psi(t), \quad (2.18)$$

and

$$\tilde{V}(t) = e^{iH_0 t/\hbar} V(t) e^{-iH_0 t/\hbar}. \quad (2.19)$$

Eq. (2.17) can be solved iteratively:

$$\tilde{\psi}(t) = \left[ 1 - \frac{i}{\hbar} \int_0^t \tilde{V}(t') dt' - \frac{1}{\hbar^2} \int_0^t \tilde{V}(t') \int_0^{t'} \tilde{V}(t'') dt'' dt' + \dots \right] \psi(0). \quad (2.20)$$

Now an effective Hamiltonian  $H_{\text{eff}}$  is introduced such that it has exactly the same effect on the ground-state manifold as the full Hamiltonian  $H$ , but does not couple this manifold to the excited states [32]. For states  $|a\rangle, |b\rangle$  within the ground-state manifold,  $H_{\text{eff}}$  must satisfy

$$\begin{aligned} \langle a | 1 - \frac{i}{\hbar} \int_0^t H_{\text{eff}}(t') dt' + \dots | b \rangle &= \langle a | 1 - \frac{i}{\hbar} \int_0^t \tilde{V}(t') dt' \\ &\quad - \frac{1}{\hbar^2} \int_0^t \tilde{V}(t') \int_0^{t'} \tilde{V}(t'') dt'' dt' + \dots | b \rangle. \end{aligned} \quad (2.21)$$

Since  $V(t)$  does not induce direct transitions between states in the ground-state manifold, i.e.  $\langle a|\tilde{V}(t)|b\rangle = 0$ , we find that

$$\begin{aligned}\langle a|H_{\text{eff}}(t)|b\rangle &= -\frac{i}{\hbar}\langle a|\tilde{V}(t)\int_0^t\tilde{V}(t')dt'|b\rangle \\ &= -\frac{i}{\hbar}\sum_n\langle a|\tilde{V}(t)|n\rangle\int_0^t\langle n|\tilde{V}(t')|b\rangle dt',\end{aligned}\quad (2.22)$$

where the sum is over all excited states  $|n\rangle$ . (The resolution of the identity  $1 = \sum_n |n\rangle\langle n|$  has been used here.)

Until now, the envelope of the perturbation,  $v(t)$ , has been completely general. To describe an electric field, it is set to

$$v(t) = e(r_+E_R - r_-E_L), \quad (2.23)$$

where the notation of Ref. [152] is used:  $r_{\pm} = \mp(x \pm iy)/\sqrt{2}$ ,  $E_R = (E_x - iE_y)/\sqrt{2}$  and  $E_L = (E_x + iE_y)/\sqrt{2}$ . Introducing the additional notation  $X_{ab} = \langle a|X|b\rangle$  for any operator  $X$  and  $\omega_{ab} = (E_a - E_b)/\hbar$  for the differences between the energies of states, and assuming that the time dependence of  $E_{R,L}$  is negligible, we find that

$$\begin{aligned}(H_{\text{eff}})_{ab} &= \frac{e^2}{\hbar}[E_RE_R^* - E_LE_L^*]\sum_n\frac{\omega}{\omega_{nb}^2 - \omega^2}[(r_+)_{an}(r_-)_{nb} - (r_-)_{an}(r_+)_{nb}] \\ &\quad + \frac{e^2}{\hbar}[E_RE_R^* + E_LE_L^*]\sum_n\frac{\omega_{nb}}{\omega_{nb}^2 - \omega^2}[(r_+)_{an}(r_-)_{nb} + (r_-)_{an}(r_+)_{nb}] \\ &\quad - 2\frac{e^2}{\hbar}E_RE_L^*\sum_n\frac{\omega_{nb}}{\omega_{nb}^2 - \omega^2}(r_+)_{an}(r_+)_{nb} \\ &\quad - 2\frac{e^2}{\hbar}E_R^*E_L\sum_n\frac{\omega_{nb}}{\omega_{nb}^2 - \omega^2}(r_-)_{an}(r_-)_{nb}.\end{aligned}\quad (2.24)$$

Eq. (2.24) is not particularly instructive. To simplify it, the assumption is made that  $J$  is a good quantum number for the ground state manifold. This is reasonable for magnetic systems (see Ref. [152] for a discussion). Thus every state in the ground state manifold can be written as  $|a\rangle = |0JM\rangle$ , where the magnetic quantum number  $M$  differentiates between states within this manifold. Making use of the Clebsch-Gordon coefficients for vector operators (see, e.g., Ref. [117]) and using the identification  $M\delta_{MM'} = \langle 0JM|J_z|0JM'\rangle$  yields

$$\begin{aligned}H_{\text{eff}} &= A[E_RE_R^* - E_LE_L^*]J_z \\ &\quad + C\left\{(E_RE_R^* + E_LE_L^*)\left[J_z^2 - \frac{1}{3}J(J+1)\right] - E_LE_R^*J_-^2 - E_L^*E_RJ_+^2\right\},\end{aligned}\quad (2.25)$$

with the coefficients

$$\begin{aligned}A &= \frac{2e^2}{\hbar}\sum_{\alpha}\left[\frac{\omega\langle 0J||r||\alpha J-1\rangle\langle\alpha J-1||r||0J\rangle}{(2J+1)2J(\omega_{\alpha J-1,0J}^2 - \omega^2)}\right. \\ &\quad - \frac{2\omega\langle 0J||r||\alpha J\rangle\langle\alpha J||r||0J\rangle}{(2J+2)(2J+1)2J(\omega_{\alpha J,0J}^2 - \omega^2)} \\ &\quad \left. - \frac{\omega\langle 0J||r||\alpha J+1\rangle\langle\alpha J+1||r||0J\rangle}{(2J+2)(2J+1)(\omega_{\alpha J+1,0J}^2 - \omega^2)}\right]\end{aligned}\quad (2.26)$$

and

$$\begin{aligned}
C = \frac{2e^2}{\hbar} \sum_{\alpha} \left[ \frac{\omega_{\alpha J-1,0J} \langle 0J || r || \alpha J - 1 \rangle \langle \alpha J - 1 || r || 0J \rangle}{(2J+1)2J(2J-1)(\omega_{\alpha J-1,0J}^2 - \omega^2)} \right. \\
+ \frac{2\omega_{\alpha J,0J} \langle 0J || r || \alpha J \rangle \langle \alpha J || r || 0J \rangle}{(2J+2)(2J+1)2J(\omega_{\alpha J,0J}^2 - \omega^2)} \\
\left. + \frac{\omega_{\alpha J+1,0J} \langle 0J || r || \alpha J + 1 \rangle \langle \alpha J + 1 || r || 0J \rangle}{(2J+3)(2J+2)(2J+1)(\omega_{\alpha J+1,0J}^2 - \omega^2)} \right]. \quad (2.27)
\end{aligned}$$

Here,  $\langle 0J || r || \alpha J' \rangle$  is a reduced matrix element, which is independent of  $M$  [117].

The first term of the effective Hamiltonian (2.25) describes the effective magnetic field of the IFE. The other terms describe quadratic effects (e.g. the optical Stark effect [124]) and cannot be related to a magnetic field. However, quadratic effects are usually smaller than linear effects [152], so the dominant effect of the effective Hamiltonian is magnetic.

This derivation places the IFE on a more rigorous theoretical footing. From it, it is possible to investigate the effect of modulating the envelope of the laser field. One can also study how a nonequilibrium situation affects the IFE. Furthermore, estimates of Eq. (2.26) can be made, allowing the approximate evaluation of the strength of the IFE in different materials. For more details on these topics, see Ref. [152].

The calculation of  $A$  from Eq. (2.26) is similar to the calculation of a dielectric constant. That this should be so is clear from the derivation of the IFE in the previous section, where the IFE is related to the Faraday effect, which is determined by the dielectric constant. Calculations of the dielectric constant are notoriously difficult [130], but the accuracy is improving. As improvements are made (largely thanks to the development of supercomputers [166]), it becomes ever more likely that the strength of the IFE will be accurately calculated. Calculations of related quantities (the strength of the Faraday effect in small diamagnetic molecules) have been performed using time-dependent density functional theory, coupled cluster expansion and the multiconfiguration self-consistent field (MCSCF) method [19, 33, 83, 114], but so far no indication has been given that such calculations can be performed for solids.

The only fundamental question concerning the IFE that remains open concerns the conservation of angular momentum. The magnetic field in Eq. (2.25) has been derived from only the interaction between light and electrons in the ground state. If the spins of these electrons precess because of the magnetic field, it seems that the light must provide the angular momentum needed for precession. But in Ref. [98], it is claimed that this is not true, that the light does not change polarization and therefore cannot change its angular momentum. The authors of Ref. [98] argue that the IFE provides an effective magnetic field in the sense of making certain states *energetically* preferable, but provides no angular momentum. The angular momentum required for precession must be found elsewhere.

The next two chapters are devoted to understanding the role of the light as an angular momentum reservoir. Before this, I shall briefly review the details of the Nijmegen experiments that inspired much of this work [99]. These experiments will often be referred to in the following chapters.

## 2.4 The ultrafast inverse Faraday effect

The first experiments that demonstrated the IFE were performed in 1965 using a ruby laser [187]. These used relatively long pulses (30 ns), partly because the aim was to demonstrate the existence of the IFE and partly because shorter pulses were not available at the time. The primary material studied was europium-doped calcium fluoride, i.e.  $\text{Eu}^{2+}:\text{CaF}_2$ , but several other glasses as well as some liquids were also investigated [187].

Surprisingly, no further experiments of the IFE in condensed matter were performed until 1995 [163]. The most similar experiments that can be found in the literature are the investigation of the influence of circularly polarized light on plasmas [22, 34] and the discovery of the “specular inverse Faraday effect”, where circularly polarized light changes the reflection properties of a solid [115, 159]. One possible reason for the lack of interest in the IFE in solids was that no practical use for a light-induced magnetic field had been conceived. Once the possibility of ultrafast optical manipulation of magnetization was demonstrated [16], it was only a matter of time before attention returned to the IFE.

However, the experiments described in Ref. [163] and in the follow-up work of Ref. [164] were not connected to ultrafast magnetization reversal. Rather, they were undertaken to investigate the possible existence of the “axial magnetostatic field of the photon”<sup>2</sup> proposed by Evans and Vigier [47]. Their results confirmed that the IFE could be attributed to the effective magnetic field described in the previous sections; the existence of a magnetostatic field of the photon in vacuum has no experimental support [164]. While these results were both interesting and important, they provided no indication that the IFE could be technologically useful. The introduction of the IFE into the field of ultrafast control of magnetization came later, with the work of Kimel *et al.* [99].

The experiments in Ref. [99] demonstrated the ultrafast IFE in  $\text{DyFeO}_3$ . This material belongs to the group of rare-earth orthoferrites and has a perovskite-type structure. The dysprosium ions occupy very high spin-states, but these are not magnetically ordered above 4 K. The iron spins are antiferromagnetically ordered. Due to the Dzyaloshinskii-Moriya interaction [41, 136], there is a slight canting of oppositely oriented iron spins above 50 K [200]. This yields an equilibrium magnetization of around 8 G.  $\text{DyFeO}_3$  exhibits a strong Faraday effect, although this is complicated by the presence of optical birefringence. (The Faraday effect in  $\text{DyFeO}_3$  is investigated in chapter 5.) The strength of the Faraday effect indicates that  $\text{DyFeO}_3$  should exhibit a strong IFE.

Samples were cut to a thickness of around  $60\ \mu\text{m}$  and irradiated with 200 fs laser pulses at a photon energy of 1.55 eV. A linearly polarized probe pulse with intensity about 1/100 that of the pump pulse was used to measure the magnetization dynamics via the Faraday effect (see chapter 5). The measured Faraday rotation indicates the instantaneous magnetization in the material. By changing the delay between the pump and probe pulses, the time evolution of the magnetization can be found. The fluence of the pump pulse was around  $30\ \text{mJ}/\text{cm}^2$  and the repetition rate was 1 kHz. For more details, see Ref. [99].

A typical result of this experiment is given in Fig. 2.2, which is taken from Ref. [99]. In this figure,  $\sigma_+$  and  $\sigma_-$  indicate the different helicities of light. The dependence on the helicity is clear. The magnetization is perturbed by the light, after which it starts

---

<sup>2</sup>According to Evans and Vigier [47], the photon has an axial magnetic field, even in vacuum, and this field causes the inverse Faraday effect. This contrasts with the classical understanding of the IFE, derived above, which shows that the IFE is caused by the interaction of the photons with the material.

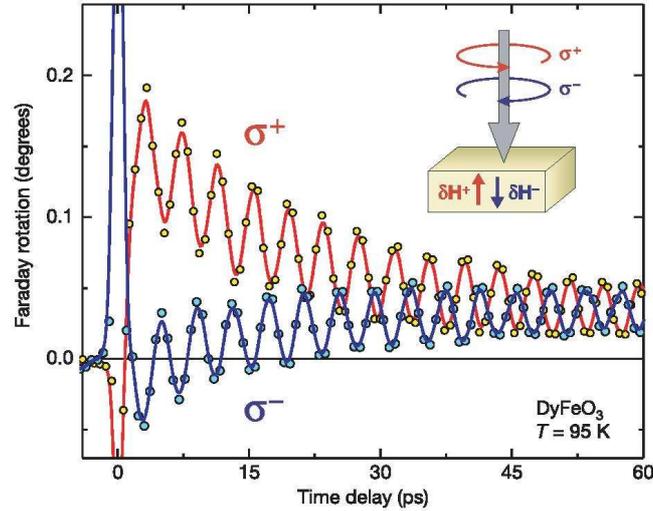


Figure 2.2: Magnetic excitations in  $\text{DyFeO}_3$  probed by the Faraday effect. This figure is taken from Ref. [99].  $\sigma_+$  and  $\sigma_-$  indicate the different helicities of light, which clearly have different effects. Both a rapid change of magnetization as well as magnetic oscillations are visible. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, *Ultrafast non-thermal control of magnetization by instantaneous photomagnetic pulses*, A. V. Kimel, A. Kirilyuk, P. A. Usachev, R. V. Pisarev, A. M. Balbashov and Th. Rasing, **435**, 655 (2005). Copyright (2005).

to decay back towards the ground state. During this decay, magnetic oscillations occur. Two types of decay are seen: decay of the amplitude of oscillations and decay of the mean magnetization.

In Ref. [99], the effective magnetic field from the IFE was estimated to be 0.3 T, with the possibility of being increased to 5 T before the damage threshold is reached (although later estimates place the maximum field at 10 T [98]). In agreement with the results from the previous sections, the strength of the effective magnetic field was shown to be linear in the intensity of the electric field. In addition, the magnetization dynamics were found to be consistent with a weakly perturbed equilibrium state: the IFE does not cause significant departures from equilibrium and is a truly *nonthermal* effect. Full reversal of the magnetization could not be attained with these pulse strengths and durations, but control of the magnetization was demonstrated [66]. Further results from recent IFE experiments can be found in the review [98]. All-optical, nonthermal magnetization reversal has not yet been demonstrated, but it seems to be just a matter of time.

## Chapter 3

# Conservation of Angular Momentum: Small Systems

### 3.1 Introduction

The importance of angular momentum transfer in magnetic switching has already been mentioned in Chapter 1, but the primary message is worth repeating: *the rate of angular momentum transfer into the magnetization places a fundamental limit on the speed of magnetic switching*. This rate is determined by the reservoir used to supply angular momentum. The maximum speed of switching can only be estimated once the active angular momentum reservoir has been identified.

As with other optical methods of magnetic switching, the source of angular momentum in the IFE has not been identified<sup>1</sup>, despite much theoretical and experimental work [66, 99, 150, 152, 179, 187]. However, the nonthermal nature of the IFE simplifies the discussion, since several of the reservoirs discussed in Chapter 1 *cannot* play a role. Electrons are not excited to higher orbital states except virtually, which means that the electrons are almost always near the ground state. Under the assumption that the ground-state manifold consists only of spin excitations (this will be assumed throughout this chapter), it follows that spin cannot be converted into orbital angular momentum. Angular momentum transfer occurs too rapidly for phonons to be relevant; the proposed mechanisms for increasing the spin-phonon interaction [103, 178] are not applicable, since the electron population is very close to equilibrium.

In this chapter, I show that the angular momentum is provided by the *photons*, i.e. the IFE both drives magnetic switching and provides the necessary angular momentum. This implies that nonthermal switching can occur within the timescale of a laser pulse, provided a material with strong enough magneto-optical response can be found. Because switching occurs via precession and is much faster than relaxation, the magnetization can be placed arbitrarily close to the switched equilibrium state. Thus the IFE should provide much faster switching than other optical methods of magnetization reversal, which require

---

<sup>1</sup> There is some confusion about this in the literature, due to an unfortunate naming convention. In plasma physics, the term “IFE” describes any magnetic field generated by circularly polarized light, even when the light is absorbed. Angular momentum transfer in this case has been investigated [65, 106], and it has been shown that the absorption of photons transfers angular momentum into the plasma. This result is not relevant to the nonresonant IFE, which is the subject of this thesis.

long times to dissipate heat and to relax to equilibrium.

Throughout this chapter, I make several assumptions concerning the system being studied. These assumptions are made to prevent other physical effects from obscuring the IFE. Specifically, I assume that circularly polarized light is scattering from a small sample, i.e. the sample is much smaller than both the wavelength of the light and the width of the laser beam. The passage to larger samples is described in Chapter 4. Furthermore, I assume that all states in the ground-state manifold have equal energy. This allows unrestricted motion between these states even for perfectly monochromatic light. Finally, I assume that the ground state manifold consists only of magnetic excitations; this allows the magnetic information to be highlighted and removes complications due to rotational and vibrational degrees of freedom. The inclusion of rotational and vibrational modes (corresponding to phonon modes in solids) is, to a first approximation, irrelevant for magneto-optical effects. These modes will give a small correction to magneto-optical effects in higher order perturbation theory, but this is beyond the scope of this thesis.

## 3.2 Angular momentum of light

It has been asserted that in the IFE, the angular momentum of the photons does not change [98]. This conclusion was reached because the IFE is caused by Raman scattering, which conserves both photon number and polarization. However, it is shown below that spin-orbit coupling causes the scattered photons to exhibit a spin-dependent change of *propagation direction*. This changes their angular momentum, which is parallel to the direction of motion. Thus the photons can provide angular momentum, even though their polarization is unchanged and they are not absorbed.

Considered classically, a change of photon propagation direction is a shift of the intensity distribution of the laser. Although this shift can be related to the angular momentum, it is simpler to calculate the radiated angular momentum directly. For an electromagnetic field interacting with a material, conservation of angular momentum is described by [82]

$$\frac{d}{dt} \int_V (\mathcal{L}_{\text{field}} + \mathcal{L}_{\text{mech}})_i dV = -\epsilon_{ilm} \int_S T_{jl} r_m dS_j, \quad (3.1)$$

where  $\epsilon_{ilm}$  is the completely antisymmetric unit tensor of third rank, repeated indices are summed over,  $\mathcal{L}_{\text{field}}$  and  $\mathcal{L}_{\text{mech}}$  are the angular momentum densities for the electromagnetic field and the sample, respectively, and  $T_{jl}$  is the Maxwell stress tensor:

$$T_{jl} = \frac{1}{8\pi} \left[ 2E_j E_l + 2H_j H_l - \delta_{jl} (E^2 + H^2) \right].$$

Because the light is nonresonant, only motion within the ground-state manifold of the material is allowed: for a small magnetic sample, the mechanical angular momentum  $\int_V \mathcal{L}_{\text{mech}} dV$  can be replaced by the spin  $\mathbf{S}(t)$ . Note that the right-hand side of Eq. (3.1) includes both the spin and orbital angular momentum of the light [4]; for problems with a single scattering site, there is no need to separate these contributions.

The electric and magnetic fields entering  $T_{jl}$  are superpositions of the homogeneous laser field

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( \mathbf{E}_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} + c.c. \right), \quad \mathbf{H}(\mathbf{r}, t) = \frac{\mathbf{k}}{k} \times \mathbf{E}(\mathbf{r}, t) \quad (3.2)$$

and the radiation from the sample, which is treated as a dipole

$$\mathbf{d}(t) = \frac{1}{2} \left[ \mathbf{d}_0(\omega, t) e^{i\omega t} + c.c. \right]. \quad (3.3)$$

Here, *c.c.* means complex conjugate. Through spin-orbit coupling,  $\mathbf{d}_0(\omega, t) \equiv \mathbf{d}_0(\omega, \mathbf{S}(t))$  depends on the magnetization, which evolves much slower than the period of the light.

Eq. (3.1) is a continuity equation, showing that the rate of change of angular momentum in the volume  $V$  is equal to the flow of angular momentum across the boundary  $S$ . By continuity, angular momentum lost by the spin  $\mathbf{S}(t)$  is not immediately radiated across the boundary, but flows first into the field within  $V$ , reaching the boundary only after some delay. To circumvent this retardation, the integrals are evaluated assuming that  $\mathbf{d}_0$  is constant. Then  $\mathcal{L}_{\text{field}}$  is periodic with period  $2\pi/\omega$ , and its contribution vanishes when Eq. (3.1) is averaged over one period of the light. After performing the surface integral over an infinitely distant surface, the time-dependence of  $\mathbf{d}_0$  is reintroduced at the nonretarded time. In this way, any change of spin angular momentum is immediately compensated by the radiation, without the need for complicated volume integrals<sup>2</sup>.

The surface integral in Eq. (3.1) is calculated using the classical expressions for the electric and magnetic fields of a dipole (see, e.g., Ref. [82]), yielding

$$\frac{d}{dt} \mathbf{S}(t) = \frac{1}{2} \text{Re} [\mathbf{d}_0(\omega, t) \times \mathbf{E}_0^*] = \langle \mathbf{d}(t) \times \mathbf{E}(\mathbf{r}, t) \rangle. \quad (3.4)$$

Here, the brackets  $\langle \dots \rangle$  indicate time-averaging over one period of the light.  $\mathbf{S}(t)$  and  $\mathbf{d}_0(\omega, t)$  evolve very slowly, and their time-averaging can be neglected. Eq. (3.4) describes the angular momentum radiated in the interference flux, i.e. only terms linear in both  $\mathbf{E}$  and  $\mathbf{d}$  are retained. Terms in  $\mathbf{E}^2$  vanish, while terms in  $\mathbf{d}^2$  are negligible because induced fields are much weaker than driving fields.

The accuracy of Eq. (3.4) in describing the IFE is limited by the validity of the dipole approximation (3.3) and the assumption that  $\mathbf{S}(t)$  has no spatial dependence; both fail for extended systems. However, the linear dependence on  $\mathbf{d}(t)$  in Eq. (3.4) implies that the radiation of angular momentum from an extended system is equal to the sum of the radiation from the constituents. Provided dipole moments and spin moments can be sensibly defined for the constituents, the results found here are applicable to large systems (see chapter 4 for more details).

It is appropriate to briefly consider the assumption that the precession of  $\mathbf{S}(t)$  and the associated change of  $\mathbf{d}_0(\omega, t)$  can be neglected. The validity of this is more easily shown in the frame where  $\mathbf{S}(t)$  is stationary and the driving electric field  $\mathbf{E}(\mathbf{r}, t)$  rotates. The spatial dependence is neglected and the electric field is written as  $\mathbf{E}(t) = e^{i\omega t} \mathbf{f}(t) + c.c.$ , where  $\mathbf{f}(t)$  changes slowly in time. The general definition of a dipole moment (see, e.g., Ref. [116]) is

$$d_i(t) = \int_{-\infty}^{\infty} \alpha_{ij}(t-t') E_j(t') dt', \quad (3.5)$$

---

<sup>2</sup>Expressed another way, the volume of integration in Eq. (3.1) is taken to be large enough that only radiation effects are seen, but small enough that the time taken for this radiation to reach the integration surface is negligible.

where the time-dependent polarizability  $\alpha_{ij}(t-t')$  vanishes for  $t < t'$  due to causality and decays rapidly to zero as  $(t-t') \rightarrow \infty$ . Then since  $\mathbf{f}(t)$  changes slowly in time, we find

$$d_i(t) = \alpha_{ij}(\omega)E_j(t) + \left\{ i e^{i\omega t} \frac{d}{dt} f_j(t) \frac{d}{d\omega} \alpha_{ij}(\omega) + c.c. \right\}. \quad (3.6)$$

Here,  $\alpha_{ij}(\omega) = \int_{-\infty}^{\infty} \alpha_{ij}(t) e^{-i\omega t}$  and terms containing higher than the first derivative of  $f_j(t)$  are neglected. Clearly, precession of  $\mathbf{S}(t)$  can be ignored provided  $|(df_j(t)/dt) (d\alpha_{ij}(\omega)/d\omega)|$  is small. In dimensionless quantities, this condition becomes

$$\left| \frac{d}{dt} f_j(t) \frac{d}{d\omega} \alpha_{ij}(\omega) \right| / \sum_j |\alpha_{ij}(\omega) f_j(t)| \ll 1. \quad (3.7)$$

Note that the sum in the denominator ensures that the components of the dipole moment are compared to the total “length” of the dipole moment. This reproduces the correct result that, e.g., the vectors  $(0, 0, 1)$  and  $(0, 10^{-6}, 1)$  are considered close to each other. (If considered component-wise, these vectors would appear very different.)

Far from a resonance, the polarizability  $\alpha_{ij}(\omega) \sim 1/(\omega - \omega_0)^2$ , where  $\omega_0$  is the nearest resonance [100, 138]. Thus Eq. (3.7) becomes  $T|\omega - \omega_0| \gg 1$ , where  $T$  is the timescale of precession. Assuming  $T \sim 100$  fs, which is extremely fast precession, we are justified in neglecting this precession as long as  $\omega$  is much more than 10 meV from resonance. This condition is easily satisfied. Note also that Eq. (3.7) arises when considering whether a pulse can be treated adiabatically or not. In this case, a realistic rise time for a pulse is  $T \sim 20$  fs and  $\omega$  must be more than 50 meV from resonance. Again, this condition is easily met, so the adiabatic limit will be assumed in all that follows.

### 3.3 Effective magnetic field

The expression  $\langle \mathbf{d}(t) \times \mathbf{E}(\mathbf{r}, t) \rangle$  in Eq. (3.4) is the time-averaged torque on a dipole in an oscillating electric field (see, e.g., Ref. [116]). The appearance of this torque is not surprising, since the total angular momentum radiated by a dipole must be compensated by the torque on that dipole<sup>3</sup>. The new feature described by Eq. (3.4) is that this torque causes magnetization dynamics. Eq. (3.4) leads to a novel interpretation of the inverse Faraday effect: light induces a dipole moment, which is deflected from the plane in which the electric field oscillates by spin-orbit coupling. This deflection leads to a torque on the dipole moment by the light, causing motion within the ground-state manifold, i.e. rotation of the spin.

Expressed in terms of an effective magnetic field, Eq. (3.4) becomes

$$g\mu \mathbf{B}_{\text{eff}} \times \mathbf{S}(t) = \frac{1}{2} \text{Re} [\mathbf{d}_0(\omega, t) \times \mathbf{E}_0^*]. \quad (3.8)$$

Here and in the next chapter, the gyromagnetic ratio is represented in the form suitable for atomic physics, i.e.  $g\mu$  [100].

---

<sup>3</sup>This is a very general statement, since torque and change of angular momentum are related by definition. However, the explicit calculation showing that the torque on a dipole in an electric field is equal and opposite to the radiation of angular momentum from the dipole was not performed until 1982 [107], and even that calculation was not for the general case.

Eq. (3.8) can only be used to recover the components of  $\mathbf{B}_{\text{eff}}$  orthogonal to the spin. This should come as no surprise: Eq. (3.8) was derived from dynamical considerations, while the component of  $\mathbf{B}_{\text{eff}}$  parallel to  $\mathbf{S}(t)$  does not influence the dynamics. To calculate this component, conservation of energy must be invoked.

The energy of a dipole in an electric field is  $U = -\frac{1}{2}\mathbf{d}(t) \cdot \mathbf{E}(\mathbf{r}, t)$  (see, e.g., Ref. [116]; the factor  $\frac{1}{2}$  arises because the dipole is induced by the field). This energy contains both a spin-dependent and a spin-independent contribution; the latter is removed by subtracting the spin-averaged dipole moment and the remainder is identified with the Zeeman energy:

$$g\mu\mathbf{B}_{\text{eff}} \cdot \mathbf{S}(t) = -\frac{1}{4}\text{Re} \left[ (\mathbf{d}_0(\omega, t) - \bar{\mathbf{d}}_0(\omega)) \cdot \mathbf{E}_0^* \right]. \quad (3.9)$$

Here  $\bar{\mathbf{d}}_0(\omega)$  is the time-independent, spin-averaged dipole moment,  $\bar{\mathbf{d}}_0(\omega) = \int_{\mathcal{M}} \mathbf{d}_0(\omega, \mathbf{S}) d\mu$ , where  $\mathcal{M}$  is the ground-state manifold and  $d\mu$  is the appropriate normalized measure,  $\int_{\mathcal{M}} d\mu = 1$ . (See, e.g., Ref. [195] for details of integration over manifolds.) As above, rapidly-oscillating terms have been removed by averaging over one period of the light.

Eq. (3.9) is not as accurate as Eq. (3.8). Eq. (3.8) describes the torque on the system, which (because of the assumptions made) can only be ascribed to an effective magnetic field. However, Eq. (3.9) describes the energy. Although the spin-independent energy has been subtracted, there will also be energetic contributions proportional to  $\mathbf{S}^2$  (the optical Stark effect [124]) and possibly even higher powers of  $\mathbf{S}$ . These terms have not been subtracted, but must be if the Zeeman energy is to be correctly identified. However, such terms are small and Eq. (3.9) is a reasonable approximation. Note that these additional terms also arose in the effective Hamiltonian formulation, i.e. in Eq. (2.25).

From Eqs. (3.8)-(3.9) the effective magnetic field is

$$g\mu\mathbf{B}_{\text{eff}} = \frac{1}{4S^2}\text{Re} \left\{ 2\mathbf{S} \times (\mathbf{d}_0 \times \mathbf{E}_0^*) - \mathbf{S} \left[ (\mathbf{d}_0 - \bar{\mathbf{d}}_0) \cdot \mathbf{E}_0^* \right] \right\}. \quad (3.10)$$

As for any effective Hamiltonian, this magnetic field should not depend on the state of the sample within the ground-state manifold, i.e.  $\mathbf{B}_{\text{eff}}$  should not depend on  $\mathbf{S}(t)$  [32]. Indeed, this is shown to be true in the next chapter, provided the polarizability  $\mathbf{d}_0$  is suitably approximated. If the full polarizability is used, including the  $\mathbf{S}^2$  and higher order effects that enter Eq. (3.9), some dependence of  $\mathbf{B}_{\text{eff}}$  on  $\mathbf{S}(t)$  will be seen. The strength of this dependence provides a measure of the accuracy of Eqs. (3.9) and (3.10).

### 3.4 Demonstration using hydrogen atom

To demonstrate the accuracy with which Eq. (3.4) describes the IFE, simulations were performed of a hydrogen atom interacting with a laser. The IFE causes the electron spin to precess as it would in a magnetic field. The integral of Eq. (3.4),

$$\mathbf{S}(t) = \mathbf{S}(0) + \int_0^t \langle \mathbf{d}(t') \times \mathbf{E}(t') \rangle dt', \quad (3.11)$$

shows that the change of angular momentum of the precessing spin will be compensated by the radiated angular momentum.

The hydrogen atom is described by the Hamiltonian

$$H_0 = -\frac{1}{2}\nabla^2 - \frac{1}{r} + \lambda \mathbf{l} \cdot \mathbf{s}, \quad (3.12)$$

where Hartree atomic units (a.u.) are used. The nucleus is treated as a static Coulomb potential and the spin-orbit coupling  $\lambda$  is a free parameter. (By varying  $\lambda$ , the simulation time is reduced, and the dependence of the inverse Faraday effect on  $\lambda$  may be investigated.) Following Ref. [152], the interaction with the laser is evaluated in the electric dipole approximation:

$$H = H_0 + \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{r}. \quad (3.13)$$

The electric field has the form (3.2) and is switched on adiabatically. The electron wavefunction is expanded using the eigenfunctions of  $h = -\nabla^2/2 - 1/r$ , with the expansion limited to the first 28 spinor eigenfunctions (i.e. up to radial quantum number  $n = 3$ ). The atom starts in the ground state  $|n, l\rangle = |1, 0\rangle$  and the initial spin is varied. The evolution of the wavefunction is determined by solving the Schrödinger equation using the implicit midpoint method to ensure unitary evolution. The rapid oscillations of the light are resolved by using a timestep of 0.01 fs.

Guided by Eq. (3.11), I have monitored the evolution of  $\mathbf{S}(t)$  and  $\int \langle \mathbf{d} \times \mathbf{E} \rangle$ . Both are calculated from the wavefunction: the components of  $\mathbf{S}(t)$  are the expectation values of the Pauli matrices and  $\mathbf{d}(t)$  is the expectation value of the position operator  $\mathbf{r}$ . Fig. 3.1 shows the calculation starting from the initial state  $\mathbf{S}(0) = (0, 0, \frac{1}{2})$  a.u., with parameters  $\lambda = 20$  meV,  $\omega = 1$  eV and  $\mathbf{E}_0 = (1, 0, i) \times 10^{-3}$  a.u. (corresponding to a laser beam with intensity of  $10^{15}$  Wm $^{-2}$  propagating in the  $y$ -direction). For clarity, only a few values of  $[\mathbf{S}(0) + \int \mathbf{d} \times \mathbf{E}]$  are shown. It is clear that the dipole radiation accounts completely for the changing angular momentum.

Further calculations revealed that the precession frequency is quadratic in  $\mathbf{E}_0$  and linear in  $\lambda$ , there is no precession for light propagating parallel to the spin direction or for linearly polarized light, and changing the sense of circular polarization reverses the precession. These are known properties of the IFE [152]. Eq. (3.4) was found to be accurate: the largest difference between  $\mathbf{S}(t)$  and  $[\mathbf{S}(0) + \int \langle \mathbf{d} \times \mathbf{E} \rangle]$  was less than 0.01 a.u. and this difference does not grow in time.

The hydrogen atom can also be used to demonstrate the accuracy of Eq. (3.10). Fig. 3.2 shows  $\mathbf{B}_{\text{eff}}$  calculated for several values of  $\mathbf{S}(0)$ , with  $\lambda = 0.02$  eV,  $\omega = 1$  eV and  $\mathbf{E}_0 = (1, i, 0) \times 10^{-3}$  a.u. (This time, the light propagates in the  $z$ -direction.) The value of  $\mathbf{d}_0$  was calculated from linear response, and  $\overline{\mathbf{d}}_0$  was determined by setting  $\lambda = 0$ . The  $x$  and  $y$  components of  $\mathbf{B}_{\text{eff}}$  are not plotted, as they are less than 1% of the magnitude of the  $z$  component. Due to symmetry,  $\mathbf{B}_{\text{eff}}$  does not depend on  $S_x$  or  $S_y$  and only the dependence on  $S_z$  is plotted.

The value of  $\mathbf{B}_{\text{eff}}$  varies by less than 3%, indicating that Eq. (3.10) is a good approximation. Furthermore, the dependence of  $\mathbf{B}_{\text{eff}}$  on  $S_z$  is almost completely linear — the deviation from linear behaviour is less than 0.0001%. This suggests that the accuracy of Eq. (3.10) can be further improved by subtracting the  $\mathbf{S}^2$  terms from the energy (3.9). This is beyond the scope of this thesis and remains a goal for future research.

To summarize, I have shown that in the inverse Faraday effect, the light both induces magnetization dynamics and provides the necessary angular momentum. The angular momentum gained by the magnetization is compensated by a change of angular momentum

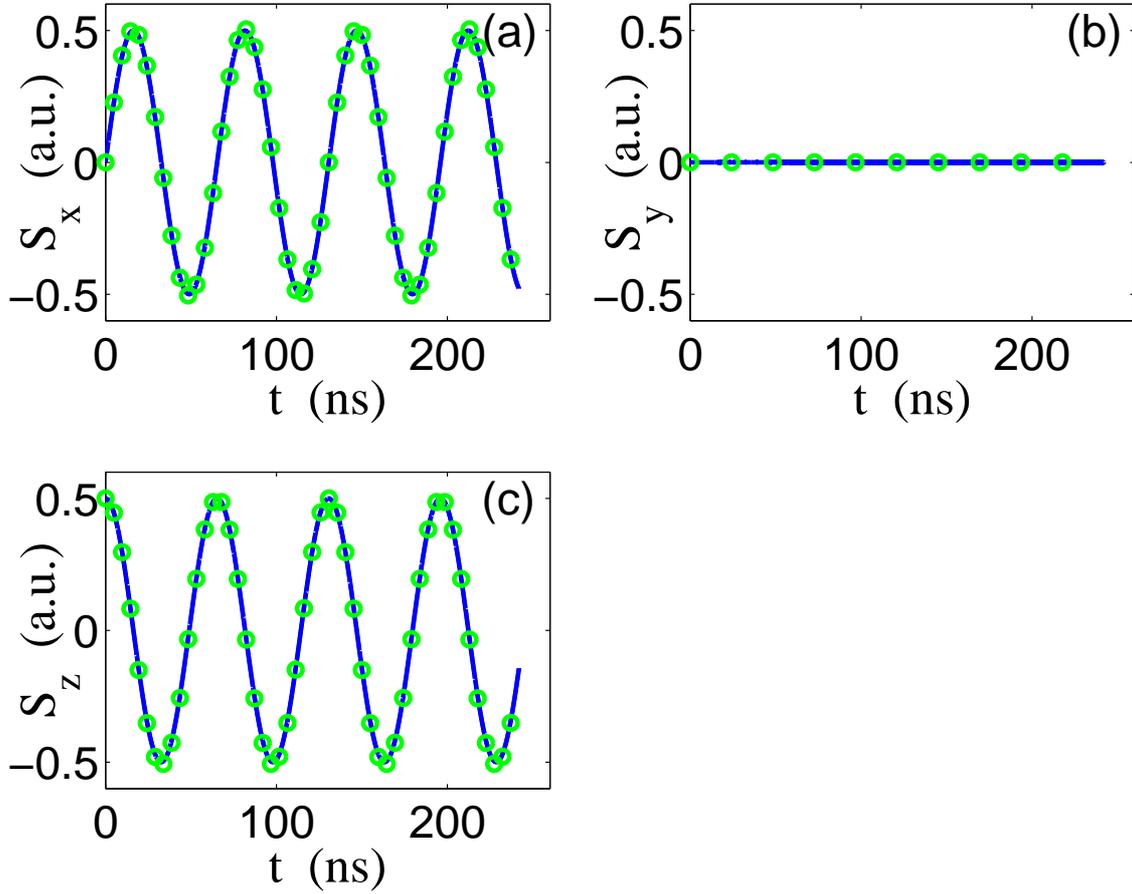


Figure 3.1: Precession of a spin due to the inverse Faraday effect. Solid line:  $\mathbf{S}(t)$ , circles:  $[\mathbf{S}(0) + \int (\mathbf{d} \times \mathbf{E})]$ , see Eq. (3.11). Parameters:  $\lambda = 0.02$  eV,  $\mathbf{E}_0 = (1, 0, i) \times 10^{-3}$  a.u. and  $\omega = 1$  eV.

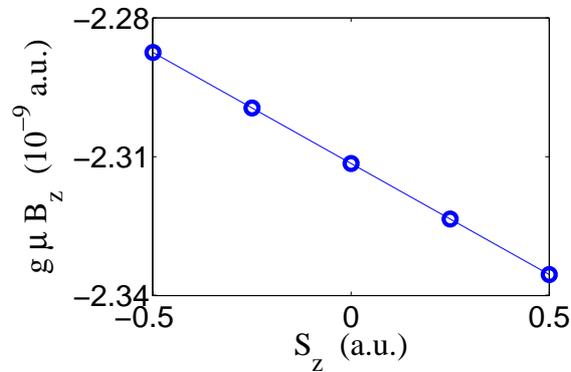


Figure 3.2: The effective magnetic field [Eq. (3.10)] as a function of  $S_z$ . A straight line has been added as a guide.

in the interference pattern between the incident and scattered light. No other angular momentum sources are required. Magnetization reversal can occur on the timescale of the laser pulse, which implies that the fundamental speed of magnetic switching could be less than 100 fs. To exploit this fast switching, magnetic materials are required that have a strong magneto-optical response and are transparent at typical laser frequencies. The search for suitable materials requires the calculation of the spin-dependent polarizability, which is possible for small molecules [19, 33, 83, 114] but remains a challenge for solids.

In the next chapter, the current results are extended to larger systems. In addition, the link between the IFE described by Eq. (3.10) and the IFE of the previous chapter [see Eq. (2.16)] is demonstrated.

## Chapter 4

# Conservation of Angular Momentum: Extended Systems

Many of the results of the previous chapter are not limited to the case of small systems. The torque and energy of a dipole in an external field [see Eqs. (3.8) and (3.9)] are locally-defined quantities, independent of their surroundings. For this reason, Eq. (3.10) is easily generalized to extended systems:

$$\mathbf{B}_{\text{eff}}(\mathbf{r}) = \frac{1}{4[m(\mathbf{r}, t)]^2} \text{Re} \left\{ 2\mathbf{m}(\mathbf{r}, t) \times [\mathbf{P}(\mathbf{r}, \omega, t) \times \mathbf{E}_0^*(\mathbf{r})] - \mathbf{m}(\mathbf{r}, t) \left[ (\mathbf{P}(\mathbf{r}, \omega, t) - \bar{\mathbf{P}}(\mathbf{r}, \omega)) \cdot \mathbf{E}_0^*(\mathbf{r}) \right] \right\}. \quad (4.1)$$

Here, in keeping with the usual definitions for an extended system, the spin  $\mathbf{S}$  has been replaced by the magnetization  $\mathbf{m}$  and the dipole moment  $\mathbf{d}_0$  by the polarization density  $\mathbf{P}$ . As in Eq. (3.10),  $\mathbf{P}$  depends on time through the time-dependent magnetization.

If Eq. (4.1) is to be used, some work is needed to ensure that the polarization density is correctly defined. For low-density gases, the atomic polarizabilities can be used, while for solids, the Clausius-Mossotti relation needs to be used to account for internal fields [100]. This is a minor change and the transition to extended systems seems well-defined. However, while this may be true from a computational viewpoint (i.e. Eq. (4.1) is the correct equation for the effective magnetic field), it is not necessarily true in terms of the physical mechanisms involved.

There are essentially two physical mechanisms considered in chapter 3. The first is the mechanism by which light rotates the spin: light induces a dipole moment, this is deflected by spin-orbit coupling and the light generates a torque on this dipole. This mechanism is still valid in a solid, provided the correct definition of the dipole moment is used.

The other physical mechanism used in chapter 3 involves the conservation of angular momentum. If the spin rotates, angular momentum must flow into the light. This angular momentum was found to be radiated in the *interference flux*, i.e. the interference between the radiation from the dipole and the original beam.

For several reasons, this mechanism is not appropriate in the context of a solid. First, quantities radiated via interference can only be summed together if their sources are incoherent. If the sources are coherent, which is likely to be true for a crystalline solid,

all interfering fields must be determined before physical quantities can be evaluated. In other words, the superposition of the radiation fields emitted by all of the dipoles must be calculated, rather than separately evaluating the angular momentum flux of each and summing these fluxes. The second objection to the use of the interference flux is more pragmatic: once light has entered a solid, there is no “original beam” remaining, so the interference of the dipole radiation with the original beam is undefined. All light entering a solid is influenced by the solid and it would be incorrect to assume otherwise.

Despite these fundamental objections, the overall picture of angular momentum flowing into the light must remain true. If light takes up angular momentum on a local scale, this angular momentum must still be present in the light when a larger scale is considered. As described in chapter 3, no other angular momentum reservoir is available. But if the angular momentum is not in the interference flux, where is it? This chapter is devoted to answering this question and determining the experimental consequences of the answer. Gaussian units are used throughout this chapter.

## 4.1 Description of the problem

The propagation of light through a solid is governed by Maxwell’s equations [82, 116]:

$$\begin{aligned}\nabla \cdot \mathbf{D} &= 0, & \nabla \times \mathbf{E} &= -\frac{1}{c} \dot{\mathbf{B}} \\ \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{H} &= \frac{1}{c} \dot{\mathbf{D}},\end{aligned}\tag{4.2}$$

where the charge density and current have been set to zero. The electric displacement  $\mathbf{D}$  and magnetic induction  $\mathbf{B}$  are assumed to be linear in the electric and magnetic fields respectively:

$$D_i = \epsilon_{ij} E_j, \quad B_i = \mu_{ij} H_j.\tag{4.3}$$

Here, the Einstein notation of summing over repeated indices is used.

The simplest medium admitting magneto-optical effects is an isotropic crystal with a magnetization. Such a medium will exhibit the Faraday effect [116] and is therefore expected to exhibit the inverse Faraday effect, too. A medium of this type has a dielectric tensor of the form [116]

$$\epsilon_{ij} = \epsilon \delta_{ij} + i\gamma \begin{pmatrix} 0 & -m_z & m_y \\ m_z & 0 & -m_x \\ -m_y & m_x & 0 \end{pmatrix}\tag{4.4}$$

where  $\mathbf{m} = (m_x, m_y, m_z)$  is the magnetization and  $\gamma$  is the magneto-optical coupling constant. Note the similarity between Eq. (4.4) and Eq. (2.1).

All magneto-optical effects have been placed in the dielectric tensor, so the permeability tensor  $\mu_{ij}$  can be set equal to the identity matrix; this is permissible at optical and higher frequencies [116, 151]. Monochromatic waves are used in the following, so the frequency dependence of  $\epsilon$  and  $\gamma$  can be neglected. I consider only the case of normal incidence; the reasons making this a convenient choice are described in the next section. The crystal is assumed to occupy the region  $z > 0$  and the incident light propagates in the  $z$ -direction.

If the magnetization is in the  $z$ -direction,  $\mathbf{m} = (0, 0, m)$ , then the inverse Faraday effect will not generate a torque. The energy of the system will change by the Zeeman energy due to the effective magnetic field, but no dynamics will occur. On the other hand, linearly polarized light passing through such a system will undergo Faraday rotation. As in chapter 2, I assume that  $|\gamma m| \ll \epsilon$ . Similarly to Eq. (2.7), the Faraday rotation angle will be

$$\beta = kd\gamma m/2\sqrt{\epsilon}. \quad (4.5)$$

Here,  $k$  is the wave-vector of the incident light ( $k = \omega/c$ ) and  $d$  is the distance through the sample that the light has propagated before being measured. Eq. (4.5) will be useful below, as it allows us to relate the physically measurable Faraday rotation angle to the magneto-optical constant  $\gamma$ .

For the IFE to have a dynamical effect, there should be a component of the magnetization orthogonal to the  $z$ -direction. For simplicity, I assume that the magnetization is constrained to the  $xy$ -plane. In this case, a simple rotation of the coordinate system makes the magnetization point along the  $y$ -direction. I therefore assume that *the magnetization is initially oriented along the  $y$ -axis*. The dielectric tensor (4.4) becomes

$$\epsilon_{ij} = \begin{pmatrix} \epsilon & 0 & i\gamma m \\ 0 & \epsilon & 0 \\ -i\gamma m & 0 & \epsilon \end{pmatrix} \quad (4.6)$$

Note that  $\gamma$  is measured in units of inverse magnetization, since  $\gamma m$  is dimensionless.

## 4.2 Orbital angular momentum of light

A similar problem to the current one was studied by Moe and Happer in Ref. [133]. They studied the problem of light passing through a transparent anisotropic atomic vapour. In their work, they showed that angular momentum is conserved, i.e. that the torque on the atoms is equal to the loss of angular momentum of the light per unit time. Their work demonstrates that other reservoirs of angular momentum are not needed for magneto-optical effects to occur. However, they focused only on what happens to the light. They did not consider the effect of the torque on the spins, nor did they make the connection to the inverse Faraday effect. Furthermore, they considered only the case of an infinitely large gas, from which the light never emerges. I show below that the emergence of the light has some important effects.

Ref. [133] was written to make certain arguments in Ref. [67] more rigorous. In Ref. [67], it was shown that the transfer of angular momentum between the atoms and the light has negligible influence on the polarization of the light, provided the atomic spin is transverse to the light propagation. Therefore, it was claimed that the orbital angular momentum,  $\mathbf{r} \times \hbar\mathbf{k}$ , of the photons must be modified if a torque is present. (See, e.g., Ref. [4] for details about the orbital momentum of a photon.) The photons are not scattered to a different  $\mathbf{k}$ -value, so the torque causes a change of  $\mathbf{r}$ , i.e. a beam of light passing through a medium in a transverse magnetic field will be laterally displaced. This was confirmed experimentally by Blasberg and Suter [20] and Schlessler and Weis [169], although both groups used light very close to resonance, so that the anisotropy of the medium was caused by absorptive processes and differs from the inverse Faraday effect.

Note that this lateral displacement occurs even in the case of normal incidence. It is well known that if light is incident on a sample at an oblique angle, light emerging on the other side of the sample will be parallel to the incident light, but displaced [84]. On the other hand, it is usual for light meeting a sample in normal incidence to pass through the sample without a change of position [84]. Magneto-optical effects cause a lateral displacement *even in normal incidence*. All results in this chapter are derived for the case of normal incidence. This is a convenient choice, because any lateral displacements can only be attributed to magneto-optical effects; in the case of oblique incidence, the non-magnetic displacement would need to be subtracted.

Returning to the problem of the IFE in solids, Refs. [67, 133] show that a change of the polarization of the light is unlikely to contain the missing angular momentum. This is not surprising, because the polarization of the light is related to its spin angular momentum, which always points in the direction of propagation. Since the IFE changes the angular momentum *transverse* to the direction of propagation, the spin angular momentum cannot contribute. Thus it is only the orbital angular momentum of the light that needs to be considered.

The contribution from the orbital momentum is easily understood in terms of Fig. 4.1, which is taken from Ref. [169]. As light enters the material, it is deflected slightly. More precisely, its Poynting vector is deflected. Its momentum, which points in the direction of  $\mathbf{k}$ , remains in the same direction. After propagating through the medium, the light is laterally displaced, but emerges parallel to the direction in which it entered. This displacement without a change of final propagation direction gives rise to an angular momentum in the direction orthogonal to both the displacement and propagation (the z-direction, in the case of Fig. 4.1). So for the IFE, we may anticipate that as the light induces a torque on the spins, it is displaced in a direction orthogonal to its linear momentum. By measuring the displacement of the beam, it is possible to estimate the angular momentum that has been deposited in the sample.

This discussion highlights one of the primary differences between the analysis for extended systems and that for small systems — the finite breadth of the laser beam. With small systems, the laser is much broader than the system. The boundary of the system must be included — it is automatically included in the definition of the dipole moment of the system — but since the boundary of the laser beam is far from the region where angular momentum transfer occurs, the beam can be treated as infinitely wide.

In the case of an extended system, transfer of angular momentum is expected to happen throughout the sample, not only at the boundaries. As such, it should be safe to neglect the boundaries and treat the system as infinite. If the light is also taken as infinitely broad, then no angular momentum transfer will be seen. That is, there will be no interference flux and there can be no displacement of the “beam” — an infinitely broad beam does not have a position. Therefore the laser beam must have a finite breadth.

The difference is clear. If the angular momentum is to be radiated in the interference flux, the system boundaries must be taken into account, so the system must be finite. If angular momentum is to be transferred into the orbital angular momentum of the laser beam, the beam must have a well-defined position and therefore must be finite. However, the local picture of light imposing a torque on an induced dipole is the same in both cases. This explains why Eq. (4.1) can be valid even in the case of a solid.

Although the beam must be finite, it need not be particularly narrow. It is generally

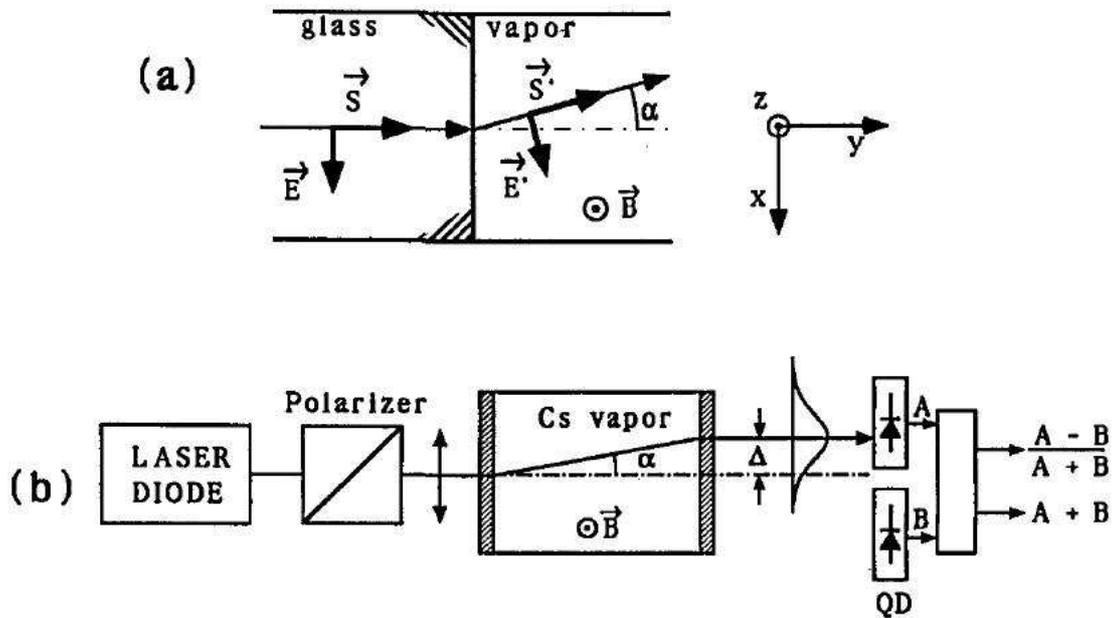


Figure 4.1: The deflection of light as it propagates through an anisotropic medium, taken from Ref. [169]. (a) shows the change of direction of the Poynting vector as the light enters the medium (in this case, an atomic vapour in a magnetic field). (b) shows the full experimental setup used in Ref. [169]. The light emerges parallel to the incident direction, but laterally displaced. Note that the choice of coordinate axes in these figures does not coincide with the axes used in the text. This figure is reused with permission from R. Schlessner and A. Weis, *Optics Letters* **17**, 1015 (1992). Copyright 1992, Optical Society of America.

true, however, that any calculations for a broad beam can be performed using an infinite plane wave, provided only those quantities are considered that do not depend on the light being finite. For example, conservation of angular momentum as expressed in Eq. (3.1) cannot be used, since it involves integrals over infinite surfaces and volumes. On the other hand, the Poynting vector and momentum density of a broad beam will coincide almost exactly with the same quantities for an infinite plane wave. Therefore, despite insisting that the beam must be finite, I have performed the following calculations using a homogeneous plane wave, with care taken to use only well-defined quantities.

### 4.3 Propagation of light

The quantity of interest is the lateral displacement of the laser beam. This displacement can be measured as the shift of the peak intensity of the beam, i.e. shift of the peak energy. Since the Poynting vector describes the flow of energy, it follows that the displacement is related to the direction of the Poynting vector in the system. This vector is calculated below. The momentum density, necessary for calculating the angular momentum, is also calculated. In this section, the solid is assumed to extend infinitely far in the  $z$ -direction. The effect of a second interface is considered in section 4.5.

At  $z = 0$ , a homogeneous plane wave

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( \mathbf{E}_0 e^{i(\omega t - kz)} + c.c. \right) \quad (4.7)$$

is incident on a material with dielectric tensor (4.6) and permeability  $\mu = 1$ . Using Maxwell's equations (4.2) and the boundary conditions

$$D_z^{out} = D_z^{in}, \quad \mathbf{E}_{\parallel}^{out} = \mathbf{E}_{\parallel}^{in}, \quad \mathbf{B}^{out} = \mathbf{B}^{in} \quad (4.8)$$

(where “in” means inside the medium and “out” means in the vacuum), it can be shown that the electric field inside the medium is [203]

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( E_+ \mathbf{n}_+ e^{i(\omega t - k_+ z)} + E_- \mathbf{n}_- e^{i(\omega t - k_- z)} + c.c. \right). \quad (4.9)$$

The normal modes and their associated wavevectors are

$$\begin{aligned} \mathbf{n}_+ &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad k_+ = k\sqrt{\epsilon} \\ \mathbf{n}_- &= \begin{pmatrix} 1 \\ 0 \\ i\gamma m/\epsilon \end{pmatrix}, \quad k_- = k\sqrt{\epsilon - \gamma^2 m^2/\epsilon}, \end{aligned} \quad (4.10)$$

and the coefficients of these modes in Eq. (4.9) are

$$E_+ = \frac{2k}{k + k_+} E_0^y, \quad E_- = \frac{2k}{k + k_-} E_0^x. \quad (4.11)$$

There is also a reflected wave in the vacuum region, but since this is not displaced relative to the incident beam, it carries no orbital angular momentum and can be ignored.

The magnetic field inside the medium can be calculated from the electric field (4.9) and Faraday's law  $\nabla \times \mathbf{E} = -\frac{1}{c}\dot{\mathbf{B}}$ , yielding

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{H}(\mathbf{r}, t) = \frac{1}{2} \left( -\frac{k_+}{k} E_+ \hat{\mathbf{x}} e^{i(\omega t - k_+ z)} + \frac{k_-}{k} E_- \hat{\mathbf{y}} e^{i(\omega t - k_- z)} + c.c \right). \quad (4.12)$$

Here I have used the notation  $\hat{\mathbf{x}}$  for the unit vector in the  $x$ -direction and similar notation for  $\hat{\mathbf{y}}$  and  $\hat{\mathbf{z}}$ . From the electric and magnetic fields, (4.9) and (4.12), the Poynting vector can be calculated [82, 116]

$$\begin{aligned} \mathbf{S} &\equiv \frac{c}{4\pi} (\mathbf{E} \times \mathbf{H}) \\ &= \frac{c}{8\pi} \hat{\mathbf{z}} \left( \frac{k_+}{k} |E_+|^2 + \frac{k_-}{k} |E_-|^2 \right) - \frac{c}{8\pi} \hat{\mathbf{y}} \frac{k_+ \gamma m}{k \epsilon} \text{Re} \left( i E_- E_+^* e^{i(k_+ - k_-)z} \right). \end{aligned} \quad (4.13)$$

The rapidly oscillating terms proportional to  $e^{\pm 2i\omega t}$  have been dropped from this equation.

Eq. (4.13) has several important features. First, since  $\gamma m/\epsilon \ll 1$ , the Poynting vector is almost parallel to  $\hat{\mathbf{z}}$  at all times. The light continues almost in the same direction that it entered, with only a small lateral displacement. Second, this displacement is in the  $y$ -direction. Since we expect the momentum to travel in the  $z$ -direction, the torque will be parallel to  $\hat{\mathbf{x}}$ . This is the correct direction for the torque on a magnetization  $m\hat{\mathbf{y}}$  in a magnetic field  $B\hat{\mathbf{z}}$ . Finally, reversing the sign of either  $E_0^x$  or  $E_0^y$  will change the direction of displacement, because of Eq. (4.11). Hence a change of the sense of circular polarization will change the sense of the torque.

These comments show that Eq. (4.13) has the correct qualitative features to explain the IFE in solids. For a quantitative analysis, the momentum density  $\mathbf{g}$  is required [82]. The calculation of this is straightforward and yields

$$\begin{aligned} \mathbf{g} &\equiv \frac{1}{4\pi c} (\mathbf{D} \times \mathbf{B}) \\ &= \frac{1}{8\pi c} \hat{\mathbf{z}} \left[ \epsilon \frac{k_+}{k} |E_+|^2 + \left( \epsilon - \frac{\gamma^2 m^2}{\epsilon} \right) \frac{k_-}{k} |E_-|^2 \right]. \end{aligned} \quad (4.14)$$

The lateral displacement of the beam is also needed. This is calculated from the Poynting vector as follows. First, the Poynting vector is normalized so that its  $z$ -component is unity. The normalized vector  $\hat{\mathbf{S}}$  is then integrated over  $z$  from the origin to  $z = z_0$ . This provides a vector describing the beam position after the beam has propagated in the sample for a distance  $z_0$  in the  $z$ -direction. The lateral components of this vector provide the displacement of the light. In equations, this reads

$$\hat{\mathbf{S}} = \hat{\mathbf{z}} - \hat{\mathbf{y}} \frac{k_+ \gamma m}{\epsilon k_+ |E_+|^2 + \epsilon k_- |E_-|^2} \text{Re} \left( i E_- E_+^* e^{i(k_+ - k_-)z} \right) \quad (4.15)$$

and

$$\begin{aligned} \mathbf{R} &\equiv \int_0^{z_0} \hat{\mathbf{S}} dz \\ &= z_0 \hat{\mathbf{z}} - \hat{\mathbf{y}} \frac{k_+ \gamma m}{\epsilon k_+ |E_+|^2 + \epsilon k_- |E_-|^2} \text{Re} \left( i E_- E_+^* \frac{e^{i(k_+ - k_-)z_0} - 1}{i(k_+ - k_-)} \right). \end{aligned} \quad (4.16)$$

Eqs. (4.14) and (4.16) allow us to determine the angular momentum density of the light after it has propagated in the sample for a distance of  $z_0$  (in the  $z$ -direction):

$$\begin{aligned} \mathbf{L} &\equiv \mathbf{R} \times \mathbf{g} \\ &= -\hat{\mathbf{x}} \frac{k_+ \gamma m}{8\pi c k (|E_+|^2 + \frac{k_-}{k_+} |E_-|^2)} \left[ |E_+|^2 + \left(1 - \frac{\gamma^2 m^2}{\epsilon^2}\right) \frac{k_-}{k_+} |E_-|^2 \right] \\ &\quad \times \operatorname{Re} \left( i E_- E_+^* \frac{e^{i(k_+ - k_-)z_0} - 1}{i(k_+ - k_-)} \right). \end{aligned} \quad (4.17)$$

At this point, it becomes useful to make several approximations. The first has already been mentioned above in the derivation of Eq. (4.5) and in the discussion following Eq. (4.13):  $\gamma m \ll \epsilon$ . This simplifies Eq. (4.17) and also yields  $k_+ \approx k_-$  [see Eq. (4.10)]. From this, it follows that both modes propagate with the same phase velocity,  $v_p = c/\sqrt{\epsilon}$ .

However,  $k_+ = k_-$  (or rather  $|k_+ - k_-| \ll 1$ ) is not a dimensionless equation and is therefore not an acceptable approximation. The only dimensionless quantity in Eq. (4.17) containing  $(k_+ - k_-)$  is the product  $(k_+ - k_-)z_0$ . This leads us to the next approximation<sup>1</sup>:  $(k_+ - k_-)z_0 \ll 1 \Rightarrow k\sqrt{\epsilon}z_0 \ll \epsilon^2/\gamma^2 m^2$ .

With these two approximations in place, Eq. (4.17) becomes

$$\mathbf{L} = -\hat{\mathbf{x}} \frac{\gamma m \sqrt{\epsilon} z_0}{8\pi c} \operatorname{Re} (i E_- E_+^*). \quad (4.18)$$

This is the angular momentum density of the light at a depth of  $z_0$ . The torque on the light at this depth can be estimated using the group velocity of the light:

$$\frac{d}{dt} \mathbf{L} = v_g \frac{d}{dz_0} \mathbf{L} = -\hat{\mathbf{x}} \frac{\gamma m}{8\pi} \operatorname{Re} (i E_- E_+^*). \quad (4.19)$$

Here, I have used the group velocity  $v_g = c/\sqrt{\epsilon}$  and neglected the transverse components of this velocity. As noted in the discussion following Eq. (4.13), these components are of order  $\gamma m/\epsilon \ll 1$ .

Eq. (4.19) is the torque on the light. An equal and opposite torque is imposed on the material. As in the previous chapter, this torque can be identified as a magnetic field. Before introducing this field, it is useful to rewrite Eq. (4.19) in coordinate-free form and in terms of the external electric field of the light,  $\mathbf{E}_0$ . Using Eq. (4.11) and the relation  $\hat{\mathbf{x}} = \hat{\mathbf{y}} \times \hat{\mathbf{z}}$ , the torque on the material  $\mathbf{T}$  can be expressed as

$$\mathbf{T} = \frac{\gamma}{4\pi(1 + \sqrt{\epsilon})^2} m \hat{\mathbf{y}} \times (i \mathbf{E}_0 \times \mathbf{E}_0^*). \quad (4.20)$$

The vector  $(i \mathbf{E}_0 \times \mathbf{E}_0^*)$  is real-valued and oriented in the  $z$ -direction for light propagating along  $z$ ; its magnitude is maximized for circularly polarized light and vanishes in the case

---

<sup>1</sup>This approximation implies that our attention is focused only on a region near the surface of the material. However, because magneto-optical effects are quite weak, this “surface region” can be quite thick, e.g. assuming  $\gamma m/\epsilon \leq 0.01$  (which is true even for materials with a very strong magneto-optical effect [31, 182, 210]) and assuming the light has an optical wavelength (i.e.  $k \approx 1/700 \text{ nm}^{-1}$ ), this approximation is still valid at a depth of several hundred micrometres. Thus this approximation holds throughout the samples used in Ref. [99], which were around  $60 \mu\text{m}$  thick.

of linear polarization (see Appendix A for details). This vector also arises in the classical derivation of the IFE, as shown by Eq. (2.16).

The torque (4.20) has been obtained directly from Eq. (4.19). It is easily shown that exactly the same torque is found using the definition  $\mathbf{T} = \mathbf{P} \times \mathbf{E}$  of the torque by an electric field on a dielectric. Here,  $\mathbf{P} = (\mathbf{D} - \mathbf{E})/4\pi$  is the macroscopic polarization density calculated from Eqs. (4.6) and (4.9). The equality of these torques shows that angular momentum has been conserved. This result is equivalent to that derived in Ref. [133].

Eq. (4.20) is almost in the required form. All that remains is to note that  $\mathbf{m} = m\hat{\mathbf{y}}$  and to introduce the effective magnetic field  $\mathbf{B}_{\text{eff}}$  and one obtains the solid-state equivalent of Eq. (3.8):

$$\mathbf{B}_{\text{eff}} \times \mathbf{m}(t) = -\frac{\gamma}{4\pi(1 + \sqrt{\epsilon})^2} (i\mathbf{E}_0 \times \mathbf{E}_0^*) \times \mathbf{m}. \quad (4.21)$$

As in the previous chapter, this equation identifies the components of the magnetic field orthogonal to the magnetization. To recover the parallel component, the Zeeman energy is needed. However, the calculation of this is not as straightforward as before. The problem is that there are many contributions to the energy that are linear in the magnetization and yet should not be included in the Zeeman energy. These terms arise because the transmission and reflection coefficients can be linear in the magnetization.

Consider two possible extremes,  $\mathbf{m} = m\hat{\mathbf{y}}$  and  $\mathbf{m} = m\hat{\mathbf{z}}$ , with the light propagating along the  $z$ -axis in both cases. For the first case, the electric and magnetic fields have already been calculated in Eqs. (4.9) and (4.12). The total energy is

$$U = \frac{1}{8\pi} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}) \quad (4.22)$$

and it is trivial to see that the term linear in  $m$  vanishes. Indeed, the only field component that is linear in  $m$  is  $E_z$  and this multiplies  $D_z = 0$ . Thus the Zeeman energy vanishes for  $\mathbf{m} = m\hat{\mathbf{y}}$ , as expected.

If, on the other hand,  $\mathbf{m} = m\hat{\mathbf{z}}$ , the normal modes and wavevectors coincide with those given in Eq. (2.3), repeated here for convenience

$$\begin{aligned} \tilde{\mathbf{n}}_+ &= \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, & \tilde{k}_+ &= k\sqrt{\epsilon + \gamma m} \approx k\sqrt{\epsilon}\left(1 + \frac{\gamma m}{2\epsilon}\right) \\ \tilde{\mathbf{n}}_- &= \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, & \tilde{k}_- &= k\sqrt{\epsilon - \gamma m} \approx k\sqrt{\epsilon}\left(1 - \frac{\gamma m}{2\epsilon}\right). \end{aligned} \quad (4.23)$$

The electric field inside the sample is now

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( \tilde{E}_+ \tilde{\mathbf{n}}_+ e^{i(\omega t - \tilde{k}_+ z)} + \tilde{E}_- \tilde{\mathbf{n}}_- e^{i(\omega t - \tilde{k}_- z)} + c.c. \right), \quad (4.24)$$

with the coefficients

$$\tilde{E}_+ = \frac{k}{k + \tilde{k}_+} (E_0^x - iE_0^y), \quad \tilde{E}_- = \frac{k}{k + \tilde{k}_-} (E_0^x + iE_0^y). \quad (4.25)$$

The wavevectors (4.23) and the coefficients (4.25) are linear in  $m$ . Therefore all of the fields entering the energy (4.22) contain terms that are linear in  $m$ . The calculation of the “Zeeman” energy is relatively straightforward, yielding

$$U_{\text{mag}} = \frac{\gamma}{2\pi(1 + \sqrt{\epsilon})^3} (i\mathbf{E}_0 \times \mathbf{E}_0^*) \cdot \mathbf{m}. \quad (4.26)$$

Comparing Eqs. (4.21) and (4.26), we see that the coefficients differ by a factor of  $-2/(1 + \sqrt{\epsilon})$ . Even the sign is incorrect!

The sign error is easily corrected by noting that Eq. (4.22) is the energy stored in the *fields*, while the Zeeman energy is the energy of the magnetization. By energy conservation, these energies are opposite in sign and Eq. (4.26) should be multiplied by  $-1$ .

A further factor of  $1/(1 + \sqrt{\epsilon})$  arises because the transmission coefficients (4.25) are linear in  $m$ , i.e. the intensity of light within the sample depends linearly on the magnetization. Clearly if the intensity of the light is changed, the strength of the inverse Faraday effect must also change. This is not a fundamental change in the IFE, merely a consequence of the fact that the reflection of light depends on the magnetization. It does not occur for small samples, where reflection is negligible and the intensity of the light is independent of the spin.

Finally, the factor of 2 enters because the magnetic contribution to the energy,  $\mathbf{B} \cdot \mathbf{H}/8\pi$ , is identical to the contribution from the electric fields,  $\mathbf{E} \cdot \mathbf{D}/8\pi$ . This is in stark contrast to the previous chapter, where the IFE was described in terms of an electric field interacting with an induced dipole and the magnetic field of the light played no role. The reason for the appearance of a magnetic energy is clear from Eqs. (4.23) and (4.25): both the wavelengths and the transmission coefficients are linear in  $m$ . A smaller wavelength implies that more periods of the light fit into a given distance, increasing the energy density. Similarly, a larger transmission coefficient increases the intensity and thus the energy density. Neither result has any connection to the IFE and both should be subtracted if the correct Zeeman energy is to be found.

Having identified the spurious factors entering Eq. (4.26), the question remains of how to correct these. Fortunately, this is easily done by setting  $\tilde{k}_+ = \tilde{k}_-$ . This corrects both the change of wavelength and the change of transmission coefficients. In this case, the Zeeman energy is

$$U_{\text{mag}} = -\frac{\gamma}{4\pi(1 + \sqrt{\epsilon})^2} (i\mathbf{E}_0 \times \mathbf{E}_0^*) \cdot \mathbf{m}. \quad (4.27)$$

This coincides with the Zeeman energy expected from Eq. (4.21). Therefore the effective magnetic field is

$$\mathbf{B}_{\text{eff}} = -\frac{\gamma}{4\pi(1 + \sqrt{\epsilon})^2} (i\mathbf{E}_0 \times \mathbf{E}_0^*), \quad (4.28)$$

The above-mentioned problem that both the reflection of light and the wavelength depend on the magnetization cannot be neglected in experiment. However, these effects only influence the energy of the system, not the precession of the magnetization. This precession will be governed by the magnetic field (4.28). The difference of energy will only be noticed when extremely long laser pulses are used and relaxation in the effective field dominates the precessional dynamics.

Finally, it is instructive to compare Eq. (4.28) with Eq. (4.1), which is the direct generalization of the results of the previous chapter to extended systems. At first glance, these

equations seem unrelated. Eq. (4.1) contains the magnetization  $\mathbf{m}$  and the polarization density  $\mathbf{P}_0$ , while Eq. (4.28) contains neither of these. However, it was noted in the previous chapter that the effective magnetic field cannot depend on  $\mathbf{m}$ ; it should be possible to cancel  $\mathbf{m}$  from Eq. (4.1). Furthermore, the polarization is related to the electric field by the dielectric tensor (4.4). This suggests that it may be possible to relate these equations.

First, it is necessary to eliminate the polarizability from (4.1). From the dielectric tensor (4.4), the polarization can be written

$$\mathbf{P}_0 = \frac{\epsilon - 1}{4\pi} \mathbf{E}_0 + \frac{i\gamma}{4\pi} \mathbf{m} \times \mathbf{E}_0. \quad (4.29)$$

Averaging the polarization over all directions of magnetization yields

$$\bar{\mathbf{P}}_0 = \frac{\epsilon - 1}{4\pi} \mathbf{E}_0. \quad (4.30)$$

Substituting this into Eq. (4.1) yields (after some straightforward manipulation)

$$\tilde{\mathbf{B}}_{\text{eff}} = -\frac{\gamma}{16\pi} (i\mathbf{E}_0 \times \mathbf{E}_0^*), \quad (4.31)$$

where a tilde has been used to distinguish this field from that in Eq. (4.28). Note that Eq. (4.31) coincides exactly with Eq. (2.16).

Eqs. (4.28) and (4.31) differ by a factor of  $4/(1 + \sqrt{\epsilon})^2$ . This is exactly the factor relating the intensity of light within the medium to the intensity of the incident light [see Eq. (4.11)]. The reason for this difference is clear. Eq. (4.31) was derived for a small system; light is scattered but reflection is extremely weak. Eq. (4.28) refers to a larger system, where reflection cannot be neglected. Such a system will experience a slightly weaker intensity of light and therefore exhibit a diminished inverse Faraday effect<sup>2</sup>. (Note that  $4/(1 + \sqrt{\epsilon})^2 < 1$ .) However, I show in section 4.5 that if the system has a finite thickness, multiple reflections cause additional contributions to the IFE and Eq. (4.31) is recovered.

Finally, it should be noted that in the above, I have neglected the precession of the magnetization, i.e. Eq. (4.28) was derived for a static magnetization. This is acceptable for two reasons. The first is that the effective magnetic field does not depend on the magnetization. The magnetization will experience the same field throughout its precession. Secondly, the timescale on which precession occurs is much longer than the time required to form a polarization density. The change of magnetization is therefore adiabatic and can be ignored<sup>3</sup>.

Nonetheless, there are two features about the precession of the magnetization that must be considered in experiments. The first is that the surface at which the light is incident will feel the magnetic field first. The magnetization nearest this surface will start precessing before the magnetization deeper in the sample, creating a small magnetic

---

<sup>2</sup>This change of intensity is independent of the magnetization, making this effect different to that discussed immediately after Eq. (4.26).

<sup>3</sup>Strictly speaking, the condition for adiabaticity is  $T d\epsilon_{ij}/d\omega \ll 1$ , where  $T$  is the period of precession. Far from resonance,  $d\epsilon_{ij}/d\omega$  is dominated by the distance to the nearest resonance peak. If the precession period is 100 fs, which is extremely fast, the adiabatic condition is satisfied provided the frequency of light is much more than 10 meV from resonance. This condition is easily achieved. For further details, see the discussion in chapter 3.

gradient. If the pulse has a slow rise-time compared to the period of magnetic precession, this gradient will be visible. (In general, this condition is unlikely to be satisfied unless the experiment has been designed with this end in mind.)

Secondly, if the pulse length is large compared to the precession period, different parts of the pulse will see different magnetizations and be displaced in different directions. Depending on the beam width and the magnitude of displacement, this will either manifest itself as the broadening of the beam or as the formation of a circle of high-intensity light with a dark spot in the middle. This effect will not be visible for femtosecond laser pulses, but if much longer pulses (e.g. of nanosecond duration) are used, it will be apparent.

#### 4.4 Change of polarization

The above results rely on the fact that  $(k_+ - k_-)z \ll 1$ . Of course, this will not be valid for very thick samples. It will also fail in the case of multiple reflections in thin samples: once the pulse has traversed the sample several times, it will have accumulated a large phase shift (although it may have negligible amplitude by this point). I consider here what happens when the assumption  $(k_+ - k_-)z \ll 1$  is relaxed. For this section, it is assumed that the magnetization is in the  $y$ -direction and that the light is circularly polarized, i.e.  $iE_-E_+^*$  is real [see Eq. (4.11)].

Consider Eq. (4.16) without assuming  $(k_+ - k_-)z \ll 1$ . For ease of notation, I introduce the real constant

$$C = \frac{k_+ \gamma m}{\epsilon k_+ |E_+|^2 + \epsilon k_- |E_-|^2} iE_- E_+^*. \quad (4.32)$$

Note that the assumption  $\gamma m \ll \epsilon$  is retained here, so that  $C \ll 1$ . The displacement vector (4.16) becomes

$$\begin{aligned} \mathbf{R} &= z_0 \hat{\mathbf{z}} - \hat{\mathbf{y}} C \operatorname{Re} \left( \frac{e^{i(k_+ - k_-)z_0} - 1}{i(k_+ - k_-)} \right) \\ &= z_0 \hat{\mathbf{z}} - \hat{\mathbf{y}} C \frac{\sin(k_+ - k_-)z_0}{k_+ - k_-}. \end{aligned} \quad (4.33)$$

For small  $(k_+ - k_-)z_0$ , the light beam is shifted in the  $y$ -direction by a small amount that grows linearly with  $z_0$ . As  $(k_+ - k_-)z_0$  approaches  $\pi/2$ , the growth of the lateral displacement starts to diminish. It reaches a maximum value and then starts to move back towards the  $y = 0$  line. The maximum value attained is very large, since  $1/(k_+ - k_-) \sim \epsilon^2/\gamma^2 m^2 \gg 1$ , but is smaller than the distance traveled in the  $z$ -direction by a factor of  $2C/\pi$ .

If the light beam starts moving back towards zero displacement, its angular momentum density must decrease. Indeed, at  $z_0 = \pi/(k_+ - k_-)$ , the light is back at  $y = 0$  and contains no angular momentum density. But while propagating through the initial part of the sample, it imposed a torque on the magnetization. What has happened to this angular momentum?

The answer is contained in Eqs. (4.9) and (4.10). After a distance  $z_0 = \pi/2(k_+ - k_-)$ , the phase of the normal mode  $\mathbf{n}_+$  has changed by  $\frac{\pi}{2}$  more than the phase of  $\mathbf{n}_-$ . The light, which started out circularly polarized, has become linearly polarized. As the light

continues to propagate beyond this point, it gains a circular polarization of the opposite sense and causes a reversed inverse Faraday effect. All the atoms between  $z_0 = \pi/2(k_+ - k_-)$  and  $z_0 = \pi/(k_+ - k_-)$  experience a magnetic field in the opposite direction to those between  $z_0 = 0$  and  $z_0 = \pi/2(k_+ - k_-)$  and their contributions to the angular momentum cancel. If the sample is large enough, regions of differently oriented precession will be visible.

In the case of multiple reflections, it will be shown below that the reflected light propagates back along the original path. The atoms that lie in this path will experience different magnetic fields from the different polarizations of multiply scattered light. If the sample is chosen to have the thickness  $z_0 = \pi/2(k_+ - k_-)$  and if losses due to reflection can be ignored (i.e. a negligible amount of light is transmitted into the vacuum) then the magnetic fields from forward and backward propagating light will cancel each other and there will be no precession. On the other hand, if a thickness  $z_0 = \pi/(k_+ - k_-)$  is used, the magnetic fields from forward and backward propagating light will have the same sign and the effective field is doubled. This magnetic field will be different at different depths in the sample and, as with an infinitely thick sample (with no reflection), regions of differently oriented precession will be visible.

Samples used in experiments will not be infinitely thick. It is clear from the above discussion that reflection from the second interface, and all subsequent reflections, have an influence on the IFE. The following section shows how the inclusion of the second interface changes the analysis.

## 4.5 The second interface

### 4.5.1 First reflection

The results of section 4.3 and Ref. [133] were derived for the light while it is still in the medium. However, as indicated in Fig. 4.1 (b), the light exits the medium before being measured. If the light is to be used to estimate the angular momentum transferred to the sample, this second interface must be taken into account. Furthermore, reflection from the second interface gives rise to additional magnetic fields that should be taken into account.

Estimates of the effect of the second interface can be made by treating the fields (4.9) and (4.12) as incident waves on the second interface and calculating the transmission and reflection coefficients. Such a calculation is not perfect, as it ignores interference effects. Nonetheless, this “first reflection” gives several insights into the problem and it is useful to consider this case before the full multiple-reflection calculation is performed.

The calculation of reflection and transmission at the second interface is very similar to the calculation given in section 4.3 for the first interface. It is straightforward to show that the transmitted electric field is

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( \mathbf{E}_T e^{i(\omega t - kz)} + c.c. \right) \quad (4.34)$$

with

$$E_T^x e^{-ikd} = \frac{2k_-}{k + k_-} E_- e^{-ik_- d}, \quad E_T^y e^{-ikd} = \frac{2k_+}{k + k_+} E_+ e^{-ik_+ d}. \quad (4.35)$$

Here,  $d$  is the thickness of the sample and  $E_{\pm}$  are given in Eq. (4.11). Likewise, the reflected field can be calculated:

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( \tilde{E}_+ \mathbf{n}_+ e^{i(\omega t + k_+ z)} + \tilde{E}_- \mathbf{n}_- e^{i(\omega t + k_- z)} + c.c \right) \quad (4.36)$$

with

$$\tilde{E}_+ = \frac{k_+ - k}{k_+ + k} E_+ e^{-2ik_+ d}, \quad \tilde{E}_- = \frac{k_- - k}{k_- + k} E_+ e^{-2ik_- d}. \quad (4.37)$$

The transmitted wave (4.34) propagates in the  $z$ -direction, as expected. Its intensity is less than that of the wave inside the medium, which should be taken into account if quantitative estimates of the angular momentum transferred into the medium are to be made by measuring the transmitted light, but again, this is the expected result.

The properties of the reflected beam are also not surprising. The Poynting vector for (4.36) can be calculated and it is found to be almost exactly the negative of the Poynting vector (4.13). The only differences are its magnitude and the phase factor  $e^{i(k_+ - k_-)z}$ , which changes to  $e^{i(k_+ - k_-)(2d - z)}$ . The phase factor is different because the light has already propagated through the sample a distance  $d$ , but since  $(k_+ - k_-)d \ll 1$ , the difference in phase is negligible for the first reflected wave. (The importance of the phase in multiple reflections has already been mentioned in the previous section and is discussed further below.) The difference of magnitudes is also clear: not all the light in the medium is reflected. Therefore the energy density of the reflected wave is less than that of the incident wave. However, this changes the amplitude of both the  $\hat{\mathbf{z}}$  and  $\hat{\mathbf{y}}$  components by the same factor and the direction of the reflected Poynting vector is almost exactly opposite that of the incident vector (4.13).

Since the Poynting vector of the reflected beam is opposite that of the incident beam, the light returns along its original path. Both the direction of travel and the helicity of the light are inverted, so the effective magnetic field due to the reflected beam will be similar to that from the incident beam, i.e. it will point in the same direction, but will be smaller in magnitude. Ignoring the phase shift (for now) and setting  $k_+ = k_- = k\sqrt{\epsilon}$ , the additional effective magnetic field is

$$\mathbf{B}_{\text{reflected}} = \left( \frac{\sqrt{\epsilon} - 1}{\sqrt{\epsilon} + 1} \right)^2 \mathbf{B}_{\text{incident}}, \quad (4.38)$$

where  $\mathbf{B}_{\text{incident}}$  is the field given in Eq. (4.28).

Eq. (4.38) allows an approximate generalization to the treatment of multiple scattering. When the reflected wave returns to the original interface, some part of it will reflect again and a similar equation will hold. Therefore the total effective magnetic field will be a geometric series:

$$\begin{aligned} \mathbf{B}_{\text{total}} &= \mathbf{B}_{\text{incident}} + \left( \frac{\sqrt{\epsilon} - 1}{\sqrt{\epsilon} + 1} \right)^2 \mathbf{B}_{\text{incident}} + \left( \frac{\sqrt{\epsilon} - 1}{\sqrt{\epsilon} + 1} \right)^4 \mathbf{B}_{\text{incident}} + \dots \\ &= \frac{(\sqrt{\epsilon} + 1)^2}{4\sqrt{\epsilon}} \mathbf{B}_{\text{incident}}. \end{aligned} \quad (4.39)$$

This is the same as Eq. (4.31), the equation describing the IFE in small systems, diminished by a factor  $1/\sqrt{\epsilon}$ . It will be shown below that this factor arises due to the incorrect treatment of interference. Eq. (4.39) also ignores the cumulative phase change from multiple reflections and, even if it correctly included interference, would only apply to extremely thin samples. A correct treatment of multiple reflection is presented below, but first, I discuss the effect of the reflected beam on angular momentum.

### 4.5.2 Loss of angular momentum

Although the reflected wave allows the approximate recovery of Eq. (4.31), it brings with it an alarming contradiction. To see this, consider a short pulse traveling through the medium. As it travels, it is laterally displaced, gathering angular momentum. Assume that immediately before colliding with the second interface, it has angular momentum density  $\mathbf{L}_0$ . After hitting the interface, some part of it is transmitted. The transmitted wave continues in the same direction with the same displacement and therefore carries off the angular momentum  $T\mathbf{L}_0$ , where  $T$  is the transmission coefficient. However, the reflected wave keeps the same displacement (at least initially) but travels in the opposite direction. Its angular momentum is therefore  $-R\mathbf{L}_0$ . The total angular momentum is now  $(T - R)\mathbf{L}_0 = (1 - 2R)\mathbf{L}_0 \neq \mathbf{L}_0$ . Angular momentum is not conserved!

Of course, angular momentum must be conserved. The resolution of this contradiction lies in the radiation pressure due to the reflected beam [82]. Part of the initial linear momentum of the beam has been transferred to the material when reflection occurs. This transfer occurs at a point that is laterally offset from the origin (the point at which the original beam was incident), so the material gains a mechanical angular momentum. The sample will start to rotate.

In the current formulation, the spatial dependence of  $\epsilon_{ij}$  (through which the position of the sample enters) is fixed. The sample cannot rotate. Therefore angular momentum transfer to the sample cannot occur and angular momentum is apparently not conserved. Maxwell's equations for a fixed sample do not conserve angular momentum<sup>4</sup>.

The rotation of the sample will be small and is unlikely to be seen in experiments. However, this extra source of angular momentum shows that *if the angular momentum transferred to the magnetization is determined by calculating the angular momentum of the light, the second interface cannot be included.*

Consider what would happen if both interfaces were included in the calculation, but the rotation of the sample was ignored. The total angular momentum transferred to the magnetization could be calculated from the angular momentum of the light. The angular momentum of the light is divided into two contributions, one from the light that is transmitted through the second interface and the other, back-reflected beam that emerges from the first interface. The latter emerges at the origin and carries no angular momentum. The former, on the other hand, carries some angular momentum. However, it has a lower intensity than the intensity of light that initially entered the sample and the estimated

---

<sup>4</sup>This is analogous to the well-known fact that Maxwell's equations do not conserve linear momentum for a fixed sample. Any light reflected by a medium has its linear momentum changed. If the motion of the sample is neglected, momentum is not conserved. The resolution of this is to note that the medium takes up the momentum and starts moving, but this motion is generally not noticeable on any reasonable timescale.

magnetic field will be even less than (4.28). This is true even if phase changes are excluded. Multiple reflections, which should *increase* the IFE, will diminish it if the rotation of the sample is excluded.

However, there is a way to determine the strength of the IFE including the effect of multiple reflections. This relies on the discussion following Eq. (4.20), that the torque on the magnetization due to the light can be calculated by using  $\mathbf{T} = \mathbf{P} \times \mathbf{E}$ . In Eq. (4.20), this torque was balanced by the torque on the light. If the second interface is included, the torque on the magnetization will be balanced by the torque on the light plus the mechanical torque on the sample; nonetheless, it will be the correct torque and yield the correct strength of the IFE. This is evaluated in the next section.

### 4.5.3 Multiple reflection

The solution to Maxwell's equations with two interfaces is more intricate than that with a single interface, but the method is the same. That is, one starts by evaluating the normal modes in each medium — the vacuum for  $z < 0$ , the material for  $0 < z < d$  and the vacuum for  $z > d$  — and then the boundary conditions are used to evaluate the coefficients [82]. Assuming that the magnetization  $\mathbf{m} = m\hat{\mathbf{y}}$ , the normal modes for the material have already been determined in Eq. (4.10). The calculation of the coefficients is straightforward and the electric field in the medium is

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( E_+ \mathbf{n}_+ e^{i(\omega t - k_+ z)} + E_- \mathbf{n}_- e^{i(\omega t - k_- z)} + \tilde{E}_+ \mathbf{n}_+ e^{i(\omega t + k_+ z)} + \tilde{E}_- \mathbf{n}_- e^{i(\omega t + k_- z)} + c.c \right) \quad (4.40)$$

with

$$\begin{aligned} E_+ &= \frac{2k(k_+ + k)}{(k_+ + k)^2 - (k_+ - k)^2 e^{-2ik_+ d}} E_0^y, & \tilde{E}_+ &= \frac{k_+ - k}{k_+ + k} e^{-2ik_+ d} E_+ \\ E_- &= \frac{2k(k_- + k)}{(k_- + k)^2 - (k_- - k)^2 e^{-2ik_- d}} E_0^x, & \tilde{E}_- &= \frac{k_- - k}{k_- + k} e^{-2ik_- d} E_- \end{aligned} \quad (4.41)$$

The polarization density  $\mathbf{P}$  is determined from the electric field (4.40) and the dielectric tensor (4.6) and the torque is calculated from  $\mathbf{T} = \mathbf{P} \times \mathbf{E}$ , yielding

$$\mathbf{T} = \hat{\mathbf{x}} \frac{\gamma m}{8\pi} \text{Re}(iw) - \hat{\mathbf{z}} \frac{\gamma^2 m^2}{8\pi\epsilon} \text{Re}(w). \quad (4.42)$$

Here, I have introduced the auxiliary variable

$$w = \left( E_+^* e^{ik_+ z} + \tilde{E}_+^* e^{-ik_+ z} \right) \left( E_- e^{-ik_- z} + \tilde{E}_- e^{ik_- z} \right) \quad (4.43)$$

for notational convenience.

The main contribution to the torque (4.42) is in the  $x$ -direction; the  $\hat{\mathbf{z}}$  component is smaller than this by a factor of  $\gamma m/\epsilon \ll 1$ . The  $\hat{\mathbf{x}}$  component has a similar form to the torque found in Eq. (4.19), except that the sign has changed (since this is the torque on the magnetization, not on the light), the backward-propagating beam is included and the coefficients  $E_+$  and  $E_-$  are different.

The  $\hat{\mathbf{z}}$  component of the torque is small but nonzero. This component arises because the light changes its polarization as it propagates through the sample, thereby changing

spin angular momentum. The comparative smallness of this effect is related to the large propagation distances that are necessary before the change of polarization is noticeable, as described in section 4.4. Since this component of the torque is quadratic in  $m$ , it cannot be thought of as being related to an effective magnetic field and it will not be included in the calculation of the magnetic field below.

In their full generality, Eqs. (4.41) and (4.43) can only be analyzed numerically. However, approximating  $k_+z = k_-z$  allows us to rewrite  $w$  in a more useful form<sup>5</sup>:

$$w = 4 \frac{(\sqrt{\epsilon} + 1)^2 + (\sqrt{\epsilon} - 1)^2 + 2(\sqrt{\epsilon} + 1)(\sqrt{\epsilon} - 1) \cos [2k\sqrt{\epsilon}(d + z)]}{(\sqrt{\epsilon} + 1)^4 + (\sqrt{\epsilon} - 1)^4 - 2(\sqrt{\epsilon} + 1)^2(\sqrt{\epsilon} - 1)^2 \cos (2k\sqrt{\epsilon}d)} (E_0^y)^* E_0^x. \quad (4.44)$$

This equation is correct up to terms of the order of  $kd\gamma^2 m^2/\epsilon^2$ , which can be neglected for thin samples. Note that for circularly polarized light,  $w$  is purely imaginary and the  $z$ -component of the torque (4.42) vanishes. Similarly, the  $x$ -component of the torque vanishes for linearly polarized light. The IFE will therefore be extremely weak, i.e. of the order of  $\gamma^2 m^2/\epsilon^2 kd$ , for linearly polarized light. (In fact, it vanishes for linearly polarized light, but gains a contribution because the light *becomes* circularly polarized after traversing the sample several times.)

Once again, the equation for the torque (4.42) is not sufficient to identify the effective magnetic field. However, if the Zeeman energy for a general magnetization is calculated, a similar problem will arise as in section 4.3 where the transmission coefficients depend linearly on  $m$ . Instead of repeating this analysis, it is easier simply to use the fact that the effective magnetic field must be independent of the magnetization [32] and to determine the Zeeman energy for  $\mathbf{m} = m\hat{\mathbf{y}}$ , which vanishes. Then by moving to a coordinate-free expression for  $(E_0^y)^* E_0^x$ , one finds the effective magnetic field

$$\mathbf{B}_{\text{eff}} = -\frac{\gamma}{4\pi} \frac{(\sqrt{\epsilon} + 1)^2 + (\sqrt{\epsilon} - 1)^2 + 2(\sqrt{\epsilon} + 1)(\sqrt{\epsilon} - 1) \cos [2k\sqrt{\epsilon}(d + z)]}{(\sqrt{\epsilon} + 1)^4 + (\sqrt{\epsilon} - 1)^4 - 2(\sqrt{\epsilon} + 1)^2(\sqrt{\epsilon} - 1)^2 \cos (2k\sqrt{\epsilon}d)} (i\mathbf{E}_0 \times \mathbf{E}_0^*). \quad (4.45)$$

The field (4.45) has a complicated dependence on  $d$  and  $z$ . This is related to the fact that internally reflected light has its polarization changed and can either enhance or reduce the IFE, and will do so differently at different depths in the sample. Although the exact dependence on  $d$  and  $z$  is difficult to visualize unless all the parameters for a given material are known, the bounds on the effective magnetic field can be determined. The field is found to satisfy

$$\frac{\gamma}{4\pi(\epsilon + 1)^2} |i\mathbf{E}_0 \times \mathbf{E}_0^*| \leq \mathbf{B}_{\text{eff}} \leq \frac{\gamma}{16\pi} |i\mathbf{E}_0 \times \mathbf{E}_0^*|. \quad (4.46)$$

The upper bound coincides with the magnetic field for a small system (4.31).

If the sample is extremely thin compared to the wavelength of the light,  $2k\sqrt{\epsilon}d \ll 1$ , the magnetic field (4.45) becomes equal to the upper bound of Eq. (4.46) for all  $z$ . In this case, Eq. (4.45) coincides with Eq. (4.31). This is the expected result, since an extremely thin sample is correctly viewed as a “small” sample.

<sup>5</sup>Note that unlike the case where multiple reflections are taken into account stepwise, extremely long path lengths of the light do not occur in the correct treatment of multiple reflection. The long path lengths are implicitly accounted for in the interference pattern. This has the effect that neither  $k_+$  nor  $k_-$  is ever multiplied by a distance larger than  $d$ . Therefore  $k_+z = k_-z$  only needs to be assumed for  $0 < z < d$ ; the validity of this approximation is easily verified.

The upper bound of Eq. (4.46) coincides with the classical expression for the IFE (2.16). The equation for a general sample, Eq. (4.45), shows that the correct internal electric field must be used to determine the magnetic field, not simply the field of the incident light. These fields are equal for thin samples.

The question can be raised, why are we now able to recover the correct result for a small system (4.31), while in Eq. (4.39), it was only recovered up to a factor of  $1/\sqrt{\epsilon}$ ? Surely the limit  $2k\sqrt{\epsilon}d \ll 1$ , the limit in which phase shifts can be neglected, is identical for the two treatments.

The answer to this question lies in interference. The forward- and backward-propagating waves in the medium interfere with each other. Eq. (4.39) was derived by ignoring interference and directly summing the effective magnetic fields. Since the magnetic fields are quadratic in  $\mathbf{E}_0$ , this is similar to ignoring interference when summing intensities of light [84]. The correct method of performing the calculation is to sum the amplitudes of the electric fields and then convert the result to an effective magnetic field. In terms of a geometric series, this becomes

$$\begin{aligned} \mathbf{B}_{\text{total}} &= \left[ 1 + \left( \frac{\sqrt{\epsilon} - 1}{\sqrt{\epsilon} + 1} \right) + \left( \frac{\sqrt{\epsilon} - 1}{\sqrt{\epsilon} + 1} \right)^2 + \dots \right]^2 \mathbf{B}_{\text{incident}} \\ &= \frac{(\sqrt{\epsilon} + 1)^2}{4} \mathbf{B}_{\text{incident}}. \end{aligned} \quad (4.47)$$

Comparing this to Eq. (4.39) shows that the extra  $1/\sqrt{\epsilon}$  has been removed and the correct result obtained.

Thus all possible contradictions have been resolved. Eq. (4.45) is the correct generalization of the results of the previous chapter (and the IFE in general) to extended systems, at least to isotropic systems, and the question of angular momentum flow has been resolved. Angular momentum is transferred between the light and the magnetization and also between the light and the material. The conclusion from chapter 3 still holds: no other sources of angular momentum are needed for precessional switching via the IFE.

The treatment for anisotropic systems is far more intricate than the current work and useful analytical results are unlikely. The method for deriving the torque is straightforward: calculate the polarization density  $\mathbf{P}$  and the electric field  $\mathbf{E}$  and use  $\mathbf{T} = \mathbf{P} \times \mathbf{E}$ . However, the magnetic field is related only to that part of the torque that is proportional to the magnetization; the torque due to crystalline anisotropy will cause deformations of the crystal, not motion of the magnetization.

There are several possible generalizations of the above work. First, we have used the fact that an effective Hamiltonian does not depend on the state of the system to ignore the precession of  $\mathbf{m}$ . While this could be rigorously argued for the case of an infinite solid using the adiabatic motion of the magnetization, it is not so clear for the case of multiple reflection. Strictly speaking, the formation of the interference pattern in the solid takes an infinitely long time. On such a timescale, the motion of the magnetization is not adiabatic. Reflected light will interact with a different magnetization. While it is expected that this will simply complicate the equations without changing the effective magnetic field (because the effective field doesn't depend on  $\mathbf{m}$ ), this should be shown more rigorously.

Secondly, the above derivation using interference patterns assumes that the light pulse is quite long compared to the thickness of the sample, i.e. that the front of the pulse

has traversed the sample several times and set up interference patterns before the end of the pulse enters the sample. If the pulse is short (e.g. a 100 fs pulse is  $30\ \mu\text{m}$  long in the vacuum, which is roughly half the thickness of the samples used in Ref. [99]) there will be less interference and the magnetic field strength will lie between (4.39) and (4.45). Furthermore, this magnetic field will not occur all within the timescale of a single pulse — magnetic fields due to reflected pulses will only occur later. If the dynamics of the magnetization are to be calculated for comparison with experiment, this delay in the magnetic field will have to be taken into account. These generalizations would certainly be interesting and useful, but they are beyond the scope of this thesis.

To conclude this chapter and to allow future comparison with experiment, it is useful to provide an estimate of the displacement of the light. It is convenient to have this in terms of other measurable quantities of the material, e.g. the Faraday rotation angle (4.5). Setting  $k_+ = k_-$  in Eq. (4.16) and assuming the condition  $(k_+ - k_-)d \ll 1$ , we find that for circularly polarized light ( $E_0^x = \pm iE_0^y$ ),

$$\mathbf{R} = d \left( \hat{\mathbf{z}} \mp \hat{\mathbf{y}} \frac{\gamma m}{\epsilon} \right). \quad (4.48)$$

Thus the lateral displacement is

$$\mathbf{R}_{\parallel} = \mp \hat{\mathbf{y}} \beta \frac{2}{k\sqrt{\epsilon}}, \quad (4.49)$$

where  $\beta$  is the Faraday rotation angle. In terms of  $\beta$ , the condition  $(k_+ - k_-)d \ll 1$  becomes

$$\beta^2 \ll \frac{1}{2} k \sqrt{\epsilon} d. \quad (4.50)$$

Assuming the wavelength of the light is  $0.5\ \mu\text{m}$ ,  $\beta \approx 1$  radian and  $\epsilon = 16$  (which is rather large, but the results do not depend strongly on  $\epsilon$ ), the condition (4.50) is valid for  $d \gg 1\ \mu\text{m}$ . Therefore, we may take an arbitrarily large sample provided the total Faraday rotation through the sample is around 1 radian. (A 3 mm sample of ZnSe under a magnetic field of 1 T yields this rotation [206].) Given that  $\beta = 1$  radian, the lateral displacement (4.49) will be roughly  $1\ \mu\text{m}$ , which is a measurable displacement (see, e.g., Ref. [169]).

## Chapter 5

# The Difficulty of Measurement

### 5.1 Measuring ultrafast magnetization dynamics

Ultrafast magnetic switching experiments require ultrafast magnetic imaging techniques. Unless time-resolved images of the magnetization are available, the timescale on which reversal occurs cannot be determined, even approximately.

Since the advent of ultrafast switching, *optical* measurement methods have been favoured. These offer several advantages: ultrafast switching experiments are already optics-based, which simplifies the setup and timing of measurements; *ultrashort* optical pulses can be made, providing an “instantaneous” measurement of the system; pulses can be made with very low intensity, so that measurement does not affect the dynamics; optical technology is in an advanced state, which means that properties of the probe pulse are reliably measured; repetition of the experiment is straightforward, enabling good data collection. However, the most significant advantage of optical measurement methods is that *no other methods of magnetic imaging provide the necessary temporal resolution*.

Of course, optical methods have some problems. The main problem is that they are an *indirect* measure of magnetization. The measured quantities are related to the effect of the sample on a probe pulse; the magnetization must be derived from these quantities. Separating the magnetization from other optical effects is nontrivial. In Chapter 1, I mentioned the problem of bleaching, where the depletion of electrons with energy similar to that of the probe pulse causes a change of the signal that has nothing to do with magnetism. This problem was important in the early development of ultrafast demagnetization and has been solved by using the two-colour method. More recently, experiments were performed where the magnetization was recovered from optical measurements using sum rules derived for an equilibrium distribution [178]; whether these sum rules are valid in highly nonequilibrium situations is still uncertain. Even close to equilibrium, non-magnetic influences on the probe pulse can make recovery of the magnetization difficult. In this chapter, I present our investigation of determining the magnetization in the presence of optical birefringence. The results of this chapter were published in Ref. [203]<sup>1</sup>

Kimel *et al.* [66, 99, 179], whose experiments inspired a large portion of this thesis, use Faraday rotation to measure magnetization dynamics. Faraday rotation is the rotation

---

<sup>1</sup>These results are reused with permission from S. R. Woodford, A. Bringer and S. Blügel, *Journal of Applied Physics* **101**, 053912 (2007). Copyright 2007, American Institute of Physics.

of linearly polarized light as it propagates through a magnetic material. (See Chapter 2 for a detailed explanation.) In isotropic media, the relationship between Faraday rotation and magnetization is well-known (see, e.g., Refs. [116, 210]). However, the material used in Ref. [99],  $\text{DyFeO}_3$ , is optically birefringent, which complicates the interpretation significantly. Below, I show how to correctly interpret the rotation. In addition, I correct several deficiencies of the analysis of Ref. [99] and I make recommendations to improve the reliability of experiments.

For convenience, I briefly repeat here the details of Ref. [99]. Using an ultrashort, circularly polarized laser pulse, magnetic oscillations were excited in  $\text{DyFeO}_3$ , which is a canted antiferromagnet with  $m_0 \approx 8$  G. This excitation is caused by the inverse Faraday effect (IFE). From the phenomenological relationship between the Faraday effect and IFE [150, 152, 187], Kimel *et al* calculated the strength of the effective magnetic field to be around 0.3 T, with a FWHM of 200 fs [99]. Magnetization dynamics were measured using Faraday rotation and they estimate the amplitude of magnetic oscillations as  $(m - m_0)/m_0 \approx 1/16$ . The effect of their calculated effective magnetic pulse has been simulated in a micromagnetic model using the Landau-Lifshitz equation [148, 149], which yields a similar amplitude of oscillation.

## 5.2 Orthoferrites

There are certain aspects of the interpretation of Faraday rotation measurements that pass unmentioned in [99] and that deserve more attention. These aspects arise in any experiments performed on orthoferrites, which are not only birefringent [182, 183] but also have enhanced magneto-optical properties due to the so-called “anisotropic orbital quenching” [87]. To obtain a reasonable estimate of the amplitude of magnetization dynamics, these effects must be taken into account.

It is well known (see, e.g., Ref. [210]) that birefringence limits the rotation angle of plane-polarized light: instead of being linearly proportional to the sample thickness  $d$ , the rotation is a periodic function of  $d$  and can vanish if  $d$  is poorly chosen. In addition, there exists an *antiferromagnetic* Faraday effect, where the antiferromagnetic vector  $\mathbf{l}$  causes rotation of the light’s polarization plane via “anisotropic orbital quenching” [87]. (Symmetry considerations show that this effect is only possible in materials with inequivalent magnetic sublattices, like the orthoferrites [156].) Since the ferromagnetic- and antiferromagnetic vectors react differently to external fields, the two rotations can oppose each other, leading to a reduced observable rotation [108].

For the remainder of this chapter, I assume that the material has the magnetic and crystal structure of the orthoferrites. More precisely, I consider an orthorhombic crystal, antiferromagnetically ordered with the antiferromagnetic vector  $\mathbf{l}$  directed along the  $a$ -axis. Due to a small canting of the antiferromagnetic sublattices, there is a weak residual ferromagnetic component  $\mathbf{m}$  directed along the  $c$ -axis [200]. Birefringence is due to the crystal structure, so the birefringence axes coincide with the crystal axes. The light beam is assumed to propagate parallel to the  $c$ -axis. Because I am interested in near-equilibrium properties, the dynamics of the antiferromagnetic vector will have a negligible effect. Furthermore, I assume that the ferromagnetic vector remains directed along  $c$  at all times and only its magnitude varies. Therefore  $\mathbf{l} = l\hat{\mathbf{a}}$  and  $\mathbf{m} = m\hat{\mathbf{c}}$ , with  $l$  a constant

and  $m$  a variable. In what follows,  $x$ ,  $y$ , and  $z$  represent the  $a$ ,  $b$  and  $c$  directions of the orthorhombic structure, respectively.

### 5.3 Faraday rotation in a birefringent weak ferromagnet

The Faraday effect in isotropic, non-magnetically ordered materials predicts that, after propagating through the material, linearly polarized light remains linearly polarized and has its polarization plane rotated by the angle (see Eq. (2.7) and Refs. [116, 187, 210])

$$\beta = V H d, \quad (5.1)$$

where  $d$  is the path length of the light,  $H$  is an external magnetic field and  $V$  is the Verdet constant. Eq. (5.1) is only valid when the material is isotropic, optically inactive and non-magnetically ordered; how Eq. (5.1) changes as birefringence and magnetic order are introduced is the subject of this section. (Optical activity is the rotation of linearly polarized light due to a structural asymmetry of the medium. Since I am most interested in the orthoferrites, which are optically inactive, I do not consider the effect of optical activity.)

Orthoferrites are orthorhombic structures. In the absence of magnetic order, the dielectric tensor is symmetric and diagonal:  $\epsilon = \text{diag}(\epsilon_x, \epsilon_y, \epsilon_z)$ . In general,  $\epsilon_x \neq \epsilon_y$ , so the material is birefringent. Since the light is assumed to propagate along the  $z$ -axis,  $\epsilon_z$  does not enter the analysis. Magnetic ordering introduces antisymmetric terms into this tensor [156]. It can be shown (see, e.g., Refs. [109, 207] and the references therein) that the antisymmetric part of  $\epsilon_{ij}$  is proportional to the dual of the ferromagnetic vector  $\mathbf{m}$ , i.e.  $\epsilon_{xy}$  contains  $m_z$ , etc. This has already been used in the previous chapters. However, it is not true for the antiferromagnetic vector  $\mathbf{l}$ :  $l_x$  transforms in the same way as  $m_z$ , so  $l_x$  enters  $\epsilon_{xy}$  and not  $\epsilon_{yz}$ , as one might naively expect [109, 207].

As noted above, the deviation from magnetic equilibrium is assumed to be small. Then the antiferromagnetic vector  $\mathbf{l}$  has fixed length and direction, while the ferromagnetic vector  $\mathbf{m}$ , which is the result of a small canting of the antiferromagnetic order, has variable length and direction. To simplify the analysis, I fix the direction of  $\mathbf{m}$  and allow only the magnitude to vary, i.e.  $\mathbf{m} = (0, 0, m)$ . Under these assumptions,  $\epsilon_{xz} = \epsilon_{yz} = 0$  and

$$\epsilon_{xy} = -\epsilon_{yx} = i(f_1 + \xi_1 \frac{m}{m_0}). \quad (5.2)$$

The phenomenological parameters  $f_1$  and  $\xi_1$  describe the antiferromagnetic and ferromagnetic contributions to  $\epsilon_{xy}$  respectively, and  $m_0$  is the equilibrium magnitude of  $\mathbf{m}$ . Introducing the ‘‘ferromagnetic deviation’’  $\kappa = (m - m_0)/m_0$  allows Eq. (5.2) to be rewritten as

$$\epsilon_{xy} = i\gamma_0(1 + \zeta\kappa) \equiv i\gamma. \quad (5.3)$$

Here,  $\gamma_0 = (f_1 + \xi_1)$  describes the equilibrium magneto-optical properties of the system and  $\zeta = \xi_1/\gamma_0$  measures the ferromagnetic contribution to magneto-optics. Experimental values of  $\gamma_0$  for the various rare-earth orthoferrites at different wavelengths of light can be found in Ref. [182]. The parameter  $\zeta$  is more difficult to find in the literature; it has been evaluated as  $\zeta \approx 0.23 \pm 0.13$  for  $\text{YFeO}_3$  at  $T = 295$  K, using light with wavelength  $\lambda = 0.6328$   $\mu\text{m}$  [207]. This value should vary with frequency and temperature, although it should

not vary significantly for different rare-earth orthoferrites, as the rare-earth species has negligible effect on the magnetic order at high temperatures [105]. Since the dependence of  $\zeta$  on temperature and frequency is not known, I fix  $\zeta = 0.23$  for the remainder of this chapter.

Eq. (5.3) is valid provided the magnetic field that caused the change of magnetization from equilibrium has stopped acting, i.e. Eq. (5.3) holds while the system relaxes towards equilibrium. If the magnetic field is still present, then a diamagnetic term should be added [207]. This was neglected in Ref. [108], which explains why their estimate  $\zeta \approx 0.13$  differs from that found in Ref. [207].

The above assumptions yield [183, 210]

$$\mathbf{D} = \begin{pmatrix} \epsilon + \eta & i\gamma & 0 \\ -i\gamma & \epsilon - \eta & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix} \mathbf{E}. \quad (5.4)$$

The strength of the birefringence is determined by  $\eta$ . Assuming no conduction current exists,  $\mathbf{D}$  and  $\mathbf{E}$  obey the wave equation

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{D}}{\partial t^2} = \nabla^2 \mathbf{E}. \quad (5.5)$$

All magneto-optical effects have been placed in the dielectric tensor and therefore the permeability  $\mu$  is a scalar quantity [116, 151]; as in the previous chapters, I set  $\mu = 1$ . Eq. (5.5) can be solved for a plane wave propagating along the  $z$ -axis:

$$\mathbf{E} = \begin{pmatrix} E_x \\ E_y \\ 0 \end{pmatrix} e^{i(\omega t - \tilde{k}z)} \quad (5.6)$$

Substituting (5.6) into Eq. (5.5) and using Eq. (5.4) leads to

$$\begin{pmatrix} \epsilon + \eta - \tilde{k}^2/k^2 & i\gamma \\ -i\gamma & \epsilon - \eta - \tilde{k}^2/k^2 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = 0. \quad (5.7)$$

This equation provides both the admissible wavevectors

$$k_{\pm} = k \sqrt{\epsilon \pm \gamma \sqrt{1 + \Gamma^2}} \quad (5.8)$$

and the normal modes in the material:

$$\begin{aligned} \mathbf{n}_+ &= \begin{pmatrix} 1 \\ i(\Gamma - \sqrt{1 + \Gamma^2}) \end{pmatrix} \\ \mathbf{n}_- &= \begin{pmatrix} 1 \\ i(\Gamma + \sqrt{1 + \Gamma^2}) \end{pmatrix}. \end{aligned} \quad (5.9)$$

Here, I have introduced  $\Gamma = \eta/\gamma$  to describe the relative strength of the birefringence compared with magnetic effects. The medium admits two normal modes, which are elliptically polarized and which propagate at different velocity (since  $k_+ \neq k_-$ ). The difference

in velocities is related to the magnitudes of  $\gamma$  and  $\eta$ , i.e. to both the magnetization and the birefringence. In general, the light will emerge from the material elliptically polarized; “rotation” of the light is therefore defined as the change of direction of the major axis of the polarization ellipse.

Now assume that plane polarized light is normally incident at  $z = 0$ , i.e.

$$\mathbf{E}_{z=0} = |E| \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} e^{i\omega t}, \quad (5.10)$$

where  $\theta$  is the angle between the polarization plane and the  $x$ -axis. The reflection of both normal modes is approximately equal and can therefore be ignored — its only effect is to reduce the intensity<sup>2</sup>. Upon entering the material, the field (5.10) decomposes into normal modes

$$\mathbf{E} = \frac{1}{2}|E| \left( \alpha_+ \mathbf{n}_+ e^{i(\omega t - k_+ z)} + \alpha_- \mathbf{n}_- e^{i(\omega t - k_- z)} \right); \quad (5.11)$$

$\alpha_{\pm}$  are complex parameters determined from Eqs. (5.9) and (5.10).

As the wave (5.11) propagates through the medium, its polarization rotates. Since the goal is to determine this rotation relative to the initial angle  $\theta$ , it is convenient to rotate to the coordinate frame ( $x'$ ,  $y'$ ) in which the light is initially polarized along the  $x'$ -axis. In this new coordinate frame, after the light has propagated a distance  $d$ , the electric field vector is

$$\mathbf{E}_{z=d} = \frac{|E| e^{i(\omega t - \bar{k}d)}}{\sqrt{1 + \Gamma^2}} \begin{pmatrix} \sqrt{1 + \Gamma^2} \cos \delta d - i\Gamma \sin \delta d \cos 2\theta \\ -\sin \delta d + i\Gamma \sin \delta d \sin 2\theta \end{pmatrix} \quad (5.12)$$

where  $\bar{k} = (k_+ + k_-)/2$  and  $\delta = (k_+ - k_-)/2$ . Note that under the assumption that  $\sqrt{\eta^2 + \gamma^2} \ll \epsilon$  (which is true in most materials, including the orthoferrites), Eq. (5.8) allows us to rewrite the wavevector mismatch  $\delta$  as

$$\delta \approx \frac{k\gamma\sqrt{1 + \Gamma^2}}{2\sqrt{\epsilon}}. \quad (5.13)$$

The polarization ellipse of the outgoing light (5.12) is described by the lengths of the major and minor axes and the orientation of the ellipse (see, e.g., Ref. [118] for details). Relative to the initial angle (i.e. relative to the  $x'$ -axis), the direction of polarization has rotated through the angle  $\beta$ , where

$$\tan 2\beta = \frac{2\text{Re}(E_x^* E_y)}{|E_x|^2 - |E_y|^2} = \frac{-\sqrt{1 + \Gamma^2} [\sin 2\delta d + \Gamma(1 - \cos 2\delta d) \sin 2\theta \cos 2\theta]}{\cos 2\delta d [1 + \Gamma^2 \sin^2 2\theta] + \Gamma^2 \cos^2 2\theta}. \quad (5.14)$$

The lengths of the major axis  $b_+$  and minor axis  $b_-$  of the polarization ellipse are:

$$\frac{b_{\pm}}{|E|} = \frac{1}{|E|} \left( \sqrt{|E|^2 + 2\text{Im}(E_x^* E_y)} \pm \sqrt{|E|^2 - 2\text{Im}(E_x^* E_y)} \right)$$

<sup>2</sup>Throughout this chapter, I am more interested in ratios between different components of light, rather than absolute intensities. Since both normal modes have similar reflection coefficients, including reflection leads to a diminishing of the total intensity without affecting ratios of components. As such, reflection is neglected in this chapter.

$$\begin{aligned}
&= \left\{ 1 + \frac{\Gamma \sin 2\delta d \sin 2\theta}{\sqrt{1 + \Gamma^2}} - \frac{\Gamma(1 - \cos 2\delta d) \cos 2\theta}{1 + \Gamma^2} \right\}^{1/2} \\
&\pm \left\{ 1 - \frac{\Gamma \sin 2\delta d \sin 2\theta}{\sqrt{1 + \Gamma^2}} + \frac{\Gamma(1 - \cos 2\delta d) \cos 2\theta}{1 + \Gamma^2} \right\}^{1/2}.
\end{aligned} \tag{5.15}$$

These equations are the key to investigating the effect of magnetization changes on the outgoing light beam. Since magnetic effects are described by  $\gamma$ , which is contained in both  $\delta$  (see Eq. (5.13)) and  $\Gamma = \eta/\gamma$ , the dependence on magnetization is complicated and analytical results are only available in certain limits.

## 5.4 Limiting cases

### 5.4.1 Zero birefringence

The first limit that should be examined is the case of zero birefringence. In this limit,  $\Gamma = 0$  and Eq. (5.14) reduces to

$$\tan 2\beta = -\tan 2\delta d, \tag{5.16}$$

i.e. the polarization plane of the light rotates through an angle  $\beta = -\delta d = -dk\gamma/2\sqrt{\epsilon}$ . Furthermore, the light emerges linearly polarized:  $b_- = 0$ . I define the angle  $\beta_0$  as the rotation caused when the sample is in magnetic equilibrium:

$$\beta_0 = -\frac{dk\gamma_0}{2\sqrt{\epsilon}}. \tag{5.17}$$

Using the definition of  $\gamma$  from Eq. (5.3) yields

$$\beta = \beta_0 + \beta_0\zeta\kappa. \tag{5.18}$$

The dependence of the angle  $\beta$  on magnetization is contained in  $\kappa = (m - m_0)/m_0$ .

Eq. (5.18) is the analogue of Eq. (5.1) for canted-antiferromagnetic materials (also called weak ferromagnets [210]). The effect of magnetic ordering is most clear if we consider a magnetic configuration in which the magnetization vanishes. In the case of Eq. (5.1), this occurs for  $H = 0$  and there is no Faraday rotation. On the other hand, for Eq. (5.18) the magnetization vanishes at  $\kappa = -1$ , but Faraday rotation still occurs:  $\beta = \beta_0(1 - \zeta)$ . Since  $\zeta$  is the ferromagnetic contribution to magneto-optics, it is clear that the factor of  $(1 - \zeta)$  implies that the residual rotation is caused by the antiferromagnetic ordering.

The Verdet constant, which describes the magneto-optical response of the material, is defined from the slope of  $\beta(m)$ . Introducing  $m = \chi H$  in Eq. (5.1) (where  $\chi$  is the magnetic susceptibility [82]),

$$V = \frac{\chi}{d} \frac{\partial \beta}{\partial m}.$$

From this definition, Eqs.(5.17)-(5.18) and letting  $\chi$  now represent the  $zz$ -component of the susceptibility in the magnetic medium,

$$V = -\frac{\chi\zeta\beta_0}{dm_0}.$$

This allows a calculation of the Verdet constant, a useful magneto-optical constant.

Finally, note that for the purposes of interpreting experimental data, it is more useful to rewrite Eq. (5.18) as

$$\frac{\beta - \beta_0}{\beta_0} = \zeta \kappa = \zeta \frac{m - m_0}{m_0}. \quad (5.19)$$

This equation shows that the change of rotation, normalized against the equilibrium rotation, is proportional to the change of magnetization, normalized to the equilibrium ferromagnetic moment. The proportionality constant is  $\zeta$ , the ferromagnetic contribution to magneto-optics. Note that for zero birefringence,  $\beta_0$  defined by Eq. (5.17) is nonzero and Eq. (5.19) is always well-defined.

### 5.4.2 Large birefringence, small $\theta$

The second limit that we considered is the relevant limit for analysis of the orthoferrites, i.e. large birefringence. Indeed, the orthoferrites generally satisfy  $\gamma_0 \ll \eta$  (see, e.g., Ref. [31]). For the particular setup described in [99], where  $\text{DyFeO}_3$  is probed by light with wavelength  $\lambda_0 \approx 0.7 \mu\text{m}$ , experimental results reveal that  $|\gamma_0/\eta| \approx 1/30$  [182], so the parameter  $\Gamma_0 \approx 30$ . It will be useful in the subsequent analysis to note that in equilibrium, this system has  $\delta_0 = \pi/20$  rads  $\mu\text{m}^{-1}$  for light with wavelength  $\lambda_0 = 0.7 \mu\text{m}$  and  $\zeta = 0.23$ , as mentioned above.

When birefringence is present, the polarization direction of the incident beam is important. Analysis of Eq. (5.14) for large  $\Gamma$  and arbitrary  $\theta$  is best performed numerically. In this section, I consider only the case where the angle  $\theta$  is small, i.e. the incident light is polarized approximately parallel to the  $x$ -axis (i.e. the  $a$  axis of the crystal.)

First, let me quantify what I mean by “small  $\theta$ ”. Expanding Eq. (5.14) in  $\kappa$ , it is clear that the  $\theta$  dependence can be neglected up to linear order in  $\kappa$  provided

$$\cot \delta_0 d \gg \Gamma_0 \sin 2\theta, \quad (5.20)$$

where  $\Gamma_0 = \Gamma(\kappa = 0)$  and  $\delta_0 = \delta(\kappa = 0)$ . Note the dependence of this condition on  $d$  — for poorly chosen sample thickness, the experimental alignment must be extremely accurate.

To illustrate the order of magnitude of the critical value of  $\theta$ , consider  $\text{DyFeO}_3$  with  $d$  chosen so that  $\cot \delta_0 d \approx 1$  for convenience. Then  $\theta$ -dependence can be neglected provided

$$\sin 2\theta \ll 1/30$$

i.e.  $\theta$  must be less than  $1^\circ$  away from the birefringence axis. This accuracy in alignment is not difficult to achieve and makes the subsequent analysis of experimental results far easier and more reliable.

For  $\theta$  small enough, Eq. (5.14) can be reduced to

$$\beta = \beta_0(1 - \zeta \kappa) - \zeta \kappa \cos 2\delta_0 d \frac{\delta_0 d}{|\Gamma_0|^3}, \quad (5.21)$$

where

$$\beta_0 = -\sin 2\delta_0 d / |2\Gamma_0|, \quad (5.22)$$

and terms of  $\mathcal{O}(\kappa^2)$  have been neglected. Eq. (5.22) shows the well known (see, e.g., Refs. [183, 210]) phenomenon that the rotation induced by the sample in equilibrium

depends on the sample thickness *periodically* when birefringence is strong, not linearly as in the zero-birefringence case described by Eq. (5.17). Note also that the equilibrium ( $\kappa = 0$ ) rotation is constrained to lie between two limiting angles ( $\beta_{max,min} = \pm(2\Gamma_0)^{-1}$ ), another well known fact [183].

Now consider the effect of magnetic deviations on the Faraday rotation angle.  $\Gamma_0^{-3}$  is assumed to be small, so the final term in Eq. (5.21) can be neglected, except when  $\beta_0$  is also small (i.e. for particular values of  $d$ ). If the thickness of the sample is not near the value where  $\beta_0$  vanishes then

$$\frac{\beta - \beta_0}{\beta_0} = -\zeta\kappa, \quad (5.23)$$

which is identical to Eq. (5.19), up to a factor of  $-1$ .

This change of sign arises because the magnetization enters Eq. (5.14) through two different parameters, the wavevector mismatch  $\delta$  and the relative strength of the birefringence  $\Gamma$ . Increasing the magnetization (i.e. increasing  $\gamma$ ) will *increase*  $\delta$  and *decrease*  $\Gamma$ . For zero birefringence, the rotation is governed by changes to  $\delta$  (see Eq. (5.19)), while when birefringence is large, the rotation is dominated by changes in  $\Gamma$ . This is clear from Eq. (5.21), where the first term in  $\kappa$  arises from expanding  $\Gamma$ , while the second, smaller,  $\kappa$ -term is due to the change in  $\delta$ . This explains the difference of sign between Eq. (5.19) and Eq. (5.23).

When Eq. (5.23) holds, i.e. when the light is accurately aligned and the equilibrium rotation  $\beta_0$  is nonzero, then magnetization dynamics are easily inferred from the measured rotation. The accuracy of this correspondence is determined by the accuracy of measurement of  $\beta$  and the accuracy with which  $\zeta$ , the ferromagnetic magneto-optical parameter, is known. It is important to note that the angle by which the rotation must be normalized is not the limiting angle  $\beta_{max}$ , but the equilibrium angle  $\beta_0$ . Since this depends both on the material and also on the sample thickness,  $\beta_0$  should be separately measured for each sample used. Furthermore, note that if  $\beta_0$  is small (but still larger than  $\Gamma_0^{-3}$ , so that Eq. (5.23) holds), the change of rotation will be extremely small, even for large changes in the magnetization.

Eq. (5.23) is valid provided

$$\sin 2\delta_0 d \gg \cos 2\delta_0 d \frac{2\delta_0 d}{\Gamma_0^2}.$$

Since  $\Gamma_0$  is large (strong birefringence), this condition fails when  $\sin 2\delta_0 d$  is extremely small, i.e.  $2\delta_0 d \approx n\pi$  (with  $n$  an integer). Writing  $2\delta_0 d = n\pi + \varphi$ ,  $|\varphi| \ll 1$ , it is clear that Eq. (5.23) is invalid if

$$|\varphi| \leq \frac{n\pi}{\Gamma_0^2}. \quad (5.24)$$

In this case, the final term in Eq. (5.21) must be included in the analysis. Depending on the sign of  $\beta_0$ , this additional term may either enhance or diminish the magnetization-dependent rotation. However, in either case it is clear from Eq. (5.21) that the rotation induced by the sample is extremely small, so that variations in the rotation as the magnetization changes will be difficult to measure reliably. The sample thickness should be chosen such that  $\beta_0$  is large to avoid this problem.

I illustrate the above analysis using numerical values relevant to the experiment described in Ref. [99]. I define the parameter  $P$  as the ratio between the change of rotation

Table 5.1: Theoretical values of  $\beta_0$  (5.22) and  $P$  (5.25) as a function of the thickness  $d$  of a dysprosium orthoferrite sample, probed with light of wavelength  $\lambda \approx 0.7 \mu\text{m}$ . Also provided are values for  $-\zeta\beta_0/P$ ; if this parameter is close to 1, then Eq. (5.23) is approximately valid. For this material,  $\beta_{max} \approx 1^\circ$ .

$d$ ( $\mu\text{m}$ )	$\beta_0$ (degrees)	$P$ (degrees)	$-\zeta\beta_0/P$ (dimensionless)
55	0.95	-0.22	1.00
58	0.56	-0.13	0.97
59	0.30	-0.07	0.94
60	0.00	-0.004	0.00
61	-0.30	0.06	1.07
62	-0.56	0.13	1.03
65	-0.95	0.22	1.00

and the normalized change of magnetization, i.e.

$$P \equiv \frac{\beta - \beta_0}{\kappa} = -\beta_0\zeta - \zeta \cos 2\delta_0 d \frac{\delta_0 d}{|\Gamma_0|^3}. \quad (5.25)$$

Table 5.1 shows the values of  $\beta_0$ ,  $P$  and  $-\zeta\beta_0/P$  for several values of the sample thickness  $d$ . Note that if  $P \approx -\zeta\beta_0$ , then Eq. (5.23) is approximately valid.

From Table 5.1, we see that small changes of the sample thickness may lead to large changes both in  $\beta_0$  and in  $P$ . However,  $-\zeta\beta_0/P$  remains approximately constant for most values of  $d$ . Finally note that  $\beta_0$  and  $P$  depend strongly on the values of  $\Gamma_0$  and  $\delta_0$ , which are known only approximately. For this reason, our estimates of the critical thicknesses for which  $\beta_0$  vanishes are only approximations; it would be more reasonable to determine these critical thicknesses from experiment. This sensitive dependence of the results on  $\delta_0$  and  $\Gamma_0$  explains why it is preferable to work in a regime where Eq. (5.23) may be used, rather than trying to work directly with Eq. (5.14).

I conclude this section with a discussion of the frequency dependence of the above results. From Ref. [182], we see that the parameter  $\delta_0$  (and, to a lesser extent,  $\Gamma_0$ ) is frequency dependent. Thus if a sample is measured using two different frequencies, one should find two different values of the angle  $\beta$ . If these are correctly normalized, taking into account the frequency dependence of  $\beta_0$ , the measurements will yield consistent results for the magnetization. If, on the other hand, the rotations are compared without normalization (or if the frequency dependence of  $\beta_0$  is neglected), the results will appear inconsistent. In an equilibrium  $\text{DyFeO}_3$  sample that is  $65 \mu\text{m}$  thick, for example, light with a wavelength  $\lambda_0 = 0.7 \mu\text{m}$  will be rotated  $\beta_0 = \beta_{max} \approx 1^\circ$ , while light with  $\lambda_0 = 0.75 \mu\text{m}$  will be rotated only  $\beta_0 \approx 0.5^\circ$ . If the wrong value of  $\beta_0$  is used to interpret experimental rotations, estimates of the amplitude of the magnetic oscillations will be wrong by a factor of 2. If  $\beta_0 \approx 0$  for one of the frequencies used, the error can be arbitrarily large.

This is of particular importance in the case where the magnetic excitation is induced by the same laser source that provides the measurement pulses (as in the inverse Faraday effect experiment described in Ref. [99]). For the same sample and fluence, but with two different laser frequencies, significantly different rotations will be observed. If these are not correctly normalized, it will appear that the induced magnetization is different between

the two experiments, i.e. that magnetic excitation via the inverse Faraday effect is strongly dependent on the excitation frequency. This is inconsistent with what is known about the inverse Faraday effect (which, being a *nonresonant* effect, has a very weak dependence on frequency). This example highlights the importance of correct normalization of the measured rotation.

Another problem that may arise due to frequency dependence is that extremely short laser pulses suffer from frequency broadening, and the above derivation should be repeated using a wave packet rather than a plane wave. However, this only becomes necessary when either the frequency broadening of the pulse is very large, or when the frequency dependence of  $\delta_0 d$  is very strong. Fortunately, laser pulses with a time-width of 200 fs do not suffer significant frequency broadening and are sufficiently short for measuring magnetic oscillations, which have a period of the order of 1 ps. However, if we want to use the Faraday effect to probe faster magnetic dynamics, e.g. dynamics during excitation, shorter pulses and a wave-packet approach must be used to interpret the results.

## 5.5 Numerical exploration of nonperturbative regions

The results of the preceding section may be summarized as follows. For incident light with  $\theta \approx 0$ , the dependence of the rotation angle on magnetization is linear for small  $\kappa$ , with the ratio given approximately by the  $\kappa = 0$  rotation. This “equilibrium rotation” depends both on the material and on the sample thickness and must be measured for each sample used.

If  $\theta$  is not extremely small (i.e. if the condition (5.20) is not satisfied), the relationship between  $\beta$  and  $\kappa$  must be determined from Eq. (5.14), which requires an accurate knowledge of all parameters. Furthermore, the analysis of the previous section does not extend to larger values of  $\kappa$ . (Recall that since both  $m$  and  $m_0$  are small near magnetic equilibrium,  $\kappa$  can be large without violating the assumption of near-equilibrium.)

To investigate the relationship between rotation and magnetization for large  $\theta$  and  $\kappa$ , numerical methods must be used. We calculated  $\beta(\kappa)$  from Eq. (5.14), where  $\kappa$  appears (through  $\gamma$ ) in both  $\delta$  and  $\Gamma$ , using the physical constants appropriate for DyFeO<sub>3</sub>, i.e.

$$\Gamma_0 \approx 30, \quad \delta_0 \approx \frac{\pi}{20} \text{rads } \mu\text{m}^{-1}, \quad \zeta = 0.23.$$

To allow easy comparison with experiment, I plot  $\Delta\beta \equiv \beta - \beta_0$  in degrees. The magnetic deviation  $\kappa$  is dimensionless; to emphasize this, I write it as  $\kappa = \Delta m/m_0$ .

Fig. 5.1 shows the rotation vs. magnetization for  $x$ -polarized light propagating through samples of thickness  $d = 55, 60$  and  $65 \mu\text{m}$ . For  $d$  significantly different to  $60 \mu\text{m}$ , the behaviour is almost linear for all  $\kappa$ , so that determining magnetization from a measured rotation is easy even for large  $\kappa$ . The proportionality “constant” between rotation and magnetization changes with thickness, even changing sign (compare the behaviour for  $d = 55 \mu\text{m}$  with  $d = 65 \mu\text{m}$ ). In fact, the ratio between changes of  $\beta$  and changes of  $\kappa$  is very well approximated by  $\beta_0$  and Eq. (5.23) is found to be valid even for large values of  $\kappa$ . This only holds provided  $d$  is kept away from the critical regions and the incident light is assumed to be  $x$ -polarized. Around  $d = 60 \mu\text{m}$ , the relationship between  $\kappa$  and  $\beta$  is no longer linear when  $\kappa$  not small. In fact the figure is not even symmetric around

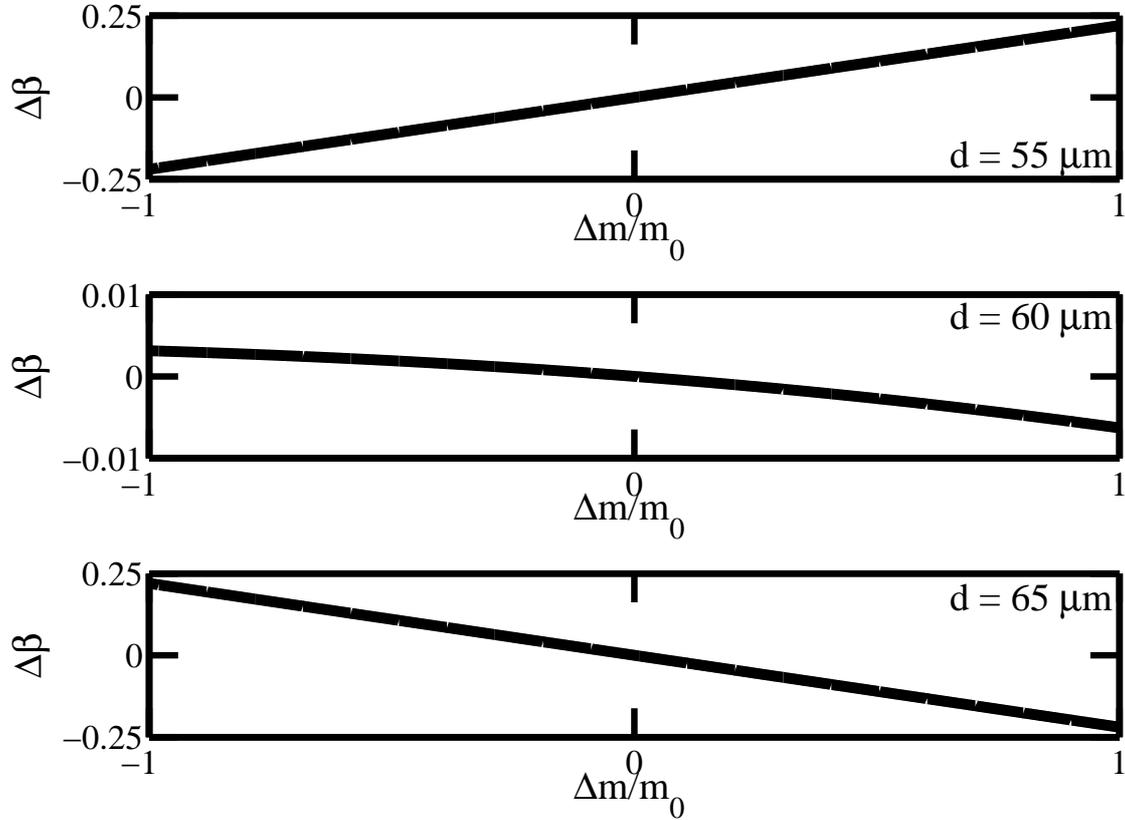


Figure 5.1: Comparison of the angle of rotation  $\Delta\beta$  (in degrees) with the deviation from magnetic equilibrium  $\kappa = \Delta m/m_0$  (dimensionless) as a function of sample thickness. Note the change of scale between different plots. In all of these, the incident light is  $x$ -polarized, i.e.  $\theta = 0^\circ$ .

$\kappa = 0$ . In this case, symmetric magnetic oscillations will give rise to asymmetric rotation measurements.

Next, I fix  $d$  and consider the effect of changing the angle  $\theta$ . I avoid the region where  $\beta_0$  vanishes by setting  $d = 65 \mu\text{m}$ . Fig. 5.2 depicts  $\Delta\beta$  as a function of  $\kappa$  for values of  $\theta$  equal to  $0^\circ$ ,  $44^\circ$  and  $60^\circ$ . Note that despite the range of  $\kappa$  being much smaller than that of Fig. 5.1, Faraday rotation is still found to be as large as  $0.5^\circ$  for  $\theta = 44^\circ$ . Note also that for  $\theta = 60^\circ$ , the Faraday rotation is nonlinear and extremely weak; for most other values of  $\theta$ , the linear approximation is reasonable, although the ratio between  $\Delta\beta$  and  $\kappa$  is dependent on  $\theta$ . In particular, Eq. (5.23) is a very poor approximation when  $\theta \neq 0$ . Measurements with nonzero  $\theta$  are useful for experimentally evaluating the values of  $\delta_0$  and  $\Gamma_0$ .

## 5.6 Summary

In this chapter, I have presented our investigation of Faraday rotation in birefringent, magnetically-ordered media and the relationship between the measurable rotation and the

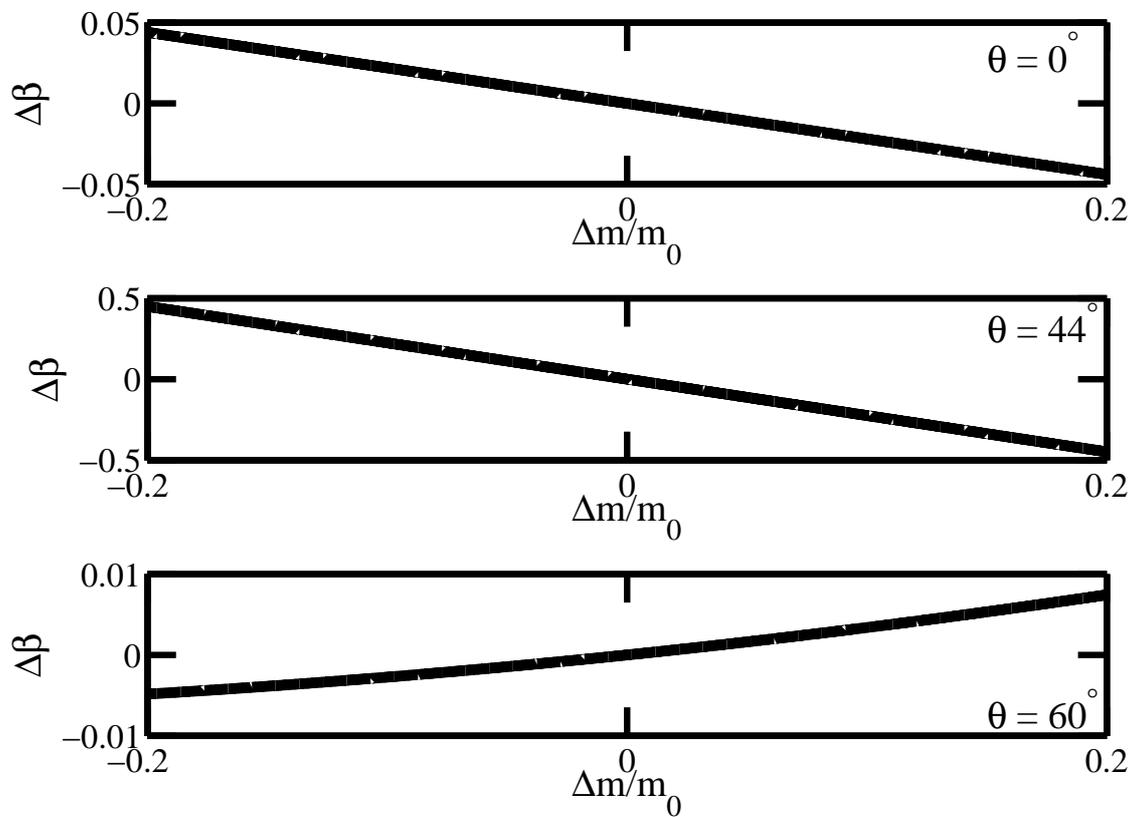


Figure 5.2: Comparison of the angle of rotation  $\Delta\beta$  (in degrees) with the deviation from magnetic equilibrium  $\kappa = \Delta m/m_0$  (dimensionless) as the incident polarization angle is changed. The angle  $\theta$  is measured in degrees counter-clockwise from the  $x$ -axis. The sample thickness is kept at  $d = 65 \mu\text{m}$  throughout. Note that the scale varies between plots.

magnetization of the sample. We found that Faraday rotation is a valid probe for magnetization dynamics, but that care must be taken when interpreting the results. Birefringence breaks the symmetry of the plane orthogonal to the light propagation direction, so that alignment of the polarization axis relative to the crystal orientation becomes an important consideration. Furthermore, results must be normalized against the equilibrium rotation, which depends on both the material used and the sample properties (most notably, sample thickness).

We restricted our investigation by neglecting magnetization dynamics in the plane orthogonal to the propagation of the probe pulse, higher order corrections to the dielectric tensor and diamagnetic terms. The first two of these should have some effect on the rotation, but can only be included at the cost of introducing a number of unknown parameters, which are difficult to estimate. On the other hand, the diamagnetic contribution will have no effect on the rotation unless an external magnetic field is present. If Faraday rotation is used to measure magnetic relaxation after an external field has been switched off, the diamagnetic term may be neglected; however, when trying to determine the magneto-optical parameters of the material using a static magnetic field, the diamagnetic term must be included.

We have focused primarily on the rotation and neglected the ellipticity of the outgoing light. Although these contain equivalent information, it is advisable to measure both and to check that the results are consistent. Inconsistency between these could indicate that the material parameters are not known accurately enough. On the other hand, inconsistency could reveal the presence of an additional effect that has been neglected in the above analysis. It was exactly this type of inconsistency that led to the discovery of optical bleaching [103]!

## Chapter 6

# Interaction of a Quantum System with a Heat Bath

### 6.1 Dephasing and dissipation

In the preceding chapters, I have considered the interactions that either excite or measure ultrafast magnetization dynamics. These are the most important interactions when considering an ultrafast magnetization reversal experiment. However, over longer timescales, other interactions will influence the dynamics. Of particular importance are interactions of the system with an external environment. The main effects of such interactions are *dephasing* and *dissipation*. This chapter and the next are devoted to the study of these effects.

Dephasing and dissipation are the quantum mechanical analogues of friction in Newtonian mechanics. On short timescales, Newton's laws show that the acceleration of particles is determined by the applied forces only. However, over longer timescales, non-applied forces due to the inevitable presence of an environment influence the dynamics and must be taken into account. These non-applied forces have two characteristics that make their study possible — the interactions are *weak*, allowing perturbation theory, and they involve a *large number of particles*, so that statistical mechanics can be used. These features also permit their study in quantum mechanical problems.

As mentioned in chapter 1, dissipation is the loss of energy of a quantum mechanical system. This manifests itself by the decay of the diagonal elements of the density matrix towards thermal occupation values. In the quantum dot experiments described in chapter 1, dissipation causes electrons in the high-energy states to scatter to lower energy states. If the states used for quantum computing or spin-based memory have different energies, dissipation leads to a loss of information, limiting the timescale on which the device is useful. However, dissipation does not only have negative consequences. One example of its positive consequences comes from magnetic switching, where the decay of small deviations ensures that each ground state has a finite *basin of attraction*. Provided the switching process places the magnetization in the correct basin of attraction, switching will be successful. In this case, dissipation makes the system more tolerant of stochastic fluctuations. This was described in chapter 1.

*Dephasing* is more difficult to define, but broadly speaking it gives an indication of the

loss of quantum-mechanical information in the system. Unlike dissipation, it cannot be defined in terms of wavefunctions, since a wavefunction is always fully quantum-mechanical. Dephasing measures the rate at which the quantum system decays away from a “pure” state (i.e. a wavefunction) into a mixed state, which can only be described in terms of a density matrix (see, e.g., Refs. [117, 194] for a description of pure and mixed states). An appropriate definition, therefore, could be that dephasing is the decay of the off-diagonal terms of the density matrix.

These definitions of dephasing and dissipation refer to decay rates of components of the density matrix. This renders these concepts somewhat ill-defined, since density matrix components are dependent on the basis in which the density matrix is expressed. A change of basis mixes diagonal and off-diagonal elements. Which basis should be chosen?

A natural choice, which links energy loss to the decay of the diagonal density matrix components as required by the above definition, is to use the basis of energy eigenfunctions, i.e. eigenfunctions of the full Hamiltonian (including the environment). However, such eigenfunctions include many variables pertaining to the environment, and cannot be calculated; indeed, calculating these eigenfunctions is equivalent to solving the dynamical problem, in which case the density matrix formalism is not needed. On the other hand, the use of the “bare” Hamiltonian, i.e. decoupled from the environment, would not be satisfactory, since the environment can lead to a *renormalization* of the energy, unrelated to the decay. Energy flows out of the system, but is stored in the system-environment interaction and does not flow into the environment.

Basis-independent definitions of dephasing and dissipation are also unsatisfactory: for dissipation, the renormalization of the energy still provides problems, while for dephasing, there is a conceptual problem that the thermodynamic state (which should be interpreted as completely dephased) cannot be distinguished from a pure wavefunction when the temperature is low and the ground state is nondegenerate. As such, basis-independent measures of dephasing (e.g. von Neumann entropy or deviation from idempotency of the density matrix [142, 192, 193]) will show that an initial wavefunction dephases initially, then “rephases” towards the thermodynamic state. This can be seen in section 6.3.1, Figs. 6.1 (c) and (d).

Thus the concepts of dephasing and dissipation are not perfectly well-defined. However, provided the interactions are weak enough, the above-mentioned problems are negligible. Just as energy can be given a reasonable definition even in the presence of decay [116], dephasing and dissipation can be approximately defined. Throughout this chapter and the next, I shall use the definition that dissipation and dephasing refer to the decay rate of the diagonal and off-diagonal components of the density matrix, respectively, where the density matrix is written relative to the basis of the eigenfunctions of the bare Hamiltonian.

## 6.2 The generalized master equation

The dephasing and dissipation rates of a quantum system in contact with a large reservoir (heat bath) are determined from the dynamics of the density matrix. The evolution of the density matrix involves not only the dynamics of the quantum system under the influence of the reservoir, but also the motion of the reservoir due to the interaction with the quantum system. This is generally an intractable problem. However, under certain

conditions, it is possible to determine the approximate evolution of the quantum system, with the dynamics of the reservoir treated perturbatively. In this section, I present the *generalized master equation method*. This is well-suited to the problem studied in the next chapter of electrons in quantum dots interacting with heat baths.

The use of master equations to describe relaxation processes was pioneered by Pauli [147] and placed on a more rigorous physical basis by van Hove [188], Nakajima [139], Zwanzig [211], Montroll [134] and Prigogine and Resibois [161]. Here I present the Nakajima-Zwanzig method; in the next chapter I show the application of this method to the study of quantum dot dephasing. A general review of this method has been given in Ref. [64], which I have followed closely in producing this summary.

The Nakajima-Zwanzig method can be applied to any system composed of two interacting parts, a “system”  $\mathcal{S}$  and “bath”  $\mathcal{B}$ , provided both  $\mathcal{S}$  and  $\mathcal{B}$  become separate physical systems when the interaction is set to zero. The method is formally exact, but relies on approximations for useful results to be obtained; therefore it is only useful when the interaction is weak enough, in a sense that is made more precise below.

I consider only the case where  $\mathcal{S} \oplus \mathcal{B}$  forms a closed quantum system. The method can be extended to describe open systems (see, e.g., Ref. [64] for details), but this was not needed for our work. The complete system  $\mathcal{S} \oplus \mathcal{B}$  is described by a density operator  $W(t)$  that obeys the Liouville-von Neumann equation

$$\dot{W}(t) = -\frac{i}{\hbar}[H, W(t)] \equiv -iLW(t), \quad (6.1)$$

where the Hamiltonian  $H$  and Liouvillian  $L$  are time-independent. These operators may be decomposed into three parts, describing the system, bath and interaction respectively:

$$\begin{aligned} H &= H_{\mathcal{S}} + H_{\mathcal{B}} + H_{\mathcal{S}\mathcal{B}} \\ L &= L_{\mathcal{S}} + L_{\mathcal{B}} + L_{\mathcal{S}\mathcal{B}}. \end{aligned} \quad (6.2)$$

The density operator  $W(t)$  includes variables from both  $\mathcal{S}$  and  $\mathcal{B}$ . A description of the subsystem  $\mathcal{S}$  alone is obtained by performing a trace operation over the  $\mathcal{B}$ -variables, yielding the “reduced density operator”

$$\rho(t) = \text{Tr}_{\mathcal{B}}W(t). \quad (6.3)$$

This operator acts only in the Hilbert space describing  $\mathcal{S}$ ; the Hilbert space of the bath has been “integrated out”.

The equation governing the evolution of  $\rho(t)$  is derived from Eq. (6.1) by introducing the projection operator

$$\mathcal{P} = B_{\text{ref}}\text{Tr}_{\mathcal{B}}. \quad (6.4)$$

Here,  $B_{\text{ref}}$  is a normalized ( $\text{Tr}_{\mathcal{B}}B_{\text{ref}} = 1$ ) state vector in the Hilbert space describing the bath  $\mathcal{B}$ . This vector may be freely chosen, making the method very flexible; judicious choice of  $B_{\text{ref}}$  can allow even strongly interacting systems to be studied. As shown below,  $B_{\text{ref}}$  is the reference state relative to which the correlation functions of bath variables are calculated. Perturbation expansions in bath variables will converge more rapidly if  $B_{\text{ref}}$  is as close as possible to the true physical state of the bath. If the system-bath interaction is

strong, this choice can be determined self-consistently<sup>1</sup>. In the work of the next chapter, the bath is assumed to be thermodynamically large and the thermodynamic equilibrium state was the most appropriate choice.

By acting with  $\mathcal{P}$  (6.4) and the associated projector  $\mathcal{P}' = 1 - \mathcal{P}$  on Eq. (6.1), we obtain

$$\begin{aligned}\mathcal{P}\dot{W}(t) &= -i\mathcal{P}L\mathcal{P}W(t) - i\mathcal{P}L(1 - \mathcal{P})W(t) \\ (1 - \mathcal{P})\dot{W}(t) &= -i(1 - \mathcal{P})L\mathcal{P}W(t) - i(1 - \mathcal{P})L(1 - \mathcal{P})W(t).\end{aligned}\quad (6.5)$$

The second equation may be formally integrated to

$$\begin{aligned}(1 - \mathcal{P})W(t) &= \exp[-i(1 - \mathcal{P})Lt](1 - \mathcal{P})W(0) \\ -i \int_0^t dt' \exp[-i(1 - \mathcal{P})Lt'] &(1 - \mathcal{P})L\mathcal{P}W(t - t').\end{aligned}\quad (6.6)$$

Substituting Eq. (6.6) into the first of Eqs. (6.5) gives an integrodifferential equation for  $\mathcal{P}W(t) = B_{\text{ref}}\rho(t)$ , with  $\rho(t)$  defined in Eq. (6.3). By taking the trace over the bath variables, we obtain an integrodifferential equation for the reduced density operator:

$$\dot{\rho}(t) = -iL_{\text{eff}}\rho(t) + \int_0^t dt' K(t')\rho(t - t') + I(t).\quad (6.7)$$

Here, I have followed Haake's notation [64] for the effective Liouvillian, the integral kernel and the inhomogeneity, which are defined as

$$\begin{aligned}L_{\text{eff}} &= \text{Tr}_{\mathcal{B}}LB_{\text{ref}} = L_{\mathcal{S}} + \text{Tr}_{\mathcal{B}}L_{\mathcal{S}\mathcal{B}}B_{\text{ref}} \\ K(t) &= -\text{Tr}_{\mathcal{B}}L_{\mathcal{S}\mathcal{B}} \exp[-i(1 - \mathcal{P})Lt](1 - \mathcal{P})(L_{\mathcal{B}} + L_{\mathcal{S}\mathcal{B}})B_{\text{ref}} \\ I(t) &= -i\text{Tr}_{\mathcal{B}}L_{\mathcal{S}\mathcal{B}} \exp[-i(1 - \mathcal{P})Lt](1 - \mathcal{P})W(0)\end{aligned}\quad (6.8)$$

The effective Liouvillian  $L_{\text{eff}}$  describes how the motion of the system in the presence of the environment would be, if the environment were in the reference state  $B_{\text{ref}}$ . The “memory kernel”,  $K(t)$ , describes how this reference state evolves in time under the influence of the system (and due to its own internal dynamics). Thus the motion of the system is “remembered” by the environment, which influences the environment's interaction with the system. Finally,  $I(t)$  takes into account the actual initial state of the environment (relative to  $B_{\text{ref}}$ ), including correlations with the system. This term will dominate the dynamics if  $B_{\text{ref}}$  is a poor approximation to the state of the environment. On the other hand, if the environment is initially close to  $B_{\text{ref}}$ ,  $I(t)$  can be neglected. Note that any deviation from  $B_{\text{ref}}$  *after* the system is initialized is accounted for in  $K(t)$ , not  $I(t)$ .

It is worth noting that the decay of  $K(t)$  gives an indication of the memory time and stochasticity of the environment. If  $K(t)$  decays slowly, the environment is always close to the state that the system induces in it. The environment remembers all dynamics of the system, and behaves deterministically. On the other hand, if  $K(t)$  decays rapidly,

<sup>1</sup>An example of this is found in laser physics, where the resonant interaction of light with atoms is strong and keeps both systems far from thermal equilibrium. In this case, an initial guess for the state of the light can be used as  $B_{\text{ref}}$ . The steady-state of the atomic subsystem is determined from this, which gives rise to a new reference state for the light. This process can be repeated until self-consistency is reached [64].

information about the system is rapidly lost, and the impact of the environment is largely stochastic.

Eq. (6.7) is the generalized master equation, also referred to in the literature as the kinetic equation [128]. It is formally exact; the bath dynamics have not been eliminated, but simply hidden in the exponential factors. Approximations of these exponentials must be made before a solution can be found. Furthermore, Eq. (6.7) is an operator equation, so it must be mapped to a matrix equation by projecting onto the basis of  $\mathcal{S}$  before it can be solved.

Before describing the methods of solution, it is appropriate to discuss several useful simplifications that we employed in our study of quantum dots interacting with large heat baths. First, we used the freedom in the state vector  $B_{\text{ref}}$  to choose

$$B_{\text{ref}} = e^{-\beta H_{\mathcal{B}}} / \text{Tr}_{\mathcal{B}} [e^{-\beta H_{\mathcal{B}}}] . \quad (6.9)$$

Here, the standard notation  $\beta = 1/k_B T$  is used, with  $k_B$  the Boltzmann constant and  $T$  the temperature [158]. The choice of vector (6.9) implies that  $L_{\mathcal{B}} B_{\text{ref}} = 0$  and the integral kernel  $K(t)$  is at least quadratic in  $L_{\mathcal{S}\mathcal{B}}$ . Note that even with this choice of  $B_{\text{ref}}$ , Eq. (6.7) is still exact. It is only when approximations are made that the choice of  $B_{\text{ref}}$  affects the results.

Next, we assumed that in the initial state, there is no correlation between  $\mathcal{S}$  and  $\mathcal{B}$  and that the bath starts in the thermodynamic state. Then

$$W(0) = e^{-\beta H_{\mathcal{B}}} \otimes \rho(0) \quad (6.10)$$

and

$$I(t) = 0 = (1 - \mathcal{P})W(0). \quad (6.11)$$

Finally, we truncated the kernel  $K(t)$  at second order in the interaction by setting

$$\exp[-i(1 - \mathcal{P})Lt] \approx \exp[-i(1 - \mathcal{P})(L_{\mathcal{S}} + L_{\mathcal{B}})t]. \quad (6.12)$$

A natural question that arises after the truncation (6.12) is whether *all* second-order terms have been included. To show that all such terms are included, we directly expanded the definition  $\rho(t) = \text{Tr}_{\mathcal{B}} \exp[-iLt]W(0)$  to second order in  $L_{\mathcal{S}\mathcal{B}}$ . This yielded the same equation as (6.7) with the truncation (6.12). The truncated equation is the time representation of the Dyson equation for  $\rho(t)$  in the second-order Born approximation [129]. Although the same equation arises, it should be noted that deriving the equation from the generalized master equation method is both faster and conceptually clearer.

Equation (6.7) preserves both the normalization and the hermiticity of  $\rho$ . It is easily shown that these important physical properties are also preserved in the truncated equation, i.e. that Eq. (6.12) is a ‘‘conservative approximation’’ [86].

### 6.3 Solving the master equation

Eq. (6.7) is an integrodifferential equation for the reduced density operator  $\rho$  of the system  $\mathcal{S}$ . Because of the trace operation  $\text{Tr}_{\mathcal{B}}$ , operators pertaining to  $\mathcal{B}$  are contained within correlation functions that are independent of the dynamics of  $\mathcal{S}$ . These correlation functions

are weighted by the state vector  $B_{\text{ref}}$ , which justifies naming  $B_{\text{ref}}$  the “reference state”. (This is quite a general statement and will be explained in more detail in the next chapter when a specific physical model is chosen.)

By expressing  $\rho$  relative to the basis of energy eigenfunctions, we obtain equations for all the components of the density matrix  $\rho_{ij}$ . These equations can be rearranged into a vector integrodifferential equation with the general form

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + \int_0^t B(t-t')\mathbf{x}(t')dt', \quad (6.13)$$

where  $A$  and  $B$  are matrices. The matrix  $B$  contains time-dependent correlation functions of the bath  $\mathcal{B}$ . Note that the time-arguments in the integrand of Eq. (6.13) are different to those of Eq. (6.7), but are related by substitution. Eq. (6.13) can be integrated once, yielding the following Volterra equation [11]:

$$\mathbf{x}(t) = \mathbf{x}(0) + \int_0^t C(t-t')\mathbf{x}(t')dt' \quad (6.14)$$

with

$$C(t-t') = A + \int_0^{t-t'} B(s)ds. \quad (6.15)$$

In this section, I present two methods of solving Eq. 6.14, with a brief explanation of their positive and negative attributes. Before introducing these methods, some comments should be made about the vector  $\mathbf{x}$ .

First, the generalized master equation (6.7) preserves the hermiticity of the density matrix. This means that if the density matrix components  $\rho_{ij}$  are directly mapped to components of  $\mathbf{x}$ , these components will not be independent (e.g. if  $x_2 = \rho_{12}$  and  $x_3 = \rho_{21}$ , then  $x_2 = x_3^*$ ). It is therefore more appropriate to form  $\mathbf{x}$  from the real and imaginary parts of the  $\rho_{ij}$ , as it renders  $\mathbf{x}$  a real-valued vector, with (nearly-)independent components. The length of  $\mathbf{x}$  is  $N^2$ , where  $N$  is the number of eigenfunctions of the system  $\mathcal{S}$ .

Second, Eq. (6.7) preserves the trace of  $\rho$ , so the elements of  $\mathbf{x}$  corresponding to the diagonal components of  $\rho_{ij}$  are not all independent. (This explains the use of the term “nearly-independent” in the preceding paragraph.) One of these components could be eliminated. However, this leads to the introduction of a time-dependent inhomogeneous term in Eq. (6.14), which makes the numerical solution more complicated<sup>2</sup>. Therefore we preferred to keep all  $N^2$  components of  $\mathbf{x}$  and solve an equation of the form (6.14).

### 6.3.1 Numerical integration

The most obvious method of solving Eq. (6.14) is to approximate the integral by a finite sum over a uniformly discretized time mesh [11]. We introduce the time-step  $\Delta t$  and assume that the time  $t$  at which Eq. (6.14) must be solved satisfies  $t = n\Delta t$  for some integer  $n$ . (The solution at intermediate times can be determined by interpolation.) Then with the notation  $\mathbf{x}_n \equiv \mathbf{x}(n\Delta t)$ ,  $C_n \equiv C(n\Delta t)$ , Eq. (6.14) becomes

$$\mathbf{x}_n = \mathbf{x}_0 + \Delta t \sum_{k=0}^n w_{n,k} C_{n-k} \mathbf{x}_k. \quad (6.16)$$

<sup>2</sup>An inhomogeneous term is not *in se* problematic, but in this case the inhomogeneity increases linearly with time and must be canceled by a corresponding term in the integral kernel  $B$ . This causes numerical inaccuracy and significant difficulty, without yielding any reasonable improvement in speed.

The integration weights  $w_{n,k}$  determine the stability and accuracy of this numerical scheme. Eq. (6.16) is an implicit equation for the unknown  $\mathbf{x}_n$ ; it can be made explicit by reorganizing it as

$$\mathbf{x}_n = (I - \Delta t w_{n,n} C_0)^{-1} \left[ \mathbf{x}_0 + \Delta t \sum_{k=0}^{n-1} w_{n,k} C_{n-k} \mathbf{x}_k \right], \quad (6.17)$$

with  $I$  the  $N^2 \times N^2$  identity matrix. Assuming that  $C_0$  is not singular, the matrix  $(I - \Delta t w_{n,n} C_0)$  is invertible for small  $\Delta t$ , so that this equation is well-defined. Given  $\mathbf{x}_k$  for all  $k < n$ , Eq. (6.17) determines  $\mathbf{x}_n$ .

A review of some of the more common choices of weights  $w_{n,k}$  is given in Ref. [11]. The  $n$ -th order Newton-Cotes scheme is unstable due to ‘‘Runge’s phenomenon’’ [168], so composite weighting schemes are preferred. Repeated trapezoidal rule is the simplest, but we used a modified form of the repeated Simpson’s rule, which is more accurate for a comparable amount of effort.

The modified scheme is applied as follows. Simpson’s rule is used when  $n$  is even:

$$w_{n,0} = w_{n,n} = \frac{1}{3}, \quad w_{n,2k+1} = \frac{4}{3}, \quad w_{n,2k} = \frac{2}{3}. \quad (6.18)$$

If  $n$  is odd and larger than  $n = 1$ , the weights (6.18) are used for  $w_{n,k}$  with  $0 \leq k \leq n - 4$ ; the remaining weights are

$$w_{n,n-3} = \frac{17}{24}, \quad w_{n,n-2} = w_{n,n-1} = \frac{9}{8}, \quad w_{n,n} = \frac{3}{8}. \quad (6.19)$$

When  $n = 1$ , we use the trapezoidal rule since this is the most accurate 2-point scheme. The weights (6.19) are derived from the ‘‘3/8-rule’’ (third order Newton-Cotes formula):  $w_{n,n-2}$ ,  $w_{n,n-1}$  and  $w_{n,n}$  are given directly by the 3/8-rule, while  $w_{n,n-3} = \frac{1}{3} + \frac{3}{8}$ , which is the sum of the endpoints of Simpson’s rule and the 3/8-rule.

The choice of weights (6.18)-(6.19) is called the ‘‘Simpson’s plus upper 3/8-rule’’ [11], since the 3/8-rule is applied to the weights  $w_{n,k}$  with  $k$  close to  $n$  (the ‘‘upper’’ limit of  $k$ ). This scheme is stable and accurate [11]. By contrast, the Simpson’s plus lower 3/8-rule, where the 3/8-rule determines  $w_{n,k}$  for  $k \leq 3$ , is unstable.

This numerical scheme works extremely well for small  $n$ . However, the method scales as  $n^2$ , since each new step requires a sum over all previous time points. It is therefore not suitable for very long simulations. If the integral kernel  $C(t - t')$  decays rapidly to zero, a cutoff ‘‘memory-time’’  $\tau$  can be introduced. By neglecting the contribution from all  $\mathbf{x}(t')$  with  $t' < (t - \tau)$ , the method becomes linear in  $n$ . A similar truncation can be performed if  $C(t - t')$  decays to a nonzero constant, although this requires more book-keeping. More precisely, assuming  $C(t - t') \rightarrow C_\infty$  as  $t \rightarrow \infty$ , Eq. (6.14) can be rewritten as

$$\mathbf{x}(t) \approx \mathbf{x}(0) + \int_{t-\tau}^t C(t - t') \mathbf{x}(t') dt' + C_\infty \int_0^{t-\tau} \mathbf{x}(t') dt'. \quad (6.20)$$

Clearly, we need to keep track of  $\int_0^{t-\tau} \mathbf{x}(t') dt'$  as the simulation progresses. Although this makes programming more complicated, it renders the method linear in  $n$ , allowing much larger timescales to be explored.

Despite this increase of speed, the method is still slow for certain applications. In particular, if the kernel  $C(t - t')$  oscillates rapidly, then  $\Delta t$  must be small enough to

resolve these oscillations. If we also find that  $\mathbf{x}(t)$  evolves slowly, extremely large values of  $n$  will be needed to show these dynamics.

We investigated the methods described in this chapter for the purpose of studying the dephasing of a quantum dot due to the interaction with a metal electrode. The characteristic oscillation time of  $C(t - t')$  is the plasma frequency of the metal,  $\hbar\omega_p \approx 10$  eV, while dissipation in quantum dots occurs on a microsecond (or longer) timescale. For such a system,  $n = 1 \times 10^{11}$  is not unreasonable. However, a simulation of  $n = 1 \times 10^8$  requires about a day of computing<sup>3</sup>, meaning that this method is not applicable to these systems.

We attempted to accelerate the dynamics by increasing the strength of the interaction. This appeared to work well even if the interaction was increased by a factor as large as  $\lambda = 1 \times 10^8$ . However, the only way to confirm the accuracy of such a method is to run simulations of the unenhanced dynamics and compare the outcome. This can only be done for short timescales, but good agreement on short timescales does not necessarily imply agreement over longer times. As such, this acceleration technique may be useful for producing estimates of the resulting decay rates and for checking the results of other methods, but is not adequate on its own.

Another drawback of the numerical integration method (6.17) is that simulations must be performed for a variety of initial conditions. Dynamical features such as the rates of dissipation and dephasing could depend on the initial conditions. A general dephasing rate can only be inferred after adequate sampling of initial conditions and only if similar decay rates are found in all simulations.

Despite these drawbacks, the time-marching algorithm is useful in providing a visualization of the dynamics. As an example of this, I include here some figures showing the evolution of a charge qubit interacting with a metal electrode. The charge qubit was modeled as a single, 1-dimensional quantum dot with infinitely high boundaries. This is a very simple model, far simpler than the physically reasonable model described in the next chapter. We studied this model only as part of a preliminary investigation into methods of solving the generalized master equation. Since its physical relevance is limited, I shall not provide a detailed analysis of the problem. Only three points need to be made about the model. First, only the lowest two eigenstates —  $\varphi_1 = \cos(\pi x/2R)$  and  $\varphi_2 = \sin(\pi x/R)$ , with  $R = 10$  nm — were used. Second, the interaction with a metal gate 100 nm away from the qubit was treated using the screened Coulomb interaction, as detailed in the next chapter. Finally, the interaction was enhanced by a factor of  $\lambda = 1 \times 10^6$  to obtain results within reasonable time. These results are plotted in Fig. 6.1.

Since Fig. 6.1 depicts results for a highly-idealized system, the numbers themselves have little physical meaning. However, Fig. 6.1 provides a good overview of how the dynamics proceed. The absolute value of the off-diagonal element decays toward 0, as seen in Fig. 6.1(b). Fig. 6.1(a) shows that the diagonal elements decay about twice as fast as the off-diagonal parts and their asymptotic value is determined by thermal equilibrium. (In this case,  $T = 100$  mK, so the system approaches the ground state.)

It is generally accepted that the diagonal density matrix components *must* decay slower than twice the decay rate of the off-diagonal components, i.e. if  $T_1$  is the characteristic

---

<sup>3</sup>Although the truncated method scales linearly with  $n$ , the coefficient of  $n$  is related to the decay rate of the kernel and can be extremely large. The numbers given are based on a kernel decaying within 600 fs, the timescale on which thermal fluctuations are damped.

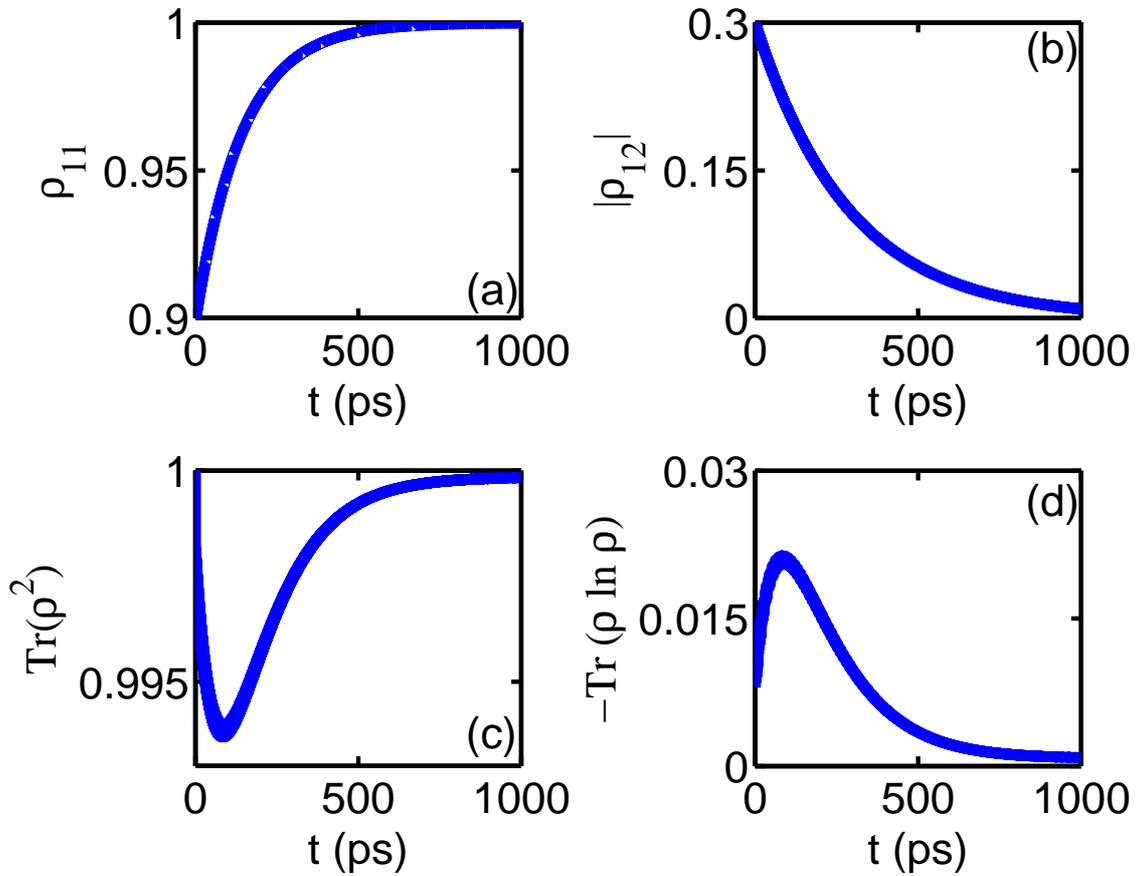


Figure 6.1: Evolution of the density matrix of a simplified charge qubit due to the interaction with a gate electrode. The interaction strength has been increased by a factor of  $\lambda = 1 \times 10^6$ . (a) and (b) show the evolution of the diagonal and off-diagonal parts, respectively, revealing a dissipation rate of  $6.88 \text{ ns}^{-1}$  and a dephasing rate of  $3.44 \text{ ns}^{-1}$ . (c) and (d) show that traditional measures of dephasing,  $\text{Tr}(\rho^2)$  and  $S(\rho) = -\text{Tr}(\rho \ln \rho)$ , are adequate over short timescales, but imply an unphysical “rephasing” of the solution over longer times.

decay time of the diagonal elements and  $T_2$  that for the off-diagonals, then

$$2T_1 \geq T_2. \quad (6.21)$$

This relationship can be easily demonstrated for a 2-state system, assuming  $T = 0$  and that the decay is exponential, i.e.

$$\rho = \begin{pmatrix} 1 - |c_2|^2 e^{-t/T_1} & c_1 c_2^* e^{iEt - t/T_2} \\ c_1^* c_2 e^{-iEt - t/T_2} & |c_2|^2 e^{-t/T_1} \end{pmatrix} \quad (6.22)$$

where  $c_1$  and  $c_2$  are the coefficients of the eigenstates in the initial wavefunction. With the evolution given in (6.22), Eq. (6.21) follows from  $\text{Tr}(\rho^2) \leq 1$ . The physical basis for Eq. (6.21) is the existence of two types of scattering: elastic and inelastic. *Inelastic* interactions lead to energy transfer, which yields dissipation. Such inelastic interactions also cause dephasing, because information flows between the interacting particles, but this occurs at half the rate of dissipation. *Elastic* collisions conserve the energy of the interacting particles and cannot yield dissipation, but information can be exchanged and dephasing can occur. This leads to Eq. (6.21). Usually, Eq. (6.21) is rewritten as an equality by introducing the “pure dephasing” time:

$$\frac{1}{T_\Phi} = \frac{1}{T_2} - \frac{1}{2T_1}. \quad (6.23)$$

Thus  $T_\Phi$  governs the dephasing rate due to elastic collisions,  $T_1$  yields the dissipation rate (due to inelastic collisions) and  $T_2$  measures the total dephasing time due to a combination of scattering processes. The results in Fig. 6.1 (a) and (b) show that the dynamics of the system under consideration are dominated by inelastic processes, as the pure dephasing rate is negligible.

Figs. 6.1(c) and (d) show the evolution of two traditional measures of dephasing,  $\text{Tr}(\rho^2)$  and the von Neumann entropy  $S(\rho) = -\text{Tr}(\rho \ln \rho)$  [142, 192, 193]. As mentioned in section 6.1, these adequately demonstrate dephasing during the initial evolution, but as the system evolves towards the thermodynamic ground state, the system appears to “rephase”. This is because the thermodynamic ground state is very similar to the zero-temperature ground state, which is a pure state. The decay of  $|\rho_{12}|$  is a more consistent definition of dephasing.

In summary, numerical integration methods are very useful for visualizing the evolution, but are not fast enough to allow an adequate study of dephasing. A faster and more general method must be used. One example of such a method is the Laplace transform, described in the next section.

### 6.3.2 Laplace transformation

Eq. (6.14) is a special type of Volterra equation — it has a convolutive kernel [57]. This means that the integrand has the form  $C(t - t')\mathbf{x}(t')$ , rather than the more general form  $C(t, t')\mathbf{x}(t')$ . Under the Fourier or Laplace transforms, convolution becomes multiplication and Eq. (6.14) becomes an algebraic equation. The Laplace transform is more appropriate, because the time interval starts at  $t = 0$ , not  $t = -\infty$  [57].

We multiply Eq. (6.14) by  $\exp(-zt)$  and integrate over  $t$ . Here,  $z$  is a complex variable with real part chosen large enough that all relevant integrals converge. Then defining

$$\tilde{\mathbf{x}}(z) = \int_0^\infty e^{-zt} \mathbf{x}(t) dt \quad (6.24)$$

and a similar definition of  $\tilde{C}(z)$ , Eq. (6.14) becomes

$$\tilde{\mathbf{x}}(z) = \frac{1}{z}\mathbf{x}(0) + \tilde{C}(z)\tilde{\mathbf{x}}(z). \quad (6.25)$$

For any  $z$ , the solution  $\tilde{\mathbf{x}}(z)$  can be found by matrix inversion. Thus it is possible to reconstruct  $\mathbf{x}(t)$  using the inverse Laplace transform [57]

$$\mathbf{x}(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{zt}\tilde{\mathbf{x}}(z)dz, \quad (6.26)$$

where  $\gamma$  is some real number chosen to ensure convergence. The integration path in (6.26) is parallel to the imaginary axis.

The reconstruction of  $\mathbf{x}(t)$  using (6.26) requires the evaluation of an integral for each value of  $t$ . As such, it requires computation time similar to the direct time-stepping algorithm presented in the previous section. A minor advantage offered by this method is that the reconstruction parallelizes well; values of  $\mathbf{x}(t)$  at different values of  $t$  can be calculated independently of each other. Another advantage is that the inverse transform (6.26) is essentially a Fourier transform, which introduces the possibility of using the Fast Fourier Transform [160] to speed up the calculation. A possible disadvantage of the Laplace transform compared to the numerical integration method defined above is that  $\tilde{C}(z)$  may be impossible to calculate analytically. In the examples that we studied (presented in the next chapter), this was not the case.

However, if certain properties of the evolution can be assumed, the Laplace transform can greatly speed up the solution of the problem. In particular, if we assume that the evolution is asymptotically dominated by exponential decay, the solution of the problem becomes a search for the poles of the Laplace transform. The assumption of exponential decay is usually valid when the dissipation is caused by interaction with a thermodynamically large bath, which is the case for most quantum computing systems [59, 132, 142]. Tests for the validity of this assumption and the consequences of its failure are described at the end of this section, but first I present how  $\mathbf{x}(t)$  is found if this assumption is made.

It is well-known that the positions of the simple poles of the Laplace transform yield the exponential decay rates. (Singularities at the origin correspond to algebraic decay.) For asymptotic analysis, only the poles with the largest (and sometimes second largest) real parts are significant; the smaller the real part, the faster the decay of that component. Therefore, the strategy for analyzing Eq. (6.25) is to find the poles of  $\tilde{\mathbf{x}}(z)$  with the largest real parts and to investigate the nature (multiplicity, residues, etc.) of these singularities.

Eq. (6.25) can be rewritten as

$$\left[ I - \tilde{C}(z) \right] \tilde{\mathbf{x}}(z) = \frac{1}{z}\mathbf{x}_0, \quad (6.27)$$

with the right-hand side finite away from  $z = 0$ . This equation must have a unique solution for all  $z$ , since  $\mathbf{x}(t)$  is uniquely determined by Eq. (6.14) and  $\tilde{\mathbf{x}}(z)$  is uniquely defined by  $\mathbf{x}(t)$ . It follows that the singularities of  $\tilde{\mathbf{x}}(z)$  occur where the determinant of  $\left[ I - \tilde{C}(z) \right]$  vanishes and at  $z = 0$ . Thus we are led to the problem of finding the roots of a scalar function of a complex variable.

This problem is most conveniently solved using the Newton-Raphson method [160]. The numerical integration method of the previous section is useful for providing estimates

of dephasing rates, which are good starting points for the root-finding algorithm. Note that this reformulation of the problem shows that the singularities of  $\tilde{\mathbf{x}}(z)$ , and therefore the decay rates, depend only on the determinant of  $[1 - \tilde{C}(z)]$  and not on the initial conditions  $\mathbf{x}(0)$ . Decay rates are a general feature of the dynamics and do not depend on the initial configuration.

Finally, once the poles have been found, the residues at these poles are calculated using finite differences:

$$\text{Res}(a) = \lim_{z \rightarrow a} (z - a) \tilde{\mathbf{x}}(z) \approx \delta \tilde{\mathbf{x}}(a + \delta), \quad (6.28)$$

where  $\delta$  is a complex number with small magnitude,  $|\delta| \ll 1$ . A residue can be either zero, finite or infinite<sup>4</sup>. If it is zero, the singularity is removable and does not contribute to the exponential decay. If a residue is infinite, the singularity is more complicated than a simple pole. This signals asymptotic behaviour different to exponential decay. If the residue is finite and nonzero, the singularity is a simple pole, with corresponding exponential decay. From the values of the residues and the location of the singularities, the approximate solution  $\mathbf{x}(t)$  can be constructed.

It is worth noting that when  $\mathbf{x}(t)$  describes the evolution of a density matrix, we expect  $\tilde{\mathbf{x}}(z)$  to have a simple pole at the origin. Such a pole would describe a constant term in the density matrix and yield the asymptotic value of  $\mathbf{x}(t)$ . Because  $\text{Tr}(\rho) = 1$ , we expect some components of  $\mathbf{x}(\infty)$  to be nonzero.

The method outlined here is used in the following chapter to determine the dephasing of qubits due to interaction with metallic electrodes. However, before presenting these results, some attention should be given to the possibility of nonexponential behaviour of  $\mathbf{x}(t)$ .

It was already mentioned above that if the residue calculated at an isolated singularity is found to be “infinite” (i.e. the singularity is a higher order pole), then the behaviour of  $\mathbf{x}(t)$  is nonexponential. More precisely, if  $\tilde{\mathbf{x}}(z) \sim 1/(z - a)^n$  near  $z = a$ , with  $n > 1$ , the corresponding evolution is  $\mathbf{x}(t) \sim t^{n-1} e^{at}$ .

For finite  $\delta$ , the finite difference evaluation of the residue presented in Eq. (6.28) will never give an infinite result. Thus the signal that a higher order pole could be present is that the residue is extremely large, at least of the order of  $1/|\delta|$ . To evaluate the multiplicity of the pole, the well-known formula

$$\int_{\mathcal{C}} \frac{f'(z)}{f(z)} dz = 2\pi i (Z - N), \quad (6.29)$$

where  $Z$  and  $N$  are the number of zeros and poles of  $f(z)$  within the closed contour  $\mathcal{C}$ . By choosing  $\mathcal{C}$  to be a circle of very small radius centered at the singularity of interest, such that other singularities and zeros are all outside  $\mathcal{C}$ , Eq. (6.29) provides the multiplicity of the pole.

A more complicated problem occurs when  $\tilde{\mathbf{x}}(z)$  has a branch cut. These can be extremely difficult to search for numerically, since they are not signaled by any easily-identifiable local behaviour of  $\tilde{\mathbf{x}}(z)$ . However, it is clear from Eq. (6.27) that branch cuts

<sup>4</sup>The residue calculated by Eq. (6.28) will appear to be infinite if the singularity in  $\tilde{\mathbf{x}}(z)$  is a higher order pole. Strictly speaking, this should not be referred to as an infinite residue, but I use this notation as it provides a more intuitive feeling for the output of a numerical residue evaluation.

in  $\tilde{\mathbf{x}}(z)$  are associated with branch cuts in  $[I - \tilde{C}(z)]$ . Thus if  $[I - \tilde{C}(z)]$  is known to be single-valued (which is true for the systems studied in the following chapter),  $\tilde{\mathbf{x}}(z)$  cannot have branch cuts.

In general, when using the singularities of  $\tilde{\mathbf{x}}(z)$  to determine an approximate solution to (6.14), it is a good idea to plot  $(z - a)\tilde{\mathbf{x}}(z)$  for several values of  $z$  near the singularity  $a$ . This is the easiest and most reliable method of locating unexpected behaviour, and is not numerically intensive.

## Chapter 7

# Dephasing of Quantum Dots: Coulomb Interaction with a Gate

### 7.1 Overview

The development of a solid-state quantum computer is hindered primarily by the fact that solid-state systems contain many subsystems that interact with the qubits. These interactions cause information loss and limit the timescale on which reliable results can be obtained, as described in the previous chapter. Examples of such subsystems are phonons, impurities and surface effects. To limit this information loss, or at least to be able to estimate the timescale on which it occurs, it is necessary to consider the individual subsystems and evaluate the relevance of each. In this chapter, I show how we used the generalized master equation method described above to determine the dephasing and dissipation rates of quantum dots due to their interaction with gate electrodes.

Gate electrodes are metallic strips that are ubiquitous in quantum dot experiments. In electronically-controlled quantum dots, gate electrodes define the quantum dot structure within a quantum well and can also be used to prepare initial states [44, 69, 154]. Self-assembled quantum dots, on the other hand, require electrodes to manipulate the charge degrees of freedom [51, 112, 137]. Even when spin is the primary degree of freedom, gate electrodes have been used to provide an electric field gradient that separates electrons and holes and prevents recombination — see the discussion in chapter 1 and Refs. [71, 113].

A brief note of caution is needed. In some quantum dot experiments, electrons tunnel into the dot from nearby metallic leads (see, e.g., Ref. [44] and the references therein). These leads exhibit a strong coherence with the quantum dot state. Such experimental setups are not considered here; we have only investigated the case where the metallic strips are far enough from the quantum dot that tunneling and exchange interactions are prohibited. To prevent confusion, I use the term “electrode” or “gate” for a metallic strip that is far from the quantum dot and “lead” for a strip that is close.

Usually, the electrodes are assumed to contribute only a classical field, but this is not necessarily true. Their internal degrees of freedom may interact with the quantum dot, leading to both dephasing and dissipation. However, until our work, dephasing due to gate electrodes has only been considered within either the classical framework of interaction with transient magnetic fields due to thermally-induced current fluctuations [96] or the

quantum-mechanical framework of the spin-boson model [121, 198], where the electrodes are treated as a noninteracting boson gas [12]. The former neglects the interaction with charge fluctuations, while the latter neglects relaxation processes within the electrode. In an attempt to improve on these results, we implemented the full Coulomb interaction with the gate. We also did not make the Markov approximation, as Refs. [12, 96] did; memory effects in the gate are included using the master equation approach described in the previous chapter.

Before presenting the details of our investigation, it is relevant to note that quantum dots in contact with a heat bath exhibit two distinct types of dephasing. The first occurs as the initial, experimentally-prepared state evolves into a quasiparticle state of the many-body system. For example, a bare exciton excited by a photon will rapidly lose coherence when “dressed” with phonons [81]. Similarly, an electron beam passing through a quantum dot dephases as it induces an image charge in a nearby point contact [5, 24, 123, 175, 177]. I shall refer to this type of dephasing as “initialization dephasing”. It is caused by renormalization and, if the environment (i.e. everything apart from the quantum dot system) is stationary, it has only limited effect: it ceases to act once a quasiparticle state is formed. If the environment is nonstationary, as in a current-based “which-path” detector [5, 24, 123, 175, 177], a true quasiparticle state cannot form. In this case, initialization dephasing may be unlimited. However, unlike the Coulomb-induced dephasing present in “which-path” detectors, our Coulomb bath has zero current. We therefore neglected the initialization dephasing and assumed that the image charge in the gate electrodes and all related renormalization effects were included in the initial conditions.

We investigated the second type of dephasing, which is caused by interactions of the quasiparticle state with the environment. This proceeds either via scattering of the quasiparticle or by the interaction of the quasiparticle with fluctuations (both quantum and thermal). It is referred to as “dynamical dephasing”. Typically, dynamical dephasing acts on a much slower timescale than initialization dephasing, but is unlimited in effect; it will eventually lead to complete decoherence of the system. For example, a bare exciton will be dressed by LA phonons within several picoseconds of excitation, causing up to 25% loss of coherence [81]. The dressed exciton is more stable and subsequent dephasing occurs on the timescale of nanoseconds [69, 180].

We have studied the dephasing of both spin and charge qubits due to the interaction with gate electrodes. Many aspects of these two investigations are identical, so only the calculation for spin qubits is presented in full detail. For charge qubits, I present only the important differences from the spin qubit investigation and the results; a more detailed account is available in Ref. [204]<sup>1</sup>. Hartree atomic units are used, except where otherwise stated.

## 7.2 Dephasing of spin qubits

### 7.2.1 Structure of qubit

The structure of the spin qubit studied here has already been described in chapter 1. It consists of a small island, roughly 30 nm wide and 5 nm high, of InGaAs between

<sup>1</sup>These results are reused with permission from S. R. Woodford, A. Bringer and K. M. Indlekofer, Physical Review B **76**, 064306 (2007). Copyright 2007 by the American Physical Society.

two layers of GaAs [71, 113]. Other layers of different semiconductors are added to this heterostructure, most notably AlGaAs to prevent the electrons tunneling out when the bias field is applied. Finally, a thin metallic layer is added to provide the necessary bias field. As described in chapter 1, this prevents the recombination of optically excited electron-hole pairs, allowing charge to be “injected” into the qubit. In addition, the system is subjected to a magnetic field along the direction of strong confinement, which separates the different optically accessible exciton states and makes angular momentum a good quantum number.

The electron in the quantum dot is described by the Kane Hamiltonian [92, 96]

$$H = \frac{\mathbf{p}^2}{2m^*} + U(\mathbf{r}) + g\mu B\sigma_z - \beta(p_x\sigma_x - p_y\sigma_y). \quad (7.1)$$

For GaAs, the effective mass is  $m^* = 0.067$ , the  $g$ -factor is  $g = -0.44$  and the Dresselhaus velocity is  $\beta = (2 \pm 1) \times 10^5 \text{ cm s}^{-1}$  [96, 180]. The confining potential  $U(\mathbf{r})$  is approximated as a square, 2-dimensional well with infinitely high walls.

Quantum computing requires a 2-state system. For a spin qubit, the relevant states are usually the up and down states of the electronic ground state. The eigenstates of (7.1) are not pure up or down states, since the Dresselhaus effect mixes the spin components. However, the true eigenstates must be used; simultaneous perturbation theory in two independent parameters (the Dresselhaus term and the interaction with the electrodes) is difficult to perform consistently.

The true eigenstates of (7.1) could not be found analytically, so we relied on numerical techniques. We expanded the unknown states over Fourier modes in the quantum well and diagonalized the Hamiltonian relative to this basis. Since only the lowest two eigenstates were needed, the number of Fourier modes could be limited to 10 modes for each spatial direction. The probability densities of the two lowest-energy eigenstates are plotted in Fig. 7.1, separated into up and down spin components. Clearly, the mixing of spin components is small enough that the states can be referred to as “up” and “down” states, which is done in what follows.

Before describing the coupling to the gate electrodes, it is relevant to mention the other studies of spin qubit dephasing that are available in the literature. This will allow comparison with our results and reveal the relative importance of the different interactions.

The most-studied mechanism is *phonon scattering* [56, 95, 96, 155, 205], in which the higher energy state decays to the lower state by the emission of a phonon. There are many different ways that a phonon can cause a spin-flip in a quantum dot. The clearest mechanism is through piezoelectric coupling [96]. In this case, fluctuations of the spin cause changes of the local electric field via the spin-orbit coupling. If the energies match, this electric field can move the nuclei and generate a phonon. For the spin qubit considered here, this process yields an energy dissipation rate of  $\Gamma_1 = 1.5 \times 10^3 \text{ s}^{-1}$  for a magnetic field  $B = 1 \text{ T}$  and  $\hbar\omega = 10 \text{ K}$  (this is the energy separation between different orbital levels in the dot, in Kelvin) [96]. This dissipation rate scales as  $B^5$  at low temperatures. The dephasing rate is approximately half the dissipation rate [56].

In addition to the piezoelectric coupling, phonons may cause a nonadiabatic change in the parameters of the Hamiltonian, i.e.  $m^*$ ,  $\beta$ ,  $g$  and  $U(\mathbf{r})$ , each of which can give rise to spin-flip scattering [205]. Generally, these processes are weak, although they become important for quantum dots smaller than about 20 nm. All except the change of the Dresselhaus parameter lead to dissipation rates that scale as  $B^5$ ; the scattering rate for

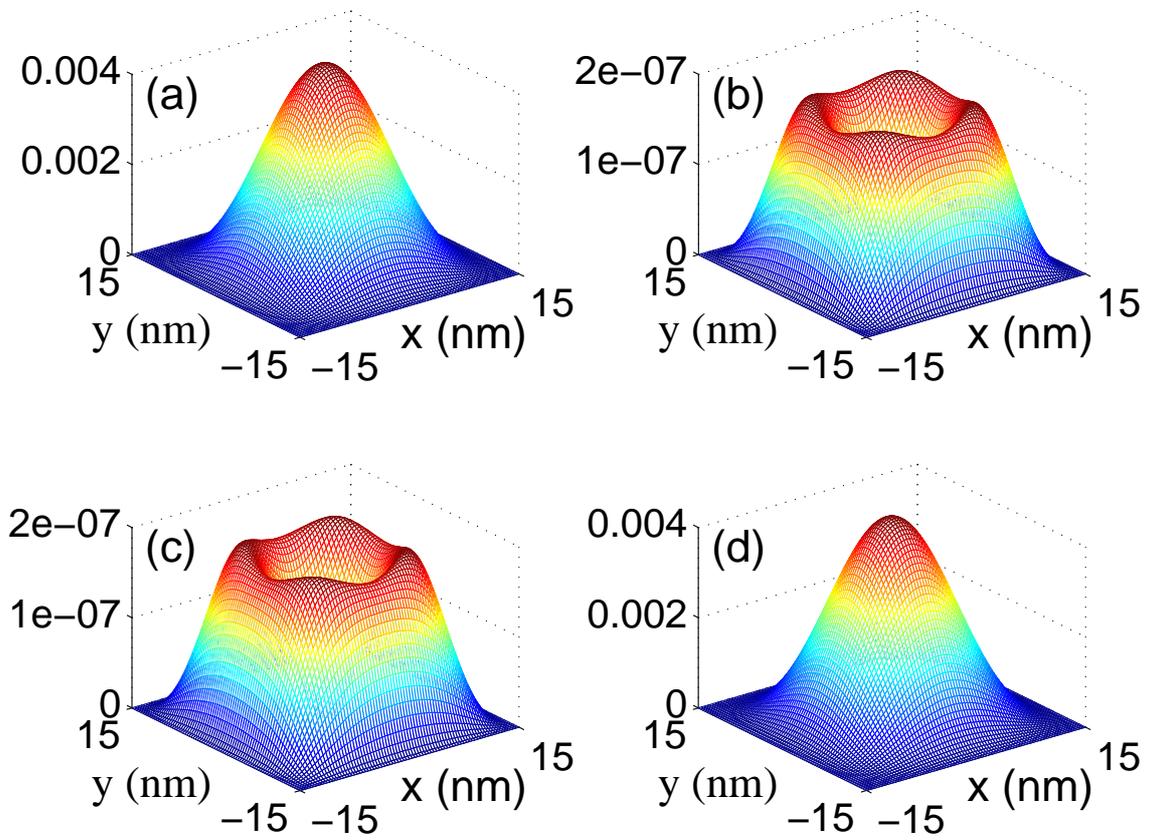


Figure 7.1: Probability densities ( $|\psi|^2$ ) of the spin components of the lowest two eigenstates of the quantum dot. (a) and (b) are the up- and down-components of the lowest eigenstate, respectively, while (c) and (d) are the up- and down-components of the second lowest eigenstate. The strong spin polarization of these low-energy states permits them to be referred to as the “up” and “down” states, respectively. Due to the strong confinement, the dependence on  $z$  takes the form of a  $\delta$ -function and is not shown.

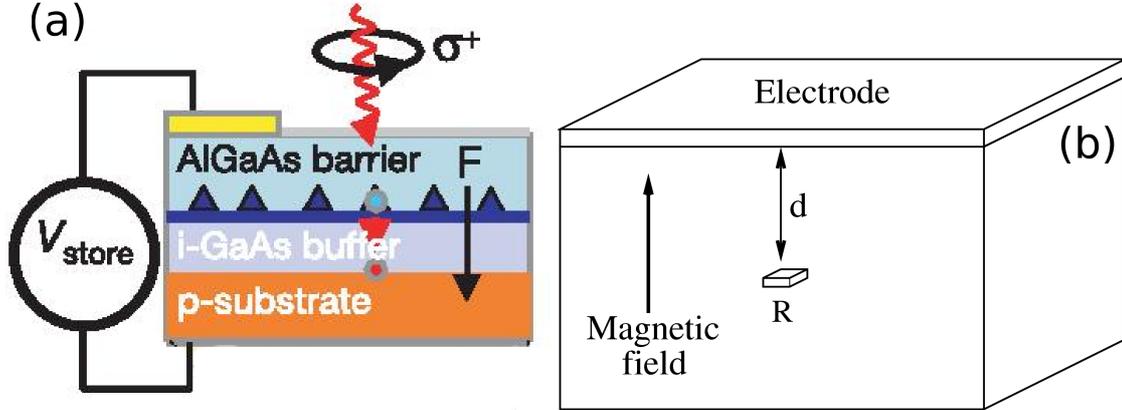


Figure 7.2: (a) The experimental setup considered here, taken from Ref. [113]. (b) An approximate model of (a), including the most important features. The quantum dot is assumed to be a square well with infinitely high walls and the gate electrode is assumed thick enough that it may be approximated as a bulk metal filling the region  $z > 0$ . (a) is reprinted by permission from Macmillan Publishers Ltd: *Nature*, *Optically programmable electron spin memory using semiconductor quantum dots*, M. Kroutvar, Y. Ducommun, D. Heiss, M. Bichler, D. Schuh, G. Abstreiter and J. J. Finley, **432**, 81 (2004). Copyright (2004).

changes of  $\beta$  is linear in  $B$ . In principle, this linear dependence will dominate at low fields, but due to the extreme weakness of this mechanism, the transition to the linear regime is unlikely to be detected.

Several other spin-flip mechanisms were discussed in Ref. [96]: phonon-induced modulation of the hyperfine coupling with nuclei, interaction with the fluctuating magnetic field of the electrodes and exchange scattering processes between the bound electron and the conduction electrons<sup>2</sup>. In general, these are all far weaker than the piezoelectric phonon scattering [89, 90, 96]. The coupling of charge fluctuations in the gate to the spin of the electron via spin-orbit coupling, which is the mechanism that we studied, has not received attention until now.

### 7.2.2 Coulomb interaction with the gate

A diagram of the experimental setup (taken from Ref. [113]) is depicted in Fig. 7.2, along with the model that we used to approximate the system. The electrode consists of a thin layer (around 5 nm) of metal located above the dot; it covers the entire structure, not just the part of it indicated in Fig. 7.2 (a). Because metals have a small skin depth and the edges of the electrode are far from the quantum dot, we have treated the electrode as a semi-infinite metal bulk occupying the entire  $z > 0$  region. This is a reasonable approximation even in the case where there are several gate electrodes charged to different potentials (e.g. the experiments described in Ref. [44]) — the dynamical properties of such a configuration will not be notably different from the single-electrode model [204].

<sup>2</sup>These pathways have also been considered previously in connection with spin-flip processes in arrays of donor atoms [1, 155].

As mentioned previously, we assumed that static renormalization effects have already been taken into account in the potential  $U(\mathbf{r})$ , i.e. that the square-well potential includes the image charge. When calculating the dynamical properties of the gate, the static image charge is therefore assumed to vanish. We also assumed that at  $t = 0$ , there is no correlation between the gate and the electron in the qubit. This means that for  $t < 0$ , the quantum dot system behaves classically, and at  $t = 0$  we “turn on” the quantum mechanics. This is equivalent to assuming that at  $t = 0$ , we prepare the dot very rapidly in a particular quantum state, thus breaking all correlations with the gate. Since spin injection occurs on a nanosecond timescale and subsequent dynamics on a microsecond scale, this is a reasonable assumption [113].

The closed quantum system consisting of the gate electrode and the quantum dot is governed by the Hamiltonian

$$\begin{aligned}
 H &= \sum_{i=1}^2 E_i a_i^\dagger a_i + H_B \\
 &+ \sum_{i,j} \int d\mathbf{r} d\mathbf{r}' \frac{1}{\epsilon_r |\mathbf{r} - \mathbf{r}'|} \varphi_i^*(\mathbf{r}) \varphi_j(\mathbf{r}) \Psi^\dagger(\mathbf{r}') \Psi(\mathbf{r}') a_i^\dagger a_j.
 \end{aligned} \tag{7.2}$$

The quantum dot is expressed in the energy eigenfunction basis:  $\varphi_1$  and  $\varphi_2$  are the numerically-constructed up and down eigenstates of Eq. (7.1), respectively, with spin densities plotted in Fig. 7.1. The creation and annihilation operators  $a_i^\dagger$  and  $a_i$  of the quantum dot are expressed relative to this basis. The position representation is used for the gate electrons.  $H_B$  is the Hamiltonian for the bare gate and  $\epsilon_r$  is the dielectric constant of the semiconductor medium between the dot and the gate.

The position representation is preferable for describing the gate, since it shows explicitly that the interaction couples the quantum dot states to the electron density of the gate [via the operator  $n(\mathbf{r}') = \Psi^\dagger(\mathbf{r}') \Psi(\mathbf{r}')$ ]. This has the advantage that the density-density correlation function used below has the form of a 2-point bosonic correlator; in any other basis, it appears as a 4-point fermionic correlator (with two creation operators and two annihilation operators). Although the interaction term in Eq. (7.2) is similar to the spin-boson interaction, i.e. it consists of a 2-state system coupled to bosonic operators, Eq. (7.2) is *not* the spin-boson model: the Hamiltonian  $H_B$  contains fermionic terms that cannot be expressed in terms of  $n(\mathbf{r})$ . In terms of the electron density operator  $n(\mathbf{r})$ , the above-mentioned vanishing of the static image charge can be written as

$$\langle n(\mathbf{r}) \rangle = \frac{\text{Tr}_B [\exp(-\beta H_B) n(\mathbf{r})]}{\text{Tr}_B [\exp(-\beta H_B)]} = 0. \tag{7.3}$$

Following the methods of chapter 6, we determined the equations of motion for the reduced density matrix of the quantum dot,  $\rho_{ij}(t)$ , expressed relative to the energy eigenfunction basis spanned by  $\{\varphi_1, \varphi_2\}$ . As mentioned in chapter 6, the generalized master equation is only tractable when approximations are made; therefore we truncated Eq. 6.7 at second order, as in Eq. (6.12). The reference vector  $B_{\text{ref}}$  was chosen as the thermodynamic state (6.9) and the initial separability of the quantum dot and the electrode, condition (6.10), was assumed.

Even with these approximations, the equations of motion are complicated and new notation is needed to write them in a reasonable form. The wave-functions of the quantum

dot enter the master equation in overlap integrals with the Coulomb interaction kernel:

$$T_{ij}(\mathbf{r}') \equiv \int d\mathbf{r} \varphi_i^*(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \varphi_j(\mathbf{r}). \quad (7.4)$$

In addition, the following density-density correlation functions of the metal gate arise:

$$\begin{aligned} \mathcal{G}^>(\mathbf{r}, \mathbf{s}, t) &= \text{Tr}_B \left( e^{-\beta(H_B - \Omega)} e^{iH_B t} \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}) e^{-iH_B t} \Psi^\dagger(\mathbf{s}) \Psi(\mathbf{s}) \right) \\ &= \langle n(\mathbf{r}, t) n(\mathbf{s}) \rangle \end{aligned} \quad (7.5)$$

$$\begin{aligned} \mathcal{G}^<(\mathbf{r}, \mathbf{s}, t) &= \text{Tr}_B \left( e^{-\beta(H_B - \Omega)} \Psi^\dagger(\mathbf{s}) \Psi(\mathbf{s}) e^{iH_B t} \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}) e^{-iH_B t} \right) \\ &= \langle n(\mathbf{s}) n(\mathbf{r}, t) \rangle. \end{aligned} \quad (7.6)$$

Here,  $\langle X \rangle$  denotes the thermal expectation value of the operator  $X$  and we follow the notation of Kadanoff and Baym [86] in which the “greater-than” function has the order of its arguments  $(\mathbf{r}, \mathbf{s})$  preserved, while the “less-than” function has the order reversed<sup>3</sup>. Finally, using the notation  $E_{ij} = (E_i - E_j)$ , the generalized master equation (6.7) becomes

$$\begin{aligned} \dot{\rho}_{ij}(t) &= iE_{ji} \rho_{ij}(t) + \\ &\frac{1}{\epsilon_r^2} \sum_{m,n} \int_0^t dt' \int d\mathbf{r}' d\mathbf{r}'' \left\{ \mathcal{G}^>(\mathbf{r}'', \mathbf{r}', t') \left[ e^{iE_{ni}t'} \rho_{mn}(t-t') T_{im}(\mathbf{r}') T_{nj}(\mathbf{r}'') \right. \right. \\ &\quad \left. \left. - e^{iE_{jm}t'} \rho_{nj}(t-t') T_{mn}(\mathbf{r}') T_{im}(\mathbf{r}'') \right] \right. \\ &\quad \left. + \mathcal{G}^<(\mathbf{r}'', \mathbf{r}', t') \left[ e^{iE_{jm}t'} \rho_{mn}(t-t') T_{nj}(\mathbf{r}') T_{im}(\mathbf{r}'') \right. \right. \\ &\quad \left. \left. - e^{iE_{ni}t'} \rho_{im}(t-t') T_{mn}(\mathbf{r}') T_{nj}(\mathbf{r}'') \right] \right\} dt'. \end{aligned} \quad (7.7)$$

Eq. (7.7) appears to be far more complicated than Eq. (6.13), which is the expected form of the master equation. It can be simplified by noting that the functions  $\mathcal{G}^{\langle, \rangle}$  and  $T_{ij}$  are independent of the density matrix and can be calculated separately. By combining these into another function

$$R_{ijkl}^{\langle, \rangle}(t) = \frac{1}{\epsilon_r^2} \int d\mathbf{r}' d\mathbf{r}'' T_{ij}(\mathbf{r}') T_{kl}(\mathbf{r}'') \mathcal{G}^{\langle, \rangle}(\mathbf{r}'', \mathbf{r}', t), \quad (7.8)$$

we can rewrite Eq. (7.7) as

$$\begin{aligned} \dot{\rho}_{ij}(t) &= iE_{ji} \rho_{ij}(t) + \\ &\sum_{m,n} \int_0^t dt' \left[ e^{iE_{ni}t'} R_{imnj}^>(t') \rho_{mn}(t-t') - e^{iE_{jm}t'} R_{mnim}^>(t') \rho_{nj}(t-t') \right. \\ &\quad \left. + e^{iE_{jm}t'} R_{njim}^<(t') \rho_{mn}(t-t') - e^{iE_{ni}t'} R_{mnnj}^<(t') \rho_{im}(t-t') \right] dt'. \end{aligned} \quad (7.9)$$

Once the functions  $R_{ijkl}^{\langle, \rangle}$  have been calculated, Eq. (7.9) can be solved using the methods of the previous chapter. In the next section, I present the method that we used to approximate  $\mathcal{G}^{\langle, \rangle}$  and  $T_{ij}$ , from which  $R_{ijkl}^{\langle, \rangle}$  is determined.

<sup>3</sup>Strictly speaking, the “less-than” correlation function of an operator  $X(\mathbf{r}, t)$  is defined by  $\mathcal{G}_X^<(\mathbf{r}t, \mathbf{r}'t') \equiv \langle X^\dagger(\mathbf{r}', t') X(\mathbf{r}, t) \rangle$ ; in writing Eqs. (7.5)-(7.6), we have used invariance in time to eliminate one of the time variables and also used the hermiticity of the density operator,  $n(\mathbf{r}, t) = n^\dagger(\mathbf{r}, t)$ .

### 7.2.3 Response function of the gate

To obtain  $\mathcal{G}^{<, >}(t)$ , we made use of the generalized reflection coefficient  $g(q, \omega)$ , which is a well known quantity in surface science [125, 153]. The function  $g(q, \omega)$  is defined as the coefficient of the reflected potential induced by a point charge near a surface. More precisely, given a potential

$$\phi_{\text{ext}}(\mathbf{r}, t) = \frac{2\pi}{q} e^{-q(z+d)} e^{i\mathbf{q}_{\parallel} \cdot \mathbf{r}_{\parallel}} e^{i\omega t}, \quad (7.10)$$

the induced potential due to the image charge will be

$$\phi_{\text{ind}}(\mathbf{r}, t) = -g(q, \omega) \frac{2\pi}{q} e^{q(z-d)} e^{i\mathbf{q}_{\parallel} \cdot \mathbf{r}_{\parallel}} e^{i\omega t}. \quad (7.11)$$

If the dot is reasonably far away from the gate (farther than about 10nm [153]), then the approximate form for  $g$  is [125, 153]

$$g(q, \omega) = \frac{1}{\epsilon_r} \frac{\epsilon(\omega) - 1}{\epsilon(\omega) + 1}. \quad (7.12)$$

Linear- and higher-order corrections in  $q$  are only relevant when the dot is so close to the gate that the electric field lines are not orthogonal to the surface [125, 153].

An alternative method of evaluating  $g(q, \omega)$  is to determine the linear response of the surface charge density to a potential of the form (7.10). The induced charge density is related by a Kubo formula to the retarded density-density correlation function [129]

$$\mathcal{G}_{\text{ret}}(\mathbf{r}, \mathbf{s}, t) \equiv i\theta(t) [\mathcal{G}^{>}(\mathbf{r}, \mathbf{s}, t) - \mathcal{G}^{<}(\mathbf{r}, \mathbf{s}, t)].$$

By the Kubo formula,  $g(q, \omega)$  can be related to the correlation functions  $\mathcal{G}^{<, >}(\mathbf{r}, \mathbf{s}, t)$ . The details of this calculation are straightforward and we found that

$$\mathcal{G}^{>}(\mathbf{q}, \omega) = e^{\beta\omega} \mathcal{G}^{<}(\mathbf{q}, \omega) = \frac{2q\epsilon_r}{(2\pi)^5 (1 - e^{-\beta\omega})} \text{Im}[g(q, \omega)], \quad (7.13)$$

where  $\mathcal{G}^{<, >}(\mathbf{q}, \omega)$  is the Fourier transform of  $\mathcal{G}^{<, >}(\mathbf{r}, \mathbf{s}, t)$  and the translational invariance of the surface has been used to write  $\mathcal{G}^{<, >}(\mathbf{q}, \omega)$  as functions of a single momentum.

By Fourier transforming the Coulomb interaction term  $1/|\mathbf{r} - \mathbf{r}'|$ , we expressed  $R_{ijkl}^{<, >}(t)$  as

$$R_{ijkl}^{<, >}(t) = \frac{1}{2\pi^2 \epsilon_r^2} I^{<, >}(t) \int d\mathbf{r}' d\mathbf{r}'' \varphi_i^*(\mathbf{r}'') \varphi_j(\mathbf{r}'') \varphi_k^*(\mathbf{r}') \varphi_l(\mathbf{r}') \int d\mathbf{q}_{\parallel} \frac{1}{|\mathbf{q}_{\parallel}|} e^{|\mathbf{q}_{\parallel}|(z'+z'')} e^{i\mathbf{q}_{\parallel} \cdot (\mathbf{r}''_{\parallel} - \mathbf{r}'_{\parallel})}. \quad (7.14)$$

The time dependence is governed by  $I^{<, >}(t)$ , with

$$I^{>}(t) = I^{<}(-t) = \int d\omega \frac{1}{1 - e^{-\beta\omega}} e^{-i\omega t} \text{Im} \left[ \frac{\epsilon(\omega) - 1}{\epsilon(\omega) + 1} \right], \quad (7.15)$$

and the integration variables  $\mathbf{r}'$ ,  $\mathbf{r}''$  in Eq. (7.14) have been resolved into components, e.g.  $\mathbf{r}' = (\mathbf{r}'_{\parallel}, z')$ .

Finally, inserting the identity

$$\frac{1}{q} e^{q(z'+z'')} = \int dk \frac{1}{q^2 + k^2} e^{ik(z'+z'')} \quad (7.16)$$

into Eq. (7.14) yields

$$R_{ijkl}^{<, >}(t) = \frac{1}{\pi \epsilon_r^2} I^{<, >}(t) \int d\mathbf{r}' d\mathbf{r}'' \frac{\varphi_i^*(\mathbf{r}'') \varphi_j(\mathbf{r}'') \varphi_k^*(\mathbf{r}') \varphi_l(\mathbf{r}')}{\sqrt{|\mathbf{r}''_{\parallel} - \mathbf{r}'_{\parallel}|^2 + (z' + z'')^2}}. \quad (7.17)$$

The spatial integral contains the qubit wavefunctions with a well-defined kernel. Since these wavefunctions were determined numerically, the integral can be computed for any choice of indices  $\{i, j, k, l\}$ .

Due to symmetry reasons,  $R_{ijkl}$  vanishes if 3 of the indices coincide and differ from the fourth. From Eq. (7.9), it can readily be checked that this causes the diagonal components of  $\rho$  to evolve independently from the off-diagonal components. This reduces the search for poles of the Laplace transform from the vanishing of the determinant of a  $4 \times 4$  matrix (see section 6.3.2) to the vanishing of two determinants of  $2 \times 2$  matrices.

Before Eq. (7.9) can be solved, the time-dependent factors  $I^{<, >}(t)$  must be calculated. We assumed that  $\epsilon(\omega)$  could be approximated by the Drude formula [40],

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + 2i\gamma)}, \quad (7.18)$$

with  $\omega_p$  the plasma frequency and  $\gamma$  the relaxation rate at the Fermi energy [100]. Eq. (7.18) allows  $I^{>}(t)$  to be evaluated explicitly using contour integration, yielding

$$I^{>}(t) \approx \begin{cases} \frac{\pi \omega_p^2 e^{-\gamma t}}{4x} e^{-ixt} - \frac{2\pi}{\beta} \sum_{n=1}^{\infty} e^{-2\pi n t / \beta} g_n & t > 0 \\ \frac{\pi \omega_p^2 e^{\gamma t}}{4x} e^{-ixt} - \frac{2\pi}{\beta} \sum_{n=1}^{\infty} e^{2\pi n t / \beta} g_n & t < 0 \end{cases} \quad (7.19)$$

Here, the function  $g_n$  is defined as

$$g_n = \frac{8\pi n \beta^3 \gamma \omega_p^2}{4(2\pi)^4 n^4 + 4(2\pi)^2 n^2 \beta^2 (\omega_p^2 - 4\gamma^2) + \omega_p^4 \beta^4} \quad (7.20)$$

and  $x = (\omega_p^2/2 - \gamma^2)^{1/2}$  is the real part of the first-quadrant pole of  $\text{Im}[(\epsilon - 1)/(\epsilon + 1)]$ . Since  $\omega_p$  is typically of the order of 10 eV and  $\gamma \approx 1$  eV,  $\exp(-ixt)$  oscillates extremely rapidly, i.e. with a period of approximately 0.1 fs. This fact was used in the previous chapter to argue that numerical integration is inadequate for solving Eq. (7.9) over long timescales.

In Eq. (7.19), terms that are a factor of  $\exp(-\beta\omega_p)$  smaller than the leading order terms have been neglected; for  $\omega_p = 10\text{eV}$ , this is a reasonable approximation for temperatures as high as 1000K and is certainly justified for the temperatures considered here.

The preferred method of solution of Eq. (7.9) is the Laplace transform technique, introduced in the previous chapter. For this, the Laplace transforms of the factors  $I^{<, >}(t)$  are needed. This calculation is complicated by the slow convergence of the sum in Eq.

(7.19) and is presented in the next section. Before describing this calculation, it is useful to make some general comments about the equations derived above.

The integral kernel in Eq. (7.17) is similar to the classically-determined interaction of a charge with its image charge [82]. This allows us to interpret the above equations as describing the interaction of the electron in the qubit with its image charge. Since we assumed that the *static* image charge vanishes (7.3), only fluctuations of the image contribute to this interaction. The time-dependent term  $I(t)$  shows that the interaction is not instantaneous; fluctuations and scattering events in the electrode make the image charge dynamics lag behind those of the electron in the qubit. This causes memory effects: through the interaction, the electron “remembers” where it was previously.

Of course, there are differences from the classical result. The classical interaction of an electron with its image charge contains only charge densities [82], while Eq. (7.17) contains wavefunctions. Furthermore, a classical image charge does not undergo quantum fluctuations. Nonetheless, this interpretation provides some physical insight into the main features of the problem and is helpful in understanding the results below.

#### 7.2.4 Laplace transform of $I^>(t)$

Here, I present the calculation of the Laplace transform of Eq. (7.19). Laplace transforms of a convergent series can be performed termwise, yielding

$$\tilde{I}^>(z) \approx \frac{\pi\omega_p^2}{4x(z + \gamma + ix)} - \frac{2\pi}{\beta} \sum_{n=1}^{\infty} \frac{g_n}{z + 2\pi n/\beta}, \quad (7.21)$$

with  $g_n$  defined in Eq. (7.20). Similarly to Eq. (7.19), this equation is only approximate since terms that are a factor of  $\exp(-\beta\omega_p)$  smaller than the leading order terms are neglected.

The series in Eq. (7.21) can be (approximately) evaluated analytically in the limits  $|z|\beta \gg 1$  (when the gate is in the ground state) and  $|z|\beta \ll 1$  (high temperature). However, we found that for most choices of physical parameters (magnetic field, temperature and separation distance between the quantum dot and the metal), the gate is neither very hot nor very cold:  $|z| \sim E_{21} \sim 1/\beta$ . For example, if the magnetic field is 10 T and the temperature is 100 mK,  $\beta z \approx 0.4$ . In this case, the series in Eq. (7.21) must be calculated numerically using a truncation at some large value  $N$ .

The factor  $g_n$  decays as  $n^{-3}$  provided  $n \gg \beta\omega_p \equiv \omega_p/k_B T$ . Since the plasma energy is much larger than the temperature, the series needs very large  $N$ . However, if  $N$  is too large, the accuracy will be reduced due to round-off errors. This problem is aggravated by the almost complete cancellation of the first term of Eq. (7.21) with the series — typically, the real part of the series sums to approximately 0.2, while  $\text{Re}[\tilde{I}^>(z)] \approx 10^{-6}$ .

We found that the optimal number of terms is  $N \sim 10^8$ . For  $T > 40$  mK, the tail of the series is negligible if  $N = 1 \times 10^8$  terms are used; for smaller temperatures,  $N = 4 \times 10^8$  terms were needed. The maximum possible round-off error (using double-precision arithmetic) for this number of terms is approximately three orders of magnitude less than  $\tilde{I}^>(z)$ ; this is sufficiently small that it does not affect the results. To numerically verify that round-off effects were negligible, we reversed the order of the sum and we also broke the term  $\pi\omega_p^2/4x(z + \gamma + ix)$  into  $N$  terms and included these terms in the summation; neither of these modifications significantly affected the result.

### 7.2.5 Results

In order to obtain numerical results, the physical constants entering Eqs. (7.1), (7.9) and (7.21) must be specified. The semiconductor surrounding the quantum dot is assumed to be GaAs, for which  $\epsilon_r \approx 13$ ,  $g = -0.44$ ,  $\beta = (2 \pm 1) \times 10^5 \text{ cm s}^{-1}$  and  $m^* \approx 0.067$ . To describe the metal, we take the “typical” values  $\omega_p = 10 \text{ eV}$  and  $\gamma = 1 \text{ eV}$  [100,138]. The side-length of the (square) quantum dot was set to  $R = 30 \text{ nm}$ , which is a typical size for self-assembled quantum dots [71,113]. The other three parameters, temperature, magnetic field and distance to the gate, were varied. When each of these was varied, the others were chosen at the following default values: temperature  $T = 1 \text{ K}$ ; magnetic field  $B = 8 \text{ T}$ ; distance to the gate  $d = 50 \text{ nm}$ . These parameters were chosen to aid comparison with the experimental results of Ref. [113].

We found qualitatively similar results for all choices of the parameters. The evolution of the diagonal elements  $\rho_{11}$  and  $\rho_{22}$  is dominated by two real, simple poles in their Laplace transforms. A pole at  $z = 0$  ensures that there is a nondecaying component and another pole at  $z < 0$  describes the decay. From the residues at these poles, we found that asymptotically (i.e. after a transient period during which poles at smaller  $z$  values influence the dynamics), the diagonal matrix elements evolve according to

$$\begin{pmatrix} \rho_{11}(t) \\ \rho_{22}(t) \end{pmatrix} = \begin{pmatrix} a + (1-a)e^{-t/T_1} & a(1-e^{-t/T_1}) \\ (1-a)(1-e^{-t/T_1}) & 1-a(1-e^{-t/T_1}) \end{pmatrix} \begin{pmatrix} \rho_{11}^{(0)} \\ \rho_{22}^{(0)} \end{pmatrix} \quad (7.22)$$

where  $a$  and  $T_1$  are functions of the parameters and I use the standard notation  $T_1$  to denote the energy relaxation time. Eq. (7.22) preserves the relation  $\rho_{11} + \rho_{22} = 1$  and shows that all initial conditions decay to the same final state,

$$\begin{pmatrix} \rho_{11}(\infty) \\ \rho_{22}(\infty) \end{pmatrix} = \begin{pmatrix} a \\ 1-a \end{pmatrix} \quad (7.23)$$

The dynamics of the off-diagonal elements  $\rho_{12}$  and  $\rho_{21}$  are governed by a complex conjugate pair of simple poles with negative real part. Many of the residues at these poles vanish and the evolution is asymptotically described by

$$\begin{pmatrix} \rho_{12}(t) \\ \rho_{21}(t) \end{pmatrix} = \begin{pmatrix} e^{-t/T_2 + i\epsilon_{21}t} & 0 \\ 0 & e^{-t/T_2 - i\epsilon_{21}t} \end{pmatrix} \begin{pmatrix} \rho_{12}^{(0)} \\ \rho_{21}^{(0)} \end{pmatrix} \quad (7.24)$$

Clearly, this evolution preserves the relation  $\rho_{12} = \rho_{21}^*$ . The oscillation frequency  $\epsilon_{21}$  is always slightly larger (around 1%) than the energy difference between the up and down states  $E_{21}$ , showing that the gate assists transitions between the two states.

Fig. 7.3 shows both the dephasing rate and the pure dephasing rate as functions of  $B$ ,  $d$  and  $T$ . The pure dephasing rate was defined in Eq. (6.23) as

$$\frac{1}{T_\Phi} = \frac{1}{T_2} - \frac{1}{2T_1}. \quad (7.25)$$

It is a general theorem that this rate should be positive (see, e.g. Ref. [142] and the discussion in chapter 6). As is clear from Figs. 7.3 (d)-(f), our calculations did not yield this result. On the contrary,  $1/T_\Phi$  was negative for all parameters.

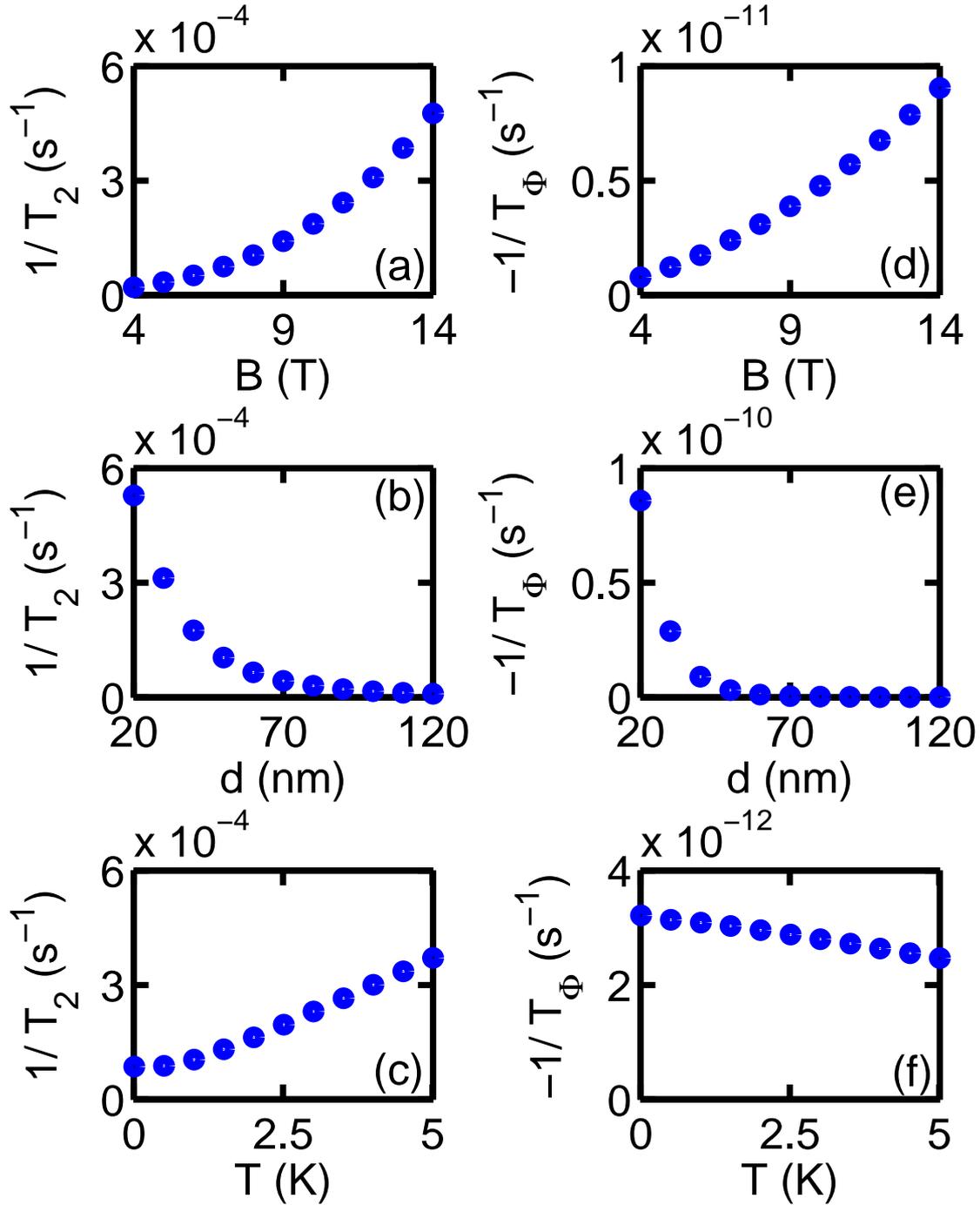


Figure 7.3: The dephasing rates (a)-(c) and pure dephasing rates (d)-(f) of a spin quantum dot due to the interaction with the gate electrode. The pure dephasing rate is defined in Eq. (7.25). These rates are functions of the magnetic field, distance between the gate and the quantum dot and temperature. In (a) and (d), we varied the magnetic field, keeping  $d = 30$  nm and  $T = 1$  K. In (b) and (e),  $B = 8$  T,  $T = 1$  K and the distance was varied. Finally, (c) and (f) show the dependence on temperature, with  $B = 8$  T and  $d = 30$  nm.

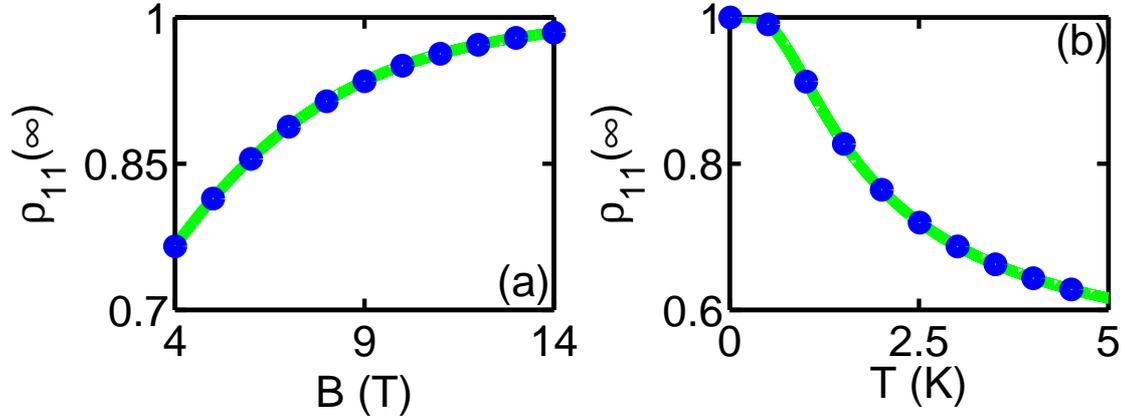


Figure 7.4: The asymptotic value of  $\rho_{11}$  as a function of  $B$  (with  $T = 1$  K) and  $T$  (with  $B = 8$  T). The solid line indicates the thermal equilibrium value, while the circles are the values calculated from the residue of the  $z = 0$  pole of the Laplace transform. In both figures,  $d = 30$  nm. Varying  $d$  changes neither the calculated asymptotic value nor the thermal equilibrium value.

The resolution of this apparent inconsistency lies in numerical round-off errors. In the previous section it was noted that the calculated values of both  $1/T_1$  and  $1/T_2$  are accurate up to three significant figures. Beyond this, errors related to finite precision arithmetic may become important. Since the calculated values for  $1/T_\Phi$  are always at least six orders of magnitude less than  $1/T_2$ , they are within this error region. As such, these values are not meaningful. All that can be concluded about the pure dephasing rate is that it is far smaller than the dephasing and dissipation rates, which is usually the case for quantum dot systems [56].

The dephasing rate due to the interaction with the gate electrode is less than  $6 \times 10^{-4} \text{ s}^{-1}$ . This is much smaller than the dephasing rate due to phonon scattering, which is of the order of  $10^3 \text{ s}^{-1}$  [96]. However, Fig. 7.3 (a) shows that the dephasing rate is proportional to  $B^3$ , while phonon dephasing rates scale as  $B^5$ ; at very low magnetic fields, dephasing due to the gate could dominate. Furthermore, the dephasing rate calculated here increases rapidly as  $d$  decreases, i.e. as the gate approaches the qubit. This will influence potential miniaturization of the device. However, for the current design of quantum dots, the gate will have very little effect.

Finally, Fig. 7.4 shows the asymptotic value of  $\rho_{11}$  compared with the thermal equilibrium value  $1/(1+e^{-\beta E_{21}})$ . It is clear from the agreement between these two quantities that the final qubit temperature is the same as that of the gate, as expected. This is a useful confirmation of the accuracy of the calculation. Note that the difference between using  $E_{21}$  (the energy difference between bare up and down states) and  $\epsilon_{21}$  (the true oscillation frequency) to evaluate the thermal equilibrium distribution is negligible. The results using  $E_{21}$  were consistently closer to the numerically determined asymptotic values, but both results were always within 0.1% of each other.

In conclusion, the gate electrodes *do* cause dephasing and dissipation of the spin in the qubit. Generally, this happens on a timescale that is much slower than other mechanisms,

at least for the experimental configuration described in Refs. [71, 113], but for smaller systems with lower magnetic fields, this dephasing mechanism could become noticeable.

## 7.3 Dephasing of charge qubits

### 7.3.1 Difference between charge and spin qubits

The analysis of dephasing of charge qubits is very similar to the analysis for spin qubits. For this reason, I discuss here only those points where the calculations differ. Further details can be found in Ref. [204]

Charge qubits have already been described in chapter 1. They consist of a pair of quantum dots separated by a tunnel barrier. Because the barrier has a finite height, the “left” and “right” states (where the electron is fully contained in one or the other quantum dot [69]) are not eigenstates. An electron initially in one of these states will move into the other and then return to the original state. By changing the height of the barrier, the rate of this oscillation can be varied.

Many experiments with charge qubits use electronically-defined quantum dots, not self-assembled quantum dots as in the spin case (see, e.g., Refs. [69, 77]). Electronically-defined quantum dots are usually larger and the network of electrodes is more complicated. A schematic diagram of a typical charge qubit is shown in Fig. 7.5(a). A 2-dimensional electron gas is formed at the interface between GaAs and AlGaAs. A double quantum dot structure [indicated by the dashed lines in Fig. 7.5(a)] is defined by charging the gate electrodes to different potentials; these are also used to influence the height of the tunnel barrier.

With spin quantum dots, the gate electrode is usually a single metallic sheet. For electronically-defined dots, as shown in Fig. 7.5(a), there are several electrodes, each of which may be at a different potential. From the point of view of the quantum dot, each electrode has two separate effects: the potential of each electrode *statically* defines the two quantum dots and each electrode exhibits a *dynamical* response to the motion of the electron. Since the potential of an electrode has a negligible effect on its Fermi level, plasma frequency and electron mean free path, each electrode will have a similar dynamical response [204]. Indeed, from the dynamical point of view, there is little difference between a system with several separate electrodes and a system with a single, large electrode; the primary differences are the increase of interaction area and the neglect of eddy currents, both of which are small effects in this system. We therefore separated the contribution from the gate electrodes into a static part, which defines the potential landscape of the quantum dots, and a dynamical part, which is taken to be equivalent to an infinite metal layer. This is shown schematically in Fig. 7.5(b) and makes the geometries for spin and charge qubits very similar [compare Fig. 7.2 (b)].

Next, consider the quantum dot system. As for the spin qubit, we assumed that the confinement in the  $z$  direction is significantly stronger than that in the  $x$  or  $y$  directions, so that the wavefunction may be written

$$\varphi_i(\mathbf{r}) = \sqrt{\delta(z)}\varphi_i(\mathbf{r}_{\parallel}). \quad (7.26)$$

In the  $x$  and  $y$  directions, we assumed that each quantum dot consists of a square-well potential with infinitely high walls except in the tunneling direction. Both wells have the

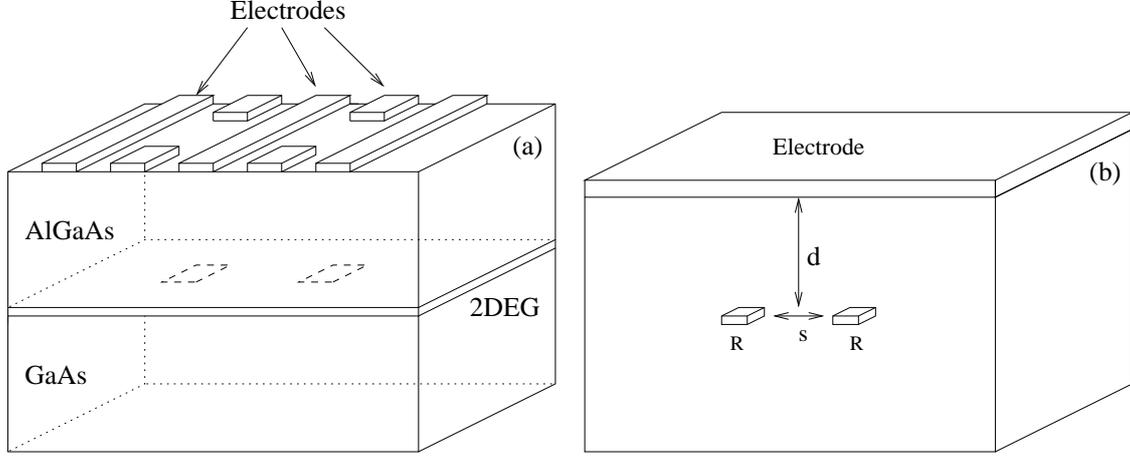


Figure 7.5: (a) A schematic drawing of a double quantum dot system. The electrodes define the quantum dots (indicated by the dashed lines), control tunneling into the source or drain electrodes (not shown) and determine the tunnel barrier. (b) An approximated system, in which the electrodes are replaced by a single, infinite electrode and the double quantum dot system is represented with a pair of square well potentials of equal height connected via a tunnel barrier. The gate electrode is assumed thick enough that it may be approximated as a bulk metal, filling the region  $z > d$ .

same depth, so the wavefunctions obey the time-independent Schrödinger equation

$$\left(-\frac{1}{2m^*}\nabla_{\parallel}^2 + V(\mathbf{r}_{\parallel})\right)\varphi_i(\mathbf{r}_{\parallel}) = E_i\varphi_i(\mathbf{r}_{\parallel}), \quad (7.27)$$

with the potential

$$V(\mathbf{r}_{\parallel}) = \begin{cases} 0 & x \in [-R - \frac{s}{2}, -\frac{s}{2}]; \quad y \in [-\frac{R}{2}, \frac{R}{2}] \\ V_0 & x \in [-\frac{s}{2}, \frac{s}{2}]; \quad y \in [-\frac{R}{2}, \frac{R}{2}] \\ 0 & x \in [\frac{s}{2}, R + \frac{s}{2}]; \quad y \in [-\frac{R}{2}, \frac{R}{2}] \\ \infty & \text{otherwise} \end{cases} \quad (7.28)$$

The height of the tunnel barrier  $V_0$  is chosen so that the tunneling energy matches that of the experiments described in Ref. [69], i.e.  $(E_2 - E_1) \equiv E_{21} \sim 10 \mu\text{eV}$ .

Similarly to the analysis for spin qubits, we considered only the two lowest energy states of Eq. (7.27); these are the bonding and antibonding wavefunctions. Using the separability of Eq. (7.27), the  $x$  component can be determined numerically by expanding  $\varphi$  in a cosine or sine series (for the bonding and antibonding orbitals, respectively) and the  $y$  component is calculated analytically. For the numerical calculation, we used 300 modes in the series expansion; the convergence of this series was confirmed by repeating the calculation using 500 modes.

As for the spin case, we assumed that the initial correlation between the qubit and the gate vanishes, the static image charge is incorporated into the definition of the potential (7.28) and that the second order Born approximation suffices. Under these conditions, the dynamics are governed by Eq. (7.9), where now the wavefunctions of the charge qubit

must be used in Eq. (7.8) to determine  $R_{ijkl}^{<, >}$ . Because the eigenfunctions of Eq. (7.27) can be chosen to be real,  $R_{ijkl}^{<, >}$  becomes symmetric with respect to some permutations of the indices  $\{i, j, k, l\}$ . This simplifies the form of the master equation, but since the integrals are determined numerically, it does not simplify the calculation. Again we find that  $R_{ijkl}$  vanishes if any 3 of its indices coincide and differ from the fourth — for charge qubits, like for spin qubits, the diagonal components of  $\rho$  evolve independently of the off-diagonal components.

To summarize: under the above assumptions, the evolution of a charge qubit is described by the same equations as the evolution of a spin qubit. The only difference lies in the calculation of  $R_{ijkl}^{<, >}$ , for which eigenfunctions of Eq. (7.27) must now be used.

### 7.3.2 Results

Many of the physical parameters for charge qubits coincide with those for spin qubits. The semiconductor surrounding the quantum dots is assumed to be GaAs/AlGaAs, for which  $\epsilon_r \approx 13$  and  $m^* \approx 0.067$ . The parameters describing the metal are taken as  $\omega_p = 10$  eV and  $\gamma = 1$  eV [100, 138]. Since electronically-defined quantum dots are usually larger than self-assembled dots, the size of the quantum dots was chosen to be  $R = 50$  nm. This time, there are four parameters that can be varied — temperature, tunneling energy, distance between the dots and distance to the gate. When each of these was varied, the others were chosen at the following default values: temperature  $T = 100$  mK; tunneling energy  $E_{21} = 8 \mu\text{eV}$ ; distance between the dots  $s = 100$  nm; distance to the gate  $d = 50$  nm. These parameters were chosen so that our results could be compared with the experimental results of Ref. [69].

The evolution was found to be qualitatively similar to that of spin qubits. The dynamics of the diagonal elements are dominated by two real, simple poles, one at  $z = 0$  and another with  $z < 0$ . After the initial transients vanish, the evolution is given by Eq. (7.22), i.e.

$$\begin{pmatrix} \rho_{11}(t) \\ \rho_{22}(t) \end{pmatrix} = \begin{pmatrix} a + (1-a)e^{-t/T_1} & a(1-e^{-t/T_1}) \\ (1-a)(1-e^{-t/T_1}) & 1-a(1-e^{-t/T_1}) \end{pmatrix} \begin{pmatrix} \rho_{11}^{(0)} \\ \rho_{22}^{(0)} \end{pmatrix} \quad (7.29)$$

(Of course, the parameters  $a$  and  $T_1$  differ from those for the spin qubit.)

The dynamics of the off-diagonal elements are governed by a complex conjugate pair of poles with negative real part. The evolution is described by Eq. (7.24), i.e.

$$\begin{pmatrix} \rho_{12}(t) \\ \rho_{21}(t) \end{pmatrix} = \begin{pmatrix} e^{-t/T_2 + i\epsilon_{21}t} & 0 \\ 0 & e^{-t/T_2 - i\epsilon_{21}t} \end{pmatrix} \begin{pmatrix} \rho_{12}^{(0)} \\ \rho_{21}^{(0)} \end{pmatrix} \quad (7.30)$$

with the dephasing given by  $T_2$ . Again, the oscillation frequency  $\epsilon_{21}$  is slightly larger than the bare tunneling energy  $E_{21}$  — the gate assists transitions between the qubit states.

The results are contained in Figs. 7.6, 7.8 and 7.9. Fig. 7.6 shows the dephasing rate as a function of  $E_{21}$ ,  $d$ ,  $s$  and  $T$ . The dephasing rate clearly increases very rapidly as  $d$  is reduced or as the tunneling energy is increased. Comparing Fig. 7.6(a) with Fig. 3(b) of Ref. [69] (reproduced in Fig. 7.7 for convenience), it is clear that the dependence of dephasing on tunneling energy is far stronger for the Coulomb interaction than for cotunneling or for the electron-phonon interaction. If this trend persists to very large

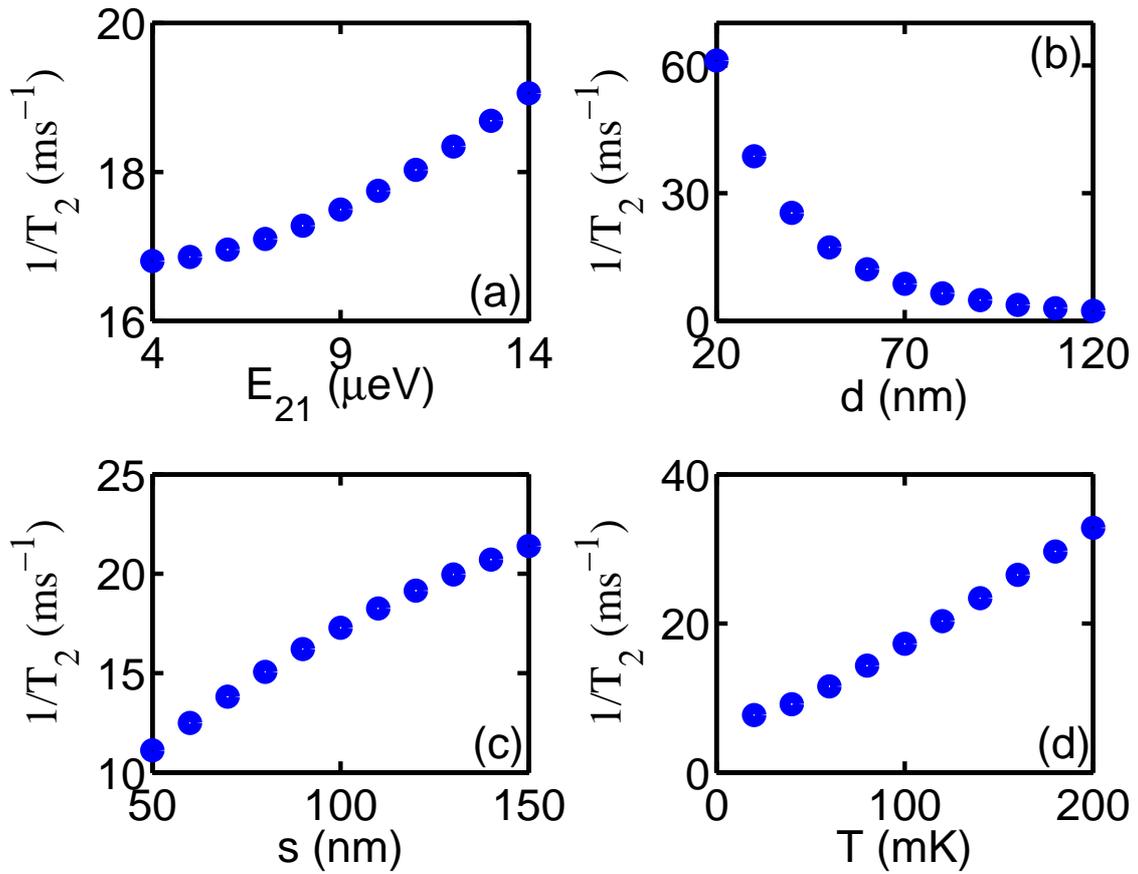


Figure 7.6: The dephasing rate of the double quantum dot system, due to the interaction with the gate electrode, as a function of tunneling energy (a), distance between the gate and the quantum dots (b), distance between the quantum dots (c) and temperature (d). In (a) we set  $d = 50 \text{ nm}$ ,  $s = 100 \text{ nm}$  and  $T = 100 \text{ mK}$ . In (b)  $E_{21} = 8 \mu\text{eV}$ ,  $s = 100 \text{ nm}$  and  $T = 100 \text{ mK}$ . In (c)  $E_{21} = 8 \mu\text{eV}$ ,  $d = 50 \text{ nm}$  and  $T = 100 \text{ mK}$ , and finally in (d) we used  $E_{21} = 8 \mu\text{eV}$ ,  $d = 50 \text{ nm}$  and  $s = 100 \text{ nm}$ .

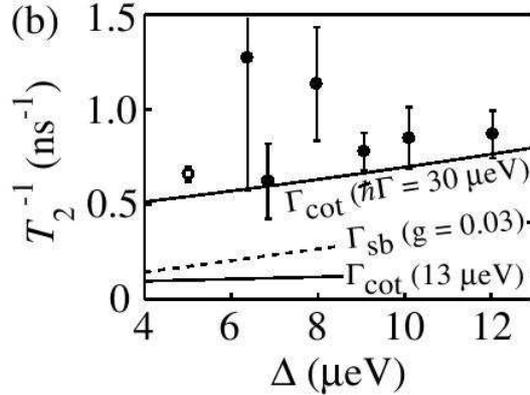


Figure 7.7: Fig. 3(b) of Ref. [69]. Experimental decoherence rate,  $T_2^{-1}$ , of an electron in a double quantum dot as a function of tunneling energy  $\Delta$ . The decoherence rates calculated from cotunneling ( $\Gamma_{\text{cot}}$ ) and the spin-boson model of phonon scattering ( $\Gamma_{\text{sb}}$ ) are shown by solid and dashed lines, respectively. For more details, see Ref. [69]. Figure reprinted with permission from T. Hayashi, T. Fujisawa, H. D. Cheong, Y. H. Jeong and Y. Hirayama, *Physical Review Letters* **91**, 226804 (2003). Copyright (2003) by the American Physical Society.

tunneling energies, there could be competition between these effects; at the tunneling energies considered here, the Coulomb interaction is negligible.

One difference between the dephasing of charge qubits and spin qubits is that the pure dephasing rate (7.25) is not negligible for charge qubits. The pure dephasing rate was found to be a little less than 1% of the dephasing rate; it is plotted in Fig. 7.8. Although small, it is larger than the numerical errors associated with double precision arithmetic, which cause inaccuracies of around 0.1% of the dephasing rate. As expected from physical grounds, the pure dephasing rate is always positive.

Finally, in Fig. 7.9 I show the comparison of the asymptotic value of  $\rho_{11}$  with the thermal equilibrium value  $1/(1 + e^{-\beta E_{21}})$ . It is clear from the agreement between these two quantities that the final qubit temperature is the same as that of the gate, i.e. the gate causes complete thermalization of the quantum dot. As for the spin qubit investigation, there is negligible difference — less than 0.1% — between using  $E_{21}$  (the bare tunneling energy) and  $\epsilon_{21}$  (the true oscillation frequency) to evaluate the thermal equilibrium distribution.

In general, the dephasing and dissipation rates for a charge qubit due to interaction with a gate electrode are of the order of  $10 \text{ ms}^{-1}$ . This is significantly slower than the typical phonon dephasing rate [69,180], which is around  $1 \text{ ns}^{-1}$  for the system parameters considered here. Gate-induced dephasing yields a negligible contribution for the current design of double quantum dots. However, the dephasing rate increases rapidly as either  $d$  is decreased or  $E_{21}$  is increased. It is possible that as systems are miniaturized or higher oscillation frequencies are sought, this form of dephasing will become increasingly relevant.

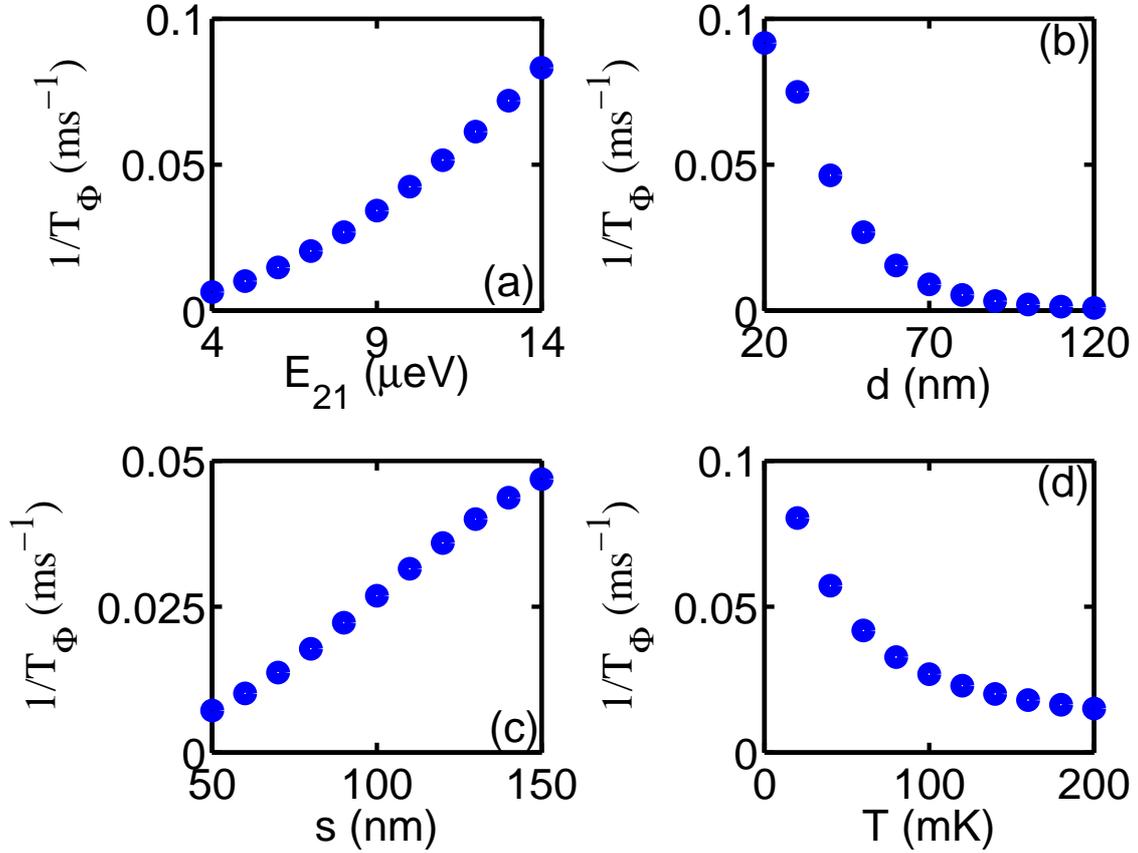


Figure 7.8: The pure dephasing rate, defined in Eq. (7.25), as a function of tunneling energy (a), distance between the gate and the quantum dots (b), distance between the quantum dots (c) and temperature (d). The other parameters are chosen as in Figs. 7.6(a)-(d), respectively.

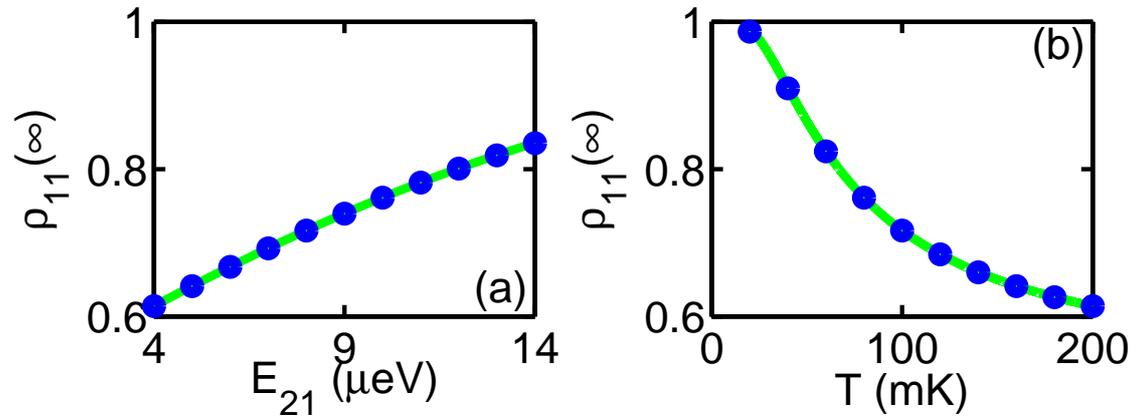


Figure 7.9: The asymptotic value of  $\rho_{11}$ , as a function of  $E_{21}$  (with  $T = 100$  mK) and  $T$  (with  $E_{21} = 8 \mu\text{eV}$ ). The solid line indicates the thermal equilibrium value, while the circles are the numerically-determined asymptotes. In both figures,  $s = 100$  nm and  $d = 50$  nm.

## Chapter 8

# Concluding Remarks

This thesis is devoted to the study of ultrafast magnetization dynamics. This very broad field includes a variety of length-scales (from macroscopic media down to individual spins), timescales (from milliseconds down to femtoseconds) and experimental techniques (optics, spin-polarized currents, magnetic fields, high-energy electron beams and many more). No thesis can address this entire range. We have investigated two topics: manipulation of magnetization using the inverse Faraday effect and relaxation of magnetization via the interaction with a metallic electrode.

Our study of the first topic was inspired by the “rediscovery” of the inverse Faraday effect (IFE) by the experimental group in Nijmegen [99]. The IFE was first derived over 40 years ago [150, 157, 187], but the idea of using it to provide nonthermal control of magnetization was only conceived in 2005 [99]. Its usefulness had gone unnoticed, so several aspects of how the IFE functions had not been clarified. The question of whether the IFE can manipulate magnetization alone or whether an additional angular momentum reservoir is required, in particular, had been unanswered.

This question has been answered in this thesis. The light beam that causes the IFE provides the angular momentum required for the magnetization to precess. No other reservoir is needed. This implies that *control* of the magnetization occurs on the timescale of a laser pulse, which can be made extremely short. Whether magnetization *reversal* can occur on this timescale is now a question of finding a material with a sufficiently strong magneto-optical response. This is a technical challenge, not a fundamental obstruction.

The light beam supplies the angular momentum in two ways. For samples that are much smaller than the beam width, most of the light is unaffected by the sample. However, some radiation will be scattered. Angular momentum is transported away from the sample in the interference between the original light and the radiation. On the other hand, if the sample is much broader than the light beam, refraction dominates the angular momentum transfer. The beam position is displaced in a direction orthogonal to its propagation, which displaces the momentum density and yields an angular momentum that balances the torque on the magnetization. This displacement occurs even when the laser beam impinges at normal incidence; it is not related to ordinary refraction.

In addition to angular momentum transfer, we also investigated the interpretation of the experimental results of Ref. [99]. The magnetization was measured using the rotation of the polarization plane of linearly-polarized light (Faraday rotation). We found that an important optical effect — birefringence — had been ignored and that the quantitative

analysis of the results was flawed. This led to several improvements to the experiments being suggested. In particular, certain critical values of the sample thickness must be avoided. Furthermore, the sample should be carefully aligned so that the plane of polarization of the light is very close to one of the crystal axes. If this is not done, the rotation will depend on many parameters that cannot be accurately determined and the results cannot be quantitatively analyzed.

The second topic that we studied was the relaxation of magnetization. Again, this is too broad a topic for a single thesis and we focused on the quantum-mechanical relaxation of a single spin within a quantum dot. This, too, was inspired by recent experiments [113] and is an important problem in the field of quantum computing. We considered the dephasing and dissipation that result from the interaction of the spin with a nearby metallic gate. In general, this was found to be negligible compared to the dephasing due to phonons, but in certain limits (small gate-dot separation, weak magnetic fields) this may be an important effect.

The methods developed to study the relaxation of a spin are also applicable to the relaxation of a charge in a double quantum dot; this is also relevant to quantum computing [69]. For this reason, some time was devoted to the interaction of a charge with a gate electrode. As for the spin, this interaction is generally much weaker than the interaction with phonons but may be important when the electrode is very close to the quantum dot and when the tunnel barrier is large.

## Appendix A

# Propagation and Polarization of Electromagnetic Waves

Propagation of monochromatic electromagnetic waves in vacuum is generally a well understood topic that is taught to most physics undergraduates. However, since these waves consist of vectors propagating through 3-dimensional space, it can be difficult to keep track of the directions in which different vectors point. This appendix is intended as a reference to simplify the visualization of these waves. I have closely followed the excellent textbooks of Landau and Lifshitz [118] and Jackson [82] in preparing this appendix; a more thorough exposition can be found in these references. Gaussian units are used in this appendix.

### A.1 Monochromatic plane waves

Electromagnetic fields in a vacuum are governed by the source-free Maxwell equations:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 0, & \nabla \times \mathbf{E} &= -\frac{1}{c}\dot{\mathbf{B}} \\ \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{B} &= \frac{1}{c}\dot{\mathbf{E}}.\end{aligned}\tag{A.1}$$

By combining the two curl equations and using the vanishing of the divergence, we find that  $\mathbf{E}$  and  $\mathbf{B}$  each satisfy the vector wave equation:

$$\frac{1}{c^2}\mathbf{E} - \nabla^2\mathbf{E} = 0, \quad \frac{1}{c^2}\mathbf{B} - \nabla^2\mathbf{B} = 0.\tag{A.2}$$

Of course, a solution of Eq. (A.2) is only physically meaningful if it also satisfies Maxwell's equations (A.1). In particular, if we solve the wave equation for the electric field  $\mathbf{E}(\mathbf{r}, t)$ , the magnetic field  $\mathbf{B}(\mathbf{r}, t)$  is uniquely determined from Eq. (A.1).

The simplest  $\mathbf{E}(\mathbf{r}, t)$  that satisfies Eq. (A.2) is the monochromatic plane wave

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \left( \mathbf{E}_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} + c.c. \right),\tag{A.3}$$

where the vector  $\mathbf{E}_0$  is complex and the complex conjugate is added to ensure the electric field  $\mathbf{E}(\mathbf{r}, t)$  is real. According to Fourier's theorem, any electric field that satisfies Eq.

(A.2) can be written as a superposition of waves of the form (A.3), which justifies the study of these simple solutions.

The wave equation (A.2) forces a relationship between the frequency  $\omega$  and the wave-vector  $\mathbf{k}$ :

$$\sqrt{\mathbf{k} \cdot \mathbf{k}} \equiv k = \frac{\omega}{c}. \quad (\text{A.4})$$

In addition, Gauss's law (the first of Maxwell's equations) yields

$$\mathbf{k} \cdot \mathbf{E}_0 = 0. \quad (\text{A.5})$$

Finally, the magnetic field  $\mathbf{B}(\mathbf{r}, t)$  associated with the electric field (A.3) is

$$\mathbf{B}(\mathbf{r}, t) = \frac{1}{2} \left( \hat{\mathbf{k}} \times \mathbf{E}_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} + c.c. \right), \quad (\text{A.6})$$

where  $\hat{\mathbf{k}} = \mathbf{k}/k$  is the unit vector in the direction of  $\mathbf{k}$ . It is useful to note that at every point  $(\mathbf{r}, t)$  in space-time, the vectors  $\mathbf{E}(\mathbf{r}, t)$ ,  $\mathbf{B}(\mathbf{r}, t)$  and  $\mathbf{k}$  are mutually orthogonal:

$$\mathbf{E}(\mathbf{r}, t) \cdot \mathbf{B}(\mathbf{r}, t) = 0, \quad \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{k} = 0, \quad \mathbf{B}(\mathbf{r}, t) \cdot \mathbf{k} = 0. \quad (\text{A.7})$$

Light is therefore a *transverse* wave.

The energy flux density associated with an electromagnetic wave is given by the *Poynting vector*

$$\mathbf{S} = \frac{c}{4\pi} \mathbf{E}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t). \quad (\text{A.8})$$

For the monochromatic plane wave described by Eqs. (A.3) and (A.6), the Poynting vector (averaged over one period of oscillation  $2\pi/\omega$ ) is

$$\mathbf{S} = \frac{c}{8\pi} |\mathbf{E}_0|^2 \hat{\mathbf{k}}. \quad (\text{A.9})$$

The energy flows in the direction of the wave-vector  $\mathbf{k}$ , which is also the direction of the group velocity of the wave:

$$\mathbf{v}_g = \frac{d\omega}{d\mathbf{k}} = c\hat{\mathbf{k}}. \quad (\text{A.10})$$

This shows the expected result that light in a vacuum propagates with the velocity  $c$ .

## A.2 Polarization of plane waves

It is clear from the above discussion that the degrees of freedom available when defining a monochromatic plane wave are  $\mathbf{k}$  and  $\mathbf{E}_0$ ; given these,  $\omega$ ,  $\mathbf{B}_0$  and  $\mathbf{S}$  are uniquely defined. Furthermore, Eq. (A.7) imposes a relationship between  $\mathbf{k}$  and  $\mathbf{E}_0$ ,  $\mathbf{k} \cdot \mathbf{E}_0 = 0$ , so that these two vectors cannot be independently chosen.

We therefore have the freedom to choose the 3 real components of  $\mathbf{k}$ , but only 2 complex components of  $\mathbf{E}_0$ . The components of  $\mathbf{E}_0$  could be chosen relative to the Cartesian

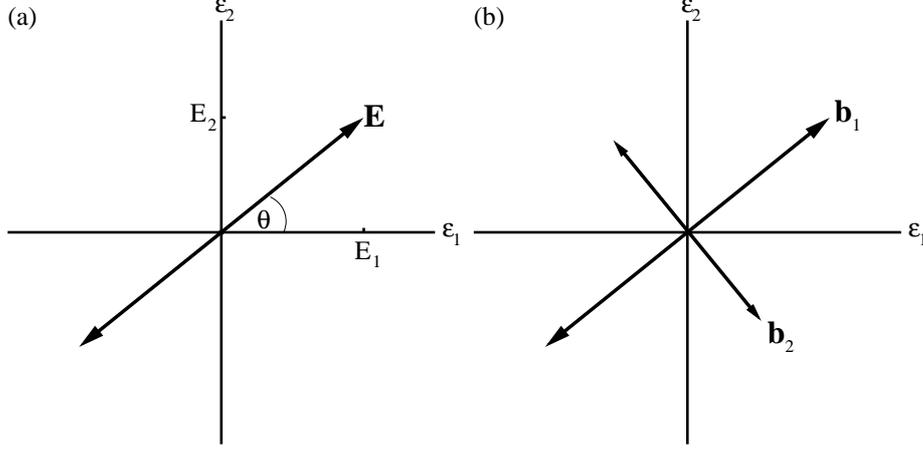


Figure A.1: The electric field of light for (a) linear and (b) elliptic polarization. The light propagates in the direction of the normal to the page. The magnetic field of the light is within the page, orthogonal to the light.

coordinates, but it is usually more convenient to choose them relative to a coordinate basis defined by  $\mathbf{k}$ . That is, introduce two real unit vectors  $\epsilon_1$  and  $\epsilon_2$  such that

$$\epsilon_1 \cdot \epsilon_2 = 0, \quad \epsilon_1 \cdot \mathbf{k} = 0, \quad \epsilon_2 \cdot \mathbf{k} = 0. \quad (\text{A.11})$$

Then  $\mathbf{E}_0$  can be expressed as

$$\mathbf{E}_0 = E_1 \epsilon_1 + E_2 \epsilon_2, \quad (\text{A.12})$$

where  $E_1$  and  $E_2$  are complex scalars. Thus the electric field (A.3) becomes

$$\mathbf{E}(\mathbf{r}, t) = \text{Re} \left[ (E_1 \epsilon_1 + E_2 \epsilon_2) e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} \right]. \quad (\text{A.13})$$

If  $E_1$  and  $E_2$  have the same phase, i.e.  $E_1 = |E_1|e^{i\alpha}$ ,  $E_2 = |E_2|e^{i\alpha}$ , then Eq. (A.13) simplifies to

$$\mathbf{E}(\mathbf{r}, t) = (|E_1| \epsilon_1 + |E_2| \epsilon_2) \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \alpha). \quad (\text{A.14})$$

At all points in space, this electric field points in the same direction, and the light is referred to as *linearly polarized*. The electric field vector of linearly polarized light is depicted in Fig. A.1 (a). This is a 2-dimensional plot, because the electric field is orthogonal to the direction of light propagation. As time progresses, or as we change position in space, the electric field changes its amplitude while maintaining the same direction.

If the phases of  $E_1$  and  $E_2$  differ, the light will be *elliptically polarized*. The physical meaning of this is easily exposed by introducing two phase angles,  $\alpha$  and  $\beta$ , such that

$$E_1 = |E_1|e^{i(\alpha - \beta)}, \quad E_2 = |E_2|e^{i(\alpha + \beta)}. \quad (\text{A.15})$$

Then the electric field (A.13) becomes

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) = & \cos \beta (|E_1| \epsilon_1 + |E_2| \epsilon_2) \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \alpha) \\ & + \sin \beta (|E_1| \epsilon_1 - |E_2| \epsilon_2) \sin(\omega t - \mathbf{k} \cdot \mathbf{r} + \alpha). \end{aligned} \quad (\text{A.16})$$

As the argument  $(\omega t - \mathbf{k} \cdot \mathbf{r} + \alpha)$  changes (either by the progression of time or by changing position), the electric field (A.16) traces out an ellipse. The axes of this ellipse can be determined from (A.16):

$$\begin{aligned}\mathbf{b}_1 &= \cos(\gamma + \beta) |E_1| \boldsymbol{\epsilon}_1 + \cos(\gamma - \beta) |E_2| \boldsymbol{\epsilon}_2, \\ \mathbf{b}_2 &= \sin(\gamma + \beta) |E_1| \boldsymbol{\epsilon}_1 + \sin(\gamma - \beta) |E_2| \boldsymbol{\epsilon}_2,\end{aligned}\tag{A.17}$$

with  $\gamma$  defined by

$$\tan 2\gamma = \frac{|E_2|^2 - |E_1|^2}{|E_2|^2 + |E_1|^2} \tan 2\beta.\tag{A.18}$$

With these definitions, the electric field (A.16) becomes

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{b}_1 \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \alpha + \gamma) + \mathbf{b}_2 \sin(\omega t - \mathbf{k} \cdot \mathbf{r} + \alpha + \gamma).\tag{A.19}$$

Elliptically polarized light is depicted in Fig. A.1 (b). Note that if the two axes  $\mathbf{b}_1$  and  $\mathbf{b}_2$  have equal lengths, we say that the light is *circularly polarized*. This occurs for  $\beta = \pi/4$  and  $|E_1| = |E_2|$ .

A useful measure of the type of polarization is the vector  $\mathbf{s} = i\mathbf{E}_0 \times \mathbf{E}_0^*/|\mathbf{E}_0|^2$ . This vector is real (due to the antisymmetry of the cross product) and points in the direction of light propagation. The notation  $\mathbf{s}$  is used, as this term is the normalized “spin angular momentum” of the light [4]. If the vectors  $\boldsymbol{\epsilon}_1$  and  $\boldsymbol{\epsilon}_2$  introduced in Eq. (A.11) are chosen such that  $\boldsymbol{\epsilon}_1 \times \boldsymbol{\epsilon}_2 = \hat{\mathbf{k}}$ , we find that

$$\begin{aligned}\mathbf{s} &\equiv i \frac{\mathbf{E}_0 \times \mathbf{E}_0^*}{|\mathbf{E}_0|^2} \\ &= \frac{2\text{Re}(iE_1 E_2^*)}{|E_1|^2 + |E_2|^2} \hat{\mathbf{k}} \\ &= \frac{2|E_1||E_2| \sin(2\beta)}{|E_1|^2 + |E_2|^2} \hat{\mathbf{k}},\end{aligned}\tag{A.20}$$

where Eq. (A.15) has been used. Linearly polarized light has  $\beta = 0$ , in which case  $|\mathbf{s}| = 0$ . Elliptically polarized light satisfies  $0 < |\mathbf{s}| < 1$ . Finally, if the light is circularly polarized,  $|\mathbf{s}| = 1$ .

# Bibliography

- [1] E. Abrahams, *Phys. Rev.* **107**, 491 (1957).
- [2] M. Aeschlimann, M. Bauer, S. Pawlik, W. Weber, R. Burgermeister, D. Oberli and H. C. Siegmann, *Phys. Rev. Lett.* **79**, 5158 (1997).
- [3] A. Aharoni, *Introduction to the Theory of Ferromagnetism*, Clarendon Press, Oxford (1996).
- [4] A. I. Akhiezer and V. B. Berestetskii, *Quantum Electrodynamics*, John Wiley and Sons, New York (1965).
- [5] A. D. Armour and M. P. Blencowe, *Phys. Rev. B* **64**, 035311 (2001).
- [6] A. S. Arrott, *J. Magn. Magn. Mater.* **258-259**, 25 (2003).
- [7] A. Asamitsu, Y. Tomioka, H. Kuwahara and Y. Tokura, *Nature* **388**, 50 (1997).
- [8] C. H. Back, R. Allenspach, W. Weber, S. S. P. Parkin, D. Weller, E. L. Garwin and H. C. Siegmann, *Science* **285**, 864 (1999).
- [9] C. H. Back, D. Weller, J. Heidmann, D. Mauri, D. Guarisco, E. L. Garwin and H. C. Siegmann, *Phys. Rev. Lett.* **81**, 3251 (1998).
- [10] M. N. Baibich, J. M. Broto, A. Fert, F. N. Vandau, F. Petroff, P. Eitenne, G. Creuzet, A. Friedrich and J. Chazelas, *Phys. Rev. Lett.* **61**, 2472 (1988).
- [11] C. T. H. Baker, *The numerical treatment of integral equations*, Clarendon Press, Oxford (1977).
- [12] S. D. Barrett and G. J. Milburn, *Phys. Rev. B* **68**, 155307 (2003).
- [13] M. Bass, P. A. Franken, J. F. Ward and G. Weinreich, *Phys. Rev. Lett.* **9**, 446 (1962).
- [14] M. Bauer, J. Fassbender, B. Hillebrands and R. L. Stamps, *Phys. Rev. B* **61**, 3410 (2000).
- [15] M. Bayer, G. Ortner, O. Stern, A. Kuther, A. A. Gorbunov, A. Forchel, P. Hawrylak, S. Fafard, K. Hinzer, T. L. Reinecke, S. N. Walck, J. P. Reithmaier, F. Klopff, and F. Schäfer, *Phys. Rev. B* **65**, 195315 (2002).

- 
- [16] E. Beaurepaire, J. C. Merle, A. Daunois and J. Y. Bigot, *Phys. Rev. Lett.* **76**, 4250 (1996).
- [17] L. Berger, *Phys. Rev. B* **54**, 9353 (1996).
- [18] G. Binasch, P. Grünberg, F. Saurenbach and W. Zinn, *Phys. Rev. B* **39**, 4828 (1989).
- [19] D. M. Bishop and S. M. Cybulski, *J. Chem. Phys.* **93**, 590 (1990).
- [20] T. Blasberg and D. Suter, *Phys. Rev. Lett.* **69**, 2507 (1992).
- [21] P. Borri, W. Langbein, U. Woggon, M. Schwab, M. Bayer, S. Fafard, Z. Wasilewski and P. Hawrylak, *Phys. Rev. Lett.* **91**, 267401 (2003).
- [22] J. Briand, V. Adrian, M. El. Tamer, A. Gomes, Y. Quemener, J. P. Dinguirard and J. C. Kieffer, *Phys. Rev. Lett.* **54**, 38 (1985).
- [23] W. F. Brown, Jr., *Ann. New York Acad. Sciences* **147**, 463 (1969).
- [24] E. Buks, R. Schuster, M. Heiblum, D. Mahalu and V. Umansky, *Nature* **391**, 871 (1998).
- [25] A. R. Calderbank and P. W. Shor, *Phys. Rev. A* **54**, 1098 (1996).
- [26] J. Chiaverini, D. Leibfried, T. Schaetz, M. D. Barrett, R. B. Blakestad, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer, R. Ozeri and D.J. Wineland, *Nature* **432**, 602 (2004).
- [27] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans and J. E. Mooij, *Science* **299**, 1869 (2003).
- [28] S. B. Choe, Y. Acremann, A. Scholl, A. Bauer, A. Doran, J. Stohr and H. A. Padmore, *Science* **304**, 420 (2004).
- [29] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung and S. Lloyd, *Nature* **393**, 143 (1998).
- [30] J. I. Cirac and P. Zoller, *Phys. Rev. Lett.* **74**, 4091 (1995).
- [31] P. Coeure and D. Challeton, *Solid State Commun.* **8**, 1345 (1970).
- [32] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Atom-photon interactions : basic processes and applications*, John Wiley and Sons, New York (1992).
- [33] S. Coriani, C. Hättig, P. Jorgensen, A. Halkier and A. Rizzo, *Chem. Phys. Lett.* **281**, 445 (1997).
- [34] J. Deschamps, M. Fitaire and M. Lagoutte, *Phys. Rev. Lett.* **25**, 1330 (1970).
- [35] D. Deutsch, *Proc. Roy. Soc. London A* **400**, 97 (1985).
- [36] D. P. DiVincenzo, *Science* **270**, 255 (1995).

- 
- [37] D. P. DiVincenzo, *Fortschr. Phys.* **48**, 771 (2000).
- [38] D. P. DiVincenzo and D. Loss, *Superlattices and Microstructures* **23**, 419 (1998).
- [39] D. P. DiVincenzo and P. W. Shor, *Phys. Rev. Lett.* **77**, 3260 (1996).
- [40] P. Drude, *Ann. der Phys.* **14**, 936 (1904).
- [41] I. E. Dzyaloshinskii, *Sov. Phys. JETP* **5**, 1259 (1957).
- [42] A. Einstein and W. J. de Haas, *Verhandl. Deut. Phys. Ges.* **17**, 152 (1915).
- [43] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
- [44] J. M. Elzerman, *Electron spin and charge in semiconductor quantum dots*, PhD thesis, Technische Universiteit Delft (2004).
- [45] H. A. Engel and D. Loss, *Phys. Rev. Lett.* **86**, 4648 (2001).
- [46] H. A. Engel and D. Loss, *Phys. Rev. B* **65**, 195321 (2002).
- [47] M. Evans and J.-P. Vigi er, *The Enigmatic Photon, Vol. 1*, Kluwer Academic Publishers, Dordrecht (1994).
- [48] M. Faraday, *Philos. Trans. R. Soc. London* **136**, 1 (1845).
- [49] M. Faraday, *Philos. Mag.* **28**, 294 (1846).
- [50] R.P. Feynman, *Int. J. Theor. Phys.* **V 21**, 467 (1982).
- [51] J. J. Finley, M. Skalitz, M. Arzberger, A. Zrenner, G. B ohm and G. Abstreiter, *Appl. Phys. Lett.* **73**, 2618 (1998).
- [52] Th. Gerrits, H. A. M. van den Berg, J. Hohlfeld, L. B ar and Th. Rasing, *Nature* **418**, 509 (2002).
- [53] N. A. Gershenfeld and I. L. Chuang, *Science* **275**, 350 (1997).
- [54] M. Getzlaff, *Fundamentals of Magnetism*, Springer, Berlin (2007).
- [55] S. Gliga, R. Hertel and C. M. Schneider, *Physica B Cond. Matt.* **403**, 334 (2008).
- [56] V. N. Golovach, A. Khaetskii and D. Loss, *Phys. Rev. Lett.* **93**, 016601 (2004).
- [57] M. D. Greenberg, *Foundations of Applied Mathematics*, Prentice Hall, New Jersey (1978).
- [58] L. K. Grover, *Phys. Rev. Lett.* **79**, 325 (1997).
- [59] J. Gruska, *Quantum Computing*, McGraw-Hill, London (1999).
- [60] D. Guarisco, R. Burgermeister, C. Stamm and F. Meier, *Applied Physics Letters* **68**, 1729 (1996).

- 
- [61] J. Gdde, U. Conrad, V. Jhnke, J. Hohlfeld and E. Matthias, *Phys. Rev. B* **59**, R6608 (1999).
- [62] A. P. Guimaraes, *From Lodestone to Supermagnets*, Wiley-VCH, Weinheim (2005).
- [63] S. Gulde, M. Riebe, G. P. T. Lancaster, C. Becher, J. Eschner, H. Hffner and F. Schmidt-Kaler, I. L. Chuang and R. Blatt, *Nature* **421**, 48 (2003).
- [64] F. Haake, *Quantum Statistics in Optics and Solid State Physics*, Volume 66 of *Springer Tracts in Modern Physics*, Springer, Berlin (1973).
- [65] M. G. Haines, *Phys. Rev. Lett.* **87**, 135005 (2001).
- [66] F. Hansteen, A. Kimel, A. Kirilyuk and T. Rasing, *Phys. Rev. B* **73**, 014421 (2006).
- [67] W. Happer and B. S. Mathur, *Phys. Rev. Lett.* **18**, 577 (1967).
- [68] U. Hartmann, *Decoherence and Measurement of Charge Qubits in Double Quantum Dots*, PhD thesis, Ludwig Maximilians Universitt, Mnchen (2005).
- [69] T. Hayashi, T. Fujisawa, H. D. Cheong, Y. H. Jeong and Y. Hirayama, *Phys. Rev. Lett.* **91**, 226804 (2003).
- [70] L. He, W. D. Doyle, L. Varga, H. Fujiwara and P. J. Flanders, *J. Magn. Magn. Mater.* **155**, 6 (1996).
- [71] D. Heiss, M. Kroutvar, J. J. Finley and G. Abstreiter, *Solid State Comm.* **135**, 591 (2005).
- [72] R. Hertel and C. M. Schneider, *Phys. Rev. Lett.* **97**, 177202 (2006).
- [73] G. Herzberg, *Atomic Spectra & Atomic Structure*, Dover Publications, New York (1944).
- [74] A. Hgele, S. Seidl, M. Kroner, K. Karrai, R. J. Warburton, B. D. Gerardot and P. M. Petroff, *Phys. Rev. Lett.* **93**, 217401 (2004).
- [75] J. Hohlfeld, Th. Gerrits, M. Bilderbeek, Th. Rasing, H. Awano and N. Ohta, *Phys. Rev. B* **65**, 012413 (2001).
- [76] J. Hohlfeld, E. Matthias, R. Knorren and K. H. Bennemann, *Phys. Rev. Lett.* **78**, 4861 (1997).
- [77] A. W. Holleitner, R. H. Blick, A.K. Httel, K. Eberl and J.P. Kotthaus, *Science* **297**, 70 (2002).
- [78] L. C. L. Hollenberg, *Phys. Rev. E* **62**, 7532 (2000).
- [79] W. Hbner and G. P. Zhang, *Phys. Rev. B* **58**, R5920 (1998).
- [80] S. Iwasaki and Y. Nakamura, *IEEE Trans. Magn.* **13**, 1272 (1977).

- 
- [81] L. Jacak, J. Krasnyj, W. Jacak, R. Gonczarek and P. Machnikowski, *Phys. Rev. B* **72**, 245309 (2005).
- [82] J. D. Jackson, *Classical Electrodynamics*, John Wiley, New York (1962).
- [83] M. Jaszunski, P. Jorgensen, A. Rizzo, K. Ruud and T. Helgaker, *Chem. Phys. Lett.* **222**, 263 (1994).
- [84] F. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGraw-Hill, New York (1957).
- [85] F. Jorgensen, *J. Mag. Mag. Mat.* **193**, 1 (1999).
- [86] L. P. Kadanoff and G. Baym, *Quantum statistical mechanics*, W. A. Benjamin Inc., New York (1962).
- [87] F. J. Kahn, P. S. Pershan and J. P. Remeika, *Phys. Rev.* **186**, 891 (1969).
- [88] S. Kaka and S. E. Russek, *Appl. Phys. Lett.* **80**, 2958 (2002).
- [89] A. Kaminski, Yu. V. Nazarov and L. I. Glazman, *Phys. Rev. Lett.* **83**, 384 (1999).
- [90] A. Kaminski, Yu. V. Nazarov and L. I. Glazman, *Phys. Rev. B* **62**, 8154 (2000).
- [91] B. E. Kane, *Nature* **393**, 133 (1998).
- [92] E. O. Kane, *J. Phys. Chem. Solids* **1**, 249 (1957).
- [93] K. Karrai, R. J. Warburton, A. Högele, B. Urbaszek, C. Schulhauser, E. J. McGhee, A. O. Govorov, J. M. Garcia, B. D. Gerardot and P. M. Petroff, *Nature* **427**, 135 (2004).
- [94] Y. Kato, R. C. Myers, A. C. Gossard and D. D. Awschalom, *Nature* **427**, 50 (2004).
- [95] A. V. Khaetskii and Y. V. Nazarov, *Phys. Rev. B* **61**, 12639 (2000).
- [96] A. V. Khaetskii and Y. V. Nazarov, *Phys. Rev. B* **64**, 125316 (2001).
- [97] D. Kielpinski, C. R. Monroe and D. J. Wineland, *Nature* **417**, 709 (2002).
- [98] A. V. Kimel, A. Kirilyuk and Th. Rasing, *Laser & Photonics Rev.* **1**, 275 (2007).
- [99] A. V. Kimel, A. Kirilyuk, P. A. Usachev, R. V. Pisarev, A. M. Balbashov and Th. Rasing, *Nature* **435**, 655 (2005).
- [100] C. Kittel, *Introduction to Solid State Physics*, John Wiley and Sons, New York (1968).
- [101] B. Koopmans, J. J. M. Ruigrok, F. Dalla Longa and W. J. M. de Jonge, *Phys. Rev. Lett.* **95**, 267207 (2005).
- [102] B. Koopmans, M. van Kampen and W. J. M. de Jonge, *J. Phys. Cond. Mat.* **15**, S723 (2003).

- 
- [103] B. Koopmans, M. van Kampen, J. T. Kohlhepp and W. J. M. de Jonge, *Phys. Rev. Lett.* **85**, 844 (2000).
- [104] Y. M. Koroteev, G. Bihlmayer, J. E. Gayone, E. V. Chulkov, S. Blügel, P. M. Echenique and P. Hofmann, *Phys. Rev. Lett.* **93**, 046403 (2004).
- [105] N. Koshizuka and K. Hayashi, *J. Phys. Soc. Japan* **57**, 4418 (1988).
- [106] I. Y. Kostyukov, G. Shvets, N. J. Fisch and J. M. Rax, *Phys. Plasm.* **9**, 636 (2002).
- [107] A. P. Kovalenko and A. A. Kolokolov, *Radiophysics and Quantum Electronics* **25**, 283 (1982).
- [108] B. B. Krichevtsov, K. M. Mukimov, R. V. Pisarev and M. M. Ruvinshtein, *JETP Lett.* **34**, 379 (1981).
- [109] B. B. Krichevtsov, R. V. Pisarev and M. M. Ruvinshtein, *Sov. Phys. Solid State* **22**, 1240 (1981).
- [110] H. Kronmüller and S. Parkin, *Handbook of Magnetism and Advanced Magnetic Materials 1: Fundamentals and Theory*, John Wiley and Sons, New York (2007).
- [111] H. Kronmüller and S. Parkin, *Handbook of Magnetism and Advanced Magnetic Materials 2: Micromagnetism*, John Wiley and Sons, New York (2007).
- [112] M. Kroutvar, Y. Ducommun, J. J. Finley, M. Bichler, G. Abstreiter and A. Zrenner, *Appl. Phys. Lett.* **83**, 443 (2003).
- [113] M. Kroutvar, Y. Ducommun, D. Heiss, M. Bichler, D. Schuh, G. Abstreiter and J. J. Finley, *Nature* **432**, 81 (2004).
- [114] M. Krykunov, A. Banerjee, T. Ziegler and H. Autschbach, *J. Chem. Phys.* **122**, 074105 (2005).
- [115] M. Kuwata, *J. Lumin.* **38**, 247 (1987).
- [116] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press, Oxford (1960).
- [117] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Butterworth-Heinemann, Oxford (1977).
- [118] L. D. Landau and E. M. Lifshitz, *The classical theory of fields*, Butterworth-Heinemann, Oxford (1985).
- [119] M. le Bellac, *Quantum and Statistical Field Theory*, Clarendon Press, Oxford (1991).
- [120] M. Ledermann, S. Schultz and M. Ozaki, *Phys. Rev. Lett.* **73**, 1986 (1994).
- [121] A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg and W. Zwerger, *Rev. Mod. Phys.* **59**, 1 (1987).

- 
- [122] D. Leonard, S. Fafard, K. Pond, Y. H. Zhang, J. L. Merz and P. M. Petroff, *J. Vac. Sci. Tech. B* **12**, 2516 (1994).
- [123] Y. Levinson, *Europhys. Lett.* **39**, 299 (1997).
- [124] P. F. Liao and J. E. Bjorkholm, *Phys. Rev. Lett.* **34**, 1 (1975).
- [125] A. Liebsch, *Electronic excitations at metal surfaces*, Plenum Press, New York (1997).
- [126] F. Dalla Longa, J. T. Kohlhepp, W. J. M. de Jonge and B. Koopmans, *Phys. Rev. B* **75**, 224431 (2007).
- [127] D. Loss and D. P. DiVincenzo, *Phys. Rev. A* **57**, 120 (1998).
- [128] D. Loss and H. Schoeller, *J. Stat. Phys.* **56**, 175 (1989).
- [129] G. D. Mahan, *Many particle physics*, Plenum Press, New York (1990).
- [130] G. D. Mahan and K. R. Subbaswamy, *Local density theory of polarizability*, Plenum Press, New York (1990).
- [131] J. C. Mallinson, *The Foundations of Magnetic Recording*, Academic Press, San Diego (1987).
- [132] N. D. Mermin, *Quantum Computer Science*, Cambridge University Press, Cambridge (2007).
- [133] G. Moe and W. Happer, *J. Phys. B Atom. Molec. Phys.* **10**, 1191 (1977).
- [134] E. W. Montroll, *Fundamental Problems in Statistical Mechanics*, North Holland Amsterdam 1962.
- [135] J. E. Mooij, T. P. Orlando, L. Levitov, L. Tian, C. H. van der Wal and S. Lloyd, *Science* **285**, 1036 (1999).
- [136] T. Moriya, *Phys. Rev.* **120**, 91 (1960).
- [137] S. Muto, *Jpn. J. Appl. Phys.* **34**, L210 (1995).
- [138] H. P. Myers, *Introductory Solid State Physics*, Taylor & Francis, London (1990).
- [139] S. Nakajima, *Progr. Theor. Phys.* **20**, 948 (1958).
- [140] Y. Nakamura, Yu. A. Pashkin and J. S. Tsai, *Nature* **398**, 786 (1999).
- [141] L. Néel, *Comptes Rendus* **224**, 1550 (1947).
- [142] M. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000).
- [143] H. Ohno, *Science* **291**, 840 (2001).
- [144] H. Ohno, D. Chiba, F. Matsukura, T. Omiya, E. Abe, T. Dietl, Y. Ohno and K. Ohtani, *Nature* **408**, 944 (2000).

- 
- [145] Y. Ohno, D. K. Young, B. Beschoten, F. Matsukura, H. Ohno and D. D. Awschalom, *Nature* **402**, 790 (1999).
- [146] P. M. Oppeneer and A. Liebsch, *J. Phys. Cond. Matt.* **16**, 5519 (2004).
- [147] W. Pauli, *Festschrift zum 60. Geburtstag A. Sommerfeld*, Hirzel, Leipzig (1928).
- [148] C. A. Perroni and A. Liebsch, *Phys. Rev. B* **74**, 134430 (2006).
- [149] C. A. Perroni and A. Liebsch, *J. Phys. Cond. Matt.* **18**, 7063 (2006).
- [150] P. S. Pershan, *Phys. Rev.* **130**, 919 (1963).
- [151] P. S. Pershan, *J. Appl. Phys.* **38**, 1482 (1967).
- [152] P. S. Pershan, J. P. van der Ziel and L. D. Malmstrom, *Phys. Rev.* **143**, 574 (1966).
- [153] B. N. J. Persson, *Sliding friction : physical principles and applications*, Springer, Berlin (2000).
- [154] J. R. Petta, A. C. Johnson, C. M. Marcus, M. P. Hanson and A. C. Gossard, *Phys. Rev. Lett.* **93**, 186802 (2004).
- [155] D. Pines, J. Bardeen and C. P. Slichter, *Phys. Rev.* **106**, 489 (1957).
- [156] R. V. Pisarev, *Sov. Phys. JETP* **31**, 761 (1970).
- [157] L. P. Pitaevskii, *Sov. Phys. JETP* **12**, 1008 (1961).
- [158] C. P. Poole, Jr., *The Physics Handbook*, John Wiley and Sons, New York (1998).
- [159] S. V. Popov, N. I. Zheludev and Yu. P. Svirko, *Opt. Lett.* **19**, 13 (1994).
- [160] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge (1992).
- [161] I. Prigogine and P. Resibois, *Physica* **27**, 629 (1961).
- [162] R. Rahm, J. Stahl and D. Weiss, *Appl. Phys. Lett.* **87**, 182107 (2005).
- [163] M. Y. A. Raja, D. Allen and W. Sisk, *Appl. Phys. Lett.* **67**, 2123 (1995).
- [164] M. Y. A. Raja, W. Sisk, M. Yousaf and D. Allen, *Appl. Phys. B* **64**, 79 (1997).
- [165] P. Recher, E. V. Sukhorukov and D. Loss, *Phys. Rev. Lett.* **85**, 1962 (2000).
- [166] L. Reining, V. Olevano, A. Rubio and G. Onida, *Phys. Rev. Lett.* **88**, 066404 (2002).
- [167] R. L. Rivest, A. Shamir and L. M. Adleman, *CACM* **21**, 120 (1978).
- [168] C. Runge, *Z. Math. Phys.* **46**, 224 (1901).
- [169] R. Schlessler and A. Weis, *Opt. Lett.* **17**, 1015 (1992).

- 
- [170] F. Schmidt-Kaler, H. Häffner, M. Riebe, S. Gulde, G. P. T. Lancaster, T. Deuschle, C. Becher, C. F. Roos, J. Eschner and R. Blatt, *Nature* **422**, 408 (2003).
- [171] A. Scholl, L. Baumgarten, R. Jacquemin and W. Eberhardt, *Phys. Rev. Lett.* **79**, 5146 (1997).
- [172] 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).
- [173] R. Schützhold, *Phys. Rev. A* **67**, 062311 (2003).
- [174] P. W. Shor, *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, IEEE Computer Society, Los Alamitos (1994).
- [175] A. Silva and S. Levit, *Phys. Rev. B* **63**, 201309(R) (2001).
- [176] J. C. Slonczewski, *J. Mag. Mag. Mat.* **159**, L1 (1996).
- [177] D. Sprinzak, E. Buks, M. Heiblum and H. Shtrikman, *Phys. Rev. Lett.* **84**, 5820 (2000).
- [178] C. Stamm, T. Kachel, N. Pontius, R. Mitzner, T. Quast, K. Holldack, S. Khan, C. Lupulescu, E. F. Aziz, M. Wietstruk, H. A. Durr and W. Eberhardt, *Nature Mater.* **6**, 740 (2007).
- [179] C. D. Stanciu, F. Hansteen, A. V. Kimel, A. Kirilyuk, A. Tsukamoto, A. Itoh and Th. Rasing, *Phys. Rev. Lett.* **99**, 047601 (2007).
- [180] V. N. Stavrou and Xuedong Hu, *Phys. Rev. B* **72**, 075362 (2005).
- [181] E. C. Stoner and E. P. Wohlfarth, *Phil. Trans. Roy. Soc. London A* **240**, 599 (1948).
- [182] W. J. Tabor, A. W. Anderson and L. G. van Uitert, *J. Appl. Phys.* **41**, 3018 (1970).
- [183] W. J. Tabor and F. S. Chen, *J. Appl. Phys.* **40**, 2760 (1969).
- [184] S. Tehrani, B. Engel, J. Slaughter, E. Chen, M. DeHerrera, M. Durlam, P. Naji, R. Whig, J. Janesky and J. Calder, *IEEE Trans. Magn.* **36**, 2752 (2000).
- [185] T. D. Truitt and A. E. Rogers, *Basics of Analog Computers*, John F. Rider Publisher Inc., New York (1960).
- [186] I. Tudosa, C. Stamm, A. B. Kashuba, F. King, H. C. Siegmann, J. Stohr, G. Ju, B. Lu and D. Weller, *Nature* **428**, 831 (2004).
- [187] J. P. van der Ziel, P. S. Pershan and L. D. Malmstrom, *Phys. Rev. Lett.* **15**, 190 (1965).
- [188] L. van Hove, *Physica* **23**, 441 (1957).
- [189] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood and I. L. Chuang, *Nature* **414**, 883 (2001).

- 
- [190] A. Vaterlaus, T. Beutler and F. Meier, *Phys. Rev. Lett.* **67**, 3314 (1991).
- [191] E. M. Verdet, *Ann. Chim. Phys.* **41**, 370 (1854).
- [192] J. von Neumann, *Gött. Nach.* **1**, 173 (1927).
- [193] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin (1932).
- [194] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton (1971).
- [195] R. M. Wald, *General Relativity*, University of Chicago Press, Chicago (1984).
- [196] W. S. Warren, N. A. Gershenfeld and I. L. Chuang, *Science* **277**, 1688 (1997).
- [197] J. K. Watson, *Applications of Magnetism*, John Wiley and Sons, New York (1980).
- [198] U. Weiss, *Quantum Dissipative Systems*, World Scientific, Singapore (1993).
- [199] W. Wernsdorfer, E. Bonet Orozco, K. Hasselbach, A. Benoit, B. Barbara, N. Demony, A. Loiseau, H. Pascard and D. Mailly, *Phys. Rev. Lett.* **78**, 1791 (1997).
- [200] R. M. White, R. J. Nemanich and C. Herring, *Phys. Rev. B* **25**, 1822 (1982).
- [201] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer, New York (2003).
- [202] S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnar, M. L. Roukes, A. Y. Chtchelkanova and D. M. Treger, *Science* **294**, 1488 (2001).
- [203] S. R. Woodford, A. Bringer and S. Blügel, *J. Appl. Phys.* **101**, 053912 (2007).
- [204] S. R. Woodford, A. Bringer and K. M. Indlekofer, *Phys. Rev. B* **76**, 064306 (2007).
- [205] L. M. Woods, T. L. Reinecke and Y. Lyanda-Geller, *Phys. Rev. B* **66**, 161318(R) (2002).
- [206] J. A. Wunderlich and L. G. Deshazer, *Appl. Opt.* **16**, 1584 (1977).
- [207] A. V. Zenkov, B. B. Krichevtsov, A. S. Moskvina, K. M. Mukimov, R. V. Pisarev and M. M. Ruvinshtein, *Sov. Phys. JETP* **69**, 792 (1989).
- [208] G. P. Zhang and W. Hübner, *Appl. Phys. B* **68**, 495 (1999).
- [209] G. P. Zhang and W. Hübner, *Phys. Rev. Lett.* **85**, 3025 (2000).
- [210] A. K. Zvezdin and V. A. Kotov, *Modern magneto-optics and magneto-optical material*, IOP, Bristol (1997).
- [211] R. Zwanzig, *J. Chem. Phys.* **33**, 1338 (1960).

# List of Publications

The results presented in this thesis have been published in the following papers:

1. *Interpreting magnetization from Faraday rotation in birefringent, magnetic media*  
S. R. Woodford, A. Bringer and S. Blügel  
J. Appl. Phys. **101**, 053912 (2007)
2. *Dephasing of a quantum dot due to the Coulomb interaction with a gate electrode*  
S. R. Woodford, A. Bringer and K. M. Indlekofer  
Phys. Rev. B **76**, 064306 (2007)
3. *Conservation of angular momentum and the inverse Faraday effect*  
S. R. Woodford  
→ submitted to Phys. Rev. B

# Acknowledgments

*Unless the LORD builds the house,  
its builders labour in vain.*  
Psalm 127:1, New International Version

Many people have been extremely helpful throughout my time in Germany and the writing of this thesis.

- Thanks to my wife, Mandy, who has been very patient and encouraging during the writing of this thesis. Thanks for agreeing to come to Germany with me — I don't know what I would have done here without you.
- Thanks to my parents who have kept me full of nougat and amusing anecdotes about the French countryside.
- Ich danke meinem Betreuer, Andreas Bringer. Du hast immer Zeit, Du weisst fast alles über Physik und Du hilfst gerne. Ohne Dich wäre meine Zeit in Jülich nicht halb so produktiv.
- Thanks to my brother and sister. My sister is always extremely well organized and was great in helping take the sting out of leaving South Africa. My brother is not so well organized — he once applied for a job using a CV that was half his and half the CV template he was using — but he's always entertaining. Thanks for keeping me laughing.
- Jono and Taryn have been very encouraging whenever we return to SA, as have Tony and Sharon.
- My friends from the South have been great — Gavin, Dave, Natalie, Andrew, Kathy, Jan, Chris, Rhiannon, Ainsley, Tarin, Igor, Peter, Neil. I'm tempted to put an *etc.* here, but I'm not sure whether it is more rude to omit someone or to make someone an "other".
- The Theory 1 coffee round has been great. You have all been very helpful, both in welcoming me to Germany and dispensing scientific advice. I thank the "official" members Andreas, Ansgar, Bo, Bob and John, as well as those latecomers like Ugo, Antonio and Chun-Yan, who were suffered to join.
- Thanks to Bob for correcting my language and improving the text of this thesis in many places.

- The Freie Evangelische Gemeinde Jülich has been fantastic. You welcomed us in, provided us with a translator (thank you Benjamin!) and have provided us with great Christian support in our time here. The quality of the teaching astonished both Mandy and me. You will be missed.
- Thanks to the Fachhochschule Gebetskreis — Benjamin, Frank, Mary and Rosemary. Thanks also to the Englische Hauskreis Jülich — Marcelo and Ariadne, Vitoria, Anna, Pavel, Gunther and Beate, Wendi and Eric. You've all been very encouraging.
- Thanks to my running partner/office mate, Marcus. You're always willing to help, and you've taught me a lot of great things about Germans.
- Ich danke auch dem Direktor der Theorie 1, Stefan Blügel. Du bist sehr ermutigend, aber trotzdem kritisch genug dass man immer Fortschritten macht.
- Thanks to Prof. Dederichs, who agreed to be the second referee of this thesis.
- Ich danke Ute Winkler, die mich immer geholfen hat. Du weisst immer was zu tun ist, wie man alles tun, und vermutlich alles über alles. Theorie 1 hat wirklich ein Schatz in Dir gefunden.
- Thanks to Klaus-Michael Indlekofer for suggesting the work on quantum dot dephasing. It was good working with you.
- Thanks to Riccardo, who has been great in giving advice about the inverse Faraday effect. It's always great talking to you and I am always impressed by the work of your group.
- Letztendlich danke ich all die Studenten und Mitarbeiter in die IFF die ich getroffen habe. Ich versuche jetzt alle zu nennen, aber wenn ich einer vergisst, tut es mir Leid: Ali, Andreas×2, Clemens, Klaus, Marcus, Markus, Arno, Gustav, Mannie, Eric, Eva, Martin, Christoph, Ersoy, Swantje, Daniel, Frank, Phivos, Samir, Nicholae... vielleicht sage ich nur "und so weiter" und schliesse ab.

# Curriculum Vitae

September 15, 1980	Born in Cape Town, South Africa
1987 - 1998	Primary and high school in Johannesburg, South Africa
December 1998	Matriculation
1999	Undergraduate studies in physics, mathematics and chemistry at University of Pretoria, South Africa
2000 - 2001	Undergraduate studies in physics, mathematics and applied mathematics at University of Cape Town, South Africa
2002	Honours course in theoretical and mathematical physics at University of Cape Town, South Africa
2003 - 2004	M.Sc. thesis “Parametrically driven dark solitons” at Department of Mathematics and Applied Mathematics, University of Cape Town, South Africa
January-June 2005	Lectured the first semester honours course “Advanced Methods of Mathematical Physics” at University of Cape Town, South Africa
August 2005 - June 2008	Ph.D. thesis at Institut für Festkörperforschung, Forschungszentrum Jülich GmbH, Jülich, Deutschland
30 June 2008	Ph.D. exam at RWTH Aachen